2005

Human Factors and Human Resources Developments for Pan-European Implementation: Achievements in the European ATM Programme

Hermann Rathje

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2005

Part of the Other Psychiatry and Psychology Commons

Repository Citation

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2005 by an authorized administrator of CORE Scholar. For more information, please contact corescholar@wwwlibraries.wright.edu, library-corescholar@wright.edu.
HUMAN FACTORS AND HUMAN RESOURCES DEVELOPMENTS FOR PAN-EUROPEAN IMPLEMENTATION: ACHIEVEMENTS IN THE EUROPEAN ATM PROGRAMME

Hermann Rathje
European Organisation for the Safety of Air Navigation (EUROCONTROL)
Human Factors Management Business Division
Brussels, BELGIUM

The European Air Traffic Management (EATM) Human Resources Programme aimed to deliver harmonised tools and a body of knowledge for the management of human issues in ATM in the three areas training, manpower (human resources management) and human factors. Products are available as Guidelines, as Technical Reference Material or Reports or as Methods and Tools for direct application. The four year work programme consisted of specific developments, testing and validation of these products available since end 2003 for implementation in the 41 European Civil Aviation Conference (ECAC) States. The implementation and use is not mandatory but products are applied increasingly with the active involvement of stakeholders from Air Navigation Service Providers (ANSPs), military and regulatory authorities and professional associations. Four different products are presented in this paper: Team Resource Management (TRM), Human Error Analysis (HERA), Critical Incident Stress Management (CISM) and the First European Air traffic controller Selection Test package (FEAST).

European ATM Framework

EUROCONTROL, the European Organisation for the Safety of Air Navigation is involved in the development of a seamless, pan-European Air Traffic Management (ATM) system to cope with the growth in air traffic, while maintaining a high level of safety, reducing costs and respecting the environment.

The EUROCONTROL Agency is tasked by the ECAC Transport Ministers with defining a common European vision and strategy and coordinating its implementation.

Under the performance enhancement programme for European ATM (EATM), the EUROCONTROL Agency produces standards and guidelines and common products / systems and tools and provides guidance and assistance to its Member States in the implementation thereof.

The European Commission, the executive body of the European Union, is now progressing with the creation of a Single European Sky that aims to enhance current safety standards and support commercial and economic growth through more efficient airspace design following operational needs rather than national frontiers, to generally optimise capacity and ensure interoperability of the ATM systems across Europe.

The Human Resources Programme (HRS)

The EATM programme consists of a wide portfolio of programmes, services and support activities and includes Human Factors, Manpower, Human Resources Management and Training activities.

The objective of the latter programme is “to ensure human involvement and commitment to support the change to future ATM so that operational, technical and support staff can operate effectively, efficiently and safely within their capabilities and obtain challenge and job satisfaction.”

ATM systems are expected to remain human-centred for the foreseeable future, and people will play a key role in achieving system safety and capacity enhancements.

People are therefore an essential element in the ability to deliver ATM services, and their co-operation and involvement in developing and effecting change is essential.

It is of high importance that all human performance and training issues are sufficiently addressed and managed as early as possible, in order to ensure new technologies and operational procedures. This will enable stakeholders to proactively plan and manage their medium and long-term goals for the management of human issues in European ATM.

The aim of the Human Resources Programme is to offer, through the production of guidance material, methods and tools, a harmonised and integrated approach to:

1 Numbered 34 Member States (in Dec 2002): Albania, Austria, Belgium, Bosnia, Bulgaria, Croatia, Cyprus, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Luxembourg, the former Yugoslav Republic of Macedonia, Malta, Moldova, Monaco, the Netherlands, Norway, Poland, Portugal, Romania, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine and the United Kingdom.

2 The European Union (EU) comprises of 25 Member States: Austria, Belgium, Cyprus, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden and the United Kingdom.
• Manage human performance proactively and to ensure the timely availability of suitable operational staff through tools and methods for manpower planning, recruitment and selection, training and staff development;
• Enhance safety in day-to-day ATM operations through human factors products and tools for Team Resources Management and Critical Incident Stress Management and through tools that support integration of human factors into the life cycle of ATM systems;
• Progress with ATS staff training towards common standards in line with the regulatory requirements for controller licences and the changing ATC systems.

Next Steps in Human Factors Developments

The cultural, social, human factors and human resources aspects related to the intended reorganisation of the current European ATM system in line with the European Union Single European Sky initiative will be even more important to be appropriately addressed in the future. It will require developing new approaches and tools to
• effectively deal with the cultural, organisational and individual change and transition issues involved;
• fully integrate human factors into safety management systems and safety culture;
• provide common European training standards and tools in line with regulatory requirements

These work areas are now addressed in an proposed new Human Performance and Training Enhancement Programme to start later this year.

Human Resources Programme Product Development, Testing and Validation

During the years 2000 – 2003 the human resources work programme proceeded in parallel in 17 training, manpower and human resources management and human factors projects. Four projects are reported in more depth in this paper:
• Human Error Analysis (HERA)
• Team Resource Management (TRM)
• Critical Incident Stress Management (CISM)
• First European Air traffic controller Selection Test package (FEAST).

The general approach followed in the Human Resources Programme was to develop the projects in close consultation with stakeholders – Air Navigation Service Provider organisations (ANSPs), military and civil authorities and professional groups as well as with external partners in a coordinated fashion to ensure a broad representation of stakeholder requirements and needs. This also helps to later facilitate practical implementation and customisation of the products.

The needs, benefits and the feasibility for development were established in early feasibility studies or business cases that took stock of what was already available in the area of work. Sound feasibility reports, cost-benefit considerations and development options aim to market the intended work, clarify deliverables and justify the work in general with the aim to gain stakeholder commitment and support.

Prototypes of the tools were tested in practice and validated against established criteria. This early feedback and data is used to refine the work programme and direct further work. In all cases in which this was applicable beta-versions of the final products were tested in different national and local environment, representing a good cover of the cultural operational and administrative working environment. The outcome of the trials was validated in terms of content ("Does what is developed represent the subject area or behaviour it intends to represent?") and / or (concurrent) criterion validity ("Do results from using the product correlate with a relevant external criterion?"). Lessons learned from the ‘live’ trials were used in final updates of the products and reported.

Product Implementation

The ultimate purpose of the new products or is to provide valuable, scientifically sound, harmonised and cost effective options to ANSPs for use in training, human resources management, human factors and safety. Implementation is done in a coordinated fashion across ECAC States based on an agreed common action plan and target dates. There is an annual follow-up on the implementation actions in regard to all products across States that participate in the European ATM Programme. Air Navigation Service Providers, that decide to implement a product or tool are given professional assistance and support by the training, human resources or human factors experts from the development unit. Implementation support consists of
• product user training and certification;
• technical advise and support;
• planning and customisation support;
• functional helpdesk;
• administration and application enhancement and improvements;
• sponsoring and facilitating user Group meetings.

3 Most deliverables are to be found on the following website: www.eurocontrol.int/humanfactors
4 CISM was a project outside the HRS Programme as part of the Human Factors Domain activities.
The follow-up and user support activities are of great importance to ensure that products are used consistently and for the correct purposes as intended. The huge cultural and administrative and/or operational diversity across European States, the differences in organisational size and ‘maturity’ as well as the availability of local expertise require a sensible and sometimes sensitive approach in implementation support. There is no ‘One size fits all’.

**HERA – A New Technique for Human Error Analysis in ATM**

Human error is a key contributor to risk for incidents (and accidents) in ATM and finding and mitigating if not avoiding the root causes or causal factors of human error is hence an important aspect in safety research and safety management. The importance of finding the root causes for human error in ATM is highlighted by recent statistics showing, in the US for example, an increase of human errors by more than 25% during the time period 1998 - 2003.

The research and development of a common approach and tools for human error analysis was jointly done by Eurocontrol and FAA in the period 1999 – 2003. The outcome was a new technique for the classification and assessment of the causal factors for human error, called JANUS5 a technique originally used to retrospectively analyse ATM incident reports during investigations only. The idea was to use the taxonomy also to diagnose the potential for human errors prospectively, for example to assess potential human errors in the design of (future) systems. The collaborative Eurocontrol – FAA development process proceeded in four stages: planning, development of JANUS, field testing and validation.

*Development* proceeded on parallel tracks in Europe and in US: Eurocontrol developed the human error taxonomy following the HERA approach using ATM task and behavioural requirements and looking at the cognitive processes that lead or could lead to an incident. The partner organisation FAA used their Human Factors Analysis and Classification System (HFACS) approach that captures the conceptual breadth and depth of the system with the individual actions, along preconditions, supervision and organisational influences. Both approaches are hence complementary to some extent and had a good track record in previous validation studies. JANUS integrates both approaches in a common taxonomy of human errors and causal (cognitive) human factors.

Table 1 provides the JANUS taxonomy categories and examples.

**Table 1. JANUS Taxonomy**

<table>
<thead>
<tr>
<th>Error Category</th>
<th>Subcategory / Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Type – How error was manifested</td>
<td>Action omitted, right action but wrong object</td>
</tr>
<tr>
<td>Error Detail – Which cognitive domain failed</td>
<td>Perception, Memory, Planning and Decision making, Response</td>
</tr>
<tr>
<td>Error Mechanism – What happened?</td>
<td>Late detection of information</td>
</tr>
<tr>
<td>Information Processing - Why did it happen?</td>
<td>‘Tunneling’, forgot to monitor</td>
</tr>
<tr>
<td>Contextual Conditions</td>
<td>Pre-conditions: airspace, teamwork, supervision, organisational factors</td>
</tr>
</tbody>
</table>

The JANUS technique itself consists of a series of flow diagrams (paper based) used in interview sessions with specially trained users (i.e. investigation experts). Investigators or researchers are systematically led through a series of questions one at the time. This reduces the occurrence of user bias and prevents jumping to conclusions.

**JANUS Testing and Validation**

Beta testing of the common taxonomy JANUS took place in seven European member States analysing a total of 60 incidents (done by Eurocontrol). The FAA independently applied JANUS to 79 incidents from 12 US facilities.

The findings from both parallel studies were analysed with a view towards five ‘validity’ questions:
- Does JANUS work?
- How well does it work?
- Is it better than current methods?
- Is it ready for implementation?
- Do results improve safety management?

The findings (objective / subjective reports) indicate that the technique works, is (moderately) consistent in identifying causal factors and helps to improve the investigation of human factors related incidents. JANUS identified on average 13 causal factors per incident compared to 2 factors from current methods used. The findings also showed that JANUS broadens the current scope of investigation substantially. It prompts investigators to causal factors in a given context situation in which a human error occurred.

---

5 Janus is the name of a mythological figure (the Roam god of gates and doors) who, with his two faces looking in opposite direction. Janus represents the beginning and end, the past and the future and the transition from a less developed towards a more advanced stage of cultural live.
In summary: JANUS is more sensitive, is useful, comprehensive and practical than current available methods.

The benefits demonstrate the value of a joint undertaking and using a wider scope of expertise and experience. This has led to more consistent, sensitive and comprehensive approach in analysing and subsequently preventing or mitigating human errors in ATM based on common terminology. An important step towards international standardisation in this field has been achieved.

In the European context JANUS is seen as a means of complying with the European Safety Regulatory Requirements (ESARR). The next planned step is to have the JANUS taxonomy included in the European Co-ordination Centre for Aviation Incident Reporting System (ECCAIRS).

Team Resource Management in European ATC: A Ten-year look-back

Airlines have it since more than 25 years now and apply it around the world: programmes to promote teamwork practices in Crew Resource Management (CRM). Wiener etc al. (1993) noticed the lack of it in ATC.

The situation has changed now in Europe. In 1994 first steps were made towards developing an air traffic services Crew Resource Management programme and in 1995 the work started with a first task force of human factors experts, active controllers and training experts from across European States to look into the feasibility of what was then already called ‘Team Resource Management’ (TRM). The task force concluded that in fact TRM was feasible and beneficial and submitted Guidelines for developing and implementing TRM (EUROCONTROL, 1996).

The key objective of TRM is to develop the attitudes and behaviour towards enhanced teamwork skills and performance in ATM. Hence TRM aims to ensure the effective functioning of operational staff by helping them to use all available resources in time and as proficient as possible to reduce team work failures as a contributing factor in ATM related incidents and accidents.

### TRM Prototype Course

The course developed with the support of active controllers provides a generic content and structure carefully selected and refined to be culturally acceptable balanced for the majority of nationalities and operational cultures. The modules are open for customisation and adaptation and including national examples i.e. from incidents used in the training course and suit the learning needs of participants.

The prototype course consists of six modules (see Table 2 below) plus an introduction and a conclusion module. The course itself lasts for three days and is designed for 8 – 12 participants.

<table>
<thead>
<tr>
<th>Table 2. TRM Prototype Modules Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Team work (TW)</strong> – Typical characteristics of ATC related TW; negative impacts of behaviour on TW; character types in teams and impact on TW; team identity; safety issues related to TW; recognition and management of diversity in teams</td>
</tr>
<tr>
<td><strong>Team Roles</strong> – Understand formal / informal hierarchies; attitudes towards authority (cultural impacts); strategies to avoid misunderstanding that leads to errors in the roles as a leader / follower; strategies to deal with submissive, aggressive and assertive behaviour</td>
</tr>
<tr>
<td><strong>Communication (COM)</strong> – Functions of COM; understanding team COM related to safety; effective COM and effective intervention in ATM related situations; strategies to give / receive feedback and constructive criticism</td>
</tr>
<tr>
<td><strong>Situational Awareness (SA)</strong> – Understand SA and the effects of high / low workload on SA; identify symptoms of team / individual loss of SA and strategies to prevent loss of SA; identify factors that have positive / negative effect on SA</td>
</tr>
<tr>
<td><strong>Decision Making (DM)</strong> – contributing factors for effective DM; appreciate importance of situation and risk assessment skills for DM; appreciate concepts of shared problem models and resource management skills in team DM; structured DM</td>
</tr>
<tr>
<td><strong>Stress Management</strong> – Identify job related stress factors; stress – what it is and how it affects work and team work; stress coping strategies; develop skills to recognise and cope with stress situation in teams</td>
</tr>
</tbody>
</table>

TRM trainers (facilitators) are trained in facilitation techniques which include self-presentation, mini-lessons, interactive lessons, introducing, summarising and conclusion techniques for discussions etc. The course material includes a facilitator and a participant handbook and video scenarios for some modules.

### TRM Customisation and Implementation

ATM organisations that want to implement TRM in their operational and training environment can customise the modules using their own resources, the support from Eurocontrol experts or from external companies of their own choice. Guidance material is available to facilitate this customisation work (Woldring & Amat, 1998.)
The cultural differences between States are often substantial and require a sensible approach. Local examples on incidents, the use of local language, humour, stories and staff are important and increase acceptance, awareness and impact in learning and actual application. This stresses the importance of TRM as learning rather than a teaching experience. TRM is a learning process that aims to positively impact on actual behaviour.

The implementation in ECAC States is still progressing. TRM users exchange their experience and expertise in a TRM User Group that consists mainly of air traffic controllers as TRM facilitators.

Recently tools complementary to TRM are available, one is the ‘Behaviour Oriented Observation Method’ (BOOM). Its aim is to train TRM facilitators and training instructors in objective, reliable and valid behaviour observation and feedback methods for non-technical skills in the TRM context. This is an additional step forward to increase the impact of TRM in practice.

In summary: TRM has proven to be widely accepted in European ATC now. It has helped to increase the awareness of human factors in general in ATM operations and increased the understanding of individual, group and cultural aspects in teamwork related behaviour. This also has helped to better understand why human errors can occur as a result of poor TRM.

TRM is now recognised as an important human factor in safety management. The Strategic Safety Action Plan (SSAP) established as a reaction to the Ueberlingen mid-air collision in 2002 recognises the importance of teamwork and team culture. The requirements set in the SSAP are mandatory for safety regulators and ANSPs in Eurocontrol member States. They need to ensure safety awareness, shall establish a safety culture, attitudes and behaviour amongst air traffic controllers through the implementation inter alia of measures in line with TRM. They are required to allocate the required resources for it and to report about implementation.

The long road for TRM to become practice would not have been achieved without the continuous and persistent efforts that human factors experts across Europe have invested in this area. This has fostered a better understanding of cultural diversity in teamwork in European ATM but also to bridge differences in local team and safety cultures. TRM has hence an important role to play in the future as a means to support change and transition and the merging of cultures in cross border Functional Airspace Blocks (FAB), the integration of teams in case of merged centres or units or centralises services.

The development of TRM is continuing on a communal and collaborative basis in the TRM User Group. Two new Modules have recently been developed. One is on ‘Error Management’ and is expanding on human error in teams and teamwork, the other is on ‘Impacts of Automation’ and addresses the cognitive impacts of current and future automation on individual and team performance, decision making and actions. More modules on the integration of teams and team cultures are planned.

**CISM – Critical Incident Stress Management**

CISM in short is a structured approach to assist people who have experienced an abnormal or traumatic critical event and react with strong personal emotions. The after-effects of critical incident stress can be substantial and long-lasting and can pose a danger to the well-being and performance of individuals and can even create a concern for safety.

The CISM work done in Eurocontrol started with a small booklet – called ‘module’ on CISM (EUROCONTROL, 1997a) which gave guidance for setting up of CISM in three phases:

- **Information Phase**: making aware and provide information on critical incidents, reactions how CISM support would come into force.
- **Training Phase**: Provide detailed information on CISM and the training of volunteers that would assist colleagues after critical incidents.
- **Support Phase**: Services and support that can be given to the persons concerned after the event.

CISM techniques involve a variety of methods and approaches and include:

- Early intervention – Don’t wait after the incident happened!
- Use group dynamics – If more people are involve, get them together to speak and moderate the impact of the critical incident.
- Verbalisation, emotion ventilation etc.
- Debriefing and Defusing – Use this method to help to relieve emotions in a constructive and structured way.

**Benefits and Lessons Learned in CISM Application**

CISM case studies from various users demonstrate the benefits of investing in CISM. As in TRM customisation of the CISM approach to local and organisational culture and the working conditions and the use of peers as CISM experts are important aspects. The reports from existing schemes and from recent cost-benefit analysis on CISM indicate that CISM helps controllers to cope with the stress and
return to work more rapidly after critical events. It also reduces the risk of post-traumatic stress disorders that could lead to long-term sickness or even incapacity to continue to work in operations. The return on investment is reported to be positive.

As a high-priority action for safety-related human resources in ATM ANSPs are required to implement CISM as an integrated element of their safety management system.

CISM’s main strength is that it is a peer support model which has the effect of changing attitudes to critical incidents and the ways these are regarded. Organisations have noted more openness to the discussion of incidents and errors as a by-product of CISM programs. As with HERA and TRM CISM has the potential for the future changing European environment in terms of making aware, address and manage critical incidents in a safety critical but fast changing ATM working environment with potentially high incident risks.

FEAST – A European Selection Test Package for Controllers

**Background**

Compared with pilot selection, the selection of candidates for controller training has fallen short in a number of major respects: Task and job analysis, selection development, test validation and use of best practice and standards. Common European or even international developments are still rare or non-existent, with only a few exceptions.

EUROCONTROL (1997b, 2000, 2001) gave detailed information on the situation in controller selection in the ECAC States based on detailed surveys done over the years continuously showing that around 50% of States could not provide appropriate results on the main psychometric properties (reliability, objectivity and validity) of tests they were using. The situation did change significantly due to improvements achieved in some States that recruit higher numbers of controllers annually. However, States that select and train only small numbers of controllers per year report that they found it difficult to comply with some of the Guidelines that Eurocontrol had issued. (EUROCONTROL, 2001, 2002). The conclusion was, that implementation, validation and maintenance of psychometric sound, complete and effective selection tests was neither practically nor financially feasible for some selection users. And that this had an impact also on training success, a low credibility of the tools used and has led to low stability and length of use of the tools leading to a lack of experience, validation possibility etc. and in fact into a vicious circle.

This situation was eventually addressed. The need for a common and advanced European development that would be based on tests that had demonstrated validity became an issue. From an initial reluctance against common standards, methods, guidelines and tools in 1995 emerged a situation of openness and support towards

- European communal efforts;
- Harmonisation of approaches;
- Establishment of enhanced quality and standards and benchmarks in test use and application;
- Common new test developments and even
- Common establishment, maintenance and validation of appropriate selection tools and methods for European wide use.

The 1999 Eurocontrol selection seminar strongly recommended to work together towards establishing a commonly-used selection system. ‘Means and options should be investigated … to acquire and / or develop a European Controller Test Battery that could be used in those ECAC States that are in need of this, especially the smaller States’ (EUROCONTROL, 1999, p. 157).

In parallel to these European activities, participants at the ‘International Air Traffic Controller Selection Conference’ organised by FAA and held in Oklahoma City in the same year 1999 proposed to create an international working group of experts in ATCO selection. This group should openly exchange experience and data and share tools and ideas for mutual benefit and advancing developments in selection and especially in the cross-cultural validation of new test developments.

**FEAST Development Objectives**

FEAST was developed as a European joint venture with the objective to provide a basic, easy to administer and manage controller selection test option which reflects future impacts (e.g. as a result of changes in technology and the work environment) and enables customised implementation and use in Member States.

The test package should be flexible enough to be customised for use in different European countries and for the current and future tasks of the ATCO. Typical users for FEAST are States that recruit only small numbers of trainees annually and lack in-house expertise for own development and validation. Under these circumstances it will take long to obtain a sufficient validation sample size. Generally users would be interested if the cost-benefit ratio for an
own fully fledged selection methodology would indicate prohibitive high efforts and costs.

**FEAST Development Milestones**

The list of milestone that had to be passed during the development cycle can be summarised as follows:

1. Demonstrate feasibility and viability of the FEAST concept.
2. Gain initial commitment from potential users across Europe.
3. Establish a controller job requirement model as the basis for development.
4. Gather, evaluate and select potential tests and methods for FEAST package in line with the model.
5. Compose a consistent test package.
7. Establish comprehensive, consistent and licence and test user agreements and privacy policy for test takers.
8. Adapt and establish tests and scores for a consistent and easy administration and use.
9. Perform Quality Assurance and Standardisation of all procedures and tools.
10. Develop, investigate and validate a common criterion for initial and long-term validation of FEAST, the ‘Behavioural Observation Scale’ (BOS) for trainees and controllers.
11. Perform initial validation on multiple samples and groups of candidates / trainees / active controllers across Europe.
12. Establish FEAST as a service (service feasibility and implementation).

This paper can only highlight some aspects of the outcome of the FEAST programme in regard to concurrent validation findings and implementing FEAST as a service. (See for further details Rathje, & Golany (2003b) and Rathje, Golany & Eissfeldt (2004a and b)).

**The FEAST Package**

The FEAST package consists of tests composed into two assessment phases plus one optional assessment module:

---

**Phase I Tests:** Six web-based cognitive and knowledge tests running on PC linked to the Internet and include an English listening test. Standardised test results and a composite score are used for screening of candidates.

**Phase II Test:** A complex, dynamic multiple-task test administered on a standalone PC. The test simulates procedural control using flight strip data. Candidates are trained before taking the test using an integrated computer based training module. Candidates for Phase II testing are pre-selected, based on their results in the Phase I.

**Optional Assessment Module:** A Situational Interview (SI) paper-and-pencil format.

**Initial Validation of FEAST in Cross-cultural Samples**

**Samples** - The initial validation trials were conducted in 2002 – 2003 using a variety of samples and groups across nine European States. A total sample of 579 applicants, trainees and controllers were tested. The variation between samples in terms of age, gender, sample size and composition, selection stage etc was big and could not be influenced. The initial validation samples had hence to be used as given.

**Criterion & Predictors** - For Trainees and controllers concurrent criterion data was gathered using the common criterion method ‘Behavioural Observation Scale’ (BOS). The predictors were FEAST test scores and composites. They were correlated with the BOS and other training performance criteria in those samples where this criterion data was made available.

**Restriction of range** - It is well known that as participants in a concurrent validity study are a selected group of those persons who are actually and/or potentially able to do the job the variance in the predictors (and in the criterion) is likely to be smaller than that in a group of applicants for the job. The effect is a reduction of the size of the correlation coefficient. To estimate the correlation for the population, the standard deviation of predictor test scores in the applicant population, respectively in a sample from the population is to be known. For FEAST some data was available from real applicants in selection testing under standard conditions. Only these results are reported here in brief.

---

6 The papers summarise the approach and methods adopted in FEAST validation in conjunction with the opportunities and challenges and problems encountered in cross-cultural validation, the development and validation of the concurrent Behavioural Observation Scale (BOS) method used to establish a common criterion and the detailed reliability and validity data both of the predictors and the criterion development and validation (the FEAST Behavioural Observation Scales, BOS) and on the controller job requirement model.

7 Concurrent validation as a means to establish the validity of test scores if a full and comprehensive long-term, predictive validation approach cannot be performed, or as a first step in a more comprehensive validation process, as was the case with FEAST. As such, concurrent validation is not an alternative to a predictive validation but can offer an independent measure of test validity.
Phase I Test Results - The coefficients reported in Rathje, & Golany (2003b) were computed on (restricted) samples of Trainees and Controllers. A comparison of the (standardised) test scores and composite scores between candidates and trainees showed that the variance of the scores of trainees for example were between 55% and 90% of the variance in the candidate group. Corrected (adjusted) correlations were computed based on the standard correction formula (Hunter & Schmidt, 1990).

The correlation between the Total Test Composite and FEAST Criterion (BOS Summary Score of a total of 35 Items) was \( r = .296 \) (\( p < .05 \)) in one sample of \( n = 55 \) Trainees. The corresponding adjusted correlation based on the correction formula for range restriction is \( r = .42 \) (\( p < .01 \)) for the same sample (n=55)\(^8\).

The Controller samples were quite different one from another - more than the trainee ones – as they varied more in their age range, level of motivation for taking the FEAST and the level of consideration given by their supervisors for completing the standard BOS criterion scales. When using one sample of \( n = 24 \) ATCOs from one location with a known selection ratio and selection methodology, age range and homogenous, reliable and complete criterion data, the restriction in the range would have been even higher. Although the correlation between the FEAST Composite Score and BOS Summary Score is significant (\( r = .46 \), \( p < .05 \)) (adjusted \( r = .57 \), \( p < .01 \) the sample size (n=24) is too small for drawing conclusions.

The correlation between FEAST predictors (composite score) and other training criteria, for example, ‘Course Overall Final Pass Mark’ (training score at the end of Initial Training) from one trainee sample (n = 46) - where the selection ratio is 12.5% - to be \( r = .36 \) (\( p < .05 \)).

Phase 2 Test Results - Regarding the complex work sample test, the correction formula for range restriction\(^9\) was applied on a trainee sample for 3 scales of the BOS: “BOS Summary Score” (35 Items, as above), “Teamwork” and “Working under Stress”. Here, for example, the corrected correlations for a predictor of this test called ‘number of correctly identified opposite conflicts’ with the BOS total score, was \( r = .41 \) (\( p < .01 \)) for a trainee sample (n=43). The same predictor’s adjusted for restriction correlation with the BOS score of ‘Working under Stress’ was \( r = .39 \) (\( p < .01 \)) (n=50 trainees).

The correlations between one composite score representing the performance in regard to updating / inserting and ordering flight strips (Total Performance in one performance criterion – Strip Order) with the three BOS scales are given in Table 3 (adjusted \( r\) in brackets):

**Table 3. Adjusted and unadjusted correlation between criterion and predictor scores in FEAST Phase Testing - trainee samples**

<table>
<thead>
<tr>
<th>Predictor Criterium</th>
<th>Unadjusted Correlation Coefficient</th>
<th>Adjusted Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOS – Summary Score</td>
<td>.44** (.54**) (n = 43 Trainees)</td>
<td></td>
</tr>
<tr>
<td>BOS – Teamwork</td>
<td>.43** (.54**) (n = 50 Trainees)</td>
<td></td>
</tr>
<tr>
<td>BOS – Working under Stress</td>
<td>.37** (.47**) (n = 50 Trainees)</td>
<td></td>
</tr>
</tbody>
</table>

All correlations are significant at the < .01 level.

FEAST Cross-cultural Validation and Testing

Conclusions

The results of the various studies clearly demonstrated the challenges in a cross-cultural, common approach in validation of controller selection tests and the impact of sample size, age and composition of the validation samples, the restriction of range due to failures in training and other aspects that have or can have detrimental effects on the results.

Use of a Common Criterion Measure (BOS) - As regards the BOS criterion measure, the studies demonstrated the reliability, validity of the BOS as well as the relevance and need of a common criterion measure. The findings in some samples however also give warnings as regards the need for appropriate training material and calibration training in using the BOS (or other measure) of assessors i.e. training instructors or supervisors. The differences in training, trainee assessment methodology and culture are big. Important items for consideration and countermeasures for the future, long-term validation of FEAST are the use of behavioural anchors for all scales of BOS and reasonable training of assessors.

FEAST Predictor Tests - The tests and the various composite indices developed for use in selection decision-making, the findings from the initial validation study showed that the scores are sufficiently reliable and stable across samples. The results also demonstrated that despite the low

---

\(^8\) This result was cross-validated using a different composite score which allowed the analysis of a sample of \( n = 81 \) Trainees. Here the correlation was \( r = .27 \) (\( p < .05 \)) (not adjusted) and \( r = .45 \) (\( p < .01 \)) adjusted.

\(^9\) The variance in the trainee sample compared to a sample of (pre-selected) candidates that took this test as the Phase II test was less restricted as expected and was between 70% - 120% of the candidate sample.
variance in the FEAST BOS criterion (especially in the Controller samples, significant and stable correlations were found between test scores and the criterion in samples of trainees and qualified controllers.

In Summary: The study efforts already now demonstrate the progress in efficiency and benefits that can be reached by the application of a proper, valid selection procedure and by a combined validation effort across various European air traffic controller training schemes and the establishment of common ‘European’ norms. It is made clear that this can only be achieved if a common, collaborative and harmonised approach is adopted and quality criteria and standards are shared and actually met. Whether this is feasible to achieve is still an important challenge in the establishment of a Pan-European FEAST service.

FEAST Service Feasibility and Implementation

FEAST since the beginning of 2004 has progressed into a ‘FEAST Service Planning and Feasibility Phase’. The aim of this phase is to establish the viability of the nature and scale of such a service offered by Eurocontrol to ANSPs that wish to use FEAST during a pilot Service. During 2004 - 2005, the viability of the service delivery is established to prepare a full business case that will assist decision-making regarding the introduction of a full FEAST service. During this period, FEAST is tested under ‘live conditions’ on real applicants and test data is gathered as an input into a longer-term predictive validation study and to establishing common norms.

The current experience in now seven different States where FEAST has been implemented is very promising. The implementation requirements for FEAST include the training of administrators in standard test administration, FEAST application and installation and technical requirements. FEAST recruiters are specially trained in the valid use and interpretation of test scores. Standards in regard to the test environment are observed during local installation visits. Local customisation and the integration of FEAST into current existing selection and recruitment methods are essential.

FEAST - Potential of Improving Selection in Europe

FEAST implementation and validation findings so far demonstrate the feasibility of a service across different cultures. FEAST

- Offers a valid and scientifically sound test battery;
- Meets agreed Eurocontrol guidelines in selection and recruitment;
- Enables quicker validation, a bigger sample and proper predictive validation in the lon-term;
- Reduces development, validation and maintenance and upkeep costs;
- Ensures high quality and standards in testing and selecting candidates for ATC training;
- Fosters efficiency and effectiveness in selection;
- Includes a built-in continuous improvement.

FEAST is a project geared to continuous, ongoing improvements and maintenance. Continuous improvement is crucial for the long-term viability and sustainability of FEAST together with users and scientific partners in development. Further developments involve parallel test versions and new tests. One recent example is the development of methods for a valid and reliable testing of English language proficiency in all performance areas of the new ICAO requirements.

References


