POSITION PAPER

Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills

Rhona Flin*, Lynne Martin*, Klaus-Martin Goeters**,
Hans-Jürgen Hörmann**, René Amalberti***, Claude
Valot*** and Herman Nijhuis****
*University of Aberdeen, Scotland
**DLR Hamburg, Germany
***IMASSA, France
****NLR, Netherlands

Abstract

Crew Resource Management (CRM) courses are designed to teach pilots about non-technical (cognitive and social) skills that are essential for effective and safe flight operations. This article presents a summary of the empirical development of the European taxonomy of pilots' non-technical skills (NOTECHS) and associated rating method. It describes the system components and the experimental validation. The system has four Categories: *Co-operation, Leadership and Managerial Skills, Situation Awareness, Decision Making,* each subdivided into Elements and behavioural markers. The latter are examples of effective and ineffective behaviours supporting the evaluation and debriefing, as well as giving indications for retraining, if this is required. Operational principles for fair and

Correspondence: Rhona Flin, Industrial Psychology Research Centre, University of Aberdeen, King's College, Old Aberdeen AB24 2UB, UK or r.flin@abdn.ac.uk

objective use of the NOTECHS system and training guidelines for raters are outlined.

Introduction

The aviation community has put considerable emphasis in the last decade on flight crew members' non-technical skills as a crucial factor for enhanced safety. Crew Resource Management (CRM) courses are designed to teach pilots about these 'cognitive and interpersonal skills needed to manage the flight within an organized aviation system' (CAA, 2003, p1). The CRM training appears to result in effective transfer of the desired behaviours to the flight deck (Gregorich and Wilhelm, 1993; O'Connor, Flin and Fletcher, 2002a; Salas, et al, 2001). Additional mandatory regulations have recently appeared in Europe, that require a more formal incorporation of non-technical (CRM) skills evaluation into all levels of training and checking flight crew members' performance.

The European Joint Aviation Authorities (JAA) require the training and assessment of Crew Resource Management (CRM) skills as set out in the CRM regulations included in JAR OPS (2001) 1.940, 1.945, 1.955, and 1.965, asking for an evaluation of flight crews' CRM skills. For example, '*the flight crew must be assessed on their CRM skills in accordance with a methodology acceptable to the Authority and published in the Operations Manual. The purpose of such an assessment is to: provide feedback to the crew collectively and individually and serve to identify retraining; and be used to improve the CRM training system.*' (1.965).

This legislation resulted from a desire by the JAA from the mid 1990s, to achieve a generic method of evaluation of non-technical skills throughout the JAA countries and JAA operators. Such a generic method would minimize cultural and corporate differences, and maximize practicability and effectiveness for airline instructors and examiners. As a consequence, in 1996, the JAA Project Advisory Group on Human Factors initiated a project group that was sponsored by four European Civil Aviation Authorities (Germany, France, Netherlands, UK). A research consortium consisting of members from DLR (Germany), IMASSA (France), NLR (Netherlands) and University of Aberdeen (UK) was established to work on what was called the NOTECHS (Non-Technical Skills) project. The group was required to identify or to develop a feasible and efficient methodology for assessing pilots' non-technical (CRM) skills. (For the purpose of the project, these were defined as the cognitive and social skills of flight crew members in the cockpit, not directly related to aircraft control, system management, and standard operating procedures - SOPs). The design requirements were (i) that the system was to be used to assess the skills of an individual pilot, rather than a crew, and (ii) it was to be suitable for use across Europe, by both large and small operators, i.e. it was to be culturally-robust. This article presents a summary of the

development of the European taxonomy of pilots' non-technical skills (NOTECHS) and describes the system components. For a more comprehensive account, see the final project report (van Avermaete and Kruijsen, 1998).

Method

After reviewing alternative methods, it became obvious for various reasons that none of the existing systems could be adopted in their original form. Nor did any single system provide a suitable basis for simple amendment that could be taken as an Acceptable Means of Compliance under the scope of JAR. The reasons for this were that the existing systems were either too complex to be used on a pan-European basis, or too specific to a particular airline, or were designed to assess crews rather than individual pilots (e.g. University of Texas LLC (Line/Line Oriented Simulation Checklist system, Helmreich et al, 1995). Therefore, the NOTECHS group decided that it would have to design a new taxonomy of non-technical skills, (see Seamster and Kaempf, 2001 for a good description of techniques that can be used for this).

The identification of the basic set of non-technical skills for the NOTECHS system consisted of three interleaved phases of work.

(i) Review of existing systems to evaluate proficiency in pilots' non-technical skills to identify common categories and elements of behaviour (NOTECHS Report, work package 2; see also Flin and Martin, 1998; 2001).

(ii) Literature search for relevant research findings relating to key categories of non- technical skills identified in existing systems (NOTECHS Report, work package 3).

(iii) Extended discussions with subject matter experts, in this case - KLM pilots experienced in evaluating non-technical skills, at NOTECHS working group meetings. Particular attention was paid to two of the principal CRM skills frameworks, namely the KLM WILSC/ SHAPE systems (Antersijn and Verhoef, 1995; KLM, 1996) and the University of Texas LLC system (Helmreich, et al, 1995). The systems in use by Air France, RLD (Dutch CAA) and Lufthansa (Quick Reference System) to evaluate pilots' non-technical skills during checks were also reviewed.

It is important to note that the required system was to be usable on a pan-European basis. This means that the majority of instructors/ examiners would not have English as their native language and would be based in very small companies, as well as in the large airlines. Consequently, the following design principles were used to guide the final choice of components and descriptor terms.

(i) The basic categories and elements should be formulated with the maximum mutual exclusivity. This is only achievable to a certain degree, given the interdependence of the various non-technical skills in flight deck operations.

(ii) A rule of parsimony was applied, in that the system should contain the minimum number of categories and elements in order to encompass the critical behaviours.

(iii) The terminology used should reflect everyday language for behaviour, rather than psychological jargon.

(iv) The skills listed at the behaviour level should be directly observable in the case of social skills or could be inferred from communication, in the case of the cognitive skills.
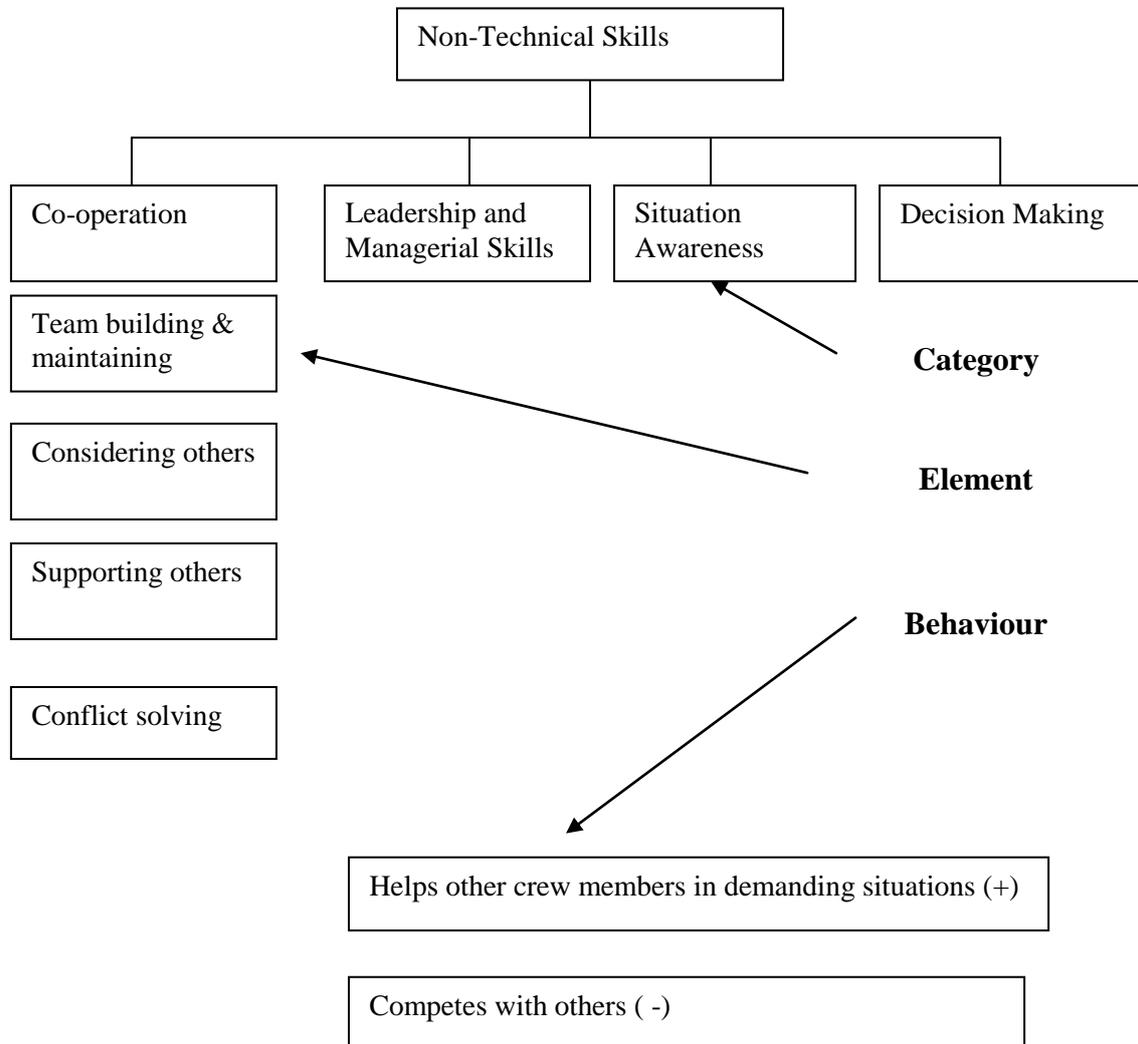
On the basis of these design principles, a prototype system was developed from non-technical skill set identified from the three work packages described above. Following extended discussions at two project meetings, and a subsequent two-day workshop involving four psychologists from the team, a draft taxonomy was prepared. This was circulated to group members for comments. Then the revised components and structure were extensively reviewed at a group meeting with the whole team, including the KLM pilots.

Structure of the NOTECHS system

The NOTECHS framework consists of four main categories: *Co-operation, Leadership and Managerial Skills, Situation Awareness, Decision Making,* each of them being subdivided into elements and behavioural markers (see figure 1). The latter are examples of effective and ineffective behaviours supporting the evaluation and debriefing, as well as giving indications for retraining, if this is required.

The four primary categories subdivide into two social skills (*Co-operation; Leadership and Management*) and two cognitive skills (*Situation Awareness; Decision Making*); it was judged to be unnecessary to add another level to the system by introducing this distinction explicitly. It should be noted that for the social skills, behaviours are generally in the form of communication (verbal and non-verbal) which can be directly observed. The cognitive skills are not directly observable since they do not directly materialise in overt behaviour but occur 'in the head of the pilot'. Hence for evaluation purposes, these cognitive processes must be inferred from observable behaviours, (e.g. specific actions or verbalisations - thus when a pilot states a decision, the observer can infer that some option selection has taken place; discussion of alternative divert airports reveals that option generation and comparison processes are being employed). Not all non-technical rating systems include cognitive skills explicitly, due to the indirect methods which must be used for their evaluation. Nevertheless, a basic tenet of CRM is that pilots should communicate in a manner that reveals their mental models and thinking processes to the other crew members. Thus it was deemed appropriate to evaluate these critical cognitive skills which have been

shown to contribute to flight safety, and which are taught to pilots as fundamental components of CRM.

```
                        ┌─────────────────────┐
                        │ Non-Technical Skills │
                        └─────────────────────┘
              ┌───────────────┬────────────┴──────┬─────────────────┐
    ┌──────────────┐  ┌──────────────┐  ┌──────────────┐  ┌──────────────┐
    │ Co-operation │  │ Leadership and│  │ Situation    │  │ Decision     │
    │              │  │ Managerial    │  │ Awareness    │  │ Making       │
    │              │  │ Skills        │  │              │  │              │
    └──────────────┘  └──────────────┘  └──────────────┘  └──────────────┘
```

┌──────────────────────┐
│ Team building &   │
│ maintaining   │
└──────────────────────┘   **Category**

┌──────────────────────┐
│ Considering others   │
└──────────────────────┘   **Element**

┌──────────────────────┐
│ Supporting others   │
└──────────────────────┘

┌──────────────────────┐   **Behaviour**
│ Conflict solving   │
└──────────────────────┘

┌──┐
│ Helps other crew members in demanding situations (+) │
└──┘

┌──┐
│ Competes with others (-)   │
└──┘

Figure 1   Basic structure of the NOTECHS system

Two further points should be noted with regard to the system components. The category 'Communication' is featured in a number of systems but is not shown in NOTECHS as a separate category. This is because communication skills are inherent in all four categories and the listed behaviours all involve communication. A category of 'Personal Awareness' skills (e.g. coping with stress or fatigue) was considered but rejected due to difficulties in observing, or inferring except in the most extreme cases.

In relation to the four categories, a number of derived elements were examined and for each element a series of indicative behaviours were identified. At the

elemental level, this set was compared against the KLM SHAPE system and the LLC system to confirm that essential elements had been encompassed. In the final version, three to four elements for each of the four categories were selected, giving a total of 15 elements (see table 1). For each element a number of exemplar behaviours were included. The exemplar behaviours were phrased as generic (e.g., closes loop for communications), rather than specific (e.g., reads back to ATC), to give an indication of type, and to avoid specifying particular behaviours which should be observed. This should also ensure that the system is as widely applicable as possible. Accompanying guidance notes to the system provide further details.

Table 1   Categories and elements of NOTECHS

Category	Elements
1. Co-operation	- Team-building and maintaining
	- Considering others
	- Supporting others
	- Conflict solving
2. Leadership and Managerial Skills	- Use of authority and assertiveness
	- Providing and maintaining standards
	- Planning and co-ordination
	- Workload management
3. Situation Awareness	- Awareness of aircraft systems
	- Awareness of external environment
	- Awareness of time
4. Decision MMaking	- Problem definition and diagnosis
	- Option generation
	- Risk assessment and option selection
	- Outcome review

Description of the four categories

Co-operation

Defined as 'the ability to work effectively in a team/ crew'. *Co-operation* requires team building and maintaining, so that co-operative actions are based on mutual agreement by crew members in a positive group climate. Such a climate is also created by factors like consideration / support of other crew members and conflict solving skills. *Co-operation* deals with the question of how people function as a working group. It does not refer to the work itself, such as the quality/quantity of output.

Good co-operation is largely dependent on active and open communication between crew members. However, communication is not a basic element of co-operation alone, as mentioned above, it is a general mediator of all four non-technical skills Categories. Some aspects of communication may fall more into one than another category, thus these particular aspects become secondary descriptors. The aspects of communication which belong to *Co-operation* deal with building a climate for open communication, sensitivity for different aspects of messages (e.g. also the emotional component) and awareness of the difference of verbal vs. non-verbal communication.

The category *Co-operation* comprises four elements (see table 2).

- *Team building and maintaining:* Establishing positive interpersonal relations and active participation of crew members in fulfilling the tasks.
- *Considering others:* Acceptance of others and understanding their personal condition.
- *Supporting others*: Giving help to other crew members in cases where they need assistance.
- *Conflict solving*: Articulation of different interpersonal positions with suggestions for solutions.

The two interpersonal categories of the NOTECHS system, namely *Co-operation* and *Leadership and Managerial Skills* are overlapping to a certain degree, since both refer to group processes. The solution for reaching a clearer distinction is that in the NOTECHS system, co-operation involves team building and maintaining (in some other systems this is subsumed in leadership and managerial skills). On the other hand, the aspects of coordination and responsibility, although often discussed as parts of co-operation, became elements of leadership and managerial skills in the NOTECHS terminology. The conceptual difference is that *Co-operation* is concerned with mutual assistance and team atmosphere during work, while *Leadership and Managerial Skills* covers all aspects of initiative, coordination and goal setting. Considering and supporting others could be grouped together as one element, since in practice both aspects are very closely interrelated, but for the sake of clarity these concepts were separated.

Leadership and managerial Skills

Defined as: 'effective leadership and managerial skills achieve joint task completion within a motivated, fully functioning team through coordination and persuasion'. The core of effective leadership is to set the highest priority on the joint completion of a given task. Leadership responsibilities include the active and goal-directed coordination of the working activities within the crew. This is always a reciprocal process. Without complementary behaviour of the crew, leadership behaviour is less effective. All crew members are expected to dedicate their efforts and initiative to the safe and efficient achievement of the flight goals.

However, the final and legal responsibility for the operation on the whole is undivided, resting with the pilot in command (PIC). Crew responsibilities include monitoring and challenging each other whenever differences in concepts or actions are perceived. Within the scope of delegated tasks crew members have the same responsibilities as the PIC.

Table 2   Co-operation category: elements and behavioural markers (examples).

Element	Good practice	Poor practice
Team building and maintaining	Establishes atmosphere for open communication	Blocks open communication
	Encourages inputs and feedback from others	Keeps barriers between crewmembers (CM)
	Does not compete with others	Competes with others
Considering others	Takes notice of the suggestions of other CM even if s/he does not agree	Ignores suggestions of other CM
	Takes condition of other CM into account	Does not take account of the condition of other CM
	Gives personal feedback	Shows no reaction to other CM
Supporting others	Helps other CM in demanding situations	Hesitates to help other CM in demanding situations
	Offers assistance	Does not offer assistance
Conflict solving	Keeps calm in interpersonal conflicts	Overreacts in interpersonal conflicts
	Suggests conflict solutions	Sticks to own position without considering a compromise
	Concentrates on what is right rather than who is wrong	Accuses other CM of making errors

The International Civil Aviation Organisation (ICAO) has defined a leader as 'a person whose ideas and actions influence the thought and the behaviour of others. Through the use of example and persuasion, and an understanding of the goals and desires of the group, the leader becomes a means of change and influence' (ICAO, 1989a). It is important to distinguish between leadership, which is acquired and authority, which is assigned. Leadership is one aspect of teamwork, and the success of a leader depends on the quality of his/her relationship in the team. The crew members should feel that they are an integral part of a well-run, well-organised operation in which their inputs are essential to reach commonly valued goals and overall success of the operation.

In comparison to the category *Co-operation, Leadership and Managerial skills* focuses more on the goal-directed initiative the respective crewmember is investing into management and coordination functions. This includes also positive influences on the motivation and morale of the crew. Whereas *Co-operation* is more focused on interactive processes without explicit appointed roles and independently of authority differences of the individuals.

The leader has a clear concept for the operation and provides general standards and directions for the completion of the different tasks. The tasks are allocated according to defined roles, specific experience, as well as to the present level of workload of the crewmembers. This concept for the operation is interactively briefed and always open for contributions from other crew members. In order to ensure proper support and the participation from all parts of the crew, active care is taken to establish and maintain closed loop communication. A leader motivates, activates, and monitors others and encourages the crew to monitor and challenge her/himself and each other in a non-degrading way.

It was decided to extract four major elements for *Leadership and Managerial skills* from the existing systems (see table 3). The skills themselves should be the same for leaders and followers. The way they are implemented in behaviour may differ according to position.

- *Use of authority and assertiveness:* Creating a proper challenge and response atmosphere. The given command authority of the PIC should be adequately balanced with assertiveness and crew member participation. If the situation requires, decisive actions are expected.
- *Providing and maintaining standards:* The compliance with essential standards (SOPs and others) for the task completion should be ensured. Therefore the crew should mutually supervise and intervene in case of deviations from standards. If the situation requires, it may be necessary to apply non-standard procedures. Such deviations should be announced and consultation should take place.
- *Planning and coordination:* In order to achieve high performance and to prevent workload peaks or dips, an appropriate concept for organised task sharing and delegation has to be established. Plans and intentions have to

be communicated so that the whole crew understands the goals and that the activities conducted by the crew do reflect proper coordination.

- *Workload management:* Clear prioritisation of primary and secondary operational tasks. Based on sound planning, tasks are distributed appropriately among the crew. Signs of stress and fatigue are communicated and taken into account as performance affecting factors. Available external and internal resources (including automation) are used to accomplish task completion in time.

Table 3   Leadership and Managerial skills category: elements and behavioural markers (examples)

Element	Good practice	Poor practice
Use of authority and assertiveness	Takes initiative to ensure crew involvement and task completion	Hinders or withholds crew involvement
	Takes command if situation requires, advocates own position	Passive, does not show initiative for decisions, own position not recognisable
	Reflects on suggestions of others	Ignores suggestions of others
	Motivates crew by appreciation and coaches when necessary	Does not show appreciation for the crew, coaches very little or too much
Providing and maintaining standards	Subscribes to SOPs, makes sure SOP compliance in crew	Does not comply to SOPs, does not monitor crew for SOP compliance
	Intervenes if task completion deviates from standards	Does not intervene in case of deviations
	With crew being consulted, deviates from standards if necessary	Deviations from standards are neither announced nor consulted
	Demonstrates will to achieve top performance	Does not care for performance effectiveness

Table 3   **Leadership and Managerial skills category: elements and behavioural markers (continued).**

Element	Good practice	Poor practice
Planning and co-ordination	Encourages crew participation in planning and task completion	Plans only for him/herself, crew not involved
	Plan is clearly stated and confirmed	Intentions not stated or confirmed
	With crew being consulted, changes plan if necessary	Changes plan without informing crew or follows plan blindly
	Clearly states goals and boundaries for task completion	Goals and boundaries remain unclear
Workload management	Distributes tasks among the crew, checks and corrects appropriately	Flying 'solo' without other crewmembers involved
	Secondary operational tasks are prioritised to retain sufficient resources for primary flight duties	Secondary operational tasks interfere with primary flight duties
	Allots adequate time to complete tasks	Workload is increased through inadequate planning
	Notifies signs of stress and fatigue	Ignores signs of stress and fatigue

Situation awareness

Situation awareness can be defined as 'one's ability to accurately perceive what is in the cockpit and outside the aircraft' (ICAO, 1989b, p13); or simply as 'knowing what is going on'; or more precisely as, 'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future' (Endsley, 1995, p36). Shrestha, et al (1995, p.52) suggest an alternative version as a result of their review of ten definitions of situation awareness: 'situation awareness is a

dynamic, multifaceted construct that involves the maintenance and anticipation of critical task performance events. Crew members must also have temporal awareness, anticipating future events based on knowledge of both the past and the present. It is crucial that individuals monitor the environment so that potential problems can be corrected before they escalate.

Table 4   Situation Awareness category: elements and behavioural markers (examples).

Element	Good practice	Poor practice
Awareness of aircraft systems	Monitors and reports changes in systems' states	Does not ask for updates
	Acknowledges entries and changes to systems	Does not signal awareness of changing systems
Awareness of external environment	Collects information about environment (position, weather and traffic)	Does not enquire about environmental changes
	Shares key information about environment with crew	Does not comment on relevant environmental factors, or is surprised by them
	Contacts outside resources when needed (to maintain situation awareness)	Operates a 'closed shop'
Awareness of time	Discusses time constraints with crew	Does not set priorities regarding time limits
	Discusses contingency strategies	Does not discuss relationship between past events and present/future
	Identifies possible future problems	Is surprised by outcomes of past events

ICAO (1989b) lists situation awareness topics that should be taught in CRM training:

- Total awareness of surrounding environment
- Reality vs. perception of reality
- Fixation
- Monitoring
- Incapacitation (partial/ total, physical, psychological)

Pilots sometimes define situation awareness as 'being ahead of your aircraft'. This implies that the pilot is prepared and is in control, has command of the aircraft and the current tasks. For all of the elements of situation awareness, the pilot needs to not only know their present state but also to be able to predict their future states, so anticipation covers this predictive aspect. Communication is the medium through which situation awareness can be assessed. It is not only *knowing something* that is crucial to good situation awareness, it is being able to fit this information into the right place in a mental model of the situation so that if it is important it triggers problem recognition. This definition stresses the ever-changing nature of the cockpit environment and the need for the crew to continually monitor and update their model of the situation by collecting information from all the sources that are available to them. The three Elements and behaviours in the NOTECHS Category *Situation Awareness* were selected to reflect this concept (see table 4).

- *Awareness of aircraft systems:* active knowledge of mode and state of systems, aircraft energy states (e.g. fuel).
- *Awareness of environment:* active knowledge of current and future position, weather, air traffic, terrain.
- *Awareness of time:* sense of available time and thinking ahead to consider future conditions and contingencies.

Decision Making

Decision Making is defined as 'The process of reaching a judgement or choosing an option'. This definition of decision making is not generally disputed in the aviation literature although it may be labelled aeronautical decision making (Kaempf and Klein, 1995) or pilot judgement. Pilot decision making does not just involve one strategy - different types of decisions are made at different times. Decision events differ enormously in what they demand of the crew, what options and supports exist in standard procedures and policies for making decisions and in features that may make the situation complex. Orasanu (1993), a NASA research psychologist, has studied the styles of decision making used by pilots in different situations. '...*crew decision making is not one thing. Crews make many different kinds of decisions, but all involve situation assessment, choice among alternatives, and assessment of risk.*' (p.138). Hence, pilots' decisions differ in the

degrees to which they call on various cognitive processes depending on the decision structure and task conditions. Whilst the captain has responsibility, he or she is supported by the other members of the team in cockpit decision making, therefore, crew decision making is managed decision making. Within the NOTECHS system, the Category *Decision Making* is narrower than pilot judgement as defined by Jensen (1996). His model encompasses decision making in its wider context, taking into account all the contributory factors of the individual as well as the situation. Rather, the focus for NOTECHS was on an information processing framework based on current research from applied cognitive psychology (e.g. Stokes, Kemper and Kite, 1997).

Table 5   Decision Making category: elements and behavioural markers (examples).

Element	Good practice	Poor practice
Problem definition and diagnosis	Gathers information to identify problem	Nature of problem not stated or failure to diagnose
	Reviews causal factors with other crew members	No discussion of probable causes
Option generation	States alternative options	Does not search for information
	Asks crew members for options	Does not ask crew for alternatives
Risk assessment and option selection	Considers and shares estimated risk of alternative options	Inadequate discussion of limiting factors with crew
	Talks about possible risks for action in terms of crew limits	No consideration of limiting factors
	Confirms and states selected option/ agreed action	Does not inform crew of decision path being taken
Outcome review	Checks outcome against plan	Fails to check selected outcome against goal

The elements which are integrated into *Decision Making* (see table 5) were derived principally from the FOR-DEC model (Hörmann, 1995) and the pilot decision making taxonomy developed by Orasanu (1993).

- *Problem definition and diagnosis:* gathering information and determining the nature of the situation. Considering alternative explanations for observed conditions.
- *Option generation:* formulating alternative approaches to dealing with the situation. The opportunity for this will depend on available time and information.
- *Risk assessment and option selection:* making a judgement or evaluation of the level of risk/ hazard in the alternative approaches and choosing a preferred approach.
- *Outcome review:* considering the effectiveness/suitability of the selected option against the current plan, once the course of action has been implemented.

The measurement of decision making follows the same lines as that for situation awareness. As decision making is a cognitive process, it is not possible to observe it directly and it has to be inferred from pilot behaviours. Again, this will be through communication - the crew telling each other what point their thinking has reached, stating diagnoses and courses of action, as well as discussing alternatives to undertake a risk assessment.

Rating system

The NOTECHS categories and elements are rated on a five-point scale – 'very good, good, acceptable, poor, very poor'. In addition an overall rating of 'acceptable/ not acceptable' is required. The five-point scale design was chosen after reviewing other systems and on the advice of subject matter experts (instructors/ examiners).

Operational principles for using the NOTECHS system

Five operational principles were established to ensure that each crewmember receives as fair and as objective an assessment as possible with the NOTECHS system.

1. *Only observable behaviour is to be assessed* - The evaluation must exclude reference to a crewmember's personality or emotional attitude and should be based only on observable behaviour. Behavioural markers were designed to support an objective judgement.
2. *Need for technical consequence* - For a pilot's non-technical skills to be rated as unacceptable, flight safety must be actually (or potentially) compromised. This requires a related objective technical consequence.

3. *Acceptable or unacceptable rating required* **-** The JAR-OPS requires the airlines to indicate whether the observed non-technical skills are acceptable or unacceptable.

4. *Repetition required* - Repetition of unacceptable behaviour during the check must be observed to conclude that there is a significant problem. If, according to the JAR-paragraph concerned, the nature of a technical failure allows for a second attempt, this should be granted, regardless of the non-technical rating.

5. *Explanation required* **-** For each Category rated as unacceptable the examiner must: a) Indicate the Element(s) in that Category where the unacceptable behaviour was observed. b) Explain where the observed NTS (potentially) led to safety consequences. c) Give a free-text explanation on each of the Categories rated unacceptable, using standard phraseology.

Judging behaviour is always more subjective than judging technical facts. NOTECHS has been designed to minimise ambiguities in the evaluation of non-technical skills. However there are several factors that can occur in the evaluation process. The first relates to the unit of observation, i.e. who is evaluated: the crew globally, the captain, or the co-pilot. The NOTECHS system is designed to be used to assess individual pilots. When an evaluation relates to individuals, a potential difficulty is to disentangle individual contributions to overall crew performance. But this difficulty already exists during checks when considering technical performance. NOTECHS does not magically solve this problem, but may serve to objectively point to behaviours that are related more to one crew member than the other, therefore allowing examiners to differentiate the judgement of the two crew members.

The second factor relates to any concern that raters are simply judging non-technical skills on the basis of intuition. Again, NOTECHS requires the instructor/ examiner to justify any criticisms at a professional level, and with a standardised vocabulary. Furthermore, a judgement should not be based on a vague global impression or on an isolated behaviour or action. Repetition of the behaviour during the flight is usually required to explicitly identify the nature of the problem.

The NOTECHS method is designed to be a guiding tool to look beyond failure during recurrent checks or training, and to help point out possible underlying deficiencies in CRM competence in relation to technical failures. NOTECHS is not intended to fail additional crew members during mandatory checks, or indeed on any other occasion, as compared to the present situation. The evaluation of non-technical skills in a check using NOTECHS should not provoke a failed (not acceptable) rating without a related objective technical consequence, leading to compromised flight safety in the short or long term. In the event of a crew member failing a check for any technical reason, NOTECHS can provide useful insights into the contributing individual human factors for the technical failure.

Used in this way, the method can provide valuable assistance for debriefing and orienting tailored retraining.

Preliminary test of NOTECHS: the JARTEL project

The prototype NOTECHS system offered a systematic approach for assessing pilots' non-technical skills in simulator and flight missions. Testing of the basic usability and psychometric properties of the NOTECHS system was then required. The Directorate General Transportation (DGTREN) of the European Community, in co-ordination with the JAA research committee tasked a consortium of five aviation research centres (NLR (N), DLR (G), University of Aberdeen (UK), DERA (UK) and IMASSA (F)) and four aviation business centres (Sofreavia (F), British Airways (UK), Alitalia (I), and Airbus) to test the NOTECHS method. A European research project JARTEL (Joint Aviation Requirements - Translation and Elaboration of Legislation) began in January 1998 and was completed in 2001. The main goals of this project were to assess:

- Usability of the NOTECHS system as an assessment tool.
- Reliability and validity of the assessment tool.
- Influence of cultural differences on the use of the NOTECHS system within Europe.

Experimental study

The experimental study was carried out using eight video scenarios filmed in a Boeing 757 simulator, with current pilots as actors. The scenarios showed simulated flight situations with predefined behaviours (from the NOTECHS elements) exhibited by the pilots of varying quality ('very poor' to 'very good' standard). The pilots' behaviours were rated using the NOTECHS systems by 105 instructors, recruited from 14 airlines in 12 European countries. The airlines represented large and smaller carriers within five different European cultural groups.

Each of the experimental sessions was conducted within an airline training centre. It began in the morning with a standard briefing on the NOTECHS method (the participants had previously been supplied with background information on the NOTECHS system), and a practice session. Questionnaires were also completed by the instructors, providing data on their background and experience. During the subsequent afternoon session, the captain's and first officer's behaviour in each of the eight cockpit scenarios was rated by the instructors using the NOTECHS score forms. At the end of the experimental session, a second questionnaire was given to the instructors for evaluating the NOTECHS rating process and material.

In summary, the results indicated that 80% of the instructors were consistent in their ratings, 88% of them were satisfied with the consistency of the method. On

average, the difference between a reference rating (established for benchmarking by consensus ratings in a set of trained expert instructors), and the instructors' ratings was less than one point on the five-point scale, confirming an acceptable level of accuracy. In the evaluation questionnaire, the instructors were very satisfied with the NOTECHS rating system, especially with the five-point scale (98%). Cultural differences (relating to five European regions) were found to be less significant than other background variables, such as English language proficiency, experience with non-technical skills evaluation, and different role perceptions of captain and first officer (see Hörmann, 2001) for details of the cultural analysis. Full details of the experimental method and the results can be found in the JARTEL project reports for work packages 2 and 3 (Sofreavia web site: sofreavia.com/jartel/) or see O'Connor, et al (2002).

A subsequent operational trial of NOTECHS was run with several airlines (see JARTEL work package 4 Report). It confirmed the applicability and feasibility of the system in real check events.

These first experimental and operational tests of the NOTECHS system showed that it was usable by instructors and appeared to have acceptable psychometric properties. These results were achieved with a minimal training period of half a day due to difficulties in recruiting experienced instructors to take part in the study, especially from the small companies. This level of training would be insufficient for using the NOTECHS system for regular training or assessment purposes. It is recommended that the basic training period is two full days or longer (depending on the level of previous experience of rating pilots' non-technical skills).

Training requirements for users of the NOTECHS system

Users of NOTECHS are certified flight instructors and authorized examiners. It is necessary to train all raters in the application of the method. NOTECHS presupposes sufficient knowledge of concepts included in the JAR-FCL theoretical program on human performance and limitations (JAR-FCL1.125/1.160/1.165-Theoretical knowledge instruction PPL/ATPL). No additional theoretical knowledge is required. Being current in CRM training and recurrent CRM is required, at least as a participant. Experience of CRM instruction is a facilitating factor for standardisation, but is not a prerequisite (see CAA 2003 for the current UK position on CRM Instructors and CRM Instructor Examiners). Most of the training effort should be devoted to the understanding of the NOTECHS methodology, the specific use of the evaluation grid, the calibration process of judgement and the debriefing phase. As the NOTECHS system is primarily used as a tool for debriefing and identification of training needs, then it is important to ensure that in debriefing an emphasis is placed on skill components, rather than more 'global' analyses of performance. (For general advice on training and other practical aspects of the use of behavioural marker

systems, such as NOTECHS, for evaluating non-technical (CRM) skills, see Baker and Dismukes, 2002; Baker, Mulqueen and Dismukes, 2001; Klampfer, Flin, Helmreich, et al., 2001.)

In the UK the CAA (2003, Chapter 7, p2) has recently published the following guidance on CRM skills assessment:

'For individual CRM skills assessment, the following methodology is considered appropriate:

- *An operator should establish the CRM training programme including an agreed terminology. This should be evaluated with regard to methods, length of training, depth of subjects and effectiveness.*
- *The CRM standards to be used (e.g. NOTECHS) have been agreed by crews, operators and regulators, and reflect best practice.*
- *The standards are clear, brief, and published (in the Operations Manual).*
- *The methodology for assessing, recording and feeding back has been agreed and validated.*
- *Training courses are provided to ensure that crews can achieve the agreed standards.*
- *Procedures are in place for individuals who do not achieve the agreed standards to have access to additional training, and independent third party appeal.*
- *Instructors and examiners are qualified to standards agreed by all parties, and are required to demonstrate their competency to the CAA or such persons as the CAA may nominate.*
- *A training and standardisation programme for training personnel should be established.'*

Conclusions

In summary, NOTECHS was designed as: (i) A professional pragmatic tool for instructors and authorised examiners; (ii) A tool to be used by non-psychologists; (iii) A tool using common professional aviation language, with the primary intention of debriefing pilots and communicating concrete directions for improvements. NOTECHS was not designed as: (i) A research tool (although it can be used for this purpose, see Goeters, 2002); (ii) A tool for judging flight-crew personality on the basis of instructors' or authorised examiners' personal opinions; (iii) A tool for introducing psychological jargon into the evaluation.

The preliminary evaluation of the NOTECHS system from the experimental and operational trials indicated that the basic psychometric properties were acceptable and that the method was usable and accepted by practitioners. Clearly a more extensive test of the psychometric quality of NOTECHS would be desirable but this would require a large data set collected under standardised conditions.

In response to the JAA requirements on evaluation of CRM skills, many airlines have now developed their own systems, several of which are at an advanced stage such as the KLM SHAPE system, the Lufthansa System 'Basic Competence for Optimum Performance' (Burger, Neb and Hörmann (2002), and the Alitalia PENTAPERF system (Polo, 2002). The Alitalia and Lufthansa systems have made use of the basic NOTECHS framework in the design of their own customised systems (Hörmann, Burger and Neb, 2002). Several other airlines are currently using NOTECHS or their own versions of it to complement their proficiency evaluation methods both in Europe (e.g. Finnair, Eastern Airways, Iberia) and beyond (e.g., Gulf Air). One important aspect of any skills training programme, such as CRM, is that the transfer of skills to the workplace should be established (Boehm-Davis, et al., 2001; O'Connor, et al., 2002b). The NOTECHS system offers one method of ascertaining whether the CRM training provided to pilots is actually enhancing effectiveness of overall crew performance on the flight deck (Goeters, 2002).

References

Antersijn, P. and Verhoef, M. (1995). Assessment of non-technical skills: is it possible? In, N. McDonald, N. Johnston and R. Fuller (Eds.) *Applications of Psychology to the Aviation System.* Aldershot: Avebury.

Avermaete, van J. and Kruijsen, E. (1998) (Eds.). *NOTECHS. The Evaluation of Non-Technical Skills of Multi-Pilot Aircrew in Relation to the JAR-FCL Requirements.* Final Report NLR-CR-98443. Amsterdam: National Aerospace Laboratory (NLR).

Baker, D. and Dismukes, K. (2002). Special issue on training instructors to evaluate aircrew performance. *International Journal of Aviation Psychology, 12*, 203-222.

Baker, D., Mulqueen, C. and Dismukes, K. (2001). Training raters to assess resource management skills. In E. Salas, C. Bowers and E. Edens (2001) (Eds.) *Improving Teamwork in Organizations. Applications of Resource Management Training.* Mahwah, N.J.: LEA.

Boehm-Davis, D., Holt, R. and Seamster, T. (2001) Airline resource management programs. In E. Salas, C. Bowers and E. Edens (2001) (Eds.). *Improving Teamwork in Organizations. Applications of Resource Management Training.* Mahwah, N.J.: LEA.

Burger, K.-H., Neb, H. and Hörmann, H.-J. (2002) Basic Performance of Flight Crew – A concept of competence based markers for defining pilots performance profile. *Proceedings of the 25th European Aviation Psychology Conference, Warsaw.* Warsaw: Polish Airforce.

CAA (2003) *Crew Resource Management (CRM) Training. Guidance for Flight Crew, CRM Instructors (CRMIs) and CRM Instructor-Examiners (CRMIEs).*

Gatwick, Sussex: Safety Regulation Group, Civil Aviation Authority. Available on www.caa.co.uk

Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32-64.

Flin, R. and Martin, L. (1998). *Behavioural Markers for CRM.* CAA Report 98005 London: Civil Aviation Authority.

Flin, R. and Martin, L. (2001) Behavioural markers for Crew Resource Management: A review of current practice. *International Journal of Aviation Psychology, 11*, 95-118.

Goeters, K.-M. (2002). Evaluation of the effects of CRM training by the assessment of non-technical skills under LOFT. *Human Factors and Aerospace Safety, 2,* 71- 86.

Gregorich, S. and Wilhelm, J. (1993). Crew resource management training assessed. In E. Wiener, B. Kanki and R. Helmreich, (Eds.) *Cockpit Resource Management.* San Diego: Academic Press.

Helmreich, R., Butler, R., Taggart, W. and Wilhelm, J. (1995). The NASA/ University of Texas/ FAA Line/ LOS checklist: A behavioural marker-based checklist for CRM skills assessment. Version 4. Technical Paper 94-02 (Revised 12/8/95). Austin, Texas: University of Texas Aerospace Research Project.

Helmreich, R. (2000). The Line Oriented Safety Audit (LOSA) system. Paper presented at the *Australian Aviation Psychology Conference*, Manly, November.

Hörmann, H.-J. (1995). FOR-DEC: A prescriptive model for aeronautical decision making. In, R. Fuller, N. Johnston, and N. McDonald, (Eds.) *Human Factors in Aviation Operations.* Aldershot: Avebury.

Hörmann, H.-J. (2001). Cultural variations of perceptions of crew behaviour in multi-pilot aircraft. *Le Travail Humain, 64,* 247-268.

Hörmann, H.-J., Burger, K-H. and Neb, H. (2002). Integration of interpersonal skills into a pilot's proficiency reporting system. First results of a usability study at Lufthansa. In O. Truszczynski (Ed.) *Proceedings of the 25th European Aviation Psychology Conference, Warsaw.* Warsaw: Polish Airforce.

ICAO (1989a). *Human factors digest no. 1. Fundamental human factors concepts.* Montreal, Canada: International Civil Aviation Organization.

ICAO (1989b). *Human factors digest no. 2. Flight crew training: Cockpit resource management (CRM) and line-oriented flight training (LOFT).* Montreal, Canada: International Civil Aviation Organization.

Jensen, R. (1996). *Pilot Judgment and Crew Resource Management.* Aldershot: Avebury.

Kaempf, G. and Klein, G. (1995). Aeronautical decision making: The next generation. In, N. Johnston, N. McDonald and R. Fuller (Eds.) *Aviation Psychology in Practice.* Aldershot: Avebury.

Klampfer, B., Flin. R., Helmreich R.,. et al (2001). *Enhancing performance in high risk environments: Recommendations for the use of Behavioural Markers.* Daimler Benz Foundation. Available as pdf file on www.psyc.abdn.ac.uk/serv02.htm.

KLM (1996). Feedback and Appraisal System. Amsterdam: KLM Internal Paper.

O'Connor, P., Flin, R. and Fletcher, G. (2002a). Techniques used to evaluate Crew Resource Management training: A literature review. *Human Factors and Aerospace Safety, 2, 217-233.*

O'Connor, P., Flin, R. and Fletcher, G. (2002b). Methods used to evaluate the effectiveness of flightcrew CRM training in the UK aviation industry. *Human Factors and Aerospace Safety, 2, 235-255.*

O'Connor, P., Hormann, H.-J., Flin, R., Lodge, M., Goeters, K.-M. and the JARTEL group (2002). Developing a method for evaluating crew resource management skills: A European perspective. *International Journal of Aviation Psychology, 12,* 265-288.

Orasanu, J. (1993). Decision making in the cockpit. In, R. Helmreich, E. Weiner and B. Kanki (Eds.) *Cockpit Resource Management.* San Diego: Academic Press.

Polo, L. (2002). Evaluation of flight crew members' performance. Is evaluation a product or a tool? In, O. Truszczynski (Ed.) *Proceedings of the 25th European Aviation Psychology Conference, Warsaw.* Warsaw: Polish Airforce.

Salas, E., Burke, S., Bowers, C. and Wilson, K. (2001). Team training in the skies. Does Crew Resource Management (CRM) training work? *Human Factors, 43,* 641-674.

Seamster, T. and Keampf, G. (2001). Identifying resource management skills for pilots. In, E. Salas, C. Bowers and E. Edens (2001) (Eds.) *Improving Teamwork in Organizations. Applications of Resource Management Training.* Mahwah, N.J.: LEA.

Shrestha, L., Prince, C., Baker, D. and Salas, E. (1995). Understanding situation awareness: Concepts, methods and training. In, W. Rouse (Ed) *Human/ Technology Interaction in Complex Systems.* Vol 7. Greenwich, CT: JAI.

Stokes, A., Kemper, K. and Kite, K. (1997) Aeronautical decision making, cue recognition and expertise under pressure. In, C. Zsambok and G. Klein. (Eds.) *Naturalistic Decision Making.* New Jersey: LEA.

Acknowledgements