

Dekker, S. W. A., Mooij, M., & Woods, D. D. (2002). Envisioned practice, enhanced performance: The riddle of future (ATM) systems. *International Journal of Applied Aviation Studies*, 2(1), 23-32.

Envisioned practice, enhanced performance: The riddle of future (ATM) systems

Sidney Dekker

Linköping Institute of Technology, Sweden

David Woods

The Ohio State University, USA

Martijn Mooij

San José State University, USA

Abstract

Future ATM (Air Traffic Management) systems are like envisioned worlds large, interconnected systems as yet unbuilt and unfielded. Assessing the real impact of design changes in such worlds is very difficult. In this paper we discuss several dimensions of envisioned worlds: plurality, underspecification and groundedness, and evaluate them for the case of future ATM. Each of the dimensions presents pitfalls for developers and designers, while at the same time offering opportunities for clearer access to visions of the future.

Enhanced performance, envisioned practice

Although nobody knows the future, systems designers certainly envision versions of it. Indeed, this is the whole purpose of design – an activity that targets the future, that envisions people's future practice, implicitly or explicitly (Beyer & Holtzblatt, 1998). Design and development is based on creating and sharing stories about prospective operational settings (Roesler *et al.*, 2000). For example, designers across the developed world envision a diversity of Air Traffic Management (ATM) systems where controllers will no longer be in charge of all the parameters of every flight in their sector (Cooper, 1994; Cotton, 1995; Baiada, 1995; Gibbs, 1995; Leslie, 1996; NRC, 1998; Eurocontrol, 1999). These visions all advance some or several aspects of

the technological substrate (e.g. dynamic density predictors (Smith *et al.*, 1998; Wyndemere, 1996); medium-term conflict detectors (Leveson, 2001) or datalinking (NRC, 1998) from which others could build specific future objects or tools. Advancing these aspects of the technological base comes with hopes and predictions of enhanced performance (e.g. Dornheim, 1995). In fact, redesign, including the push towards new forms and levels of computerization, is commonly thought to be a solution to previous limits on performance. A core argument in air traffic control is that the human controller is a limiting factor in traffic growth (Wise *et al.*, 1991; Nordwall, 1995; Baiada, 1995; Scardina *et al.*, 1996). Decoupling controllers from all individual flights in their sectors, through greater computerization and automation on the ground and greater autonomy in the air, is assumed to be the way around this limit (Cooper, 1994).

Assessing the actual impact of design changes on envisioned practice is hard. There are traditional human factors tools for the verification, validation and (lately) certification of isolated aspects of new systems. (e.g. Wise & Wise, 1993; Marti *et al.*, 2001). Yet such piecemeal "V & V" quickly amounts to an oversimplification fallacy that understanding multiple-factor, interconnected processes, let alone anticipating longer-term evolutions of tool tailoring and task transformation, can derive from the momentary assessment of the state of a few independent things, objects, artifacts or sub-processes (Woods & Dekker, 2000). In reality, design changes and their effects are never insular, they always interact with other system aspects (see e.g. Leveson, 2001) and become tangled up in organizational, procedural changes as well, affecting areas not predicted by designers. Design changes usually create unintended complexities as well as offering new capabilities. The nature of people's practice gets transformed through the emergence of new human roles, new judgments, new forms of coordination. In turn, people always endeavour to work around design complexities by tailoring objects to function as actual, workable tools in their newly shaped operating context. As a result, what is operationally canonical or routine changes; what is expertise changes; and paths to breakdown and system failure change as well. Can design actually anticipate the full range of *such* changes? History is not on the side of designers – see for instance the unanticipated effects of commercial flight deck automation (e.g. Billings, 1996; Dekker & Hollnagel, 1999). Indeed, there is always a gap between the envisioned impact of a new development and the actual impact and reverberations of that change. To some this gap is so daunting that they concede that "an ATC system can be validated only in operation" (Smoker, 1993, p. 524).

Envisioned worlds

The problem is one of envisioned worlds (Dekker & Woods, 1999a) – large,

comprehensive, interconnected operational systems that have not yet been built and/or fielded. Even more modular changes to existing systems can create fundamental reverberations and, in effect, project envisioned worlds with uncharted transformations in people's roles. For example, the "mere" introduction of medium-term conflict detection tools in ATC (see Leveson, 2001), coupled with greater autonomy for pilots, creates classical alarm problems and signal detection trade-offs for controllers (Dekker, 2002) and role conflicts between pilots and controllers (Howard *et al.*, 2002).

There are several dimensions to envisioned operation concepts. Each of these have negative and positive poles with respect to how designers envision the future. We discuss these dimensions here, and their relevance for envisioned ATM systems.

Plurality

There are always multiple versions of how the proposed changes will affect the field of practice in the future. Different stakeholders (e.g. in ATM: air carriers, pilots, dispatchers, controllers, ATM managers) have different perspectives on the impact of new objects on the nature of practice. The downside of plurality is a kind of parochialism where people mistake their partial, narrow view for the dominant view of the future of practice, and are unaware of the plurality of views across stakeholders (e.g. one pilot's claim that greater autonomy for airspace users is "safe, period" see Baiada, 1995). The upside of plurality is the triangulation that is possible when the multiple views are brought together. In examining the relationships, overlaps, and gaps across multiple perspectives we are better able to cope with the inherent uncertainty built into looking into the future.

A number of future incident studies (Smith *et al.*, 1998b; Dekker & Woods, 1999b) examined controllers' anomaly response in envisioned ATM worlds precisely by capitalizing on this plurality. To study anomaly response under envisioned conditions, groups of practitioners (controllers, pilots and dispatchers) were trained on proposed future rules (see RTCA, 1995; NRC, 1998). They were brought together to try to apply these rules in solving difficult future airspace problems that were presented to them in several scenarios. These included aircraft decompressions and emergency descents; clear air turbulence; frontal thunderstorms; cornerpost overloading (too many aircraft going to one entry point for airport area); priority air-to-air refuelling and consequent airspace restrictions and communication failures. These challenges, interestingly, were largely rule- or technology-independent: they happen in ATM systems of any generation. The point was not to test the anomaly response performance of one group against that of another, but to use triangulation of multiple stakeholder viewpoints anchored in the task

details of a concrete problem to discover where the envisioned system would crack, where it would break down. Validity in such studies 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 way in which real problem-solving expertise is brought to bear by the study participants (see Orasanu & Connolly, 1993; Klein, 1993; Woods, 1993).

Underspecification

As hypothesis and prediction about the impact of new objects on the nature of practice, each envisioned concept is of necessity vague on many aspects of what it would mean to function in that field of practice in the future. The upside of underspecification is the freedom to explore new possibilities and new ways to relax and recombine the multiple constraints to innovate and improve. For example, visions of future ATM systems typically included datalinking as an advance that avoids the narrow bandwidth problem of voice communications thus putatively enhancing system performance (NRC, 1998). In one future incident study, reported in Dekker & Woods (1999) a communications failure affected an aircraft that had also suffered problems with its altitude reporting and was headed for crossing traffic. The underspecification of how datalink would be implemented allowed one controller to suggest that ATC should contact the airline's dispatch office to see whether the aircraft had been climbing or descending, since "If dispatch monitor powersettings, they could tell us." (p.) Others objected because of the coordination overheads this would create (Dekker & Woods, 1999b). Capitalizing on underspecification requires a search for leverage points (datalink and other resources in the system) and a sensitivity to the fact that envisioned objects only become tools through use imagined or real (datalinks to dispatch as backup ATC tool).

The downside of underspecification is the risk to remain trapped in a disconnected, shallow, unrealistic view of what it means to practice then and now. When the view of practice is disconnected from the pressures, challenges, and constraints operating in that world, the view of practice is inherently distorted from the beginning, and misses how the strategies of practice are adapted to these constraints and pressures. For example, one common assumption is that human controllers will make good "exception managers", since they can handle situations that machines cannot (see e.g. Cooper, 1994). Following this logic, controllers have been envisioned to become traffic managers, waiting for problems to occur in a type of standby role (McClellan, 1998). The view of controller practice is one of passive observer, ready to act when necessary.

But intervening effectively from a position of disinvolvement has proven to be

singularly difficult. For example, Endsley *et al.* (1997) point out in their study of direct routings that allowed aircraft deviations without negotiations, the more freedom of action was granted to individual aircraft, the more difficult it became for controllers to keep up with traffic. Controllers were less able to predict how traffic patterns would evolve over a foreseeable timeframe. In other studies too, passive monitors of traffic seem to have trouble to maintain a sufficient understanding of the traffic under their control (Galster *et al.*, 1999), and are more likely to overlook separation infringements (Metzger and Parasuraman, 1999). In one study, controllers effectively gave up control over an aircraft with communications problems, leaving it to other aircraft and their collision avoidance systems to sort it out among themselves (Dekker & Woods, 1999b). This turned out to be the controllers' only route out of a fundamental double bind: if they would intervene early they would create a lot of workload problems for themselves (suddenly a large number of previously autonomous aircraft would be under their control). Yet if they waited with intervention (in order to gather more evidence on the aircraft's intentions), they would also end up with an unmanageable workload and very little time to solve anything in. This confirms that controller disinvolvement can create more work rather than less (cf. Hilburn *et al.*, 1998)

These results indicate that starting from a distorted view of practice guarantees predictions that will prove wildly wrong as objects-to-be-designed come into contact with actual practice. The difficulty is that in envisioning, designers are pursuing a moving target so that it is difficult to say how the current vectors of practice play into a changing future.

Groundedness

Envisioned modes of operation are predictions about the effects of change on people, technology and work. As predictions, envisioned concepts can vary on a dimension of groundedness. Are the predictions grounded in patterns derived from the empirical research base on how technology change effects the interplay of people, technology and work? The downside is when envisioned concepts remain ungrounded from the research base on the actual consequences of the changes on people, technology and work. This is often the case as advocates recruit continued resource investment needed to develop objects-to-be-designed. Their claims about future impact are often at odds with or even contradict the empirical base.

For example, developers of future air traffic control architectures have been envisioning a number of pre-defined situations that call for controller intervention (e.g. RTCA, 1995). For example: potentially dangerous aircraft manoeuvres, local traffic density (which would require some density index), or other conditions that compromise safety. These rules, however, do not reduce

uncertainty about whether to intervene. They are all a form of threshold crossing – intervention is called for when a certain dynamic density has been reached (Smith *et al.*, 1998b) or a number of separation miles has been transgressed. But threshold crossing alarms are very hard to get right – they come either too early or too late. If too early, a controller will lose interest in them. If too late their contribution to flagging or solving the problem will be useless. The way in which problems in complex, dynamic worlds grow and escalate and the nature of collaborative interactions (Patterson *et al.*, 1998) all indicate that recognizing exceptions in how others are handling anomalies is complex. The disappointing history of automating problem diagnosis (Woods, 1994) inspires little further hope. Threshold crossing alarms cannot make up for a disinvolvement – they can only make a controller acutely aware of those situations in which it would have been nice to have been involved from the start.

Future incident studies also allow us to extend the empirical and theoretical base. For example, supervisory control literature makes no distinction between anomalies and exceptions (see e.g. Edwards & Lees, 1974; Sheridan, 1976; 1987; 1992; Umbers, 1979; Sanderson, 1989; Moray *et al.*, 1994; Kerstholt *et al.*, 1996). This indistinction results from the source of supervisory control work: how do people control processes over physical distances (time lag, lack of access, etc.). As Corker (2001) suggests, however, ATC augments the issue of supervisory control with a cognitive distance: airspace participants have some system knowledge and operational perspective, as do controllers. But there are only partial overlaps and many, many gaps. Indeed, the studies in Smith *et al.* (1998b) and Dekker & Woods (1999b) force us to make a distinction between anomalies in the process, and exceptions from the point of view of the supervisor (controller). Exceptions can arise in cases where airspace participants are dealing with anomalies (for example: an aircraft with pressurization or communications problems) in a way that forces the controller to intervene. An exception is a judgement about how well others are handling or going to handle disturbances in the process. Are airspace participants handling things well? Are they going to get themselves in trouble in the future? Judging whether airspace users are going to get in trouble in their dealings with a process disturbance would require a controller to recognize and trace a situation over time – countervailing the argument that human controllers make good "stand-by" interveners.

Management by obstacle

The empirical investigation of envisioned worlds can even generate novel directions for solving performance limits and problems. For example, the future incident studies reported in Smith *et al.* (1998) and Dekker & Woods (1999b) allowed practitioners and researchers to develop revised ways of

intervening that would be sensitive to the pressures and demands of actual practice. In future ATM systems with more aircraft autonomy, it could be easier for a controller to broadly tell aircraft where *not* to go, instead of telling each in detail where to go. This would open a new avenue to system management and airspace representation on flight decks: management by obstacle. Controllers would have to be given the opportunity to quickly create and communicate to airspace users an airspace obstacle – a piece of airspace set apart because of the need to deal with a local problem in it. Controllers in the studies reported here suggested the use of a track ball to swiftly designate a piece of airspace off limits. These newly created obstacles could be hard (no other aircraft allowed in until problem solved) or soft (other aircraft allowed in only if willing to submit to tougher restrictions), and different categories of obstacles could be automatically communicated to, and graphically represented in aircraft flight decks. Such airspace management would largely preserve the principle on which many future airspace ideas are founded: more user freedom and less controller involvement.

Conclusion

Envisioning a co-evolving, dynamic and future process of change and adaptation is highly uncertain. Advocates can easily become miscalibrated and overconfident that, if the systems envisioned can be realized, the predicted consequences and only the predicted consequence will occur. Designers' views of the future are tentative hypotheses. As such they need to remain open to revision and subject their hypotheses to empirical jeopardy in the face of feedback.

All who would design are susceptible to the fragility of envisioned world stories. To design one must speculate about the impact of the object-to-be-created as a source of change in a field of practice. The envisioned world problem demands that we develop means to ground predictions on relevant empirical results abstracted from observations in context. Understanding the dynamic process of change and adaptation will lead to better control of the innovation process at the intersection of people, technology and work. Armed with knowledge about the dynamics of change and adaptation, we can address potential side effects at a time when intervention is less difficult and less expensive (because the field of practice is already in a period of change and systems development is in the process of creating tangible objects).

References

Baiada, R. M. (1995). ATC biggest drag on airline productivity. *Aviation Week*

- and Space Technology*, 31, 51-53.
- Billings, C. E. (1996). *Aviation Automation: The search for a human-centered approach*. Hillsdale, NJ: Erlbaum.
- Cooper, G. (1994). Euro flow control. *Royal Aeronautical Society Aerospace*, 21 (3), 8-11.
- Corker, K. M. (2001). Human Factors of Automation: Human Performance Representation in Large Scale Dynamic Transportation Systems. *Human Factors in Aviation course*, June 5-6, 2001. Linköping.
- Cotton, W. B. (1995). Free flight in domestic ATM. *Journal of Air Traffic Control*, 3, 10-17
- Dekker, S. W. A. (2002). *The human factors of medium term conflict detection (MTCD) in ATC* (HF Tech Rep. No. 2002/03). Linköping Institute of Technology: Human Factors Group.
- Dekker, S. W. A., & Woods, D. D. (1999a). Extracting data from the future: Assessment and certification of envisioned systems. In S. W. A. Dekker and E. Hollnagel (Eds.), *Coping with computers in the cockpit*, 131-144. Aldershot, UK: Ashgate.
- Dekker, S. W. A., & Woods, D. D. (1999b). To intervene or not to intervene: The dilemma of management by exception. *Journal of Cognition, Technology and Work*, 1, 86-96.
- Dornheim, M. A. (1995, July 31). Equipment will not prevent free flight. *Aviation Week and Space Technology*, 31, 46.
- Edwards, E., & Lees, F. P. (1974). *The human operator in process control*. London: Taylor & Francis.
- Endsley, M. R., Mogford, M., Allendoerfer, K., and Stein E. (1997). *Effect of free flight conditions on controller performance, workload and situation awareness: a preliminary investigation of changes in locus of control using existing technologies*. Texas Tech University, Lubbock, TX.
- Eurocontrol (1999b). *Operational Concept Document of the strategy 2000+*. Brussels, Belgium: Author.
- Galster, S. M., Duley, J. A., Masolanis, A.J., and Parasuraman, R. (1999). Effects of aircraft self-separation on controller conflict detection and workload in mature Free Flight. In Scerbo, M. W., and Mouloua, M. (Eds.), *Automation technology and human performance: current research and trends*, pp. 96-101. Mahwah, NJ: LEA.
- Gibbs, W. W. (1995, December). Free-for-all flights: The FAA plans a revolution in air traffic control. *Scientific American*, 12, 11-12.
- Hilburn, B.G., Bakker, M.W.P. & Pekela, W.D. (1998) *Free Flight and the Air Traffic Controller: An exploratory analysis of human factors issues*. NLR Report NLR-TP-98237.
- Howard, M. V., Dekker, S. W. A., & Woods, D. D. (2002). Traffic information in free flight operations: The pilots' perspective. *Journal of Human Factors and Aerospace Safety* (submitted for publication).
- Kerstholt, J. H., Passenier, P. O., 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, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood & C. E. Zsombok (Eds.), *Decision making in action: Models and methods*. Norwood, NJ: Ablex.
- 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.
- Leveson, N. et al. (2001). A safety and human-centered approach to developing new Air Traffic Management tools. In proceedings of the 4th USA / Europe Air Traffic Management R&D Seminar Santa Fe, 3-7 December 2001.
- Marti, P., Lanzi, P., & Pucci, F. (2001). Evaluating safety and usability of ATM systems. In proceedings of the 4th USA / Europe Air Traffic Management R&D Seminar Santa Fe, 3-7 December 2001.
- McClellan, J. M. (1998). Is the sky overcrowded? *Flying* (7), 13-16.
- Metzger, U., and Parasuraman, R. (1999). Free Flight and the air traffic controller: active control versus passive monitoring. *Proceedings of the human factors and ergonomics society 43rd annual meeting*.
- Moray, N., Lee, J., & Hiskes, D. (1994, April). Why do people intervene in the control of automated systems? *Proceedings of the First Automation Technology and Human Performance conference*, Washington, DC.
- National Research Council (1998). *The future of air traffic control: Human operators and automation*. Washington, DC: National Academy Press.
- Nordwall, B. D. (1995, July 31). Free flight: ATC model for the next 50 years. *Aviation Week and Space Technology*, 31, 38-41.
- Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood & C. E. Zsombok (Eds.), *Decision making in action: Models and methods*. Norwood, NJ: Ablex.
- RTCA (1995). *Free flight implementation (Final Report of RTCA Task Force 3)*. Washington, D.C.: RTCA.
- Sanderson, P. M. (1989). The human planning and scheduling role in advanced manufacturing systems: An emerging human factors domain. *Human Factors*, 31, 635-666.
- Scardina, J. A., Simpson, T. R., Ball, M. J. (1996, March). ATM: The only constant is change. *Aerospace America*, 3, 20-23.
- Sheridan, T. B. (1976). Toward a general model of supervisory control. In T. B. Sheridan, & J. Johannesen (Eds.), *Monitoring behavior and supervisory control*. New York, NY: Plenum, pp. 271-282.
- Sheridan, T. B. (1987). Supervisory control. In: I. Salvendy (Ed.), *Handbook of human factors*. New York, NY: Wiley.
- Sheridan, T. B. (1992). *Telerobotics, automation and human supervisory control*. Cambridge, MA: MIT Press.
- Smith, K., Scallen, S. F., Knecht, W., & Hancock, P. (1998a). An index of

- dynamic density. *Human Factors*, 40(1), 69-78.
- Smith, P. J., Woods, D. D., McCoy, C. E., Billings, C. E., Sarter, N. B., Denning, R., Dekker, S. W. A. (1998b). Using forecasts of future incidents to evaluate future ATM system designs. *Air Traffic Control Quarterly*, 61(1), 71-85.
- Smoker, A. (1993). Simulating and evaluating the future: Pitfalls or success? In: J. A. Wise, V. D. Hopkin, & P. Stager (Eds.): *Verification and validation of complex systems: Human factors issues*, pp. 521-528. NATO ASI series F, Vol. 110. Berlin: Springer-Verlag.
- Umbers, I. G. (1979). Models of the process operator. *International Journal of Man-Machine Studies*, 11, 263-284.
- Wise, J. A., Hopkin, V. D., & Smith, M. L. (1991). *Automation and system issues in air traffic control*. NATO ASI Series F, Vol. 73. Berlin: Springer-Verlag.
- Wise, J. A., & Wise, M. A. (1993). Basic considerations in verification and validation. In: J. A. Wise, V. D. Hopkin, & P. Stager (Eds.): *Verification and validation of complex systems: Human factors issues*, pp. 87-96. NATO ASI series F, Vol. 110. Berlin: Springer-Verlag.
- Woods, D. D. (1993). Process-tracing methods for the study of cognition outside of the experimental psychology laboratory. In G. A. Klein, J. Orasanu, R. Calderwood & C. E. Zsombok (Eds.), *Decision making in action: Models and methods*. Norwood, NJ: Ablex.
- Woods, D. D. (1994). Automation: Apparent simplicity, real complexity. *Proceedings of the First Automation Technology and Human Performance Conference*, Washington, D.C., April 7-8.
- Wyndemere (1996). *An evaluation of air traffic control complexity*. Final report, contract no NAS 2-14284. Boulder, CO: Wyndemere.