Cognition, Technology & Work (1999) 1:86–96

Cognition
Technology &
Work

To Intervene or not to Intervene: The Dilemma of Management by Exception

S. W. A. Dekker[1] and D. D. Woods[2]

[1]Centre for Human Factors in Aviation, Linköping Institute of Technology, Linköping, Sweden; [2]Cognitive Systems Engineering Laboratory, Department of Industrial and Systems Engineering, Ohio State University, Columbus, OH, USA

Abstract: Future air traffic management architectures propose to give aircraft more flight path autonomy and turn the air traffic controller into a manager of exceptions. This article reports on one experiment in a series of studies that empirically explored the cognitive work underlying management by exception in air traffic control. Active practitioners (controllers, pilots, dispatchers) were prepared on the rules of the envisioned system and presented with a series of future incidents, each of which they were required to jointly resolve. Management by exception turns out to trap human controllers in a double bind, where intervening early seems appealing but is difficult to justify (airspace throughput) and carry out (controller workload problems). Late interventions are just as difficult, since controllers will have to take over in the middle of a potentially challenging or deteriorating situation. Computerised decision support that flags exceptions migrates the decision criterion into a device, creating a threshold crossing that is typically set either too early or too late. This article lays out the intertwined trade-offs and dilemmas for the exception manager, and makes recommendations for cooperative human–machine architectures in future air traffic management.

Keywords: Air traffic control; Cooperative architectures; Future incidents; Management by exception; Protocol analysis; Supervisory control

1. INTRODUCTION

1.1. Traffic Pressures in Air Traffic Control

Air traffic control (ATC) organisations throughout the world are facing the challenge to absorb ever greater throughput pressures (National Research Council 1998). One key to accommodating growing air traffic is flexible routings, where aircraft get to 'pick their own course, altitude and speed' (McClellan 1998). The basic idea of flexible routes is to disperse air traffic and offload so-called 'airways' that connect city pairs and other waypoints. With electronic alert zones around their aircraft, pilots could themselves have more responsibility for maintaining separation with other air traffic (Baiada 1995; Cotton 1995; Gibbs 1995). Developments in satellite navigation (GPS), automatic dependent surveillance (ADS-B), digital communication (datalink) and collision avoidance technologies on the ground and in the air are making flexible routes ever more feasible (Dornheim 1995). Flexible routes are known under different labels in different countries. In the USA, the concept of 'free flight' gained momentum from the mid-1990s onwards (RTCA 1995; Gibbs 1995) and has now evolved in part into 'Safe Flight 21' (Nordwall 1999). In Europe meanwhile, the harmonisation of different ATC systems across nations has opened opportunities for flexible routings as well (Cooper 1994; Van Ghent 1994).

Different groups have begun to explore the human factors implications of flexible air traffic routings. For example, Smith and colleagues reported on the development of a measure of dynamic air traffic density. ATC can use this measure to gauge the need to intervene for reasons of aircraft separation or other airspace management activities (Smith et al 1998). Dynamic density is seen as one of the key components of managing aircraft separation in a more flexible ATC architecture (RTCA 1995).

1.2. Management by Exception in ATC

Another important premise of flexible routings – and a prerequisite for increasing airspace capacity – is that controllers step away from controlling every aircraft individually and instead become traffic managers, responsible for resolving exceptional situations. 'To permit unrestricted ATC growth, we should first determine how to eliminate one-to-one coupling between a proactive sector controller and every aircraft in flight. The basic requirement is to minimise human control involvement in routine events, freeing controllers to concentrate on the key areas where human skills have most to offer: traffic management, system safety assurance, and dealing with the exceptional occurrence' (Whicher, as cited by Cooper 1994). Indeed, the idea of controllers as supervisors at a larger distance has become entrenched as a stereotypical industry solution to traffic growth problems. An example from one aviation magazine states: 'Controllers – renamed traffic managers – would stand by to intervene to resolve conflicts' (McClellan 1998). In future ATC architectures, aircraft should largely be left to manage their own routings and separation. Controller intervention will be necessary only when potential manoeuvres might interfere with other aircraft, when traffic density precludes route flexibility, or when flight restrictions are considered necessary for safety (RTCA 1995; Nordwall 1995). In other words, the exception manager will have to resolve conflicts when aircraft themselves are unable to, and take over control when airspace gets too busy or when other critical parameters are exceeded.

Technological and procedural innovations are often introduced for their putative quantifiable benefits (e.g., less workload, fewer human errors, higher precision – see Woods 1996) and future ATC architectures are no exception. With flexible routes and pilot responsibility for separation, there will be *less* workload for controllers in terms of routine tasks (they will only manage exceptional situations and ATC facilities might even be able to do with *fewer* controllers), there will be *larger* traffic throughput, and *higher* accuracy in navigation (GPS) and flight path monitoring (ADS-B) (Cotton 1995; RTCA 1995; Baiada 1995; RMB 1996; Scardina et al 1996; Leslie 1996).

Quantifiable benefits are often offset by qualitative side effects of the introduction of technological innovations. Previous research has documented the profound changes that occur in human roles and human work when humans become more distant supervisors of their monitored process (Hollnagel and Woods 1983; Hollnagel 1992; Stix 1991; Hughes 1992; Sarter, Woods and Billings 1997). Indeed, experiences from other application worlds suggest that different patterns of function allocation have wide reverberations for the entire human–system ensemble and how people co-ordinate activities (Hollnagel and Woods 1983; Dekker and Wright 1997). Turning the human into a higher-level supervisor often does not reduce human task demands, but changes them in nature. The kind of workload and its distribution over time may change too (Wiener and Curry 1980; Stix 1991). In addition, some forms of human error and vulnerability may disappear, only to open the doors for new and unanticipated types of human–system breakdowns (Wiener 1989; Woods et al 1994).

1.3. The Scientific Basis of Management by Exception

The general idea of management by exception can be traced back more than a century. Authors such as Towne (1886) and Gilbreth (1916) in part discussed management by exception as a supervisory strategy – specifically because it could counterbalance the contemporary 'control-minded management' (Bittel 1964). Trying to get managers away from the details of every aspect of the work they supervised, Taylor was explicit in recommending that in management by exception.

> The manager should receive only condensed, summarized, and invariably comparative reports covering, however, all of the elements entering into management, and even these summaries should all be carefully gone over by an assistant before they reach the manager, and have all of the exceptions to the past averages or standards pointed out, both the especially good and the especially bad exceptions, thus giving him in a few minutes a full view of the progress which is being made, or the reverse, and leaving him free to consider the broader lines of policy. (Taylor 1911, p. 86)

With increasing industrial automation and the fragmentation of supervised tasks during the twentieth century, advice of this kind was generally heeded. Management by exception became reputed to be successful in monitoring processes as wide-ranging as quality control, employee morale and financial ratios (Bittel 1964; Mackintosh 1978).

The idea of management by exception in human–machine systems was born out of the supervisory control paradigm in the mid-1970s (Edwards and Lees 1974; Sheridan 1976; Umbers 1979). The essential features of management by exception remained: a greater supervisory distance that was enabled by the interposition of other agents, that is, automated subsystems, and receipt of partially processed or summarised data about the monitored system. In the slightly later words of Wiener (1988, p. 456): 'The exception principle states that as long as things are going well or according to plan, leave the managers alone. Don't clutter their world with reports, warnings, and messages of normal conditions.' Taylor's idea of an assistant going over all the data before they would reach the exception manager conveniently translated into computer support given the technologically now available (the Human Interactive System, or HIS, in supervisory control language. Sheridan 1976). Once again, Wiener: 'lower-level managers or computers flag exceptions, which are routed to the manager' (Wiener 1988, p. 456). The goal was to simplify the supervisor's cognitive work in situation assessment and monitoring (Wiener, 1988) – themes that hark back to the roots of management by exception as a 'simplification and systemization' of the managerial job (Bittel 1964). The idea that an exception manager would intervene in the supervised process only when the reports reaching him or her demanded so, has remained in modern treatments as well.

There are a number of ways in which intervention in an ongoing process can take place in management by exception. To describe the intervention options, supervisory control theorists adopted a continuum of increasing human involvement with the details of the process

(Sheridan 1992; Stein 1992). In other words, the human operator can intervene deeply and take much control away from subordinates, or intervene less deeply and leave subordinates relatively free in their control of the monitored process. Thus, a series of levels of supervisory control is created, whereby human and subordinate control over the process becomes symmetrically apportioned as viewed from top to bottom (Fig. 1).

The list of levels indicates the varying degrees of possible supervisor involvement and alludes to the nature of the human task at each of the levels. What is problematic is that neither the list nor much of the accompanying supervisory control literature represents the cognitive work that might be involved in deciding how and when to intervene or how to switch from level to level (instructive here is the work on adaptive automation, which is in a sense the mirror image of management by exception: here too defining the basis for switching levels of automation support to the human remains difficult – see, for example, Parasuraman et al 1992). The list of supervisory control levels leaves unspecified how the human should decide when and whether to intervene or when to back off. Note that in management by exception the answer to the question about when and how to intervene depends upon the answer to another question: What is an exception?

Levels of Supervisory Control

The subordinate:

1. offers no assistance: human supervisor must do it all;
2. offers a complete set of action alternatives, and
3. narrows the selection down to a few, or
4. suggests one, or
5. executes that suggestion if the supervisor approves, or
6. allows the supervisor a restricted time to veto before automatic execution, or
7. executes automatically, then necessarily informs the supervisor, or
8. informs him after execution only if he asks, or
9. informs him after execution if the subordinate decides to
10. decides everything and acts autonomously, ignoring the supervisor.

Fig. 1. A list of levels of supervisory control (after Sheridan, 1976, 1992).

According to Sheridan (1987, p. 1249), human monitoring and intervention involve activities such as observing displays, looking for signals of abnormal behaviour, making minor adjustments of system parameters when necessary, and deciding when continuation of automatic control would cease to be satisfactory. These activities are all relevant of course, but their description remains vague. What is 'satisfactory', for example? The typical description of supervisory work assumes that the kinds of deviations supervisors look for in their observations of the process and of the activities of his or her subordinate agents are known; that evidence on developing anomalies is unambiguous; and that the level to which intervention is necessary is clear cut. However, none of these are likely to be true at least in more difficult situations. Furthermore, conventional descriptions of management by exception have implied that the supervisor passively waits for reports of difficulties to flow to them. Yet results in supervisory control and cooperative work have shown the opposite: the effective supervisor actively searches out information and assesses the state of the process and the state of how other agents are managing the process (Patterson et al. 1998). Harking back to a metaphor derived from earlier generations of technology – one can refer to this finding as the "directed telescope" role of supervisory function (van Creveld 1985 p. 75; 255–257).

Differing interpretations of the role of the supervisor hinge on one's view of what counts as an exception. Is it any anomaly in the process itself predicting a loss of separation in ATC)? Or does an exception that warrants possible intervention by the supervisor occur only when a potential loss of separation situation is not being handled well by the other agents involved (the flight crews of the various aircraft)? The supervisory control literature has blurred the difference between a disturbance in the process, which we would call an anomaly, and a concern about how the situation is being handled by other agents, for which we would reserve the label – an exception. A clear distinction between these two senses is essential. An anomaly is behaviour of the underlying process being monitored and controlled that deviates from standard or from expectations (Woods 1994), while an exception is a judgment about how well others are handling or are going handle an evolving situation. Thus, we would use exception to refer to a relationship between developing events in the process and the ongoing and future activities of other agents. Are the agents to whom authority has been delegated currently handling the situation well; are they at risk of being unable to handle the situation in the future; is there some reason to anticipate that they may be unable to handle the situation as it develops further? Furthermore, what is enough evidence that an exception is occurring or will occur on which to base a weaker or stronger intervention?

Developers and researchers have tried to avoid considering the complications of judging/anticipating exceptions, as defined above, by reducing supervisory interventions to fixed rules about how to respond to specific situations. Pre-defined situations (often via threshold crossings on process parameters that can be measured) become the triggers for supervisory intervention and are compiled in procedures and policies for the supervisory controller (Wiener 1988). For future ATC architecture, developers are determining a set of pre-defined situations calling for supervisory intervention by considering potentially dangerous aircraft manoeuvres (measured in separation miles), traffic density (a complicated hypothetical measure mixing traffic numbers and flow patterns), or other conditions that compromise safety (a vague, underspecified directive) (RTCA 1995). Computers will play a major role in detecting/predicting areas of developing trouble, for example predicting separation conflicts or (as they do today) forecasting areas with saturating traffic densities. However, taking into account how well the activities of other agents will resolve the trouble is much more difficult.

Many other aspects of such a control architecture remain to be explored (see Smith et al. 1998). It is, however, not hard to imagine that circumstances may arise where evidence on developing anomalies is too uncertain and ambiguous for pre-engineered algorithms to flag everything for the human manager. The nature of how problems develop and escalate (Woods and Patterson, in press), the repeated disappointing history of automating diagnosis (Woods 1994), and the nature of collaborative interactions (Patterson et al. 1998) all indicate that recognizing exceptions in how others are handling a developing situation is quite complicated. For example, triggers based on threshold crossings (as in 'dynamic density is now so much, now you are in control, traffic manager') may fail to capture the more subtle cases that trigger congestion problems, or foreclose the possibility of a bumpless transfer of authority from one agent to another. In many cases, air traffic managers themselves will have to trace the development of situations where others are having difficulty coping with active or potential disturbances.

Most theories on supervisory control acknowledge that the supervisor's decision to intervene can be expressed as a trade-off between gathering more evidence and intervening in time (Edwards and Lees 1974; Sheridan 1987; Sanderson 1989; Kerstholt et al 1996). The more evidence collected that an anomaly is developing and not being handled well, the more accurate the manager's assessment. However, waiting longer may allow the situation to worsen and make intervention more difficult, or even impossible.

That this can be a tricky trade-off was shown by the introduction of automated process control in steel manufacturing. The 1976 Hoogovens report details how supervisors of newly commissioned process control systems in steel plants had little ability to see the autonomous process in action, and the complexity of process activities obscured sources of problems and malfunctions. To supervisors it was often unclear what kind of intervention should take place, and when or whether they should intervene at all. Unsure of how their manual intervention would interfere with automatic anomaly compensation, supervisors frequently left anomalies to escalate (European Coal and Steel Community 1976).

Establishing the need for intervention and identifying the best ways to intervene in a (partially) autonomous process is difficult – not only theoretically but also practically (Moray et al 1994). For management by exception (in air traffic control, but in other applications as well), we still need to explore the following intertwined questions:

- What evidence must the operator gather to establish the need for intervention? Can computers flag exceptions or situations that may correlate with exceptions?
- Is early intervention difficult (because not enough evidence may be available to justify taking back authority that had been delegated)? Is late intervention difficult (because trouble has escalated and time has run out)?
- Is the decision of *how* to intervene in any way dependent on *when* the operator decided to intervene?

2. METHOD

In the context of a larger NASA funded project to study future air traffic control architectures (see Smith et al 1997) we set out to investigate the kinds of cognitive work an air traffic controller would have to do as manager of exceptions. As discussed, the nature of the controller's cognitive tasks and challenges was hard to predict in detail on the basis of existing literature on management by exception and supervisory control. A series of empirical studies was conducted over a period of more than a year, using active air traffic controllers, airline pilots and flight dispatchers in simulated 'future incidents'. Each probed different aspects of the cognitive work of (and coordination between) various practitioners in the future ATC architecture in order to characterise management by exception and eventually develop models that may capture its underlying psychological mechanisms.

2.1. Investigating Envisioned Worlds

The challenge in generating results about cognitive work in future air traffic control is that such future architectures exist nowhere in the world. The envisioned world represents a radical departure from existing practice – rendering conclusions on the basis of today's world inapplicable and generally making it attractive to wait for further developments, requirements specification and hardware in order to generate valid results about operator problem solving in the future system. But waiting for a more developed world is predicated on the desire to generate valid results. But an equally strong imperative is to learn about what would be useful early, in order to influence the development process. We decided not to wait for more future ATC specifications, as experience in almost all domains of human–machine interaction has taught that leverage for change decreases as time goes on. This need for early results when new systems are not fully specified shifts the dominant source of validity in research on envisioned worlds as compared to more developed or existing operational environments. In the latter, face validity of a simulation tool or experimental setup is often thought to provide much of the requisite mapping between test situation and target world. In contrast, in research on envisioned worlds, validity (one could say ecological validity) derives from (1) the extent to which problems-to-be-solved in the test situation represent the vulnerabilities and challenges that exist in the target world, and (2) the extent to which real problem-solving expertise is brought to bear by the study participants (see Orasanu and Connolly 1993; Klein 1993; Woods 1993). Our studies rated high on these latter dimensions by (1) creating future incidents, and (2) involving real practitioners who had been prepared for their future roles. In other words, these studies investigated real practitioners caught up in solving real domain problems.

To get empirical data, a test situation had to be created that mapped onto critical aspects of the future target world. The target world we aimed our studies towards is a large set of future rules and procedures and proposed technologies that together make up the 'free flight' proposal (RTCA 1995). Other target situations on flexible routes share crucial characteristics with the free flight proposal (notably more separation responsibility on the part of pilots and a larger supervisory distance for controllers), so the results from these studies are relevant to many future air traffic control architectures in development today.

2.2. Future Incidents

The studies were built around 'future incidents' (Smith et al. 1997; 1998). The future incident method is based on developing a failure or near miss that could happen given one view of how the envisioned world might work. These incident-contained critical events that could happen in air traffic control systems of any vintage, because of their technology- or time-independent nature. For example, we introduced communication system failures, clear air turbulence, frontal thunderstorms, a cabin depressurisation and a priority air-to-air refuelling request into the future world to probe the problem-solving activities of its exception managers. This method was used to explore ways in which a future architecture could break down. We elected to turn to vulnerabilities that a future world could be exposed to as experimental probes, and investigated the cognitive demands on supervision and coordination that would have to be met in order to handle and contain them. Indeed, exception managers are justified in their role and existence in the system precisely because of their presumed ability to deal with unexpected permutations of problem factors and circumstances. We identified and specified sources of vulnerability in close collaboration with many different domain experts (air traffic controllers, pilots, airline safety officers, dispatchers) from different areas and countries. These practitioners did not participate in subsequent studies.

For the studies reported in this paper, we invited four active air traffic controllers (mean age 34, mean professional experience 13.75 years), a flight dispatcher (34 years; 10 years' experience) and a pilot (57 years, 32 years' experience) to participate, representing the different user perspectives of the future system. The participants were invited to come to the laboratory over several days, where a representation of their future problem-solving environment had been built (also on the basis of proposed RTCA 1995 rules). This consisted of airspace maps (where flexible route airspace had been drawn in to represent possible future airspace layouts) and static representations of future radar displays (including aircraft symbols). In advance of the study, all participants were prepared for the procedures of the future (also according to RTCA 1995), using the kinds of materials that would normally constitute their procedural and policy guidance (for example, in this case, faked pages out of 'future' air traffic control handbooks and aviation Advisory Circulars).

The four controllers and other practitioners were all gathered in front of one future radar representation, which portrayed the starting situation of the future incident. The initial conditions were given and explained to the participants, much in the same way as in a regular hand-over from one controller (going off duty) to another (coming on duty). Then the anomaly was introduced (e.g., a comm failure), from which point on the participants were asked to solve the problem together using the rules of the future. Thus, participants themselves became the engine of action. We decided not to have our radar representation mimic the event-driven and time-limited nature of their operating environment at this stage. The emphasis in this research was exploratory, so it was more important to have practitioners consider as many potential problem permutations, circumstantial developments and solution paths as possible, without a hard time limit. Participants could mark up the radar representation to suggest proposed aircraft movements. This generated a heavily annotated series of radar representations (or snapshots of the various stages in the problem as driven by the participants) which we used in our subsequent analysis. Participants were also encouraged to voice their proposals to the other participants for consideration, which generated protocols based on verbal and motor behaviour that occurred as part of the participants' natural task behaviour. The entire sessions were videotaped and later transcribed, creating a verbal and visual process trace that not only documented the various ways in which the incident could have unfolded, but also captured the cues, tools and rules (which document was grabbed and referred to, for example) necessary to address the situation in the envisioned architecture.

2.3. Highlighting One Study: A Communications Failure in Future ATC

Although our studies on future ATC problem solving expose different aspects of the future problem-solving world (see Dekker 1996), we have chosen to highlight one study in this article because it most fully brings out the diversity of cognitive challenges associated with management by exception (i.e., how to gather evidence on a developing situation; how to intervene). The future incident in this case involved two crossing streams of traffic, with one aircraft suddenly squawking 7600 on the mode A of its transponder. This signified a communications failure of that aircraft, as well as possible problems with its altitude encoding equipment and collision avoidance technology (Fig. 2). Uncertainty about the non-comm aircraft's intentions was increased by making the aircraft report initiating a climb to a higher flight level just before the altitude-reporting and radio equipment failures occurred. Also, the flight plan given to participants showed that the aircraft was headed for its home base, which rendered others unsure whether it would press on (attracted by its own maintenance facilities at base) or choose to land at the nearest suitable airport instead.

3. RESULTS

3.1. Overview

After Hollnagel et al (1981), we elected to describe observed performance at multiple levels of analysis (see also Ericsson and Simon 1993). The context-dependent, concept-independent level of analysis captures central portions of the raw verbal and visual data trace. This is used to step up to subsequent levels of more concept-dependent description of the underlying phenomena. Jumping ahead to give a broad idea of what will be encountered in this section: the participants' performance matches basic results from previous supervisory control work. In this future incident (as in all of the studies), they searched for evidence on how the situation could develop over time based in part on how other agents would respond to aspects of their environment, in effect asking themselves, 'Is this a situation that I need to get involved in?' They also reviewed ways in which they could intervene to stop the situation from deteriorating. This is where the results begin to add to current descriptions of cognitive work supervisory control: these two questions are fundamentally interdependent. Deciding whether to intervene over time is deeply

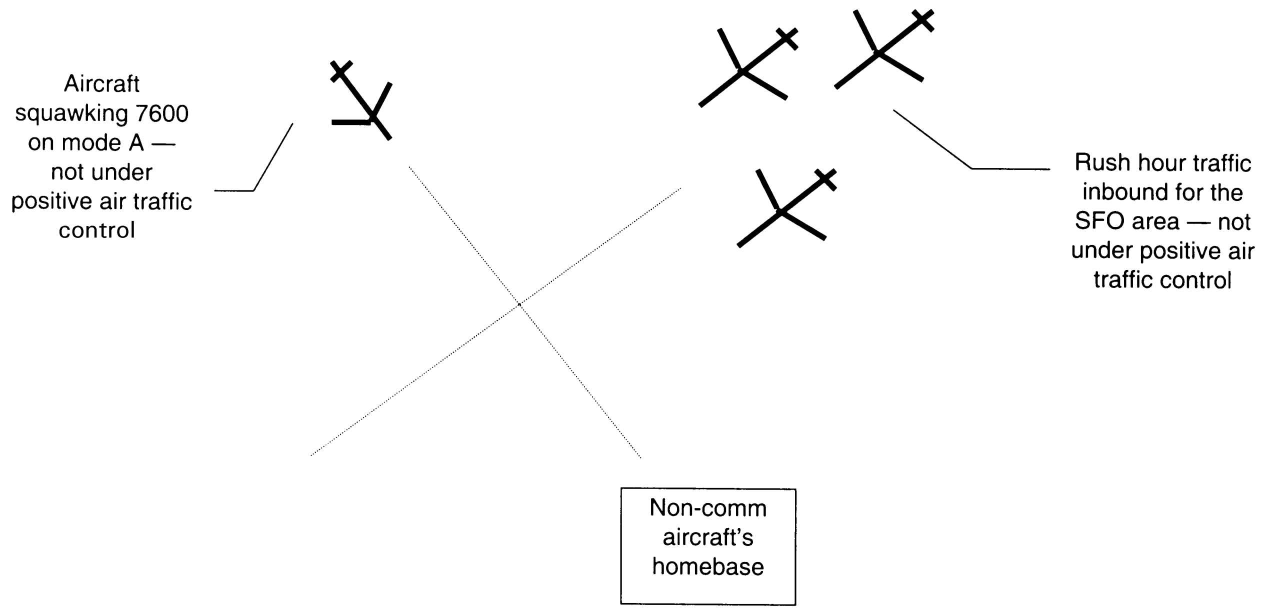

Fig. 2. Simple graphic of future incident situation.

intertwined with the question of what intervention is possible.

3.2. Is this an Exception?

With respect to the non-comm aircraft, participants first asked among themselves what evidence they could gather on the development of this situation. 'What is the non-comm aircraft's last reported altitude – 330?', asked one. Looking at the rules for flexible routes, another commented, 'Well, that doesn't count', while another added, 'For all you know he could be at 370.' Since its altitude could not be established, participants reviewed the flight routes the non-comm aircraft could possibly take. 'He'll go to the nearest suitable airport', asserted one controller. 'No', countered another, 'he's going to the approach fix and hold till his estimated time of arrival'. This the participating pilot did not believe: 'This airline would go to home base? Most guys wouldn't do that.' It was interesting to note that even though procedural guidance exists today about what aircraft can be expected to do in case of communications failure, there is only loose coupling between procedure and practice. The procedure indicates what an aircraft or other kind of participant might do, not necessarily what it will do (Woods et al 1994). In future ATC architectures where movements of other aircraft in the area may not be directly governed by the controller, such uncertainty creates an even more challenging situation. Searching the rest of the system for clues about the non-comm aircraft's intentions, one controller suggested calling the airline's flight dispatch and finding out from them, because 'There is going to be more downlinking. If dispatch monitor power settings [e.g., for purposes of maintenance monitoring], they could tell us if it's climbing or descending', which prompted a variety of reactions such as 'You're talking a lot of coordination now.'

Indeed, future ATC architectures that put more flight path autonomy in the air could enhance the fragmentation of knowledge. With partial representations about an aircraft's whereabouts distributed across system participants, one can imagine future incidents where all the knowledge to resolve the problem was available somewhere in the system, but none of the participants was able to put it together in time. Increasing reliance on the distribution and sharing of knowledge means that coordination problems can smoothly translate into system failure (Billings and Cheaney 1981). This shifts the dominant source of vulnerability as compared to more centrally governed architectures which characterise the air traffic control role today (Nolan 1994; Serfaty et al 1994; Pawlak et al 1996).

3.3. How to Intervene?

So, should controllers (traffic managers) intervene in this developing situation? One controller suggested a firm and deep intervention, targetting the wave of incoming rush hour traffic. These aircraft were not under positive control (yet). 'All incoming aircraft should immediately be put under positive control with five-mile separation.' This suggestion met with immediate protest from other participants. 'You can't do that', said one. 'You mean, I'm sitting here and all of a sudden I'm in control over all those aircraft?', asked another. 'It would be a mad scramble.' Alluding to concerns over airspace utilisation and throughput, one controller added, 'Five miles separation might be twice what you need.'

The alternative is a less deep intervention, leaving more autonomy in the air, and then 'I want everybody to tell me before they do something.' In other words, the proposal is (and indeed is according to RTCA 1995) that every aircraft should state its intentions before executing their manoeuvres. But this too was judged to lead to data overload in circumstances such as these. Said one controller, 'Personally, when I'm busy and I'm working over here, and I've got everybody and their brother downlinking or yelling on my thing, 'I'm climbing, I'm turning, I'm doing this, I'm doing that' – I couldn't care less! Mind yourself , I'm taking care of an imminent situation here.'

Becoming less and less deeply involved in the details of the airspace process, controllers now switched to taking only the front-runners of the incoming rush hour wave under positive control, after having been able to establish clearly who might by under greatest threat. Even this was considered difficult: 'How do you know what the others [coming up from behind] will do?', asked one controller. Mixing levels of control in a limited piece of airspace was determined to be problematic. 'By taking only a few under positive control, you can actually make it worse', said one, while another controller added, 'You can end up putting them together.'

Pulling themselves out of the deep intervention even further, one controller suggested, 'I would issue an advisory [to the incoming traffic]: 'there is a nordo [no radio] airplane out there, last reported 330...'. But this suggestion was referred back to earlier discussions about the non-comm aircraft's unknown intentions: 'You don't know what it's going to do...', which would make a meaningful advisory rather difficult. Apparently out of options, one controller concluded, 'Since we don't have all the parameters, do we *have* to do anything?' Effectively detaching themselves from the problem, another controller offered, 'Incoming traffic would have a better view of the non-comm aircraft on their collision avoidance display then we.'

3.4. Three Trade-Offs

It turned out that controllers faced *three* interrelated judgements in their decision to intervene: (1) deciding *when* to intervene; (2) determining *how many* aircraft should be taken under tighter control (how broadly to intervene, in other words); and (3) *how much* authority should be taken away from those aircraft (how deeply to intervene in the monitored process). In attempts to deal with the non-comm aircraft, controllers negotiated a variety of paths through these three trade-offs. Remarkably, they reached consensus on none. Instead they encountered a number of dilemmas and finally arrived at an option where they effectively excluded themselves from any control over the situation.

4. DISCUSSION

Other studies in our series where different kinds of domain problems were introduced into the flow of events (clear air turbulence, thunderstorms, aircraft cabin depressurisations – see Dekker 1996) confirmed the interdependence of the three trade-offs. Early intervention would of necessity have to be broad, because it was often unclear which airspace users would be involved in the situation or affected by its implications. But by being broad, early intervention was often impossible or at least difficult – it could create throughput problems, elicit potential resistance from pilots unaware of the larger picture, and could lead to controller overload. The alternative would be to wait for more evidence on who would be affected and how. But postponing intervention meant that the situation could deteriorate and leave controllers no other option than intervening deeply. Of necessity these interventions would have to be narrow (involve only a few aircraft) because of the limited time available to get the detailed instructions through to affected aircraft. But the mixed levels of control that would remain in one piece of airspace were generally judged to be undesirable.

Amplifying the Hoogovens experience, this study reveals how the interdependence between when to intervene and how to intervene creates a profound dilemma for exception managers. There turn out to be various pressures not to intervene early in ATC, for reasons that include controller taskload, downstream repercussions on system throughput, pilot and airline reluctance to give up delegated authority unless ‘necessary’, and even flight safety. But the decision not to intervene early is not without consequences. Results from this study form one more testimony to findings on dynamic, sequential decision making in which earlier decisions constrain or can even pre-ordain later ones, stringing them together into a thread of sometimes hard-to-forecast interdependence (Brehmer 1991). By the time enough evidence is gathered on whether a situation really warrants intervention, controllers hardly have an option left to usefully intervene or to contribute meaningfully to a resolution of the unfolding incident. Thus, not only are there various pressures against intervening late (as well as early); among them flight safety considerations (e.g., the issue of mixed-level control) complexities in changing levels of authority across agents in a dynamic situation and controller task load under increasing time pressure. The more profound issue turns out to be whether a controller can meaningfully step in and do something by the time he or she has actually figured out what is going on. Some might argue that these difficulties are attributable to a learning curve, that initial exposure to a new set of rules and techniques would naturally lead to performance degradation but that such performance will improve over time. This may turn out to be true in a marginal sense. But management by exception traps human controllers in a central dilemma that cannot be erased by training or experience: intervening early provides only thin ground for justifying restrictions (and compromises larger air traffic system goals). But intervening late leaves little time for actually resolving the problem, which by then will be well under way (thereby compromising larger air traffic system goals). In summary, intervening early would be difficult, *and* intervening late would be difficult, although for different reasons. Based on this research, management by exception puts the future controller in a fundamental double bind.

4.1. Computerised Decision Support in Management by Exception

Automating a variety of detection and alerting tasks (such as conflict detection, airspace density calculations and predictions) and providing a controller with computer-generated resolution suggestions does little to alleviate the fundamental dilemma of whether and how to intervene. Asking a machine to do the conflict detection migrates the intervention criterion into a machine, in effect creating a threshold crossing alarm. The typical problem with threshold crossing alarms is that they are set either too early or too late (Woods and Sarter in press). When the controller takes over in case of a high threshold, the human intervention may land in the middle of a deteriorating or challenging situation. On the other hand, with low thresholds the alert may come too early to be meaningful: the automation flags conflicts without benefiting from the controller’s contextual experience, information and knowledge. Such computerisation produces a human–machine system that effectively operates in one of two modes: fully automated (before threshold crossing) or fully manual (after). Controllers might be able to say what is wrong with the machine’s decision, but remain powerless to influence it in any way other than through manual take-over. Research on collaboration has shown that architectures where either one agent does the whole task or the other agent does the whole task are not cooperative. Cooperation occurs only when the agents interact in the process of handling the situation.

4.2. Toward a Cooperative Architecture

How do we make progress toward a more cooperative human–machine architecture in future ATC architectures? Drawing on experiences in other domains as well as air traffic control (Woods and Roth 1988; Woods 1994; Layton et al 1994; Billings 1996), the active partner in a well-coordinated human–machine team (which in management by exception would often be the machine) would not sound threshold-crossing alarms to signal the end of its problem-solving capability. It would instead continuously comment on the difficulty or increasing effort needed to keep relevant parameters on target. The (human) supervisor could ask about the nature of the difficulty, investigate the problem and perhaps finally intervene to achieve overall safety goals (Sarter, Woods and Billings 1997; Woods and Sarter, in press). These types of cooperative interaction also specify the kinds of feedback that decision support in ATC (such as conflict detection and resolution advisory technologies) would require. For example, the machine partner would have to show when (and why) it is having increasing trouble handling a situation. Displays must be future-oriented to highlight significant sequences that reveal what could happen next and where (Woods and Sarter, in press). This is consistent with other findings of supervisory control demands in dynamic domains: the more a supervisor is distanced from the details of his or her monitored process, the more his or her judgements, assessments and decisions will have to be about the future (see Brehmer 1991). Displays should also be pattern-based, enabling controllers to scan at a glance and quickly pick up expected or abnormal conditions relative to airspace loadings or conflict areas (see Klein 1993).

In order to build such a cooperative architecture (and inspired by earlier supervisory control work), we should start with determining what levels and modes of interaction will be meaningful to controllers in which situations. In some cases, controllers may want to take very detailed control of some portion of a problem, specifying exactly what decisions are made and in what sequence, while in others the controllers may want to make very general, high-level corrections to the course of events. We have begun to explore some of these levels with our studies. In one situation, controllers suggested that telling aircraft in general where *not* to go was an easier (and sufficient) intervention than telling each individual aircraft where to go. These ideas also begin to specify future system requirements (for instance, the ability to communicate to aircraft the unavailability of a piece of airspace because of a particular problem in it). We believe that integrating all these possibilities of intervention within the future ATC system is not easy: it will require careful iterative analysis of the interactions between controller goals, situational factors and the nature of computer support (see Woods and Sarter in press).

5. CONCLUSION

5.1. To Intervene or not to Intervene

What seemed like an easy assignment of activities to future air traffic managers ('wait for enough evidence and then decide how to intervene') in fact turned out to be an extremely hard cognitive problem when this role was simulated for future incidents. The decision to intervene is an intricate trade-off between multiple interacting goals that are simultaneously active (e.g., separation safety, system throughput, controller workload, user economic concerns). In practical terms for the future system, early interventions are likely to create various misunderstandings with local airspace users who are uncertain about the reasons for sudden restrictions or revisions. Early intervention can also create problems downstream (in terms of throughput, efficiency, backlogs) and controller workload. Late intervention would leave more time for gathering evidence: the exception manager can establish with more certainty that problems are indeed afoot and which aircraft are going to be affected. But every second spent assessing a situation will have been lost resolving it. Management by exception traps human controllers in a dilemma: intervening early provides little justification for restrictions (and compromises larger air traffic system goals). But intervening late leaves little time for actually resolving the problem, which by then will be well under way (thereby compromising larger air traffic system goals). In conclusion, intervening early is difficult, *and* intervening late is difficult, putting the exception manager in a double bind.

It may also be problematic for air traffic management by exception to become smarter with experience. Consultations with practitioners inside and outside the studies reported here indicate that late interventions may frequently trigger a 'that was close' reaction, and drive a swing back over to early interventions. This will disrupt learning, since early interventions resolve anticipated problems even if there was no actual problem developing. They eradicate any evidence on whether the intervention itself was actually warranted, possibly allowing slippage back to late interventions until the awareness of close calls once again shifts the decision criterion, and recycling the pattern.

When computerised decision support provides threshold-crossing alarms (e.g., exceeding pre-set closure rates or dynamic density figures) to flag possible exceptions, they do nothing to resolve the fundamental double bind between early and late interventions. To assist the human supervisor in making the call to intervene, the machine portion of a more cooperative human–machine architecture would instead inform the human about the difficulty or increasing effort needed to keep relevant parameters on target, allowing the supervisor to probe the nature of the difficulty and assess the requirement of his involvement.

Despite the intentions of those involved in ATC improvements, many developments remain fundamentally technology-centred. Developing the technology (automatic dependent surveillance, conflict probes, digital communications) remains the primary activity around which all else is organised. The focus is on pushing the technological frontier; on creating the technological system in order to influence human cognition or human activity. These efforts are likely to produce ideas (for instance, that controllers will become exception managers) which are based on generous assumptions about human performance and collaborative activity. Similarly, these efforts are likely to produce computerised support that is not cooperative from the human controller's perspective. Indeed, it can make the relatively easy problems in a controller's life go away, but make the hard ones even harder.

Acknowledgements

This work was funded by the Advanced Air Transportation Technologies program at NASA Ames Research Center under Grant Number NAG2-995. We would like to express our appreciation to Judith Orasanu and to the controllers, dispatchers, pilots and traffic managers who assisted us in completing the work. The research program of which this theme was a part was a collaborative effort of Phillip Smith, Elaine McCoy, Nadine Sarter, Rebecca Denning, David Woods, Sidney Dekker, and Charles Billings.

References

Baiada RM (1995). ATC biggest drag on airline productivity. Aviation Week and Space Technology 31:51–53

Billings CE (1996). Aviation automation: the search for a human-centered approach. Erlbaum, Hillsdale, NJ

Billings CE, Cheaney ES (1981). Information transfer problems in the aviation system. In The information transfer problem: summary and comments. NASA Technical Paper 1875,Washington, DC, September 1981, pp 89–90

Bittel LR (1964). Management by exception: systematising and simplifying the managerial job. McGraw-Hill, New York

Brehmer B (1991). Modern information technology, time scales and distributed decision making. In Rasmussen J, Brehmer B, Leplat J (eds). Distributed decision making: cognitive models for co-operative work. Wiley, Chichester

Cooper G (1994). Euro flow control. Royal Aeronautical Society. Aerospace 21(3):8–11

Cotton WB (1995). Free flight in domestic ATM. Journal of Air Traffic Control 3:10–17

Dekker SWA (1996). Cognitive complexity in management by exception: deriving early human factors requirements for an envisioned air traffic management world. In Harris D (ed). Engineering psychology and cognitive ergonomics. Vol I: Transportation systems. Ashgate, Aldershot, UK, pp 201–210

Dekker SWA, Wright PC (1997). Function allocation: a question of task transformation. In Proceedings of ALLFN '97: revisiting the allocation of functions issue conference, Galway, Ireland

Dornheim MA (1995). Equipment will not prevent free flight. Aviation Week and Space Technology 31(31 July):46

Edwards E, Lees FP (1974). The human operator in process control. Taylor & Francis, London

Ericsson KA, Simon HA (1993). Protocol analysis: verbal reports as data. MIT Press, Cambridge, MA

European Coal and Steel Community (1976). Human factors evaluation at Hoogovens no. 2 hot strip mill. Report of the meeting on 26–27 October 1976, IJmuiden, Netherlands. Secretariat of Community Ergonomics Action, European Coal and Steel Community, Luxembourg

Gibbs WW (1995). Free-for-all flights: the FAA plans a revolution in air traffic control. Scientific American 12(December):11–12

Gilbreth FB (1916). Graphical control on the exception principle for executives. Paper 1537a delivered to the American Society of Mechanical Engineers, New York

Hollnagel E (1992). Coping, coupling and control: the modelling of muddling through. In 2nd interdisciplinary workshop on mental models, Robinson College, Cambridge, UK, March 1992

Hollnagel E, Woods DD (1983). Cognitive systems engineering: new wine in new bottle. International Journal of Man–Machine Studies 18:583–600

Hollnagel E, Pederson OM, Rasmussen J (1981). Notes on human performance analysis (tech rep Riso-M-2285). Riso National Laboratory, Denmark

Hughes D (1992). Automated cockpits: keeping pilots in the loop. Aviation Week and Space Technology 12(23 March):48–70

Kerstholt JH, Passenier PO, Houttuin K, Schuffel H (1996). The effect of a priori probability and complexity on decision making in a supervisory control task. Human Factors 38(1):65–78

Klein GA (1993). A recognition-primed decision (RPD) model of rapid decision making. In Klein GA, Orasanu J, Calderwood R, Zsambok CE (eds). Decision making in action: models and methods. Ablex, Norwood, NJ

Layton C, Smith PJ, McCoy CE (1994). Design of a cooperative problem-solving system for en-route flight planning: an empirical evaluation. Human Factors 36(1):94–119

Leslie, J (1996). Free flight: the solution to the antiquated air traffic control system? Make pilots their own air traffic controllers! Wired 3:92–97

Mackintosh, DP (1978). Management by exception. Prentice-Hall, Englewood Cliffs, NJ

McClellan JM (1998). Is the sky overcrowded? Flying (7):13–16

Moray N, Lee J, Hiskes D (1994). Why do people intervene in the control of automated systems? In Proceedings of the first automation technology and human performance conference, Washington, DC, April 1994

National Research Council (1998). The future of air traffic control: human operators and automation. National Academy Press, Washington, DC

Nolan MS (1994). Fundamentals of air traffic control. Wadsworth, Belmont, CA

Nordwall BD (1995). Free flight: ATC model for the next 50 years. Aviation Week and Space Technology 31(30):38–41

Nordwall BD (1999). FAA shifts to 'slow and steady'. Aviation Week and Space Technology 150(15):59

Orasanu J, Connolly T (1993). The reinvention of decision making. In Klein GA, Orasanu J, Calderwood R, Zsambok CE (eds). Decision making in action: models and methods. Ablex, Norwood, NJ

Parasuraman R, Bahri T, Molloy R, Singh I (1992). Adaptive automation and human performance. II. Effects of shifts in the level of automation on operator performance (report no. NAWCADWAR-92036-60). Catholic University of America, Washington, DC

Patterson ES, Woods DD, Sarter NB, Watts-Perotti, J (1998). Patterns in cooperative cognition. COOP '98, Third International Conference on the Design of Cooperative Systems, Cannes, France, 26–29 May 1998

Pawlak WS, Brinton CR, Crouch K, Lancaster KM (1996). A framework for the evaluation of air traffic control complexity. Paper presented to the American Institute of Aeronautics and Astronautics

RMB Associates (1996). Free flight: reinventing air traffic control. Vol 2: The 'minimalist' solution (study and research report jointly produced by RMB Associates and Aviation Systems Research Corp). RM Baiada, Evergreen, CO

RTCA (1995). Free flight implementation (final report of RTCA Task Force 3). RTCA, Washington, DC

Sanderson PM (1989). The human planning and scheduling role in advanced manufacturing systems: an emerging human factors domain. Human Factors 31:635–666

Sarter N, Woods DD, Billings CE (1997). Automation surprises. In G. Salvendy (ed) Handbook of Human Factors/Ergonomics. Wiley, New York

Scardina JA, Simpson TR, Ball MJ (1996). ATM: the only constant is change. Aerospace America 3(March): 20–23

Serfaty D, Entin EE, Deckert JC (1994). Implicit co-ordination in command teams. In Lewis AH, Levis IS (eds). Science of command and control. Part II: Coping with change. AFCEA International Press, Fairfax, VA

Sheridan TB (1976). Toward a general model of supervisory control. In Sheridan TB, Johannesen J (eds). Monitoring behavior and supervisory control. Plenum, New York, pp 271–282

Sheridan TB (1987). Supervisory control. In G. Salvendy (ed). Handbook of human factors. Wiley, New York

Sheridan TB (1992). Telerobotics, automation and human supervisory control. MIT Press, Cambridge, MA

Smith K, Scallen SF, Knecht W, Hancock P. (1998). An index of dynamic density. Human Factors 40(1):69–78

Smith et al (1997). Human-centered technologies and procedures for future air traffic management. Columbus, OH: Department of Industrial and Systems Engineering, Cognitive Systems Engineering Laboratory, Report CSEL 97-TR-04

Smith PJ, Smith, Woods DD, McCoy CE, Charles Billings C, Sarter N, Denning R, Dekker S (1998). Using forecasts of future incidents to evaluate future ATM system designs. Air Traffic Control Quarterly, 61(1),71–85

Stein W (1992). Models of the human operator in the control of complex systems. Research Institute for Human Engineering, Wachtberg-Werthoven, Germany

Stix, G (1991). Along for the ride? Scientific American 7(July):94–105

Taylor FW (1911). The principles of scientific management. Harper, New York

Towne HR (1886). The engineer as economist. Transactions of the American Society of Mechanical Engineers 7:428–441

Umbers IG (1979). Models of the process operator. International Journal of Man–Machine Studies 11:263-284

Van Ghent RNHW (1994). Results of the 1994 FAA datalink project conducted at NLR's research flight simulator (National Aerospace Laboratory report). NLR, Amsterdam

Wiener EL (1988). Cockpit automation. In Wiener EL, Nagel DC (eds). Human factors in aviation. Academic Press, San Diego, CA

Wiener EL (1989). Human factors of advanced technology (glass cockpit) transport aircraft (NASA contractor report 177528). NASA–Ames Research Center, Moffett Field, CA

Wiener EL, Curry RE (1980). Flight-deck automation: promises and problems (NASA technical memorandum 81206). NASA–Ames Research Center, Moffett Field, CA

Woods DD (1992). Cognitive activities and aiding strategies in dynamic fault management (series of three Cognitive Systems Engineering Laboratory reports). Department of Industrial and Systems Engineering, Cognitive Systems Engineering Laboratory, Columbus, OH

Woods DD (1993). Process-tracing methods for the study of cognition outside of the experimental psychology laboratory. In Klein GA, Orasanu J, Calderwood R, Zsambok CE (eds). Decision making in action: models and methods. Ablex, Norwood, NJ

Woods DD (1994). Automation: apparent simplicity, real complexity. In Proceedings of the first automation technology and human performance conference, Washington, DC, 7–8 April 1994

Woods DD, Hollnagel, E (1987). Mapping cognitive demands in complex problem-solving worlds. International Journal of Man–Machine Studies 26:257–275

Woods DD, Roth EM (1988). Cognitive systems engineering. In Helander M (ed). Handbook of human–computer interaction. North-Holland, Amsterdam

Woods DD, Sarter NB (1993). Evaluating the impact of new technology on human–machine cooperation. In Wise JA, Hopkin VD, Stager P (eds). Verification and validation of complex systems: human factors issues. Springer-Verlag, Berlin

Woods DD, Sarter NB (1998). Learning from automation surprises and 'going sour' accidents: progress on human-centered automation (report ERGO-CSEL-98-02). Institute for ergonomics, Ohio State University, Columbus, OH

Woods DD, Johannesen LJ, Cook RI, Sarter NB (1994). Behind human error: cognitive systems, computers and hindsight. CSERIAC, Dayton, OH

Correspondence and offprint requests to: S. W. A. Dekker, Centre for Human Factors in Aviation, IKP, Linköping Institute of Technology, S-581 83 Linköping, Sweden. Fax: +46-13-282479; e-mail: sidde@ikp.liu.se