DRAFT INTERNATIONAL STANDARD ISO/DIS 11064-7

ISO/TC **159**/SC **4** Secretariat: **BSI**

Voting begins on: **2004-09-23**

Voting terminates on: **2005-02-23**

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

Ergonomic design of control centres —

Part 7:
Principles for the evaluation of control centres

Conception ergonomique des centres de commande —

Partie 7: Principes pour l'évaluation des centres de commande

ICS 13.180

ISO/CEN PARALLEL ENQUIRY

The CEN Secretary-General has advised the ISO Secretary-General that this ISO/DIS covers a subject of interest to European standardization. **In accordance with the ISO-lead mode of collaboration as defined in the Vienna Agreement, consultation on this ISO/DIS has the same effect for CEN members as would a CEN enquiry on a draft European Standard.** Should this draft be accepted, a final draft, established on the basis of comments received, will be submitted to a parallel two-month FDIS vote in ISO and formal vote in CEN.

To expedite distribution, this document is circulated as received from the committee secretariat. ISO Central Secretariat work of editing and text composition will be undertaken at publication stage.

Pour accélérer la distribution, le présent document est distribué tel qu'il est parvenu du secrétariat du comité. Le travail de rédaction et de composition de texte sera effectué au Secrétariat central de l'ISO au stade de publication.

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing international Standards is normally carried out through ISO technical committees. Each member body interested in a subject for whom a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in ISO/IEC Directives, Part 3.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this International Standard may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

International Standard ISO 11064-7 was prepared by Technical Committee ISO/TC 159, *Ergonomics*, Subcommittee SC 4, *Ergonomics of Human-system Interaction*, Working Group 8, *Ergonomic Design of Control Centres*.

ISO 11064 consists of the following Parts under the general title *Ergonomic Design of Control Centres:*

— Part 1: Principles of the design of control centres

— Part 2: Principles of control suite arrangement

— Part 3: Control room layout

— Part 4: Workstation layout and dimensions

— Part 5: Displays and controls

— Part 6: Environmental requirements for control rooms

— Part 7: Principles for the evaluation of control centres

Introduction

This part of ISO 11064 establishes ergonomic requirements, recommendations and guidelines for evaluation of control centres.

User requirements are a central theme of this part of ISO 11064 and the processes described are designed to take account of needs of users at all stages. The overall strategy for dealing with the user requirements is presented in ISO 11064-1.

ISO 11064-2 provides guidance on the design and planning of the control centre in relation to its supporting areas. ISO 11064-3 gives all the requirements and guidance on control room layout. Requirements for the design of workstations, displays and controls and the physical working environment are presented in ISO 11064-4 to ISO 11064-6.

ISO 11064-1 to ISO 11064-7 covers general principles of ergonomic design appropriate to a range of industries and service providers.

The ultimate beneficiaries of this part of ISO 11064 will be the control centre operator and other users. It is the needs of these users that provide the ergonomic requirements used by the developers of International Standards. Although it is unlikely that the end user will read this part of ISO 11064, or even know of its existence, its application should provide the user with interfaces that are more usable and a working environment which is more consistent with operational demands. It should result in a solution that will minimize error and enhance productivity.

Ergonomic Design of Control Centres — Part 7: Principles for the Evaluation of Control Centres

1 Scope

This part of ISO 11064 establishes ergonomic principles for the evaluation of control centres. It includes requirements, recommendations and guidelines on evaluation of the different elements of the control centre, i.e., control suite, control room, workstations, displays and controls, and work environment.

It covers all types of control centres, including those for the process industry, transport systems and dispatching rooms in the emergency services. Although, this part of ISO 11064 is primarily intended for non-mobile control centres, many of the principles could be relevant / applicable to mobile centres, such as those found on ships and aircraft.

2 Normative references

The following normative documents contain provisions that, through reference in this text, constitute provisions of this part of ISO 11064. Parties to agreements based on this part of ISO 11064 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

ISO 9000, *Quality Management Systems — Fundamentals and Vocabulary*

ISO 9241-11, *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), Part 11: Guidance on Usability*

ISO 11064-1, *Ergonomic Design of Control Centres, Part 1: Principles for the Design of Control Centres*

IEC 61771, *Nuclear Power Plants, Main Control Room, Verification and Validation of Design*

IEEE *Standard 845, Guide to Evaluation of Human-System Performance in Nuclear Power Generating Stations*

3 Definitions

For the purpose of this part of ISO 11064, the following terms and definitions apply.

3.1
Evaluation process
evaluation processes or Evaluation is the combined effort of all verification and validation activities in a project using selected methods and recording the results

3.2
human engineering discrepancy (HED)
a departure from some benchmark of system design suitability for the roles and capabilities of the human operator and/or user. This may e.g., include a deviation from an operator/user preference or need that is required for an operator's or user's task but is not provided to the operator or user

3.3
resolution
the identification and implementation of solutions to the deviations identified during the verification and validation activities

3.4
situation awareness
the relationship between the operator's/user's understanding of the controlled system's and/or process's condition and its actual condition at any given time

NOTE Originally defined by Endsley, 1988, in an aircraft pilot context 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".

3.5
validity
the degree to which an instrument or technique can be demonstrated to measure what it is intended to measure

NOTE 1 Face validity is concerned with how a measure or procedure appears. It answers the question: Does it seem like a reasonable way to gain the information the evaluator(s) is attempting to obtain?

NOTE 2 Predictive validity examines whether the human factors engineering outcome measure predicts some other measure. Predictive validity will tell whether it is possible to predict from the studied performance measure to the real environment.

NOTE 3 To assist with the interpretation of the following definitions, Figure 1 is included in this clause.

3.6
validation
confirmation by examination and tangible evidence that the particular requirements for a specific intended use are fulfilled, (ISO 9000)

NOTE 1 In design and development, validation concerns the process of examining a product to determine conformity with user needs.

NOTE 2 Tangible evidence is regarded as being information that can be proved to be true, based on facts obtained through observation, measurement, test or any other means.

3.7
verification
confirmation by a systematic examination and tangible evidence that specified requirements have been fulfilled, (ISO 9000)

NOTE In design and development, verification concerns the process of examining the result of a given activity to determine conformity with the stated requirements for that activity.

3.8
verification and validation plan
a plan specifically developed to govern the evaluation process

3.9
verification and validation process
see evaluation process

3.10
workload
the physical and cognitive demands placed on the system user(s) and/or staff

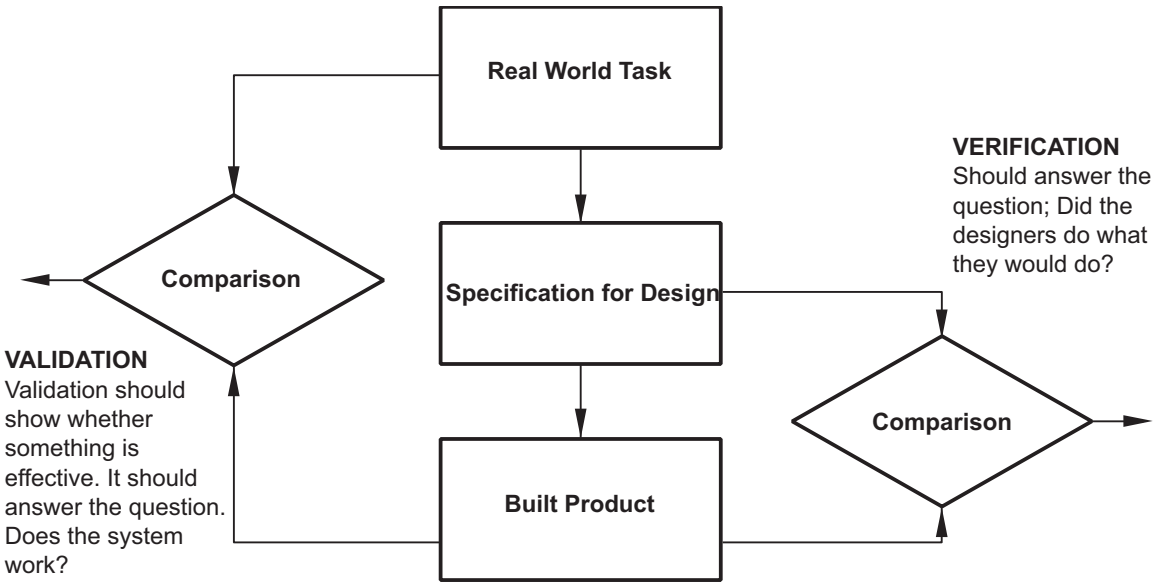

Figure 1 — The Role of Verification and Validation, (V&V)

4 Requirements and Recommendations for the Evaluation Process

The following sections present general requirements and recommendations for the human factors evaluation process. (See annex A for a checklist.)

4.1 General Verification and Validation (V&V) Issues

1) The verification and validation (V&V) activities shall be an integrated part of the design process, cf. with ISO 11064 Part 1, Figure 2, and Figure 2 below.

2) The V&V activities shall take place throughout the life of a project.

3) Tests shall be done as early in the design process as possible, to allow modifications to be made.

NOTE For evolutionary changes (modernisation, upgrading, etc.) a comprehensive V&V programme is not necessary to carry out for every upgrade. Previous V&V work can be reused under certain conditions. Final determination of what form of V&V is acceptable for evolutionary changes must be decided in each particular case. For further information, see annex B.

4.2 Verification and Validation Plan

1) A V&V plan shall be prepared early in the project and before the V&V work is carried out.

NOTE The plan would be expected to contain, as a minimum, details of:

— The objectives for V&V.

— The mandate and terms for V&V.

— The relationship and interfaces of V&V to other elements both within and outside that of the project, for example, the design process and the quality assurance programme.

— The V&V team, its primary responsibilities, and resources available to it.

— A description of approach taken to V&V.

— How the process will be applied.

2) The plan should detail the time requirements, relations and dependencies between the tasks within the evaluation process.

3) The plan for evaluation should have an entry for each topic being reviewed.

4) The plan should document all the criteria, the techniques and tools to be utilised in the evaluation process.

5) The plan shall describe the activities to be performed, and for the verification case describe each phase to show whether the requirement specification is met.

6) For the validation case, the project should develop performance and safety objectives for the topic under review if applicable.

7) Estimates of the resources required to undertake V&V tasks shall be prepared and include staff, equipment, accommodation and subjects for trials.

4.3 Verification and Validation Scope

1) The evaluation scope should be appropriate for the stage of the project at which it is performed.

2) The validation process should challenge the design and ascertain that the system will perform acceptably under a broad range of operating conditions. The validation should include consideration of appropriate scenarios or working sequences that should cover normal operation, a mix of multiple failure events and disturbances, and emergency conditions.

3) There should be written description of appropriate operating situations, adapted to the chosen verification / validation method and the stage of the project.

4) The general scope of the V&V should include all essential facilities defined in the project plan.

NOTE The V&V scope might cover, amongst other items, the following topics:

— Hardware having a human-system interface (HSI).

— HSI software.

— Communications facilities.

— Procedures (written or electronic form).

— Workstation and console configurations.

— Design of the overall work environment.

— Training and selection of personnel.

— Team working.

— Auxiliary shutdown rooms and panels, etc.

— Local control rooms.

— Local control panels or stations.

— The needs of maintenance personnel.

— Other needs of the operators, (e.g., storage, pause area, rest rooms, etc.).

4.4 Verification and Validation Criteria

1) The criteria developed shall cover the complete set of human factors topics that are relevant to a project.

2) Criteria should be defined for the evaluations of each human factors topic and for the objectives that the evaluation is intended to reach.

NOTE The criteria can be derived from the source documents in use for the project:

— Performance aspects.

— Safety principles.

— Availability and reliability requirements.

— Operator interface and display principles.

— Requirements from applicable standards and guidelines.

— Recommendations and requirements from human factors literature.

Performance criteria can be classified into several types, e.g.:

— Requirement-referenced criteria; the comparison of the performance of the system to an accepted performance requirement.

— Benchmark-referenced criteria — the comparison of the performance of the system to a benchmark system that is defined as acceptable.

— Normative referenced criteria — the comparison of the performance of the system to norms established for the performance based on many system evaluations.

— Expert-judgement referenced criteria — the comparison of the performance of the system to criteria established through the judgement of subject-matter experts.

4.5 Verification and Validation Input Documents

1) The design project's evaluation team should collect all important documentation related to the topic under consideration and used in the design process.

NOTE The documentation will be the basis for the human factors V&V process.

2) A design project's evaluation team should have access to appropriate documentation.

3) The evaluation team should have access to those members of the team that are responsible for design and documentation.

4) The evaluation team should have access to a human factors operating experience review.

4.6 Verification and validation team

1) The human factors evaluation team should be independent of, but have access to, the design team.

2) The communication between the independent human factors evaluation team and the designers should be supported and stimulated.

3) The human factors evaluation team should be suitably placed in the project organisation, i.e., have responsibility, authority and placement within organisation, such that the commitment to human factors V&V is achieved.

4) The specific expertise represented in a human factors evaluation team should be based on the scope of the evaluation.

NOTE A team might include the following areas of expertise:

— Systems engineering.

— Architectural design and civil engineering.

— Systems analysis.

— Instrumentation and control systems.

— Information and computer systems.

— Human factors engineering.

— Facility operation and training, (user representatives).

4.7 Verification and Validation Resources

1) The design project shall supply suitable resources for the evaluation team.

2) Suitable working materials for the conduct of V&V should be prepared.

NOTE Working materials might include:

— Documentation control.

— Control centre components and features.

— Measurements – noise, lighting, heating.

— Questionnaire and interview records.

— Records of operator responses to specific tests (e.g., simulator based tests or assessments).

— Human engineering discrepancies (HEDs) — to identify their location and nature so that follow-up action can be taken.

— Resolution of HEDs.

4.8 Verification and Validation Methods and Measures

1) The evaluation method(s) and/or technique(s) used should be systematic and well documented.

2) The evaluation process should, as far as possible, include quantitative measures of the required features and performance.

NOTE The following are examples of performance measures for dynamic evaluations:

— Systems performance measures relevant to facility safety, (e.g., by keeping specific process parameters within a certain range).

— Crew primary task performance, e.g., task times, procedure violations.

— Crew errors.

— Situation awareness.

— Workload.

— Crew communications and co-ordination.

— Dynamic anthropometry evaluations.

— Physical positioning and interactions.

4.9 Verification and Validation Results

1) The results from the evaluation should be recorded and documented, including any deviations from criteria.

2) The process for assessing deviations found in the evaluation should be systematic and documented.

3) All deviations found in the evaluation should be acted on.

4) The evaluation team should check for any risk of side effects of any design changes made because of deviations or non-conformities.

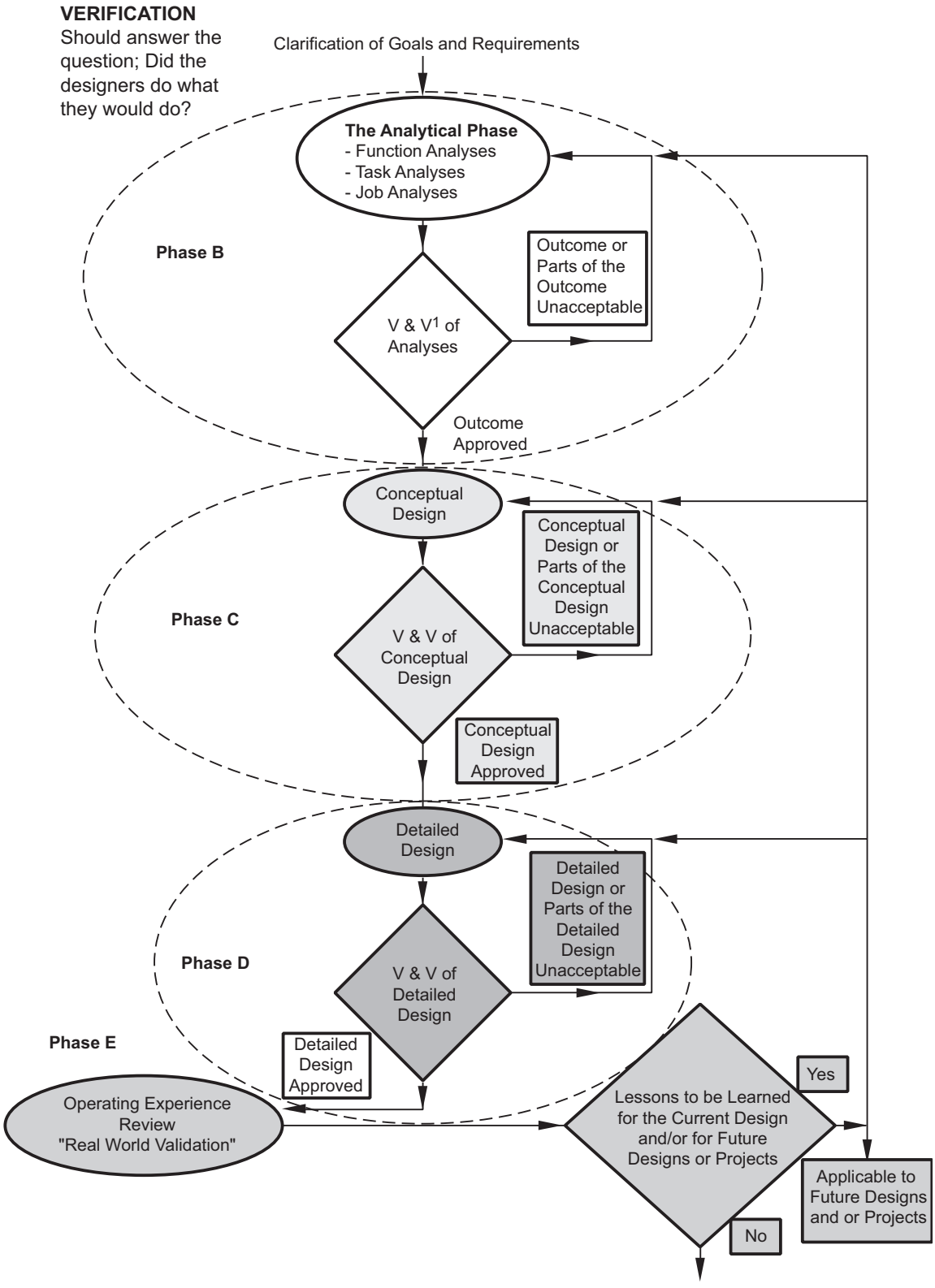

Figure 2 — Integrated V & V in the design process

5 Evaluation (Verification and Validation) Measures

The evaluation process should be design oriented, practical, and effective.

Fast and inexpensive evaluation methods should be used wherever possible and the more sophisticated and expensive methods restricted to those evaluations that require them.

Since overall goals such as safety and availability are often hard to measure, other aspects may need to be addressed during evaluation of control centres and human-system interfaces. These might include: compatibility, understandability, situation awareness, controllability, workload, teamwork, learnability, effectiveness, efficiency and satisfaction.

'Compatibility' means that the nature of physical presentations to the operators, and responses to be expected from the operators, should be compatible with human input-output abilities and limitations. Regardless of overall system objectives, operators should be able to read displays, reach controls, etc.

'Understandability' means that the structure, format and content of the human-system dialogue should result in meaningful communication. The information displayed should be easily understood, and the manual control actions should achieve the desired system response.

'Situation awareness' means that the situation is understood and based on knowledge of the history and the present status; it is possible to predict future developments to a certain extent.

'Controllability' means to have a certain control of the present situation, knowledge of the history that has led up to the existing status, and based on this know what to do next.

'Workload' in the control room context focuses mainly on mental workload. O'Donnell and Eggemeier (1986) formulated that "The term workload refers to that portion of the operator's limited capacity actually required to perform a particular task". The theoretical basis behind this definition is that the operator has limited processing capacity.

'Teamwork' - the major factors usually listed when describing effective team processes concern its 'potency'. This includes social support for team members by helping each other. Other factors include positive social interactions, sharing of workload, communication and cooperation within the team. All these factors are positively related to team effectiveness, productivity and satisfaction.

'Learnability' means that naive users can easily learn how to use the system with little or no need to consult manuals.

A 'human-system environment' is effective if it supports the operator (or crew) to improve their performance. Improved performance means to make a difficult task easier or enable an operator to accomplish a task that might otherwise be impossible.

'Effectiveness', 'efficiency' (i.e., the resources expended in relation to the accuracy and completeness with which users achieve goals), and 'satisfaction' (i.e., freedom from discomfort and positive attitudes with which users achieve goals), together form the three measures of usability. ISO Standard 9241-11 gives more details on how to measure usability.

5.1 Applicable techniques

Many human factors evaluation techniques are applicable in a control centre context. A few of the most commonly used techniques are briefly described in annex C, (for more information, see IEEE Std 845). The evaluation techniques may be divided into different categories that are related to the way each technique is used. The following categories are included in annex C:

1) Paper and pencil techniques.

2) Observational techniques.

3) Expert opinion techniques.

4) Experimental techniques.

Annex A
(informative)

Checklist Related to the Evaluation Process of Requirements and Recommendations

No.	Requirements and Recommendations	Yes	No	N/A	Comments
General Verification and Validation (V&V) Issues					
1	Are the V&V activities an integral part of the design process?				
2	Do the V&V activities take place throughout the life of a project?				
3	Are tests performed early in the design process?				
Verification and Validation Plan					
4	Is a proper V&V plan prepared early in the project?				
5	Does the V&V plan detail items such as time requirements, relations and dependencies between the tasks within the evaluation process, and does this plan extend throughout the entire project's duration?				
6	Does the V&V plan have an entry for each topic being reviewed?				
7	Does the plan document all the criteria, the techniques and tools to be utilised in the evaluation process?				
8	Does the plan describe the activities to be performed, and for the verification case describe each phase to show whether the requirement specification is met?				
9	Does the project define specific objectives for the V&V of a topic under review?				
10	Have estimates of the resources required undertaking V&V tasks including staff, equipment, accommodation, subjects, etc. been prepared?				
Verification and Validation Scope					
11	Is the V&V scope appropriate for the stage of the project at which it is performed?				
12	Does the V&V include consideration of all appropriate operating conditions?				
13	Are there written descriptions of appropriate operating situations, adapted to the chosen V&V method and the stage of the project?				

No.	Requirements and Recommendations	Yes	No	N/A	Comments
14	Does the general scope of the V&V include all essential facilities and locations defined in the project plan?				
Verification and Validation Criteria					
15	Do the criteria that are developed include a complete set of human factors topics that are relevant to the project?				
16	Are criteria developed for the evaluations of each human factors topic?				
Verification and Validation Input Documents					
17	Are all important documentation related to the topic under consideration and used in the design process collected by the project's human factors V&V team?				
18	Do the design project's V&V team have access to appropriate documentation?				
19	Do the V&V team have access to members of the team that is responsible for design and documentation?				
20	Do the V&V team have access to a human factors operating experience review?				
Verification and Validation Team					
21	Is the V&V team independent of the design team?				
22	Is the communication between the independent V&V team and the designers supported and stimulated?				
23	Is the V&V team suitably placed in the project organisation such that the commitment to human factors V&V is achieved?				
24	Is the specific expertise represented in a V&V team based on the scope of the evaluation?				
Verification and Validation Resources					
25	Do the design project supply suitable resources for the V&V team?				
26	Are suitable working materials for the conduct of V&V prepared?				
Verification and Validation Methods and Measures					
27	Are the evaluation method(s) and/or technique(s) used systematic and well documented?				
28	Does the V&V process include quantitative measures of the required features and performance?				
Verification and Validation Results					
29	Are results from the evaluation recorded and documented, including any deviations from criteria?				

No.	Requirements and Recommendations	Yes	No	N/A	Comments
30	Is the process for consideration of deviations found in the evaluation systematic and documented?				
31	Are all deviations found in the evaluation acted on?				
32	Are there checks for side effects of any design changes made because of deviations or non-conformities?				

Annex B
(informative)

The Process of Evaluation

B.1 Use of Existing V&V Information

Where evolutionary changes are being made, information often exists already, such as analyses from previous design documents, procedures, and operation experience. Together these can constitute an important pre-validated data set. This data set can be used to meet some of the requirements of the verification and validation (V&V) process, although issues such as the degree of change and the quality of existing material must obviously also be taken into account. IEC 61771 notes that the V&V activities need to be tailored to the particular needs and circumstances of individual projects. The basic framework for carrying out a V&V is, however, constant; i.e., the stages of preparation, evaluation and resolution are retained.

a) Preparation (to prepare the V&V).

b) Evaluation (to actually perform the V&V).

c) Resolution (to identify and implement solutions to the deviations identified during the V&V).

The additional work that does, or does not, take place under these headings should be justified and documented.

Two important aspects when deciding the V&V requirements for projects of this nature are the ‘degree of innovation’ and the possibility of ‘qualification by similarity’. The *degree of innovation* relates to those areas of innovation in the change and concentrates V&V activities on them. The degree of innovation varies along a continuum from a replica of an existing design, which would require very little V&V, to an evolutionary design requiring selected V&V activities, to an advanced design required the full scope of V&V activities. For evolutionary changes, V&V activities can be concentrated on the areas of change and their integration with existing, proven features of the design.

Besides this, the potential to affect or influence risk levels should be considered. Existing safety analyses can help here.

B.2 New V&V Information

In an upgrade, there is a need to verify and validate new and innovative aspects, including their interaction with the existing facility. A number of issues relevant to the V&V process for evolutionary changes can be identified including:

— The use and consideration of current and previous change programmes and their objectives and philosophies. (Use of existing documentation.)

— Consideration of the possible effects of the change on other aspects of work and organisational factors.

— The effect of the changes on training requirements, simulators, procedures and other relevant elements.

— The way changes will be introduced and whether parallel use of old and new system is desirable for V&V.

— The implementation of modifications in a facility simulator where appropriate V&V can take place.

B.3 The changing nature of facility design and control room tasks

Changes in control centre systems and equipment may affect the role of operators and their tasks both during normal operations and during emergencies. There are, for example, changes in the interface, tasks and functions allocated to the operator, including:

— Greater use of automation.

— A shift of the operator's role from active involvement to monitoring, supervision, and backup of automated systems.

— Greater centralisation of controls and displays, both on a station/facility basis and within the control room.

— Use of large displays in the control centre that allows for a shared viewing of high-level or summary information and critical parameters.

— A shift of the operator's primary interface, from direct interaction with components to interaction with a data based system.

— increased use of integrated displays and graphical displays.

— Greater use of information-processing aids and decision-support aids.

NOTE If the operator's role has changed in this way, it will be more difficult to argue for qualification by similarity or to claim that the degree of innovation is small.

These trends affect the design, and equipment, in both new facilities and existing control centres. There may be a range of technologies and solutions to the design of the human-system interface at any one location, even if it is a new control centre. In an existing facility there may be a range of degrees of upgrading. These changes mean that any human factors programme, and V&V of it, must allow for a diversity of approaches to control and display. It must be particularly sensitive to new problems created.

New problems can arise because there is a potential to affect human performance, to create new types of human error and to reduce human reliability in new ways. Because these new effects on human performance tend to be of a different kind from those found in conventional control centre, they are at first less obvious and less likely to be understood, or even recognised. The human factors programme must address these issues and resolve them in some way. Some of these new threats to human reliability are mentioned below:

— *Lack of Knowledge* — Cognitive issues are emerging as more significant than the physical ergonomic considerations of control centre design that have heretofore dominated the design of conventional interfaces, and indeed human factors as a subject.

— *Changes in Function Allocation* — Increases in automation have tended to result in a shift from physical workload to cognitive workload. As a result, there are dangers such as loss of vigilance, loss of situation awareness, and eventually, loss of a full understanding of the processes as the operator is taken more and more 'out of the loop'.

— *Changes in Cognitive Implications of Designs* — Systems have changed in several ways. Information tends to be more pre-digested, information is resident on a workstation or computer system rather than physically located in a room, there is a greater quantity of information, and there is an additional burden of operating the interface equipment. These lead to a greater need to specify system requirements in cognitive rather than physical terms. This requires techniques, such as cognitive task analysis.

— *Changes in Skill Demands* — Although systems are increasingly automated, they also create new, usually highly skilled tasks for operators. Operators must understand and evaluate the performance of automatic systems, or even take over from them when they fail. It is difficult to see how this level of skill can reasonably be expected of operators, when the same automation has made their daily tasks more boring and monotonous.

These points make clear that the changing nature and equipment in control rooms itself changes the roles, functions and tasks of the control centre and the staff within it. This in turn puts requirements on the kind of human factors work that is needed.

In response to these problems, many organisations have begun to look more seriously at the implications of advanced control centre systems. It is often difficult to set pass/fail criteria or to prescribe methods in advance for some of these new problems. There has consequently been an increased emphasis that the user organisations/utilities should give evidence of a design process and a V&V process that can stand up to scrutiny and create confidence that a design is satisfactory.

B.4 Sources of Confidence in a Design

When it comes to human factors, it is important that:

— The design follows accepted human factors principles.

— The design supports the performance of the operators.

— The design supports the reliability of operators.

— The design is resilient to changes and upgrades.

V&V of the human factors aspects of a design is just one source of confidence that a design is satisfactory. There are several sources of evidence for the efficacy of the human factors design as shown in Table 1.

Further confidence in a design can be gained by a detailed test programme of the actual facility and through successful operation of it. The record of operation can also be a source of validation early in the design process for the next similar design or upgrade.

Table B.1 — Types of Information for Assessment of Human Factors Engineering (HFE) Adequacy

Type of evidence	Minimal evidence	Best evidence
Planning of human factors activities	An HFE design team, a programme plan and methods for doing the work	A qualified HFE design team with all the skills and resources required, using an acceptable HFE programme plan
Design analysis work	Function requirement analysis, task analysis, task synthesis, assessments of alternative technologies	Results of appropriate HFE studies, analyses that provide accurate and complete inputs to the design process and V&V assessment criteria
Record of the design	Specifications and descriptions of designs	Designed using proven technology based on human performance and task requirements incorporating accepted HFE standards and guidelines
Verification and validation of the project	Compliance with HFE guidelines and project specifications, operation of the integrated system under actual or simulated conditions	Evaluated with a thorough V&V test programme throughout the project
Use of feedback from other systems	Simple collection of operational experiences from earlier projects or systems	Performance of a comprehensive operational experience review

B.5 Timing of V&V within the Design Process

It is difficult to find guidance on when in the design process V&V is best applied. Historically there has been a tendency to focus on V&V at the end of a project - after the design work has been completed. More recently, there has been general agreement that V&V should be more iterative and integrated into the design process, although guidance as to exactly when and how often V&V should be carried out is less clear.

It is proposed that lower fidelity test-beds are used for addressing human performance issues much earlier in the design process to allow modifications to be made with minimal effect on the overall HSI system. It is suggested that use is made, for example, of different grades of modelling technology and part-task simulators comprising both individual and partially integrated sets of prototype components. Also dynamic simulations of selected parts of the process may be used. These simulations should be performed as soon as they are available.

Annex C
(informative)

Evaluation (Verification and Validation) Methods

C.1 Applicable techniques

In a control centre context many different human factors evaluation techniques are available though no single technique can usually handle the full problem. This leads to the use of a combination of techniques.

Some of the most commonly used techniques are briefly described in this annex C, (for more information, see IEEE Std 845). annex C contains only a few examples and is not intended to be a complete list. The evaluation techniques listed below are divided into four categories that are related to the way each technique is used. The following categories are included:

— Paper and pencil techniques.

— Observational techniques.

— Expert opinion techniques.

— Experimental techniques.

C.1.1 Paper and pencil techniques

No actual performance observation is required when using paper and pencil techniques. No prototypical hardware/software is required for most of these methods and the outcome can be a simple accept/reject decision or a ranking.

C.1.1.1 Human factors checklist

A very common technique is the use of a checklist to verify that a design meets certain criteria. A checklist can be used to best evaluate issues related to compatibility. This approach is most applicable during the design process but can be used in a confirmatory fashion.

The technique is easy to use and has high face validity when applied properly. It is very sensitive to those characteristics of systems with easily measurable parameters such as height, colour, etc. The cost of using a checklist is low but its output tends to be categorical.

C.1.1.2 Historical review

This technique involves the examination of historical records related to the performance of systems that are identical or similar to the system under evaluation. In certain application areas this technique typically involves the use of significant event reports or incident reports, trip reports, operational logs, interviews with operators, etc.

Historical review is most useful for evaluating issues related to system effectiveness in the real setting where the system performance can be evaluated during operation. The face validity is high, however, the predictive validity is dependent on the data available for review and the similarities between historical system applications and the proposed new application. Its output tends to be qualitative.

C.1.1.3 Task analysis

Task analysis is the name given to any process that identifies and examines the tasks that must be performed by users when they interact with the system(s) to be evaluated. Task analysis is often considered to be a primary system's design tool, but can also be used as an evaluative instrument. However, it is necessary to thoroughly consider the suitability of using task analysis for evaluation if the technique has been used as a design tool.

Task analysis should be used early in the design process to address issues of compatibility and understandability. The technique may be time consuming if done in enough detail to be useful. It is manpower intensive and, thus, moderate to highly expensive. The output of task analysis is normally used as the basis for further evaluative analysis. It is sensitive to most design concerns, with the exception of crew interactions and time dependencies. Task analysis is capable of identifying subtle human-system interaction problems, and its output is a mixture of qualitative and quantitative data.

C.1.1.4 Logic tree application

Probabilistic risk assessment (PRA) and human reliability assessment (HRA) belongs to this category. They are used primarily to estimate human error rates, and in particular to identify dominant or the most likely accident sequences.

Their application is to evaluate the effectiveness of existing or planned systems, and is most useful in the design process when prediction of the safety effects of a new or revised system is of interest. The techniques have moderate face validity.

C.1.2 Observational techniques

The evaluator(s) examine operator performance on the system to be evaluated using observational techniques. One major factor to be considered is the setting in which the observation takes place. The setting has an impact on the cost and ease of use of the technique. Three examples of basic settings are:

— Mock-ups.

— Full-scope simulators.

— Real environment.

C.1.2.1 Walk-through/Talk-through

The most widely used observational technique is the walk-through/talk-through technique. The technique consists of having potential users of the system under test walking and talking through (in the sense of physically showing and verbally describing) one or more of the tasks that will be done using that system when it is operational. A walk-through/talk-through is used during the design process, after a prototypical system is available.

The face validity is very high, and the predictive validity is limited depending on the similarity between the test condition and the real environment. Likewise, the cost of this technique is dependent on the facilities in which it is conducted. The technique provides qualitative output.

C.1.2.2 Time line analysis

Time line analysis is very similar to the walk-through technique, but it is used to determine the time required to perform the tasks related to the system under test and their interdependence. Time line analysis can be conducted in any of the settings described earlier. The data-taking requirements are more stringent, because every action must be timed to a reasonable degree of accuracy. The

technique can be used either in the design environment or in an operational setting. The best condition is, however, in a realistic and dynamic setting.

Time line analysis is resource intensive and expensive. It has high face validity and the predictive validity is usually high. The output is quantitative and is as precise as the method used to measure the times involved. The technique is valuable since it can identify situations where the operator has to perform two or more actions simultaneously. It is not particularly easy to use due to the precise data collection.

C.1.2.3 Automated performance tracking

This technique combines some features of walk-through and time line analysis. Automated performance tracking requires a part-task or full-scope simulator to collect performance data while users interact with the system under test. The automated performance tracking system records all control and switch manipulations with time tagging. Their weakness is the vast amount of recorded data must be analysed to provide any meaningful evaluative information. The technique can be used only with functional hardware. Its face validity is high and it is a non-intrusive technique.

The face validity of automated performance tracking is normally high, again dependent on the similarity between the test environment and the real task situation. This technique is the most expensive due to the test system requirements.

C.1.3 Expert opinion techniques

Expert opinion consists of soliciting opinions from individuals with expertise in a particular content area that relates to the system being evaluated. The individual techniques described below are all designed to make expert opinions more objective and precise.

C.1.3.1 Delphi technique

The Delphi method is a consensus technique that should be used to estimate whether a human-system environment is adequate. The method develops a scale that can be used as a continuum for any performance related characteristic (e.g., error probability; display readability; understandability; etc.) of the system. The technique uses surveys (questionnaires) that are filled out by the experts at their own locations. The analyst collects and collates the responses, then sends out a second round of surveys that detail the opinions that were received, (without revealing which expert made each comment). This process is repeated until a consensus is reached.

The technique is useful for a design process when no equipment exists. The output is both qualitative and imprecise, however, the technique can quickly highlight likely extreme values in a performance measure. The predictive validity is not particularly high, although this depends very much on the experience of the experts used. This is also valid for the face validity. The technique has low costs.

C.1.3.2 Nominal group technique

This technique is similar to the Delphi technique since multiple experts estimate how a particular human-system environment falls on a scale of one or more continuums that are related to the system's design characteristics.

The major difference between this technique and the Delphi method is that this technique strives for a consensus of judgements in face-to-face meetings. Special precautions have to be taken to ensure that the consensus is not swayed by one individual. Otherwise the features of this technique are very similar to those of the Delphi method.

C.1.3.3 Paired comparisons

There are several variants of paired comparison techniques, however, the basis for these types of techniques is to present each expert with two of something and ask which one is more (larger, brighter, probable, etc.) This procedure is repeated for all possible pairs of the items of interest in the evaluation. The judgements are then sorted and compared to determine the judgement scale and dimension. Paired comparison techniques are limited to relatively imprecise, qualitative data. The outcome in terms of data is more reliable than data from the Delphi and nominal group techniques, but paired comparison is more expensive.

C.1.3.4 Ratio estimation

In ratio estimation technique, the experts are not asked to make absolute numerical estimates or to compare two things on a relative scale, but to judge whether one thing is one-half or twice one item of the set that has been designated as the standard. The instructions to the experts require them to make ratio judgements. The data collected by the ratio estimation technique are much more quantitative compared to other expert opinion techniques, however, the shortcomings are the same.

C.1.4 Experimental techniques

All experimental techniques may be used to measure statistically significant differences among candidate human-system environment designs in a laboratory setting or a simulator. They can also be used for validation of a single human-system environment.

The distinguishing feature of experimental techniques is the requirement to tightly control all sources of performance variation. This necessitates rather large and well-balanced pools of potential users, very well controlled settings, and sophisticated data recording. The techniques are preferably applied during design, but needs careful planning, are time consuming and very expensive. Their face validity is generally low because of the requirement to reduce extraneous sources of performance variation, task complexity and realism. For the performance that is measured, these techniques are usually sensitive. Their output is quantitative, but usually categorical.

Bibliography

1. ANSI, ANSI/AIAA G-035-1992, Guide to Human Performance Measurements, Washington D.C., 1993.

2. Endsley, M.R. (1988). Design and Evaluation for Situation Awareness Enhancement. In Proceedings of the Human Factors Society 32nd Annual Meeting)pp. 97-101). Santa Monica, CA, Human Factors and Ergonomics Society.

3. Fink, A., Ed., (1995*).* How to measure survey reliability and validity, v. 7*.* Thousand Oaks, CA, Sage.

4. Handbook of Human Factors and Ergonomics, (1997). Second Edition, G. Salvendy, Ed., John Wiley & Sons, Inc., New York, NY.

5. Handbook of Human Factors Testing and Evaluation, (1996). T.G. O'Brien and S.G. Charlton, Ed., Lawrence Erlbaum Ass. Publishers, Mahwah, New Jersey.

6. IEC 60964, Design for Control Rooms of Nuclear Power Plants.

7. IEEE Standard 1023, Guide for the Application of Human Factors Engineering to Systems, Equipment, and Facilities of Nuclear Power Generating Stations.

8. Interim DEF STAN 00-25, (Part 4, Part 12)/Issue 1, Ministry of Defence, UK, 1989, *Human Factors for Designers of Equipment. Part 4: Work Place Design.* Glasgow: Ministry of Defence, Directorate of Standardisation.

9. NUREG/CR-6393, 1996, Integrated System Validation: Methodology and Review Criteria. Washington: U.S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research. Authors O'Hara, J., Stubler, W., Higgins, J., Brown, W.

10. NUREG-0700, Rev.2, 2002, Human System Interface Design Review Guidelines. Washington DC: U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research.

11. NUREG-0711, Rev.1, 2002, Human Factors Engineering Program Review Model. Washington DC: U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research.

12. O'Donnel, R.D. and Eggemeier, F.T. (1986). Workload assessment methodology. In K.R. Boff et al Eds., Handbook of Perception and Human Performance: Volume II. Cognitive Processes and Performance. New York: John Wiley, Chapter 42.