VALIDATING DIAGNOSTIC DESIGN KNOWLEDGE FOR AIR TRAFFIC MANAGEMENT : A SUCCESSFUL CASE-STUDY

Becky Hill and John Long

Ergonomics & HCI Unit

Now at: UCL Interaction Centre
University College London
31-32 Alfred Place
London WC1E 7DP

This paper reports research, validating design knowledge for air traffic management (ATM). The knowledge is applied to an ATM simulation to diagnose design problems, associated with controller planning horizons. The case-study is judged a success. The design knowledge is correctly operationalised, tested and generalised to a simulation, more complex than that used to develop the knowledge.

Introduction

Cognitive Ergonomics researchers have been criticised for not building on each other's work (Newman, 1994). Elsewhere, Long (1996) claimed that poor discipline progress resides partly in the failure of research to validate its design knowledge. This paper reports a successful case-study, validating diagnostic design knowledge, applied to air traffic management (ATM).
Reconstructed Air Traffic Management: ATM is the planning and control of air traffic. Operational ATM manages air traffic, e.g., Manchester Ringway Control Centre (UK). The Centre manages a terminal manoeuvring area configured by: 9 beacons; more than 2 airways; 1 stack; and 2 exits. The management involves track and vertical separation rules. Planning is supported by paper flight progress strips and controlling by radar. Dowell (1998) developed a simulation of the Centre – termed 'reconstructed air traffic management' (rATM). It comprised: 5 beacons; 2 airways; and no stack. Traffic was limited to 8 aircraft and sector entry was staggered. There was a single controller. Dowell (1993) also developed a domain model, comprising airspace and aircraft objects, consisting of attributes having values. Transformation of these attribute values results in aircraft 'safety' and 'expedition', which express performance as 'task quality'.
Diagnostic Design Knowledge: Timmer (1999) has developed a Theory of Operator Planning Horizons (TOPH). It consists of: a set of frameworks (domain; interactive worksystem (operator and devices); and performance (Dowell, 1998)) and a method for diagnosing design problems, associated with operator planning horizons. Timmer applied TOPH to rATM to produce a set of models, which he used to diagnose design problems. TOPH is to be validated here.
Design Knowledge Validation: following Long (1996), design knowledge validation

comprises: conceptualisation; operationalisation; test; and generalisation. Here, Timmer's conceptualised TOPH is operationalised, tested, and generalised over a more complex ATM simulation.
Features of a Correct Operationalisation: following Stork, Middlemass and Long (1995), the features of a correct operationalisation of TOPH are: 1. Diagnosis completeness; 2. Diagnosis consistency; 3. Application of domain, worksystem and performance models; 4. Rationale for model application; 5. Features of diagnostic method, embodied in diagnosis.
Case-study Success: following Middlemass, Stork and Long (1999), case-studies of design knowledge can be successful or unsuccessful. These two types of outcome together establish the scope of the design knowledge. Design scenarios are considered to vary in their: definition; complexity; and observability.
Case-study Scenario: TOPH was applied to an ATM simulation, 'reconstructed validation air traffic management' (rvATM) (Debernard and Crevits, 2000). rvATM simulates an en-route sector in the region of Bordeaux (France). It is configured by: 21 beacons; multiple airways and multiple exits. The traffic is heavy, up to 40 aircraft on sector at any one time and flight patterns are very varied. Track and vertical separation rules are close to operational. There are two controllers – planning, responsible for electronic flight strips and radar, responsible for control. The validator (first author here), applying TOPH, was trained in HCI. She had no previous experience in TOPH application. The validation study was 'managed' by the second author. All difficulties in applying TOPH were documented. rATM and rvATM are similar as concerns definition and observability. However, rvATM is more complex, having a more complex sector configuration and a more extensive and varied traffic profile. The flight strips are electronic. There are two controllers. These differences are the basis for the generalisation process.

Design Knowledge Application and Evaluation

An observational study was conducted, using four video cameras. Videos recorded both planning and radar displays and their controllers, including verbal communications. The controllers were practised in rvATM. The validator later produced a protocol of the synthesised data (PSD) and constructed a table of controller interventions (TCI). Ambiguities were resolved with the controllers.
Examples of interventions follow: 'Worried about conflict between KLM051 and N7225U. Plan change KLM051 after AFR543. KLM051 turn right. IBE712 change heading direct TERNI'. The TCI includes: the aircraft; beacons; controller plans; and the validator's comments.
An integrated model for rvATM (rvIM) is now constructed, using the TOPH frameworks. Table 1 shows extracts from the rvIM for aircraft IBE550. It integrates work system-related models (Columns 1-5) with domain related-models (Columns 6 and 7). Column 1 models the goals of the worksystem ((planning and radar) and devices (flight progress strips and radar)). Column 2 models the controllers' behaviours. The model uses the TOPH operator architecture - physical ('head' and 'hands') and mental ('working and long-term memory, and goal store'). It also includes 'process structures' ('search for' and 'form goal') and 'representation structures' (categories of aircraft – 'active', 'expeditious' and goals – 'establish', 'amend', and 'intervene'). Physical behaviours can be observed on the video recording (a controller head movement towards the radar, indicating a 'search for' (aircraft) behaviour). Mental behaviours are inferred. Column 3 shows a model of the controllers' representation of the domain. The model uses TOPH mental categories for managed aircraft ('incoming/ active'; 'safe/ unsafe'; 'expeditious/ unexpeditious'). Categories in turn derive from domain attribute

values, such as aircraft; radar position; altitude; speed; heading etc. Column 4 shows the controllers' representation of the devices, i.e. flight strips and radar. Column 5 shows a model of device behaviours, with which the controller's behaviours (Column 2) interact. A comparison between Column 2 and Column 5 indicates appropriateness of the interactions for achieving the goals (Column 1). Column 6 shows a model of the product goal achievement, expressing the effect of an intervention on an aircraft. The achievement relates to the worksystem's goals. Column 7 shows a domain model of the state of each aircraft. The two highest states are 'safe' (not in conflict with other aircraft) and 'expeditious' (moving through the sector in a timely manner). The rvIM is now complete.

Table 1 Extracts from rvATM Integrated Model for aircraft IBE550

Worksystem goals	Controller behaviour	Controller rep (domain)	Controller rep devices	Device behaviour	Product goal	Aircraft transformation
(A) Intervention IBE550 Heading 39 at ENSAC	HIGHLIGHT: IBE550 FPS PULLDOWN: change heading SELECT: 39 CATEGORISE: IBE550	IBE550, heading 45 changing IBE550, (from) active safe expeditious to active safe unexpeditious (heading) aircraft	IBE550 FPS selected IBE550 FPS, heading 39	IBE550 FPS highlighted IBE550 FPS heading 39 Radar BTZ, IBE550, heading 39		IBE550 Progress worse Fuel use worse Safety same Exit worse
Planning/ Execution	POP GOAL: (B)					

Before diagnosing design problems, the controller's planning horizons need to be constructed. Following TOPH, controller tasks comprise: administration; monitoring; and planning/execution. Planning horizons can be constructed only for planning /execution tasks. A plan is a mental representation structure, associated with mental process structures ('form'; 'discard'; 'decay' etc), giving rise to planning behaviours. Plans can have three different outcomes: 'plan and decay'; 'plan and discard'; and 'plan and execute'. Planning horizons are constructed on information, associated with: the controller; the devices; the plan; its extension (over time) and its adequacy (to achieve worksystem's goals). The data are extracted from the PSD and the rvIM. The planning horizon for IBE550 is shown is Table 2.

Table 2 Planning horizon for aircraft IBE550

Encode	Intervention	Category	Plan/Execution
FPS		Incoming aircraft	
Radar trace		Active aircraft	
Heading 45 Altitude 310		Active safe expeditious aircraft	Change heading 39 at ENSAC
Position Ensac	IBE550 Heading change 39	Active safe aircraft unexpeditious (heading)	Leave IBE550
Position SAU Alt 310 Heading 39		Active safe unexpeditious (heading)	**Give Heading Terni after KLM358 passed**
		Lapse	Lapse
Position Velin Alt 310 Heading 39	IBE550 Heading change POI	Active safe expeditious (heading)	Change heading to POI
		Active aircraft exit	

Column 1 shows the controller's encoding of IBE550. Column 2 shows the controller's interventions. Column 3 shows the aircraft category. Column 4 shows the plan/execution.

Given the PSD, the rvIM and the planning horizon, the diagnosis method can be applied in its four stages: 1. Identify problem; 2. Analyse planning horizon; 3. Extract data from the integrated model; and 4. Generate causal theory. A design problem exists, when actual and desired performance differ. The planning horizon provides an overview of the design problem. It supports causal theory generation. The theory suggests possible design solutions. In the case of IBE550: an intervention has produced poor quality of work (unexpeditious aircraft with respect to heading). The problem arose, due to an intervention to make the aircraft safe. The controller later made a plan to rectify the unexpeditious problem by changing heading. This plan decays (or is discarded). To solve this design problem the controller would have to take more account of aircraft progress and fuel use, when planning. In rvATM, safety is of primary importance and little attention is paid to fuel use etc. Information on fuel use could be displayed and a prompt issued, indicating such reductions in performance. The prompt might be to direct the controller to return an aircraft to its original airway more quickly, thus enabling performance parameters to become as desired and so solving this particular design problem.

The rvATM diagnosis is evaluated here analytically. First, the diagnosis is considered to be of rvATM, as supported by the controller interventions, observed by video and documented in the PSD and rvIM. Second, the diagnosis is of design problems, as identified by Column 7 of Table 1. Last, the diagnosis relates to planning, as supported by the planning horizon and the causal theory. The rvATM diagnosis meets the requirements of being a design problem, associated with controller planning.

TOPH is judged to be correctly operationalised. First, the diagnosis is complete, as it corresponds to the application of the diagnosis method. Second, the diagnosis is consistent with the planning horizon, which is consistent with the rvIM, which is in turn consistent with the controller's interventions and the PSD. Third, the domain, worksystem and performance models of the rvIM are applied to the planning horizon construction and so to diagnosis formulation. Fourth, the rationale for the application of the models has been (selectively) exposed. Last, features of the diagnostic method are embodied in the diagnosis (plan extension and adequacy).

The TOPH application is considered to meet the validation requirements. First, the design knowledge was operationalised, that is, the already conceptualised TOPH was applied in the case-study to a more complex simulation. Second, the knowledge was tested, in that it resulted in the identification of design problems, associated with operator planning. The test, however, also identified difficulties in the application of the knowledge, experienced by the validator, which must count to some extent, against the validation. For example, the syntax for representing interventions in the rvIM was found difficult to apply (Table 4). The validation can, then, be considered only partial.

The case-study is considered a success. That is, rvATM, more complex; but equally well-defined and observable as rATM, is judged to fall within the scope of TOPH. Although the case-study is successful, the validation of TOPH is only partial, because of the validator's difficulties in its application.

Table 4 Difficulties, experienced by the validator in the application of TOPH

From Page	From section /paragraph	Diagnosis of problem	Solution to problem	Comments	Speculations
85	6.3.1.1.2 3rd paragraph	In rATM Operator physical behaviour hand movements correspond to radar (highlight; pulldown; select) whereas in rvATM, these behaviours correspond to both radar and FPSs.	Analyse hand movements corresponding to Radar (highlight; pulldown; select) and to FPS (highlight; pulldown; select)	Implemented as the flight strips are electronic and thus the corresponding hand movements in rATM for FPS (move, delete; write) do not apply here	Warn users of the method that the physical architecture will change with changes in the simulation being analysed

Discussion and Conclusions

The aim of this paper is to report a successful case-study, validating TOPH diagnostic design knowledge, as applied to ATM, in the form of rvATM. The case-study is considered a success, as rvATM, more complex than rATM, is judged to fall within the scope of TOPH. The latter is partially validated, inasmuch as it was operationalised, tested and generalised. The validation, however, was only partial, because the validator experienced difficulties in the application of TOPH. These difficulties constitute design problems for TOPH and their solution is a requirement for future research.

Last, this paper began with a critique of Cognitive Ergonomics researchers for not building on each other's work (Newman, 1994; Long, 1996). It is hoped that the research, reported here, of a successful case-study, which partially validated design knowledge for ATM, suggests how this criticism may be met.

Acknowledgments

This research was carried out in collaboration with the Laboratoire d'Automatique et de Mécanique Industrielles et Humaines (LAMIH) (Valencienne, France), under its director of research Jean-Michel Hoc and his team. CNRS (France) and the Royal Society (UK) funded the collaboration.

References

Debenard, S., and Crevits, I. 2000, Projet Amanda - note intermediaire 1.2 CENA/N12v1/.

Dowell, J. 1993, Cognitive Engineering and the rationalisation of the flight strip. Unpublished PhD Thesis, University of London.

Dowell, J. (1998). Formulating the design problem of air traffic management. *International Journal of Human-Computer Studies*, **49**, 743-766.

Long, J. 1996, Specifying relations between research and the design of human-computer interactions. *Int. Jnl. of Human-Computer Studies*, **44 (6)**, 875-920.

Newman, W. 1994, A preliminary analysis of the products of HCI research, using pro forma abstracts. In Proc. CHI '94, Boston, Mass., 278-284.

Middlemass, J., Stork, A., and Long, J. 1999, Successful case study and partial validation of MUSE, a structured method for usability engineering. In Proc. INTERACT 99, Edinburgh UK, 399-407.

Stork, A., Middlemass, J. and Long, J. 1995, Applying a structured method for usability engineering to domestic energy management user requirements: a successful case study. In Proc. HCI '95, Huddersfield, UK, 367-385.

Timmer P. 1999, Expression of operator planning horizons: a Cognitive Approach. Unpublished PhD Thesis, University of London.