Validating Diagnostic Design Knowledge for Air Traffic Management : a Successful Case-study

John Long
Ergonomics & HCI Unit
Now at: UCL Interaction Centre
University College London
31-32 Alfred Place
London WC1E 7DP
j.long@ucl.ac.uk

Becky Hill
Ergonomics & HCI Unit
Now at: UCL Interaction Centre
University College London
31-32 Alfred Place
London WC1E 7DP
b.hill@ucl.ac.uk

ABSTRACT
This paper reports research, aimed at validating design knowledge (Timmer, 1999) for airtraffic 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 shown to be correctly operationalised, tested and generalised to an ATM simulation, more complex than that used to develop the knowledge. However, problems with application are reported. The validation is, thus, only partial. Solution of these problems constitutes a requirement for future research. More general problems of applying design knowledge from research are identified and discussed.

Keywords
Validation; design knowledge; diagnosis; planning horizon; air traffic management.

INTRODUCTION
Cognitive Ergonomics researchers have been criticised for not building on each other's work. Newman (1994) claimed that only 30% of such work enhanced modelling techniques, solutions and design tools, as against 90% for Engineering more generally. Elsewhere, Long (1996) claimed that poor discipline progress resides partly in the failure of research to validate its design knowledge. This paper has two aims. The primary aim is to report a successful case-study, which attempts to validate diagnostic design knowledge, both substantive and methodological, as applied to air traffic management (ATM). The secondary aim is to identify general problems of applying design knowledge derived from research.

Reconstructed Air Traffic Management
ATM is here understood as the planning and control of air traffic. Operational ATM manages air traffic, for example, Manchester Ringway Control Centre in the UK. The Centre manages a terminal manoeuvring area, as configured by: 9 beacons; more than 2 airways; 1 stack; and 2 exits. Its traffic is: departing; arriving; over flying; 'low and slow'; and high-level bunching. The management involves track and vertical separation rules. Planning is supported by paper flight strips and controlling by radar. Dowell (1998) developed a simplified simulation of the Centre – termed 'reconstructed air traffic management' (rATM). The sector was configured by: 5 beacons; 2 airways; and no stack. Its traffic did not include 'low and slow' aircraft. Track and vertical separation rules were also simplified. Traffic was typically limited to 8 aircraft and entry to the sector was staggered. There was a single controller, using the paper flight strips to plan and the radar to control. Dowell (1993) also developed a domain model, comprising airspace objects 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). The theory consists of a set of frameworks (domain; interactive worksystem (operator and devices); and performance), as proposed by Dowell (1998) and a method for diagnosing design problems, associated with operator planning horizons. Timmer applied TOPH to rATM, producing a set of models, corresponding to each of the frameworks and a set of design problem diagnoses. TOPH is the diagnostic design knowledge to be validated here.

Design Knowledge Validation
Following Long (1996), design knowledge validation comprises: conceptualisation; operationalisation; test; and generalisation. Conceptualisation requires the design knowledge to be made explicit and so to be communicable to others (researchers and practitioners). Operationalisation requires the design knowledge to be applied correctly and demonstrably. Test requires the design knowledge to be evaluated against its own aims or claims, that is, its fitness-for-purpose. Generalisation requires design knowledge to be applied successfully

over a number of different instances (design scenarios) to establish its scope. Here, Timmer's conceptualised TOPH is operationalised, tested, and generalised over an ATM simulation, more complex than rATM.

Features of a Correct Operationalisation

Following Stork, Middlemass and Long (1995), the features of a correct operationalisation of TOPH diagnostic design knowledge are taken to be: 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. In successful case-studies, the design scenario is considered to fall within the scope of the knowledge. In unsuccessful case-studies, the design scenario is not considered to fall within the scope of the design knowledge. Successful and unsuccessful case-studies, thus, together establish the scope of the design knowledge. Design scenarios are considered to vary in: their definition, that is, how well they are specified; their complexity (that is, how simple or complicated they are); and their observability, that is, the access accorded the validators of the design knowledge.

Case-study Scenario

In the case-study, the TOPH diagnostic design knowledge (Timmer, 1999 – see earlier) was applied to an ATM simulation, which differed from rATM in a number of respects, termed 'reconstructed validation air traffic management' (rvATM) (for full details see Debernard and Crevits, 2000). rvATM simulates an en-route sector in the region of Bordeaux, France. The sector is configured by: 21 beacons; multiple airways and multiple exits. The traffic is heavy, up to 40 aircraft in the sector at any one time and flight patterns are very varied, requiring changes of heading, rather than changes of speed or altitude. Aircraft entry to the sector was not staggered. Track and vertical separation rules were close to operational practice. There are two controllers – the planning controller, responsible for flight strips and the radar controller, responsible for the radar. The flight strips are electronic and are displayed. They can be grouped by the planning controller and offered to the radar controller as decision support. Aircraft headings can be changed by means either of the flight strips or of the radar.

The validator, who applied the design knowledge, was trained in HCI and had considerable experience as an HCI researcher. She was familiar with the TOPH research, through seminars, given by Timmer, the developer of the theory. However, she had no detailed knowledge of the method or any experience in its application, prior to the validation study. The validator is the second author of this paper. The source, for the design knowledge to be validated, was Timmer (1999) – the most complete version of TOPH in print.

The validation study was 'managed' by the first author of this paper. He had been involved with Timmer in the development of the design knowledge (Timmer and Long, 1996; and 2002). His role here was to check the correct operationalisation of the TOPH theory. Any difficulties, experienced by the validator, in applying the design knowledge, were referred to the case-study manager for clarification. The manager also monitored the diagnostic products (models; method application etc) for errors. These interventions were necessary to ensure the correct operationalisation of TOPH. However, these interventions count against the validation of the design knowledge, since they indicated it was not entirely fit-for-purpose, either because of problems of conceptualisation, of expression or of both, as concerns that knowledge. All difficulties, associated with case-study manager interventions, were documented as design problems with TOPH, along with any difficulties in application, experienced, but correctly resolved by the validator, in the absence of case-study manager interventions (see later).

If rATM and rvATM are compared in terms of the design scenario dimensions (see earlier), they can be judged similar as concerns definition and observability.

In both cases, the definition of the studies was the same, that is, to diagnose design problems, associated with operator planning horizons, using the TOPH diagnostic design knowledge. The two studies are equally observable, both developer/validators having video access to the controllers. However, rvATM is much more complex than rATM. rvATM has a more complex sector configuration (more beacons; more pathways; and more exits) and a more extensive and varied traffic profile (number of aircraft). Traffic entry to the sector was not staggered, and vertical and separation rules were close to operational practices. The flight strips are electronic, with greater functionality. There are two controllers, who share management and who communicate verbally, as well as via the devices. Although the same in definition and observability, the rvATM design scenario differs from that of rATM, in that it is more complex. This difference is as required by the generalisation process of design knowledge validation – see earlier. A successful case-study would, thus, extend the scope of the TOPH diagnostic design knowledge to a more complex design scenario.

DESIGN KNOWLEDGE APPLICATION AND EVALUATION

An observational study was conducted, using four video cameras, in which the design knowledge validator recorded a rvATM simulation session. Four videos were made, recording both planning and radar displays and their associated controllers. The latters' verbal communications were also recorded. In addition, the validator made notes; but did not otherwise intervene. The simulation comprised a full rvATM design scenario (as described earlier). Although not professional controllers, the two controllers were well versed and practised in ATM and on the sector simulated. They had acted as controllers in numerous, earlier observational and system development studies.

The validator later observed the simulation videos and produced a protocol of the synthesised data and also constructed a table of controller interventions, to aid interpretation of the data. Ambiguities in controller behaviours or intentions were identified, discussed and resolved with the controllers, as required. The protocol data are not shown here due to space limitations. Examples of controllers' interventions are shown in Table 1.

IBE712 at ENSAC change heading VELIN not BDX
AFR543 at LMG change heading via FOUCO
(Believes conflict btw IBE539 and MON904 there isn't)
IBE550 heading change to 39
SAW105 change heading
Worried about conflict between KLM051 and N7225U
Plan change KLM051 after AFR543
KLM051 turn right
IBE712 change heading direct TERNI
KLM051 gone left not right has to change again (bug in the system)
N7225U change heading Direct BTZ
Plan want PAAA11 to go behind SAW105
KLM358 change heading direct TERNI
PLAN MUST NOT FORGET IBE550 (Iberia for TERNI)
SAW105 change heading direct PEROT
SAB6338 transfer to next sector
PAAA111 change heading to the left
IBE615 transfer to the next sector
N7225U change heading direct to LESGA
FUKCH change heading
IBE550 change heading direct to POI

Table 1 Table of controller interventions

Table 1 includes the aircraft traversing the sector, for example, IBE550 and KLM358. Also, shown are the beacons, over which the aircraft pass, for example, 'POI' (Poitiers) and BDX (Bordeaux). Controller plans are identified, for example, 'Plan change KLM051 after AFR543. Controller's references to plans are also recorded, for example, 'Plan must not forget IBE550'. Last, the validator's comments are shown, for example, 'Believes there is a conflict between IBE539 and MON904 – there isn't'.

The protocol data along with the table of interventions include all the information, expected to be required to construct (for, example, all the TOPH models) and to apply (by means of the diagnosis method), the diagnostic design knowledge.

Before the TOPH diagnosis method can be applied to diagnose design problems, associated with operator planning horizons, an integrated model for rvATM needs to be constructed. The integrated model includes models, corresponding to all the TOPH frameworks, identified earlier and derived by means of the TOPH theory from the protocol data and the table of controller interventions. Table 2 shows extracts from the rvATM integrated model for aircraft IBE550. It integrates work system-related models (Columns 1-5) with domain related-models (Columns 6 and 7).

Column 1 shows a model of the goals of the rvATM worksystem (that is, controllers (planning and radar) and devices (flight progress strips (FPSs) and radar)) For example: '(A) Intervention IBE550 heading 39 at ENSAC', that is, change aircraft IBE550's heading to 39. Letters denote different worksystem goals (and so tasks) and numbers denote the sequence of sub-tasks (and so sub-goals).

Column 2 shows a model of the controllers' behaviours. The model is expressed in terms of the TOPH operator architecture, which comprises physical architecture (for example, 'head' and 'hands') and mental architecture (for example, 'working memory', long-term memory, and goal store'). The architecture also includes 'process structures' (for example, 'search for' and 'form goal') and 'representation structures' (for example, categories of aircraft – 'active', 'expeditious' and goals – 'establish', 'amend', and 'intervene'). Behaviours occur when process structures are activated in conjunction with representation structures (for example, the physical behaviour – 'search for (aircraft) and the mental behaviour 'encode' (aircraft category)). Physical behaviours can be observed on the video recording, for example, a controller head movement towards the radar, indicating a 'search for' (aircraft) behaviour. Mental behaviours are inferred, for example, 'form goal'. Applying this architecture to the protocol data and the table of controller interventions identifies the controller behaviours, intended to achieve the worksystem goals of Column 1, for example, 'HIGHLIGHT IBE550) FPS; PULLDOWN: change heading; SELECT: 39; CATEGORISE: IBE550; POPGOAL: (B).

Column 3 shows a model of the controllers' representation of the domain of rvATM. The model derives from the TOPH framework of mental categories for the aircraft being managed, applied to the protocol data and the table of controller interventions. Categories include: 'incoming/ active'; 'safe/ unsafe'; 'expeditious/ unexpeditious'; etc. Categories in turn derive from domain attribute values, such as aircraft: radar position; altitude; speed; heading etc. Examples from Column 3 are: 'IBE550 (from) active safe expeditious to active safe unexpeditious (heading) aircraft'. The controller's representation of the domain may be true or false. For example, here, the representation is true.

Column 4 shows a model of the controllers' representation of the devices. The main rvATM devices are the FPSs (electronic, that is, displayed) and the radar. For, example, 'IBE550 FPS, Heading 39'. The controller's representation of the domain may be true or false. Here, if the controller's representation of the FPS is 45, then the representation is false.

Worksystem goals	Controller behaviour	Controller rep (domain)	Controller rep devices	Device behaviour	Product goal	Aircraft transformation
(A) Intervention IBE550 Heading 39 at ENSAC Planning/Execution	HIGHLIGHT: IBE550 FPS PULLDOWN: change heading SELECT: 39 CATEGORISE: IBE550 POP GOAL: (B)	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
7 minutes later						
(C1) Establish future intervention IBE550 (P5) Give IBE550 heading TERNI when KLM358 has passed Planning/Execution	SEARCH FOR: IBE550 radar/SAU ENCODE/CAT: IBE550 SEARCH FOR: KLM358 radar/SAU ENCODE/CAT: KLM358 PROBLEM SOLVE: IBE550 IBE550 Position SAU IBE550 Altitude: 310 IBE550 Heading: 39 KLM358 Position LFBA KLM358 Altitude 310 KLM358 Heading: Terni ** Give IBE550 heading TERNI when KLM358 has passed FORM GOAL: (P5) Intervention IBE550 heading TERNI when KLM358 has passed SUSPEND GOAL: (P5) POP GOAL: (C1)	IBE550 Position SAU IBE550 Altitude: 310 IBE550 Heading: 39 KLM358 Position LFBA KLM358 Altitude 310 KLM358 Heading: TERNI				
3 minutes later						
(F) Establish future safety of IBE550 Monitoring	SEARCH FOR: IBE550 radar/VELIN ENCODE/CAT: IBE550 SEARCH FOR: KLM358 radar/VELIN ENCODE/CAT: KLM358 PLAN DECAY (P5) POP GOAL: (F) FORM GOAL: (F1) Intervention IBE550 heading POITIERS	IBE550 Position VELIN IBE550 Altitude: 310 IBE550 Heading: 39 KLM358 Position LFBE KLM358 Altitude 310 KLM358 Heading: TERNI				
(F1) Intervention, IBE550 Heading Poitiers Planning/Execution	HIGHLIGHT: IBE550 FPS PULLDOWN: change heading SELECT: POI CATEGORISE: IBE550 POP GOAL: (F1)	IBE550, heading 39 changing IBE550, (from) active safe unexpeditious to active safe expeditious (heading) aircraft	IBE550 FPS selected IBE550 FPS, Heading POI	IBE550 FPS highlighted IBE550 FPS Heading POI Radar Velin, IBE550, Heading POI	(F) Intervention Improve QPM for IBE550 event vector	IBE550 Progress better Fuel Use better Safety same Exit better

Table 2 Extracts from rvATM Integrated Model for aircraft IBE550

Column 5 shows a model of the rvATM device behaviours, with which the controller's behaviours (Column 2) interact. A comparison between Column 2 and Column 5, that is, the interaction between the controller's behaviours and the device behaviours, indicates their appropriateness for achieving the worksystem goals (Column 1). For example, 'SELECT 39' behaviour by the controller, interacting with 'IBE550 FPS heading 39' behaviour is appropriate for the '(A) Intervention IBE550 heading 39 at ENSAC'. Column 5 completes the models, related to the worksystem. Columns 6 and 7 relate to the domain.

Column 6 shows a model of the product goal achievement, which expresses the effect of an intervention on the state of the associated aircraft. The product goal achievement relates to the worksystem's goals, which appear in Column 1. For example, the worksystem goal 'F1' of 'Intervention IBE550 heading Poitiers' (Column 1), associated with the controller's behaviour of 'PULLDOWN: change heading' (Column 2), reflects the controller's representation of the domain, 'IBE550 heading 39' (Column 3) and the devices, 'IBE550 FPS, heading POI' (Column 4) and achieves the product goal, 'Intervention improves QPM (Quality Progress Management' (Column 5). QPM refers to progress through the sector and fuel use.

Column 7 shows a model of data from the domain framework about the state of each aircraft, corresponding to the TOPH performance framework. The two highest level states are 'safe' (for example, not in conflict with other aircraft) and 'expeditious' (for example, moving through the sector in a timely manner). In the case of IBE550, Intervention F1 (Column 1), changing its heading to beacon Poitiers, in turn transforms its perceived state from 'active, safe, unexpeditious' to 'active, safe, expeditious' (Column 3) and its aircraft transformation to 'Fuel use better; safety same; exit better'.

This completes a description of the integrated rvATM model of worksystem and domain. The model is a pre-requisite for applying the TOPH diagnostic design knowledge to rvATM and so to its validation.

Before the TOPH diagnosis method can be applied to diagnose rvATM design problems, associated with operator planning horizons, there is a need to construct the controller's planning horizons. The latter are derived from the rvATM integrated model (Table 2).

Following TOPH, controller tasks can be divided into those of: administration; (for example, updating the FPSs); monitoring (for example, checking whether two planes are in conflict); and planning/execution (for example, specifying a future intervention, which is to be carried out) (see Table 2, Column 1). Planning horizons can be constructed only for planning/execution tasks. A plan is a mental representation structure, associated with mental process structures (for example, 'form'; 'discard'; 'decay' etc), giving rise to planning behaviours. Plans can have three different outcomes: 'plan and decay' (that is, plan with unintended non-execution); 'plan and discard' (that is, plan with intended non-execution); and 'plan and execute' (that is, plan with intended execution). Planning horizons are constructed on the basis of information, associated with: the controller; the devices; the plan; its extension (over time) and its adequacy (to achieve worksystem's goals). The data required to construct the planning horizon are extracted from the protocol data and the integrated model (Table 2). The rvATM planning horizon for aircraft IBE550 (see Tables 1 and 2) is shown in Table 3.

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	

Table 3 Planning horizon for aircraft IBE550

Column 1 of Table 3 shows the controller's encoding of IBE550, for example, 'Heading 45 altitude 310' and 'Position ENSAC'. Column 2 shows the controller's interventions, for example, 'IBE550 Heading change POI' (Poitiers). Column 3 shows the aircraft category, for example, IBE550, as 'Active, safe, aircraft, unexpeditious heading'. Column 4 shows the plan/execution, for example, 'Give Heading TERNI after KLM358 has passed' – a plan and 'change heading to POI' (Poitiers) – an execution.

A plan's 'extension' expresses its projection into future time. In the case of IBE550 from 'Give heading TERNI, after KLM358 passed' to 'Change heading to POI (Poitiers) (Column 4, Table 4). The extension can also be identified in terms of the controller's behaviours (Column 2) and representations of the domain (Column 3) – both in Table 2. A plan's adequacy expresses the quality of work, that is effected by the worksystem, if the plan is executed, that is, in terms of 'safe' and 'expeditious' aircraft. In the case of IBE550, it is a 'Safe unexpeditious (heading) aircraft', when the plan is formed, which would have resulted in a 'safe expeditious (heading) aircraft', had the plan been carried out. Thus, the plan was adequate. However, the plan was not executed. It was either forgotten or discarded. Effectiveness expresses the adequacy of a plan's extension. Here, the controller's plan for IBE550 was effective, when formed. According to TOPH, controller (resource) costs (that is, mental workload) are incurred, if plans are formed, but not executed. The controller costs, associated with the IBE550 plan of 'Give Heading TERNI after KLM358 passed' serve no purpose (in the event, the plan is either forgotten or discarded, the aircraft being changed to POI (Poitiers) and not TERNI) and so are wasted.

Given the rvATM protocol data and the integrated model (Table 2) and the planning horizon (Table 3), it is now possible to proceed with the application of the diagnosis method. The latter comprises four stages: 1. Identify problem; 2. Analyse planning horizon; 3. Extract required data from the integrated model; and 4. Generate causal theory. A design problem is considered to exist, when actual and desired performance differ, for example, if aircraft safety is violated or if aircraft expedition is too low (see Table 2, Column 5). The planning horizon provides an overview of the design problem and analysis thereof supports identification of whether planning is involved or not. For example, safety violation may have been rectified, or an adequate plan may have been formed, but have decayed in memory; re-planning may have taken place; or no plan may have been constructed. The planning horizon identifies relevant information from the integrated model to support causal theory generation. Such information may include particular behaviours and patterns of involvement of mental structures, both representation and process. Causal theory generation identifies the reasons for the design problem, as it relates to planning, for example, involvement of the mental structures and behaviours of the controller the forgetting of a plan or the failure to update the FPSs) or the structures and behaviours of the worksystem devices (,aircraft display on the radar and flight plan display on the FPSs). The causal theory would be expected to suggest possible design solutions to the design problems or to provide the basis for reasoning about such problems and solutions.

In the case of rvATM:

Problem Identification: An intervention with aircraft IBE550 has produced poor quality of work (the performance parameters for progress, fuel use and exit have become worse – see Table 2, integrated model, Column 7).

Analysis of Planning Horizon: Consultation of the planning horizon for IBE550 and establishing, that the problem arose, due to an intervention changing its heading to 39 (the aircraft was predicted to be unsafe). The planning horizon shows that the controller later made a plan to rectify this problem by changing IBE550 to a different heading. This plan decays (or is discarded). However, another plan is formed to rectify the situation and is executed to produce better quality for progress, fuel use and exit values.

Data extraction from Integrated Model: 'IBE550 heading change 39. IBE550 from active safe expeditious aircraft to active safe unexpeditious (heading) aircraft'.

'Problem solve IBE550. Give IBE550 Heading TERNI when KLM358 has passed'.

'Intervention IBE550 heading POI' (Poitiers).

Causal Theory Generation: The first intervention with IBE550, changing heading to 39, is carried out in response to a predicted conflict with KLM358, at a later time. This new IBE550 heading affects the performance of the aircraft with regards to progress and fuel use, as the aircraft is changed to a heading that takes the aircraft on a longer route, than that originally specified by the FPS. The controller knows, that at some stage in the scenario, he must change the heading of IBE550 back to its original airway, as IBE550 must exit at the beacon specified by the FPS. Later in the integrated model, the controller changes the heading of KLM358, which reminds him that IBE550 is not on its specified heading, and that he must change the heading back to that on the FPS, so that IBE550 will leave the sector at the correct exit beacon, once aircraft KLM358 is safely passed. The controller states that he will send IBE550 to TERNI at this time. TERNI is the last beacon on IBE550's original airway. However, when the controller realises that KLM358 has now passed and decides to carry out an intervention on IBE550 to put it back on its airway, and thus to increase its performance, he decides to send it to POI. When the controller makes the heading change he has obviously forgotten that his original plan was to send IBE550 to TERNI. POI is the last beacon specified on the FPS of IBE550. Looking at the FPS for IBE550 has prompted the controller to send IBE550 to POI. The original plan to send IBE550 to TERNI was made when the controller was looking at the radar. The aircraft information shown on the radar does not display the exit beacons. The intervention to send IBE550 to POI, will improve performance: fuel use better, progress better and exit better. To solve this design problem the controller would have to take more account of aircraft progress and fuel use, when planning interventions. At this time, what is of primary import to the controllers is safety and little notice is taken of other parameters. Information on an aircraft's fuel use could be displayed and a prompt issued to controllers to remind them of such reductions in performance. The prompt might be to direct the controllers to return an aircraft to its original airway more quickly, thus enabling performance parameters to become as desired.

Diagnosis Evaluation

The rvATM diagnosis is evaluated here analytically. First, the diagnosis is considered to be of the rvATM system (see earlier), as supported by the controller interventions, observed by video and shown in Table 1 (and elsewhere, in the protocol data). Second, the diagnoses is of a rvATM design problem, as identified by Column 7 of Table 2, specifically and the table more generally. Last, the diagnosis relates to planning, as supported by the planning horizon (Table 3) and the causal theory, developed on the basis of the integrated model and the planning horizon. The rvATM diagnosis is, thus, considered to meet the requirements of being a design problem, associated with controller planning.

Correct Operationalisation of the Design Knowledge

The TOPH design knowledge is judged to be correctly operationalised, according to the requirements, set out earlier. First, the diagnosis is complete, inasmuch as it corresponds to the complete application of the diagnosis method, that is, the four stages. Second, the diagnosis is consistent with the planning horizon (Table 3), which is in turn consistent with the integrated model (Table 2), which is in turn consistent with the controller's interventions (Table 1) and the protocol data. Third, the domain, worksystem and performance models of the integrated model (Table 2) are applied to the planning horizon construction and to the 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, for example, plan extension and adequacy. Last, the case-study manager, familiar with TOPH, checked its operationalisation for correctness.

Design Knowledge Validation

The application of the TOPH diagnostic design knowledge is considered to meet the validation requirements, set out earlier. First, the design knowledge was operationalised, that is, the already conceptualised design knowledge was applied in the case-study to an ATM simulation, rvATM, more complex than rATM. 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 against the validation. The latter can, then, be considered only partial. Some examples of the difficulties are shown in Table 4. Last, TOPH was generalised over rvATM, a more complex simulation than rATM (see earlier). One might be tempted to claim that TOPH is sufficiently scaled up to accommodate rvATM. However, in the absence of well specified relations between ATM, rATM and rvATM, it is perhaps safer to claim that TOPH has been 'scaled across' from rATM to rvATM.

Case-study Success

The case-study is considered a success, in the terms set out earlier. That is, rvATM, more complex; but equally well-defined and observable as rATM, is judged to fall within the scope of TOPH diagnostic design knowledge. Although the case-study is successful, the validation of TOPH is only partial, because of the validator's difficulties in its application (Table 4).

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
125	7.2.1.1 whole section	When forming the integrated model, the syntax for representing an intervention is difficult to apply, as FPSs in rvATM are not updated manually as in rATM so an intervention may have taken place and not been identified in the data	Check protocol with electronic printout of the scenario, being analysed and construct a table identifying all interventions	Solution implemented. Table of interventions constructed (see Table 1 earlier)	Possibly include in new version of TOPH for use when simulation uses electronic strips

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

DISCUSSION AND CONCLUSIONS

The primary aim of this paper is to report a successful case-study, which attempts to validate diagnostic design knowledge, both substantive and methodological, in the form of the TOPH theory, as applied to ATM, in the form of the rvATM simulation. The case-study is considered successful, inasmuch as rvATM, a more complex ATM simulation than rATM, is judged to fall within the scope of TOPH. The latter is, thus, considered to be partially validated, inasmuch as it was operationalised, tested and generalised (over rvATM). The validation, however, was only partial, because the validator experienced a number of difficulties in the application of TOPH, some of which needed the support of the case-study manager for their resolution. These difficulties can best be understood as design problems for TOPH and their solution constitutes a requirement for future research. A more effective version of TOPH would result.

The secondary aim of the case-study is to identify general problems of applying design knowledge, derived from research. The main problem, encountered by the validator, was the very complexity of the diagnostic design knowledge itself. The complexity resided in: the rvATM simulation; the associated controllers' interventions; the number and variety of models, required by the integrated model; and the reasoning, involved in applying the diagnosis method. The problem occurred, inspite of the details provided by TOPH (unusually fulsome for research) and the provision of methodological knowledge at all (most research omits how to apply frameworks, models, experimental results for design). The problem also occurred, despite the validator being a trained and experienced HCI researcher and the case-study manager, being involved in the development of TOPH.

Initial reaction to this general problem of applying research prompts the following suggestions for its possible alleviation. First, researchers should be encouraged to include methodological knowledge. If Cognitive Ergonomics is about designing for effectiveness, which requires the identification and solution of design problems, then these latter practices need to be supported by research (Long and Dowell, 1989; Dowell and Long, 1998).

Second, a distinction might usefully be made between research reports for researchers and reports for the application of the knowledge by practitioners. Rather than thinking in terms of a dichotomy, between research and its application, we might better think in terms of a gradient, between research and practice.

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.

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.

Dowell, J., and Long, J. (1989) Towards a conception for an engineering discipline of human factors. Ergonomics, 32, 1513-35.

Dowell, J. and Long, J. (1998). Conception of the cognitive engineering design problem. *Ergonomics*, 41(2), 126-139.

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

Long, J. and Dowell, J. (1989). Conceptions for the Discipline of HCI: Craft, Applied Science & Engineering. In *Proc. BCS HCI SIG Conference*, Nottingham, UK, 9-32.

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.

Timmer, P., and Long, J. (1996). Integrating domain and worksystem models: an illustration from air traffic management. In Proc. Conference on Domain Knowledge in Interactive System Design, 194-207.

Timmer, P., and Long, J. (2002). Expressing the effectiveness of planning horizons. Le Travail Humain, 65(2), 103-126.