Human Factors Design Methods Review

Authors

P. Salmon **Brunel University**

N.A. Stanton **Brunel University**

G. Walker **Brunel University**

Reference HFIDTC/WP1.3.2/1

Version 1

Date ..... 28 November 2003

The work described in this document has been undertaken by the Human Factors Integration Defence Technology Centre, part funded by the Human Sciences Domain of the U.K. Ministry of Defence Scientific Research Programme.

Abstract

This report describes the human factors (HF) design methods review that was conducted as part of work package 1.3.2 - Design methods review, and is part of work package 1, 'Human Factors Integration for C4i Systems'. The overall aim of work package 1.3.2 was to review and evaluate HF methods and techniques suitable for use in the design lifecycle for C4I systems. This document is intended to act as a guide for HF practitioners in the selection and use of appropriate HF techniques.

A great number of HF techniques exist. A survey of existing techniques identified over 200, including human error identification, human reliability analysis, task analysis, situation awareness measurement, mental workload measurement, usability evaluation and charting techniques. A shortlist of 48 design techniques were selected for further review. 58 further HF techniques were also short listed for review in work package 1.3.3 – Evaluation method review. Each of the 48 techniques were evaluated against a set of fourteen criteria designed to determine the suitability of the techniques use in the design of C4 systems. The resultant output is presented in a user manual, offering guidelines and advice on how to use each technique.

Contents

Executive Summary……
Introduction……..4
Introduction to HEI……10
1. SHERPA……15
2. HET – Human Error Template……..23
3. TRACEr ……30
4. TAFEI – Task Analysis For Error Identification……...37
5. Human Error HAZOP……47
6. THEA – Technique for Human Error Assessment……54
7. HEIST – Human Error Identification in Systems Tool…….63
8. The HERA framework……69
9. SPEAR - System for Predictive Error Analysis and Reduction…….75
10. HEART - Human Error Assessment and Reduction Technique……79
11. CREAM – Cognitive Reliability Analysis Method……...86
Introduction to charting techniques……92
12. Process Charts……94
13. Operator event sequence diagrams……...100
14. Event tree analysis……104
15. DAD – Decision Action Diagrams……...108
16. Fault tree analysis…….112
17. Murphy Diagrams…….116
Introduction to Task Analysis……...120
18. HTA – Hierarchical Task Analysis……...122
19. CPA – Critical Path Analysis……...127
20. GOMS……...134
21. VPA – Verbal Protocol Analysis……..138
22. Task Decomposition ……144
Introduction to Cognitive Task Analysis……..150
23. ACTA – Applied Cognitive Task Analysis……..152
24. Cognitive Walkthrough……159
25. CDM – Critical Decision Method……165
26. Critical Incident Technique……..171
Introduction to Team Task Analysis……175
27. Comms Usage Diagrams……..177
28. Social Network Analysis……..182
29. Groupware Task Analysis……187
30. Team Task Analysis……191
31. HTA (T)……195
Introduction to design techniques…….200
32. Link Analysis……202
33. Layout Analysis……207
34. TCSD – Task-Centred System Design…….212
35. Scenario analysis……..218
36. Checklists…….222
37. Heuristic analysis…….226
38. Focus groups……229

39. Walkthrough analysis……...234
Introduction to data collection techniques……..237
40. Interviews……238
41. Questionnaires…….243
42. Observation analysis……249
Introduction to performance time prediction techniques…….254
43. KLM……256
44. Timeline Analysis……261
45. CPA – Critical Path Analysis……...265
Introduction to Mental Workload prediction……266
46. Cognitive Task Load Analysis…….268
47. Pro-SWAT – Subjective Workload Assessment Technique……273
48. Pro-SWORD – Subjective Workload Dominance……273

List of Appendices

Appendix 1 – Methods database
Appendix 2 – Rejected

1. Introduction

The purpose of this document is to present a review of human factors (HF) design methods that could potentially be used in the design of C4i systems. This document represents work package 1.3.2 'Design methods review' and is part of work package 1, 'Human Factors Integration for C4i Systems'. The overall aim of work package 1.3.2 was to review and evaluate HF methods and techniques suitable for use in the design lifecycle for C4I systems. Each HF technique was evaluated against a set of pre-determined criteria and presented in a user manual, offering guidelines and advice on how to use each method.

The following methods review was conducted in order to evaluate the potential use of techniques in the design and analysis of C4i systems. Work package 1 involves the analysis of current C4 systems in a number of industries, such as air traffic control, railway and gas. HF techniques are required in order to record data regarding existing C4 systems and procedures and also to and represent existing C4 practices. The evaluation of existing C4 systems is also required. This methods review aims to contribute to the specification of HF techniques used for such purposes. Work Package 3 involves the analysis of the current practice of HFI in military and civilian domains. It is proposed that a number of the techniques outlined in this review will be used for this process and also to interpret, evaluate and understand the processes described in subsequent work package 3 outputs. Work package 1.4 involves the design of a C4 system. It is intended that the most suitable techniques outlined in this review will be used during the design process. The techniques used will also be evaluated as part of work package 1.3.4.

Stage 1 – Initial Literature Review of Existing HF Methods

A literature review was conducted in order to create a comprehensive list of existing HF methodologies. The purpose of this literature review was to provide the authors with a comprehensive database of available HF methods and their associated author(s) and source(s). The literature review was based upon a survey of standard ergonomics textbooks, relevant scientific journals and existing HF method reviews. At this stage, none of the HF methods were subjected to any further analysis and were simply recorded by name, author(s) and class of method (e.g. mental workload assessment, Human Error Identification, Data collection, Task analysis etc). In order to make the list as comprehensive as possible, any method discovered in the literature was recorded and added to the database. The result of this initial literature review was a database of over 200 HF methods and techniques, including Human Error Identification (HEI) techniques, Human Reliability Analysis (HRA) techniques, Mental Workload Assessment techniques, Task analysis techniques, Interface analysis techniques, Data collection techniques, usability evaluation techniques and design techniques. The HF methods database is presented in appendix 1 of this report.

Stage 2 – Initial Methods Screening

Before the HF techniques were subjected to further analysis, a screening process was employed in order to remove any techniques that were not suitable for review with respect to their use in the design of C4i systems. The list of rejected techniques can be found in appendix 2 of this report. Techniques were deemed unsuitable for review if they fell into the following categories.

- Availability – the technique should be freely available in the public domain. The techniques covered in this review included only those that were freely

available. Due to time constraints, techniques that could be obtained only through order were rejected.

- Evaluation techniques – Work package 1.3.2 entails a review of design techniques. Work package 1.3.3 entails a review of evaluation techniques. Evaluation techniques were rejected, as they are reviewed in work package 1.3.3
- Software – Software based techniques are time consuming to acquire (process of ordering and delivery) and often require a lengthy training process. Any HF software tools (e.g. PUMA) were rejected.
- Applicability – The applicability of each technique to C4i systems was evaluated. Those techniques deemed unsuitable for the use in the design of C4i systems were rejected e.g. anthropometric techniques, physiological techniques.
- Replication – HF techniques are often re-iterated and presented in a new format. Any techniques that were very similar to other techniques already chosen for review were rejected.
- Limited use – Often HF techniques are developed and not used. Any techniques that had not been applied in an analysis of some sort were rejected.

As a result of the method screening procedure, a list of 48 design HF methods suitable for use in the C4i design process was created. This HF design methods list was circulated internally within the HFI-DTC group to ensure suitability of the methods chosen for review, and the comprehensiveness of the HF design methods list. The HF design list was also reviewed independently by Peter Wilkinson of BAE systems. A second list of 58 evaluation HF methods was also created. The second list of methods will form the basis of the evaluation methods review in work package 1.3.3. Table 2 shows the 48 HF design methods subjected to further evaluation in this review. The methods review is divided into nine sections, each section representing a specific category of method or technique. The sequence of the sections and a brief description of their contents are presented in table 1. The nine sections are intended to represent the different categories of human factors methods and techniques that will be utilised during the C4i design process.

Stage 3 – Methods Review

The 48 HF design methods were then analysed using the set of pre-determined criteria outlined below. The criteria were designed not only to establish which of the methods were suitable for use in the design of C4i systems, but also to aid the HF practitioner in the selection and use of the appropriate method(s). The output of the analysis is designed to act as a methods manual, aiding practitioners in the use of the HF design techniques reviewed.

1. Name and acronym – the name of the technique and its associated acronym.
2. Author(s), affiliations(s) and address(es) – the names, affiliations and addresses of the authors are provided to assist with citation and help in using the method.
3. Background and applications – This section introduces the method, its origins and development, the domain of application of the method and also application areas that it has been used in.
4. Domain of application – describes the domain that the technique was originally developed and applied in.

5. Procedure and advice – This section describes the procedure for applying the method as well as general points of expert advice.
6. Flowchart – A flowchart is provided, depicting the methods procedure.
7. Advantages – Lists the advantages associated with using the method in the design of C4i systems.
8. Disadvantages - Lists the disadvantages associated with using the method in the design of C4i systems.
9. Example – An example, or examples, of the application of the method are provided to show the methods output.
10. Related methods – Any closely related methods are listed, including contributory and similar methods.
11. Approximate training and application times - Estimates of the training and application times are provided to give the reader an idea of the commitment.
12. Reliability and Validity - Any evidence on the reliability or validity of the method are cited.
13. Tools needed – Describes any additional tools required when using the method.
14. Bibliography - A bibliography lists recommended further reading on the method and the surrounding topic area.

Table 1. Methods categories

Method category	**Description**
HEI/HRA techniques	HEI techniques are used to predict any potential human/operator error that may occur during a man-machine interaction. HRA techniques are used to quantify the probability of error occurrence
Charting techniques	Charting techniques are used to depict graphically a task or process using standardised symbols. The output of charting techniques can be used to understand the different task steps involved a particular scenario, and also to highlight when each task step should occur and which technological aspect of the system interface is required.
Task Analysis techniques	Task analysis techniques are used to represent human performance in a particular task or scenario under analysis. Task analysis techniques break down tasks or scenarios into the required individual task steps, in terms of the required human-machine and human-human interactions
Cognitive Task analysis techniques	Cognitive task analysis (CTA) techniques are used to describe and represent the unobservable cognitive aspects of task performance. CTA is used to describe the mental processes used by system operators in completing a task or set of tasks.
Team Task Analysis techniques	Team task analysis techniques are used to represent team performance in a particular task or scenario. Tasks are broken down into those steps requiring individual performance and those steps requiring teamwork. Knowledge requirements are also typically described using TTA.
Design techniques	Design techniques represent techniques that are typically used during the early design lifecycle by design teams, including techniques such as focus groups and scenario-based design.
Data collection techniques	Data collection techniques are used to collect specific data regarding a system or scenario. According to Stanton (2003) the starting point for designing future systems is a description of a current or analogous system.
Performance time prediction techniques	Performance time prediction techniques are used to predict the execution times associated with a task or scenario under analysis.
Mental Workload prediction techniques	Mental workload represents the proportion of operator resources demanded by a task or set of tasks. A number of MWL assessment techniques exist, which allow the HF practitioner to evaluate the MWL associated with a certain task.

Table 2. Design methods reviewed

Technique	**Category**	**Author**
CREAM – Cognitive Reliability Error Analysis Method	Human Error Identification	Hollnagel (1998)
HEART – Human Error Assessment and Reduction Technique	Human Error Identification	Williams (1986)
HEIST – Human Error Identification In Systems Tool	Human Error Identification	Kirwan (1994)
HET – Human Error Template	Human Error Identification	Marshall et al (2003)
Human Error HAZOP	Human Error Identification	Whalley (1988)
SHERPA – Systematic Human Error Reduction and Prediction Approach	Human Error Identification	Embrey (1986)
SPEAR - System for Predictive Error Analysis and Reduction	Human Error Identification	CCPS (1993)
TAFEI – Task Analysis For Error Identification	Human Error Identification	Baber & Stanton (1996)
THEA – Technique for Human Error Assessment	Human Error Identification	Pocock et al (2000)
The HERA Framework	Human Error Identification	Kirwan (1998a, 1998b)
TRACer - Technique for the Retrospective and Predictive Analysis of Cognitive Errors in Air Traffic Control (ATC)	Human Error Identification	Shorrock & Kirwan (2000)
DAD – Decision Action Diagram	Charting technique	Kirwan and Ainsworth (1992)
Event Tree analysis	Charting technique	Kirwan and Ainsworth (1992)
Fault Tree analysis	Charting technique	Kirwan and Ainsworth (1992)
Operational Sequence Diagrams	Charting technique	Various
Process Charts	Charting technique	Kirwan and Ainsworth (1992)
CPA – Critical Path Analysis	Task Analysis	Baber
HTA – Hierarchical Task Analysis	Task Analysis	Annett, Duncan & Stammers (1971)
GOMS – Goals, Operators, Methods and Selection rules	Task Analysis	Card, Newell & Moran (1983)
Task Decomposition	Task Analysis	Kirwan & Ainsworth (1992)
VPA – Verbal Protocol Analysis	Task Analysis	Various
ACTA – Applied Cognitive Task Analysis	Cognitive Task Analysis	Militello & Hutton (2000)
Cognitive Walkthrough	Cognitive Task Analysis	Polson et al (1992)
CDM – Critical Decision Method	Cognitive Task Analysis	Klein (2003)
CIT – Critical Incident Tecgnique	Cognitive Task Analysis	Flanagan (1954)
HTA (T)	Team Task Analysis	Annett (2003)
Comms Usage Diagram	Team Task Analysis	Watts & Monk (2000)
Social Network Analysis	Team Task Analysis	Driskell & Mullen (2003)
Team Task Analysis	Team Task Analysis	Burke (2003)
Checklists	Design technique	Various
EHFA – Early Human Factors Analysis	Design technique	
Focus Groups	Design technique	Various
Goupware Task Analysis	Design technique	Van Welie & Van Der Veer (2003)
Heuristic Analysis	Design technique	Stanton & Young (1999)
Layout Analysis	Design technique	Stanton & Young (1999)
Link Analysis	Design technique	Drury (1990)
TCSD – Task Centred System Design	Design technique	Greenberg (2003) Clayton & Lewis (1993)
Walkthrough Analysis	Design technique	Various
Interviews	Data Collection	Various
Observation	Data Collection	Various
Questionnaire	Data Collection	Various
KLM – Keystroke Level Model	Performance time assessment	Card, Moran & Newell (1983)
Timeline Analysis	Performance time assessment	Various
CTLA – Cognitive Task Load Analysis	Mental Workload prediction	Neerincx (2003)
Pro-SWAT	Mental Workload prediction	Reid & Nygren (1988)
Pro-SWORD	Mental Workload prediction	Vidulich (1989)

Summary
As a result of the methods review described in this document, a number of the techniques reviewed are currently being used or have been selected for use in other HFI/DTC work packages. Table 3 describes which of the techniques are being used in which work packages.

Table 3. HF techniques used or selected for use in DTC work packages.

	Work Package		
Method	**1.1**	**1.2**	**1.4**
CDM – Critical Decision Method	■		■
Checklists	■		■
Comms Usage Diagram	■		■
CREAM – Cognitive Reliability and Error Analysis Method		■	
Focus groups			■
HEART – Human Error Assessment and Reduction technique		■	
HEIST – Human Error Identification in Systems Tool		■	
The HERA framework		■	
HET – Human Error Template		■	
Heuristics			■
HTA – Hierarchical Task Analysis	■	■	■
KLM – Keystroke Level Model		■	
Observation	■		■
Operator Sequence Diagrams	■		■
SACRI*		■	
SAGAT*		■	
SART*		■	
Scenario based design			■
SHERPA – Systematic Human Error Reduction and Prediction Approach		■	■
SNA – Social Network Analysis	■		
SPEAR		■	
TAFEI – Task Analysis For Error Identification		■	
TCSD – Task centred System Design			■
Timeline Analysis		■	
TRACEr		■	
TTA – Team Task Analysis	■		■
VPA – Verbal Protocol Analysis	■		■

The techniques used in work packages 1.1, 1.2 and 1.4 are to be evaluated In work package 1.3.4.

Human Error Identification (HEI) Techniques

Human Error Identification (HEI) techniques are used to predict potential human or operator error in complex, dynamic systems. Originally developed in response to a number of human (operator) error related high profile catastrophes in the nuclear and chemical power domains (Three Mile Island disaster, Bhopal, Chernobyl) the use of HEI techniques is now widespread, including applications in Nuclear power and petro-chemical processing industry (Kirwan 1999), air traffic control (Shorrock & Kirwan 2000), aviation (Marshall et al 2003), naval operations, military systems, and public technology (Baber & Stanton 1996). HEI techniques can be used either during the design process to highlight potential design induced error, or to evaluate error potential in existing systems and are typically conducted on a task analysis of the task or scenario under analysis. The output of HEI techniques typically describes potential errors, their consequences, recovery potential, probability, criticality and offer associated design remedies or error reduction strategies. A number of different variations of HEI techniques exist, including error taxonomy based techniques (SHERPA, HET), which offer error modes linked to operator behaviours, error identifier-prompt techniques (HEIST, THEA), which use error identifier prompts linked to error modes, and error quantification techniques (HEART), which offer a numerical probability of an identified error occurring.

Taxonomic based HEI techniques typically involve the application of error modes to task steps identified in a HTA, in order to determine credible errors. Techniques such as SHERPA, HET, TRACEr, and CREAM possess domain specific error mode taxonomies. Taxonomic approaches to HEI are typically the most successful in terms of sensitivity and also the quickest and easiest to use. However, these techniques place a great amount of dependence upon the judgement of the analyst. Different analysts often make different predictions for the same task using the same technique. Similarly, the same analyst may make different judgements on different occasions (inter-analyst reliability). This subjectivity of analysis may weaken the confidence that can be placed in any predictions made.

SHERPA (Embrey 1986) uses hierarchical task analysis (HTA) (Annett, Duncan, and Stammers 1971) together with an error taxonomy (action, retrieval, check, selection and information communication errors) to identify potential errors associated with human activity. The SHERPA technique works by indicating which error modes are credible for each bottom level task step in a HTA. The analyst classifies a task step into a behaviour and then determines whether any of the associated error modes are credible. For each credible error the analyst describes the error, determines the consequences, error recovery, probability and criticality. Finally, design remedies are proposed for each error identified.

The HET technique is a checklist approach and comes in the form of an error template. HET works as a simple checklist and is applied to each bottom level task step in a hierarchical task analysis (HTA) (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) of the task under analysis. The HET technique works by indicating which of the HET error modes are credible for each task step, based upon analyst subjective judgement. The analyst simply applies each of the HET error modes to the task step in question and determines whether any of the modes produce any credible errors or not. The HET error taxonomy consists of twelve error modes that were selected based upon a study of actual pilot error incidence and existing error

modes used in contemporary HEI methods. For each credible error (i.e. those judged by the analyst to be possible) the analyst should give a description of the form that the error would take, such as, 'pilot dials in the airspeed value using the wrong knob'. Next, the analyst has to determine the outcome or consequence associated with the error and then determine the likelihood of the error (Low, medium or high) and the criticality of the error (Low, medium or high). If the error is given a high rating for both likelihood and criticality, the aspect of the interface involved in the task step is then rated as a 'fail', meaning that it is not suitable for certification.

HAZOP (Kletz 1974) is a well-established engineering approach that was developed in the late 1960s by ICI (Swann and Preston 1995) for use in process design audit and engineering risk assessment (Kirwan 1992a). HAZOP involves a team of analysts applying guidewords, such as 'Not Done', 'More than' or 'Later than' to each step in a process in order to identify potential problems that may occur. Human Error HAZOP uses a set of human error guidewords (Whalley 1988). These guidewords are applied to each step in a HTA to determine any credible errors. For each credible error, the team should describe the error, determine the associated causes, consequences and recovery steps. Finally, design remedies for each identified error are offered by the HAZOP team.

TRACEr is a human error identification (HEI) technique developed specifically for use in air traffic control (ATC). TRACEr is represented in a series of decision flow diagrams and comprises eight taxonomies or error classification schemes: Task Error, Information, Performance Shaping Factors (PSF's), External Error Modes (EEM's), Internal Error Modes (IEM's), Psychological Error Mechanisms (PEM's), Error detection and error correction.

SPEAR (CCPS 1993) is another taxonomic approach to HEI that is extremely similar to the SHERPA approach. SPEAR uses an error taxonomy consisting of action, checking, retrieval, transmission, selection and planning errors. SPEAR operates on a HTA of the task under analysis. The analyst considers performance-shaping factors for each bottom level task step and determines whether or not any credible errors can occur. For each credible error, the analyst records an error description, its consequences and determines any error reduction measures.

The Cognitive Reliability and Error Analysis Method (CREAM) (Hollnagel 1998) is a recently developed HRA technique that can be used either predictively or retrospectively. CREAM uses an error taxonomy containing phenotypes (error modes) and genotypes (error causes). CREAM also uses common performance conditions (CPC'c) to account for context.

Error identifier based HEI techniques, such as HEIST and THEA provide error identifier prompts to aid the analyst in identifying potential human error. Typical error identifier prompts are, 'could the operator fail to carry out the act in time?' 'Could the operator carry out the task too early?' and 'could the operator carry out the task inadequately?' (Kirwan 1994). The error identifier prompts are linked to a set of error modes and reduction strategies. Whilst these techniques attempt to remove the reliability problems associated with taxonomic based approaches, they add considerable time to the analysis, as each error identifier prompt must be considered.

The Human Error Identification in Systems Tool (HEIST) (Kirwan 1994) uses a set of error identifier questions or prompts designed to aid the analyst in the identification of potential errors. There are eight sets of error identifier prompts including Activation/Detection, Observation/Data collection, Identification of system state, Interpretation, Evaluation, Goal Selection/Task Definition, Procedure selection and Procedure execution. The analyst applies each error identifier prompt to each task step in a HTA and determines whether any of the errors are credible or not. Each error identifier prompt has a set of linked error modes. For each credible error, the analyst records the system causes, the psychological error mechanism and any error reduction guidelines.

The Technique for Human Error Assessment (THEA) is a highly structured one that employs cognitive error analysis based upon Norman's (1988) model of action execution. THEA uses a series of questions in a checklist style approach based upon goals, plans, performing actions and perception/evaluation/interpretation. THEA also utilises a scenario-based analysis, whereby the analyst exhaustively describes the scenario under analysis before any analysis is carried out.

Error quantification techniques are used to offer a numerical probability of an error occurring. Identified errors are assigned a numerical probability value that represents the probability of occurrence. Performance Shaping factors (PSF's) are typically used to aid the analyst in the identification of potential errors. Error quantification techniques, such as JHEDI and HEART are typically used in the probabilistic safety assessment (PSA) of Nuclear processing plants. For example, Kirwan (1999) reports the use of JHEDI in a HRA risk assessment for the BNFL Thermal Oxide Reprocessing Plant at Sellafield, and also the use of HEART in a HRA risk assessment of the Sizewell B pressurised water reactor.

HEART (Williams 1986) is a HEI technique that attempts to predict and quantify the likelihood of human error or failure. The analyst classifies the task under analysis into one of the HEART generic categories (such as a) Totally familiar, performed at speed with no real idea of the likely consequences). Each HEART generic category has an associated human error probability associated with it. The analyst then identifies any error producing conditions (EPCs) associated with the task. Each EPC has an associated HEART effect. Examples of HEART EPCs include 'Shortage of time available for error detection and correction', and 'No obvious means of reversing an unintended action'. Once Any EPCs have been assigned, the analyst has to determine the assessed proportion of effect of each EPC (between 0 and 1). Finally all values are put into a formula and a final human error probability is produced.

A more recent development within HEI is to use a toolkit of different HEI techniques in order to maximise the comprehensiveness of the error analysis. The HERA framework is a prototype multiple method or 'toolkit' approach to human error identification that was developed by the Kirwan (1998a, 1998b). In response to a review of HEI methods, Kirwan (1998b) suggested that the best approach would be for practitioners to utilise a framework type approach to HEI, whereby a mixture of independent HRA/HEI tools would be used under one framework. In response to this conclusion, Kirwan (1998b) proposed the Human Error and Recovery Assessment (HERA) system, which was developed for the UK nuclear power and reprocessing

industry. Whilst the technique has yet to be applied to a concrete system, it is offered in this review as a representation of the form that a HEI 'toolkit' approach may take.

Task Analysis for Error Identification (TAFEI) (Baber & Stanton 1996) combines HTA with state space diagrams (SSDs) in order to predict illegal actions with a device. In conducting a TAFEI analysis, the analyst requires a description of the co-operative endeavour between the user and the product under analysis. The plans from the HTA are mapped onto an SSD for the device and a TAFEI diagram is produced. The TAFEI diagram is then used to highlight any illegal transitions. Once all illegal transitions have been identified, solutions or remedies are proposed.

In terms of performance, the literature consistently suggests that SHERPA is the most promising of the HEI techniques available to the HF practitioner. Kirwan (1992b) conducted a comparative study of six HEI techniques and reported that SHERPA achieved the highest overall rankings in terms of performance and ranking. In conclusion, Kirwan (1992b) recommended that a combination of expert judgement together with SHERPA would be the best approach to HEI. Other studies also show encouraging reliability and validity data for SHERPA (Baber & Stanton 1996, 2001; Stanton & Stevenage 2000). In a more recent comparative study of HEI techniques, Kirwan (1998b) used fourteen criteria to evaluate 38 HEI techniques. In conclusion it was reported that, of the 38 techniques, only nine are available in the public domain and are of practical use (Kirwan 1998b). These nine techniques are THERP, Human Error HAZOP, SHERPA, CMA/FSMA, PRMA, EOCA, SRS-HRA, SRK and HRMS.

HEI techniques suffer from a number of problems. The main problem associated with HEI techniques is the issue of validation. Few studies have been conducted in order to evaluate the reliability and validity of HEI techniques. A number of validation/comparison studies are reported in the literature (Williams 1985; Whalley & Kirwan 1989; Kirwan 1992a, 1992b, 1998a, 1998b, Kennedy 1995; Baber & Stanton 1996, 2002; Stanton & Stevenage 2000). However, considering the number of HEI techniques available and the importance of their use, this represents a very limited set of validation studies. Problems such as cost, time spent and access to systems under analysis often affect attempts to validate HEI techniques.

Stanton (2002) suggests that HEI techniques suffer from two key problems. The first of these problems relates to the lack of representation of the external environment or objects. Typically, human error analysis techniques do not represent the activity of the device and material that the human interacts with, in more than a passing manner. Hollnagel (1993) emphasises that Human Reliability Analysis (HRA) often fails to take adequate account of the context in which performance occurs. Second, HEI techniques place a great amount of dependence upon the judgement of the analyst. Different analysts, with different experience may make different predictions regarding the same problem (called intra-analyst reliability). Similarly, the same analyst may make different judgements on different occasions (inter-analyst reliability). This subjectivity of analysis may weaken the confidence that can be placed in any predictions made. The analyst is required to be an expert in the technique as well as the operation of the device being analysed if the analysis has a hope of being realistic.

The HEI techniques reviewed in this document are shown below:

1. SHERPA
2. HET – Human Error Template
3. TRACEr
4. TAFEI – Task Analysis For Error Identification
5. Human Error HAZOP
6. THEA – Technique for Human Error Assessment
7. HEIST – Human Error Identification in Systems Tool
8. The HERA framework
9. SPEAR - System for Predictive Error Analysis and Reduction
10. HEART - Human Error Assessment and Reduction Technique
11. CREAM – Cognitive Reliability Analysis Method

It is hypothesised that HEI techniques will be used during the design process of C4i systems by the DTC. The most suitable HEI techniques will be used throughout the C4i design process in order to evaluate system design concepts, highlight potential design induced human error and to offer error remedy design strategies.

SHERPA - Systematic Human Error Reduction and Prediction Approach

Neville A. Stanton, Department of Design, Brunel University, Runnymede Campus, Egham, Surrey, TW20 OJZ, UK

Background and Applications

SHERPA was developed by Embrey (1986) as a human error prediction technique that also enabled tasks to be analysed and potential solutions to errors to be presented in a structured manner. The technique is based upon a taxonomy of human error, and in its original form specified the psychological mechanism implicated in the error. The method is subject to ongoing development, which includes the removal of this reference to the underlying psychological mechanism.

SHERPA was originally designed to assist people in the Process Industries (e.g. conventional and nuclear power generation, petrochemical processing, oil and gas extraction, and power distribution, Embrey, 1986). An example of the application of SHERPA applied to the procedure for filling a chorine onto a road tanker may be found in Kirwan (1994). A recent example of SHERPA applied to oil and gas exploration may be found by consulting Stanton & Wilson (2000). The domain of application has broadened in recent years, to include ticket machines (Baber & Stanton, 1996), vending machines (Stanton and Stevenage, 1998), and in-car radio-cassette machines (Stanton & Young, 1999).

Domain of application

Process industries e.g. nuclear power generation, petro-chemical industry, oil and gas extraction and power distribution.

Procedure and advice

There are 8 steps in the SHERPA analysis, as follows:

Step 1: Hierarchical Task Analysis (HTA)

The process begins with the analysis of work activities, using Hierarchical Task Analysis. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Step 2: Task classification

Each operation from the bottom level of the analysis is taken in turn and is classified from the error taxonomy, into one of the following types:

- Action (e.g., pressing a button, pulling a switch, opening a door)
- Retrieval (e.g., getting information from a screen or manual)

- Checking (e.g., conducting a procedural check)
- Selection (e.g., choosing one alternative over another)
- Information communication (e.g., talking to another party)

Step 3: Human Error Identification (HEI)
This classification of the task step then leads the analyst to consider credible error modes associated with that activity, using the error taxonomy below. For each credible error (i.e. those judged by a subject matter expert to be possible) a description of the form that the error would take is given.

Action Errors

A1 - Operation too long/short

A2 – Operation mistimed

A3 – Operation in wrong direction

A4 – Operation too little/much

A5 – Misalign

A6 – Right operation on wrong object

A7 – Wrong operation on right object

A8 – Operation omitted

A9 – Operation incomplete

A10 – Wrong operation on wrong object

Checking Errors

C1 – Check omitted

C2 – Check incomplete

C3 – Right check on wrong object

C4 – Wrong check on right object

C5 – Check mistimed

C6 – Wrong check on wrong object

Retrieval Errors

R1 – Information not obtained

R2 – Wrong information obtained

R3 – Information retrieval incomplete

Communication Errors

I1 – Information not communicated

I2 – Wrong information communicated

I3 – Information communication

Selection Errors

S1 – Selection omitted

S2 – Wrong selection made

Step 4: Consequence Analysis
Considering the consequence of each error on a system is an essential next step as the consequence has implications for the criticality of the error. The analyst should describe fully the consequences associated with the identified error.

Step 5: Recovery Analysis
Next, the analyst should determine the recovery potential of the identified error. If there is a later task step at which the error could be recovered, it is entered next. If there is no recovery step then "None" is entered.

Step 6: Ordinal probability Analysis
Once the consequence and recovery potential have been identified, the analyst is required to rate the probability of the error occurring. An ordinal probability value is entered as low, medium or high. If the error has never been known to occur then a low (L) probability is assigned. If the error has occurred on previous occasions the medium (M) probability is assigned. Finally, if the error occurs frequently, a high (H) probability is assigned. This relies upon historical data and/or a subject matter expert.

Step 7: Criticality Analysis
If the consequence is deemed to be critical (e.g. is causes unacceptable losses) then a note of this is made. Criticality is of assigned in a binary manner. If the error would lead to a serious incident (this would have to be defined clearly before the analysis) then it is labelled as critical. Typically a critical consequence would be one that would lead to substantial damage to plant or product and/or injury to personnel.

Step 8: Remedy Analysis
The final stage in the process is to propose error reduction strategies. These are presented in the form of suggested changes to the work system that could have prevented the error from occurring, or at the very least reduced the consequences. This is done in the form of a structured brainstorming exercise to propose ways of circumventing the error or to reduce the effects of the error. Typically, these strategies can be categorised under four headings:

- Equipment (e.g. redesign or modification of existing equipment)
- Training (e.g. changes in training provided)
- Procedures (e.g. provision of new, or redesign of old, procedures)
- Organisational (e.g. changes in organisational policy or culture)

Some of these remedies may be very costly to implement. Therefore they needed to be judged with regard to the consequences, criticality and probability of the error. Each recommendation is analysed with respect to four criteria: Incident prevention efficacy, cost effectiveness, user acceptance and practicability.

Advantages
- Structured and comprehensive procedure, yet maintains usability
- The SHERPA taxonomy prompts analyst for potential errors
- Encouraging validity and reliability data
- Substantial time economy compared to observation
- Error reduction strategies offered as part of the analysis, in addition to predicted errors
- SHERPA is an easy technique to train and apply.
- The SHERPA error taxonomy is generic, allowing the technique to be used in a number of different domains.
- According to the HF literature, SHERPA is the most promising HEI technique available.

Disadvantages
- Can be tedious and time consuming for complex tasks
- Extra work involved if HTA not already available
- Does not model cognitive components of error mechanisms
- Some predicted errors and remedies are unlikely or lack credibility, thus posing a false economy
- Current taxonomy lacks generalisability

Example
The following example is a SHERPA analysis of programming a VCR (Baber & Stanton 1996). The process begins with the analysis of work activities, using

Hierarchical Task Analysis. HTA (see Annett, this volume) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). An example of HTA for the programming of a videocassette recorder is shown in figure one.

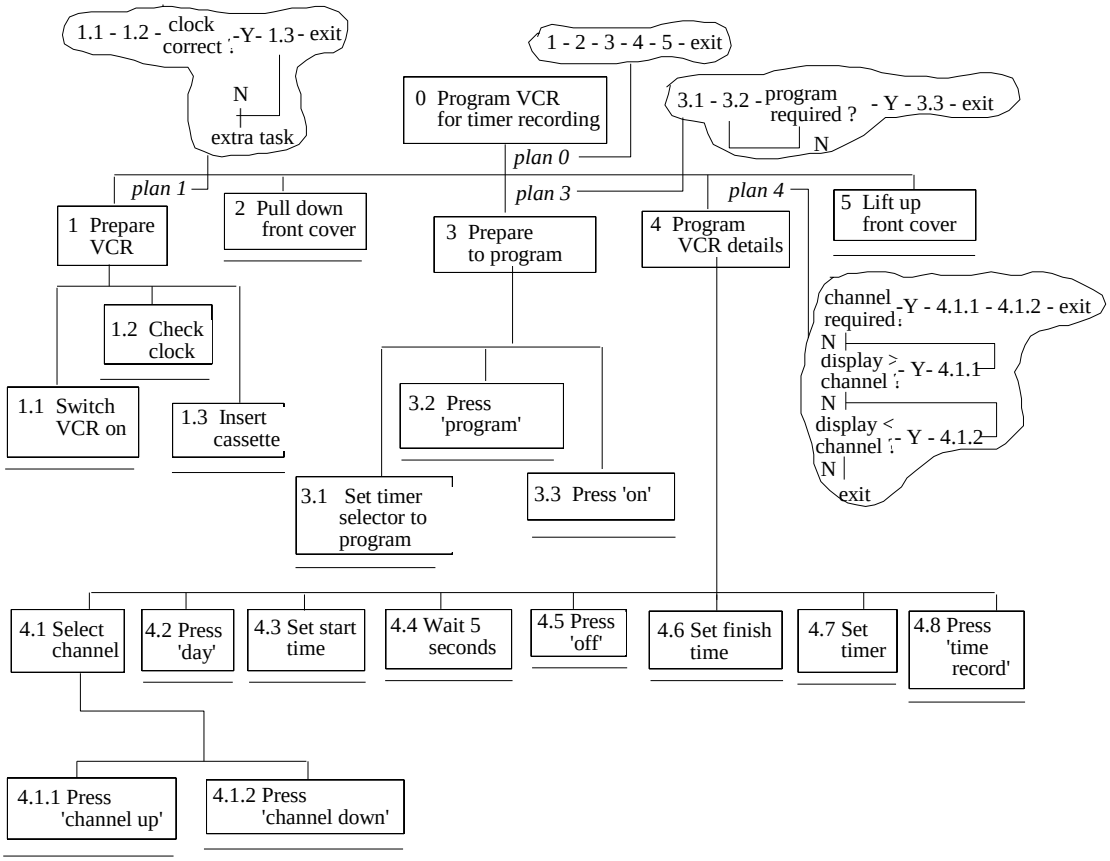


Figure 1. HTA for programming a VCR

For the application of SHERPA, each task step from the bottom level of the analysis is taken in turn. First each task step is classified into a type from the taxonomy, into one of the following types:

- Action (e.g. pressing a button, pulling a switch, opening a door)
- Retrieval (e.g. getting information from a screen or manual)
- Checking (e.g. conducting a procedural check)
- Information communication (e.g. talking to another party)
- Selection (e.g. choosing one alternative over another)

This classification of the task step then leads the analyst to consider credible error modes associated with that activity, as shown in step three of the procedure. For each credible error (i.e. those judged by a subject matter expert to be possible) a description of the form that the error would take is given as illustrated in table 1. The consequence of the error on system needs to be determined next, as this has implications for the criticality of the error. The last four steps consider the possibility for error recovery, the ordinal probability of the error (high, medium of low), its criticality (either critical or not critical) and potential remedies. Again these are shown in table 1.

Table 4. The SHERPA description

Task Step	Error Mode	Error Description	Consequence	Recovery	P	C	Remedial Strategy
1.1	A8	Fail to switch VCR on	Cannot proceed	Immediate	L		Press of any button to switch VCR on
1.2	C1 C2	Omit to check clock Incomplete check	VCR Clock time may be incorrect	None	L	!	Automatic clock setting and adjust via radio transmitter
1.3	A3 A8	Insert cassette wrong way around Fail to insert cassette	Damage to VCR Cannot record	Immediate Task 3	L L	!	Strengthen mechanism On-screen prompt
2	A8	Fail to pull down front cover	Cannot proceed	Immediate	L		Remove cover to programming
3.1	S1	Fail move timer selector	Cannot proceed	Immediate	L		Separate timer selector from programming function
3.2	A8	Fail to press PROGRAM	Cannot proceed	Immediate	L		Remove this task step from sequence
3.3	A8	Fail to press ON button	Cannot proceed	Immediate	L		Label button START TIME
4.1.1	A8	Fail to press UP button	Wrong channel selected	None	M	!	Enter channel number directly from keypad
4.1.2	A8	Fail to press DOWN button	Wrong channel selected	None	M	!	Enter channel number directly from keypad
4.2	A8	Fail to press DAY button	Wrong day selected	None	M	!	Present day via a calendar
4.3	I1 I2	No time entered Wrong time entered	No programme recorded Wrong programme recorded	None None	L L	! !	Dial time in via analogue clock Dial time in via analogue clock
4.4	A1	Fail to wait	Start time not set	Task 4.5	L		Remove need to wait
4.5	A8	Fail to press OFF button	Cannot set finish time				Label button FINISH TIME
4.6	I1 I2	No time entered Wrong time entered	No programme recorded Wrong programme recorded	None None	L L	! !	Dial time in via analogue clock Dial time in via analogue clock
4.7	A8	Fail to set timer	No programme recorded	None	L	!	Separate timer selector from programming function
4.8	A8	Fail to press TIME RECORD button	No programme recorded	None	L	!	Remove this task step from sequence
5	A8	Fail to lift up front cover	Cover left down	Immediate	L		Remove cover to programming

As table four shows there are six basic error types associated with the activities of programming a VCR. These are:

A. Failing to check that the VCR clock is correct.
B. Failing to insert a cassette.
C. Failing to select the programme number.
D. Failing to wait.
E. Failing to enter programming information correctly.
F. Failing to press the confirmatory buttons.

The purpose of SHERPA is not only to identify potential errors with the current design, but also to guide future design considerations. The structured nature of the

analysis can help to focus the design remedies on solving problems, as shown in the remedial strategies column. As this analysis shows, quite a lot of improvements could be made. It is important to note however, that the improvements are constrained by the analysis. This does not address radically different design solutions, i.e., those that may remove the need to programme at all.

Related methods

SHERPA relies heavily upon Hierarchical Task Analysis, which must be conducted before SHERPA can be carried out. The taxonomic approach is rather like a human version of a Hazard and Operability study. Kirwan (1994) has argued that more accurate predictions of human error are produced by using multiple methods, so SHERPA could be used in conjunction with TAFEI (see Task Analysis for Error Identification in this volume). Our research suggests that more accurate predictions are also found by pooling the data from multiple analysts using the same method.

Approximate training and application times

Based on the example of the application to the radio-cassette machine, Stanton & Young (1998) report training times of around 3 hours (this is doubled if training in Hierarchical Task Analysis is included). It took an average of 2 hours and 40 minutes for people to evaluate the radio-cassette machine using SHERPA.

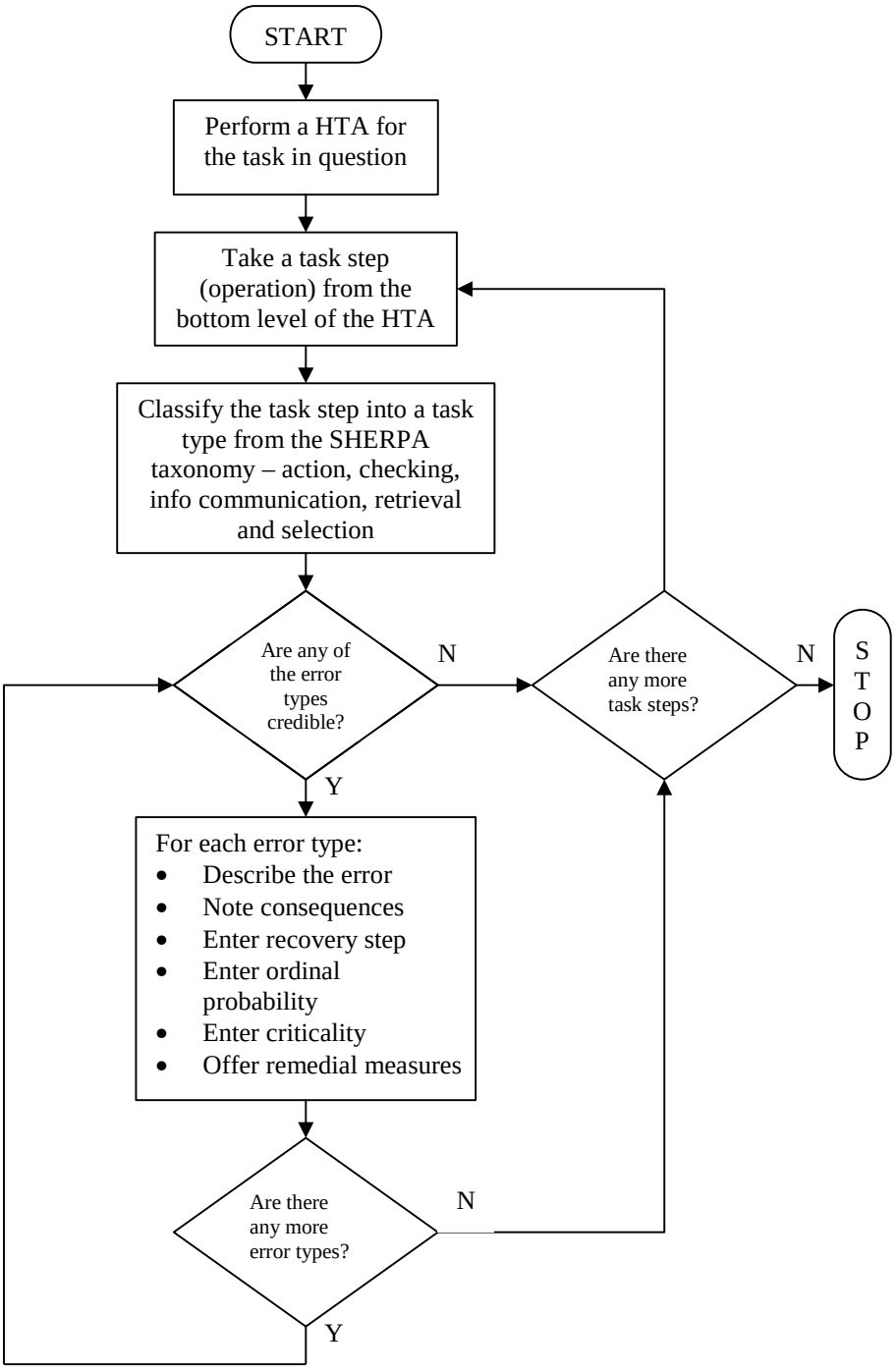
Reliability and Validity

Kirwan (1992) reports that SHERPA was the most highly rated of 5 human error prediction techniques by expert users. Baber & Stanton (1996) report a concurrent validity statistic of 0.8 and a reliability statistic of 0.9 in the application of SHERPA by two expert users to prediction of errors on a ticket vending machine. Stanton & Stevenage (1998) report a concurrent validity statistic of 0.74 and a reliability statistic of 0.65 in the application of SHERPA by 25 novice users to prediction of errors on a confectionery vending machine. Stanton & Young (1999) report a concurrent validity statistic of 0.2 and a reliability statistic of 0.4 in the application of SHERPA by 8 novice users to prediction of errors on a radio-cassette machine. It is suggested that reliability and validity is highly dependent upon expertise of the analyst and the complexity of the device being analysed.

Tools needed

At its simplest, SHERPA can be conducted with just a pen and paper. This can become slightly more sophisticated with the use of a computerised spreadsheet or table on a computer. The latter has the advantage of making the process less tedious when reorganising the material. Finally, some companies offer software specialist software for conducting the analysis. These systems are more labour saving and offer prompts to aid novice users (e.g. Bass et al, 1995).

Flowchart


START
Perform a HTA for the task in question
Take a task step (operation) from the bottom level of the HTA
Classify the task step into a task type from the SHERPA taxonomy – action, checking, info communication, retrieval and selection
Are any of the error types credible?
N
Y
Are there any more task steps?
N
STOP
For each error type:
Describe the error
Note consequences
Enter recovery step
Enter ordinal probability
Enter criticality
Offer remedial measures
Are there any more error types?
N
Y

Bibliography

Baber, C. & Stanton, N. A. (1996). Human error identification techniques applied to public technology: predictions compared with observed use. Applied Ergonomics, 27(2), 119-131.

Bass, A. Aspinal, J., Walter, G. & Stanton, N. A. (1995) A software toolkit for hierarchical task analysis. Applied Ergonomics. 26 (2) pp. 147-151.

Embrey, D. E. (1986). SHERPA: A systematic human error reduction and prediction approach. Paper presented at the International Meeting on Advances in Nuclear Power Systems, Knoxville, Tennessee.

Embrey, D. E. (1993). Quantitative and qualitative prediction of human error in safety assessments. Institute of Chemical Engineers Symposium Series, 130, 329-350.

Hollnagel, E. (1993). Human Reliability Analysis: context and control. London: Academic Press.

Kirwan, B. (1990). Human reliability assessment. In J. R. Wilson & E. N. Corlett (eds.), Evaluation of human work: a practical ergonomics methodology (2nd ed., pp. 921-968). London: Taylor & Francis.

Kirwan, B. (1992). Human error identification in human reliability assessment. Part 2: detailed comparison of techniques. Applied Ergonomics, 23, 371-381.

Kirwan, B. (1994). A guide to practical human reliability assessment. London: Taylor & Francis.

Stanton, N. A. (1995). Analysing worker activity: a new approach to risk assessment? Health and Safety Bulletin, 240, (December), 9-11.

Stanton, N. A. (2002) Human error identification in human computer interaction. In: J. Jacko and A. Sears (eds) The Human-Computer Interaction Handbook. New Jersey: Lawrence Erlbaum Associates.

Stanton, N. A. & Baber, C. (2002) Error by design: methods to predict device usability. Design Studies, 23 (4), 363-384.

Stanton, N. A. & Stevenage, S. V. (1998). Learning to predict human error: issues of acceptability, reliability and validity. Ergonomics, 41(11), 1737-1756.

Stanton, N. A. & Wilson, J. (2000) Human Factors: step change improvements in effectiveness and safety. Drilling Contractor, Jan/Feb, 46-41.

Stanton, N. A. & Young, M. (1998) Is utility in the mind of the beholder? A review of ergonomics methods. Applied Ergonomics. 29 (1) 41-54

Stanton, N. A. & Young, M. (1999) What price ergonomics? Nature 399, 197-198

HET - Human Error Template

Neville Stanton, Paul Salmon and Mark Young, Department of Design, Brunel University, Englefield Green, Surrey, TW20 0JZ
Don Harris and Jason Demagalski, Human Factors Group, Cranfield University, Cranfield University, Bedfordshire, MK43 0AL
Andrew Marshall, Marshall Associates
Thomas Waldmann, University of Limerick, Department of Psychology
Sidney Dekker, Centre for Human Factors in Aviation, Linkoping University, Sweden

Background and Applications

HET is a human error identification (HEI) technique that was developed by the ErrorPred consortium specifically for use in the certification of civil flight deck technology. The impetus for such a methodology came from a US Federal Aviation Administration (FAA) report entitled 'The Interfaces between Flight crews and Modern Flight Deck Systems' (Federal Aviation Administration, 1996), which identified many major design deficiencies and shortcomings in the design process of modern commercial airliner flight decks. The report made criticisms of the flight deck interfaces, identifying problems in many systems including pilots' autoflight mode awareness/indication; energy awareness; position/terrain awareness; confusing and unclear display symbology and nomenclature; a lack of consistency in FMS interfaces and conventions, and poor compatibility between flight deck systems. The FAA Human Factors Team also made many criticisms of the flight deck design process. For example, the report identified a lack of human factors expertise on design teams, which also had a lack of authority over the design decisions made. There was too much emphasis on the physical ergonomics of the flight deck, and not enough on the cognitive ergonomics. Fifty-one specific recommendations came out of the report. The most important in terms of this study were the following:

- 'The FAA should require the evaluation of flight deck designs for susceptibility to design-induced flightcrew errors and the consequences of those errors as part of the type certification process', and

- 'The FAA should establish regulatory and associated material to require the use of a flight deck certification review process that addresses human performance considerations'

In response to these findings, the ErrorPred consortium was established and set the task of developing and testing a HEI technique that could be used in the certification process of civil flight decks. The finished methodology was to be used in the flight deck certification process to predict potential design induced pilot error on civil flight decks. Beyond this, it was stipulated that the methodology should be: easily used by non-human factors/ergonomics professionals, relatively easy to learn and use, easily auditable, reliable and valid. The final criterion was that the method would fit in with existing flight deck certification procedures.

The HET methodology was developed from a review of existing HEI method external error mode (EEM) taxonomies and an evaluation of pilot error incidence. An EEM classifies the external and observable manifestation of the error or behaviour exhibited by an operator i.e. the physical form an error takes. An EEM taxonomy was created based on an analysis of EEM's used in a selection of existing HEI methods.

The error modes were then compared to a number of case studies involving civil flight decks and pilot error. The key pilot error in each case study was converted into an EEM e.g. the error 'pilot fails to lower landing gear' was converted into the EEM 'Fail to execute', and the error 'pilot dials in airspeed value of 190Kn using the heading knob' was converted into the EEM 'Right action on wrong interface element'. Furthermore, the errors reported in a questionnaire surrounding the flight task, "Land A320 at New Orleans International Airport using the Autoland system" were compared to the external error mode list. This allowed the authors to classify the errors reported by pilots into the external error modes currently in use in HEI methods. As a result of this error mode classification, it was possible to determine which of the existing HEI error modes would be suitable for predicting the types of EEM's that pilots exhibit. As a result of this process, the HET error mode taxonomy was created.

The HET technique is a checklist approach and comes in the form of an error template. HET works as a simple checklist and is applied to each bottom level task step in a hierarchical task analysis (HTA) (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) of the task under analysis. The HET technique works by indicating which of the HET error modes are credible for each task step, based upon analyst subjective judgement. The analyst simply applies each of the HET error modes to the task step in question and determines whether any of the modes produce any credible errors or not. The HET error taxonomy consists of twelve error modes that were selected based upon a study of actual pilot error incidence and existing error modes used in contemporary HEI methods. The twelve HET error modes are shown below:

- Fail to execute
- Task execution incomplete
- Task executed in the wrong direction
- Wrong task executed
- Task repeated
- Task executed on the wrong interface element
- Task executed too early
- Task executed too late
- Task executed too much
- Task executed too little
- Misread Information
- Other

For each credible error (i.e. those judged by the analyst to be possible) the analyst should give a description of the form that the error would take, such as, 'pilot dials in the airspeed value using the wrong knob'. Next, the analyst has to determine the outcome or consequence associated with the error e.g. Aircraft stays at current speed and does not slow down for approach. Finally, the analyst then has to determine the likelihood of the error (Low, medium or high) and the criticality of the error (Low, medium or high). If the error is given a high rating for both likelihood and criticality, the aspect of the interface involved in the task step is then rated as a 'fail', meaning that it is not suitable for certification.

Domain of application
Aviation.

Procedure and advice
Step 1: Hierarchical Task Analysis (HTA)
The process begins with the analysis of work activities, using Hierarchical Task Analysis. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Step 2: Human Error Identification
The analyst takes each bottom level task step from the HTA and considers each HET error mode for the task step in question. Any error modes that are deemed credible by the analyst are analysed further. At this stage, the analyst ticks which error mode is deemed credible for the task step under analysis and provides a description of the error e.g. pilot dials in the airspeed using the heading/track selector knob instead of the speed/mach knob.

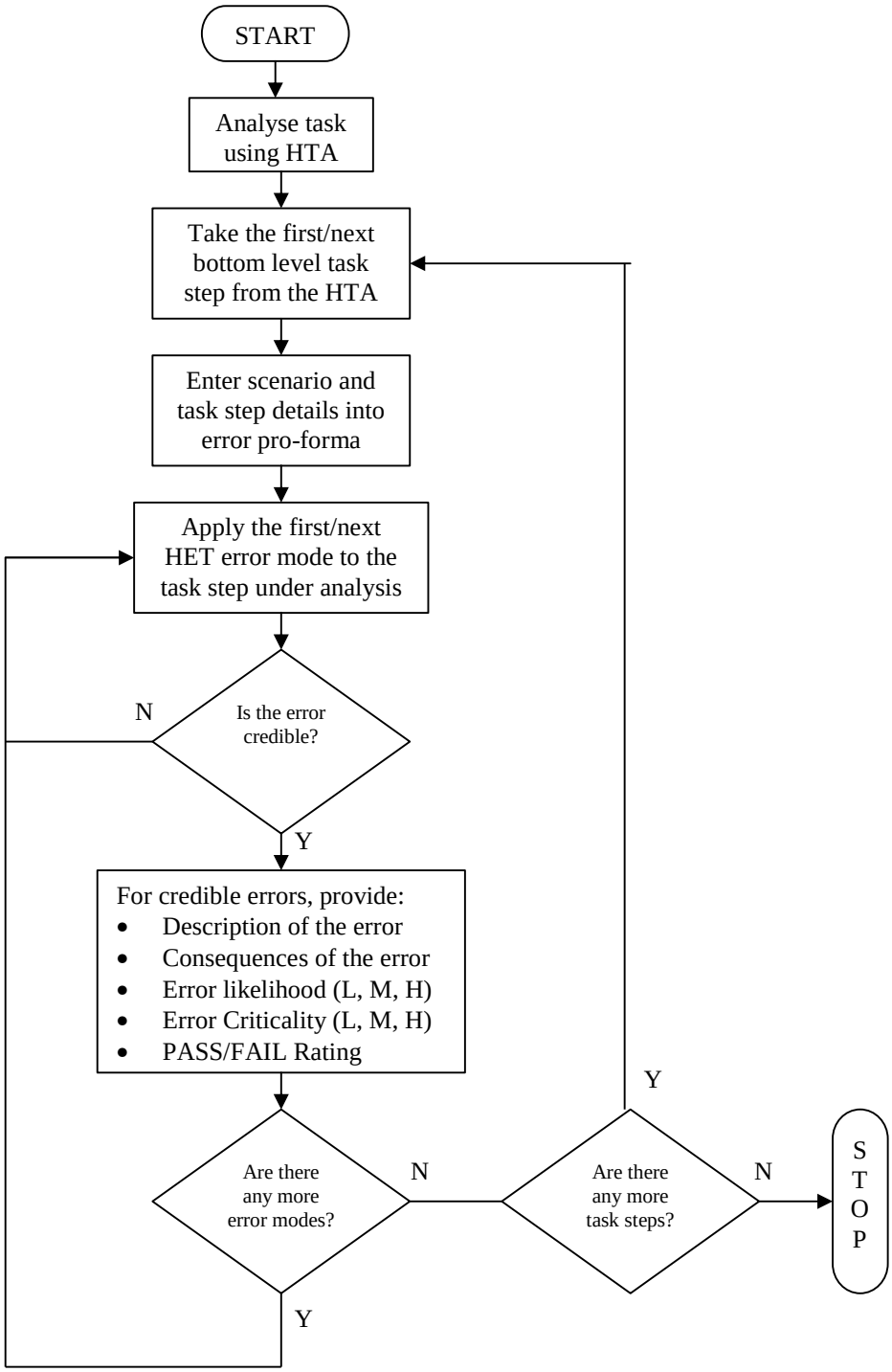
Step 3: Consequence Analysis
The analyst considers the consequence of the error and provides a description of the consequence. For example, the error, 'Pilot dials in airspeed of 190Kn using the heading knob' would have a consequence of 'aircraft does not slow down as required and instead changes heading to 190'.

Step 4: Ordinal Probability Analysis
An ordinal probability value is entered as low, medium or high. This based upon the analysts subjective judgement. If the analyst feels that chances of the error occurring are very small, then a low (L) probability is assigned. If the analyst thinks that the error may occur and has knowledge of the error occurring on previous occasions then a medium (M) probability is assigned. Finally, if the analyst thinks that the error would occur frequently, then a high (H) probability is assigned.

Step 5: Criticality Analysis
The criticality of the error is assigned next. Criticality is entered as low, medium or high. If the error would lead to a serious incident (this would have to be defined clearly before the analysis) then it is labelled as high. Typically a high critical consequence would be one that would lead to substantial damage to the aircraft or injury to crew and passengers. If the error has consequences that still have a distinct effect on the task, such heading the wrong way or losing a large amount of height or speed, then it is labelled medium. If the error would have minimal consequences such as a small loss of speed or height, then it is labelled as low.

Step 6: Interface Analysis
The analyst then has to determine whether or not the part of the interface under analysis (dependent upon the task step) passes or fails the certification procedure. If a high probability and a high criticality were assigned previously, then the interface in question is classed as a 'fail'. Any other combination of probability and criticality and the interface in question is classed as a 'Pass'.

Flowchart



Advantages

- The HET methodology is quick, simple to learn and use and requires very little training.
- HET utilises a comprehensive error mode taxonomy based upon existing HEI EEM taxonomies, actual pilot error incidence data and pilot error case studies.
- HET is easily auditable as it comes in the form of an error-proforma.
- Taxonomy prompts analyst for potential errors.
- Reliability and Validity data exists.
- Although the error modes in the HET EEM taxonomy were developed specifically for the aviation domain, they are generic, ensuring that the HET technique can potentially be used in a wide range of different domains, such as command and control, ATC, and nuclear reprocessing.

Disadvantages

- For large, complex tasks it may become tedious to perform
- Extra work involved if HTA not already available

HET Example – Land A320 at New Orleans using the Autoland system

A HET analysis was conducted on the flight task 'Land A320 at New Orleans using the Autoland system.

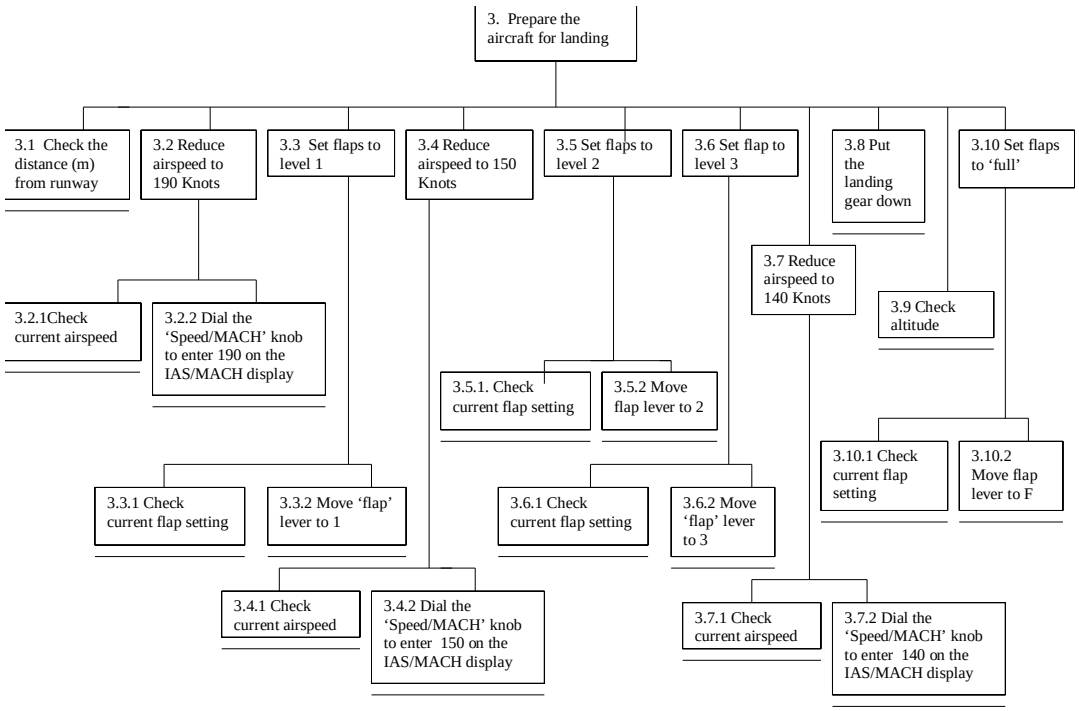


Figure 2. Extract of HTA 'Land at New Orleans using auto-land system'

Table 5. Example of HET output

Scenario: Land A320 at New Orleans using the Autoland system			**Task step:** 3.4.2 Dial the 'Speed/MACH; knob to enter 150 on IAS/MACH display								
Error Mode		**Description**	**Outcome**	**Likelihood**			**Criticality**			**PASS**	**FAIL**
				H	**M**	**L**	**H**	**M**	**L**		
Fail to execute											
Task execution incomplete											
Task executed in wrong direction	✓	Pilot turns the Speed/MACH knob the wrong way	Plane speeds up instead of slowing down		✓			✓		✓	
Wrong task executed											
Task repeated											
Task executed on wrong interface element	✓	Pilot dials using the HDG knob instead	Plane changes course and not speed	✓			✓				✓
Task executed too early											
Task executed too late											
Task executed too much	✓	Pilot turns the Speed/MACH knob too much	Plane slows down too much	✓				✓		✓	
Task executed too little	✓	Pilot turns the Speed/MACH knob too little	Plane does not slow down enough/Too fast for approach	✓				✓		✓	
Misread information											
Other											

Related Methods

HET is a taxonomic approach to HEI. A number of error taxonomy techniques exist, such as SHERPA, CREAM and TRACer. A HET analysis also requires an initial HTA (or some other specific task description) to be performed for the task in question.

Approximate Training and Application Times

In HET validation studies Marshall et al (2003) reported that with non-human factors professionals, the approximate training time for the HET methodology is around 90 minutes. Application time varies dependent upon the scenario under analysis. Marshall et al (2003) reported a mean application time of 62 minutes based upon an analysis involving a HTA with 32 bottom level task steps.

Reliability and Validity

Salmon et al (2003) reported SI ratings between 0.7 and 0.8 for subjects using the HET methodology to predict potential design induced pilot errors for the flight task 'Land A320 at New Orleans using the auto-land system'. Furthermore, it was reported that subjects using the HET method were more successful in their error predictions than subjects using SHERPA, Human Error HAZOP and HEIST.

Tools needed.

HET can be carried out using the HET error Proforma, a HTA of the task under analysis, functional diagrams of the interface under analysis, a pen and paper. In the example HET analysis given, subjects were provided with an error pro-forma, a HTA of the flight task, diagrams of the auto-pilot panel, the captain's primary flight

display, the flap lever, the landing gear lever, the speed brake, the attitude indicator and an overview of the A320 cockpit.

Bibliography

Marshall, A., Stanton, N., Young, M., Salmon, P., Harris, D., Demagalski, J., Waldmann, T., Dekker, S. (2003) Development of the Human Error Template – A new methodology for assessing design induced errors on aircraft flight decks.

Salmon, P. M, Stanton, N.A, Young, M.S, Harris, D, Demagalski, J, Marshall, A, Waldmann, T, Dekker, S (2003) Predicting Design Induced Pilot Error: A comparison of SHERPA, Human Error HAZOP, HEIST and HET, a newly developed aviation specific HEI method, In D. Harris, V. Duffy, M. Smith, C. Stephanidis (eds) Human-Centred Computing – Cognitive, Social and Ergonomic Aspects, Lawrence Erlbaum Associates, London

TRACEr – Technique for the Retrospective and Predictive Analysis of Cognitive Errors in Air Traffic Control (ATC)

Steven Shorrock, Det Norske Veritas (DNV), Highbank House, Exchange Street, Chesire, SK3 OET, UK
Barry Kirwan, EUROCONTROL, Experimental Centre, BP15, F91222, Bretigny Sur Orge, France

Background and Applications

TRACEr is a human error identification (HEI) technique developed specifically for use in air traffic control (ATC). TRACEr was developed as part of the human error in European air traffic management (HERA) project. Under the HERA project remit, the authors were required to develop a human error incidence analysis technique that conformed to the following criteria (Isaac, Shorrock & Kirwan, 2002).

- Flowchart based for ease of use.
- Should utilise a set of inter-related taxonomies (EEM's, IEM's, PEM's, PSF's, Tasks and Information and equipment).
- Technique must be able to deal with chains of events and errors.
- PSF taxonomy should be hierarchical and may need a deeper set of organisational causal factor descriptors.
- Must be comprehensive, accounting for situation awareness, signal detection theory and control theory.
- Technique must be able to account for maintenance errors, latent errors, violations and errors of commission.

TRACEr can be used both predictively and retrospectively and is based upon a literature review of a number of domains, including experimental and applied psychology, human factors literature and communication theory (Isaac, Shorrock & Kirwan, 2002). Existing HEI methods were reviewed and research within ATM was conducted in the development of the method. TRACEr is represented in a series of decision flow diagrams and comprises eight taxonomies or error classification schemes: Task Error, Information, Performance Shaping Factors (PSF's), External Error Modes (EEM's), Internal Error Modes (IEM's), Psychological Error Mechanisms (PEM's), Error detection and error correction.

Domain of application

Air Traffic Control.

Procedure and advice (Predictive analysis)

Step 1: Hierarchical Task Analysis (HTA)

The process begins with the analysis of work activities, using Hierarchical Task Analysis. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate

stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Step 2: PSF and EEM consideration
The analyst takes the first bottom level task step from the HTA (operation) and considers each of the PSF's for the task step in question. The purpose of this is to identify any environmental or situational factors that could influence the air traffic controller's performance. Once the analyst has considered all of the relevant PSF's, the EEM's are considered for the task step under analysis. Based upon subjective judgement, the analyst determines whether any of the TRACEr EEM's are credible for the task step in question. Figure 3 shows the TRACEr EEM taxonomy. If there are any credible errors, the analyst proceeds to step 3. If there are no errors deemed credible, then the analyst goes back to the HTA and takes the next task step.

Selection and Quality	**Timing and Sequence**	**Communication**
Omission	Action too long	Unclear Info transmitted
Action Too much	Action too short	Unclear info recorded
Action Too little	Action too early	Info not sought/obtained
Action in wrong direction	Action too late	Info not transmitted
Wrong action on right object	Action repeated	Info not recorded
Right action on wrong object	Mis-ordering	Incomplete info transmitted
Wrong action on wrong object		Incomplete info recorded
Extraneous act		Incorrect info transmitted
		Incorrect info recorded

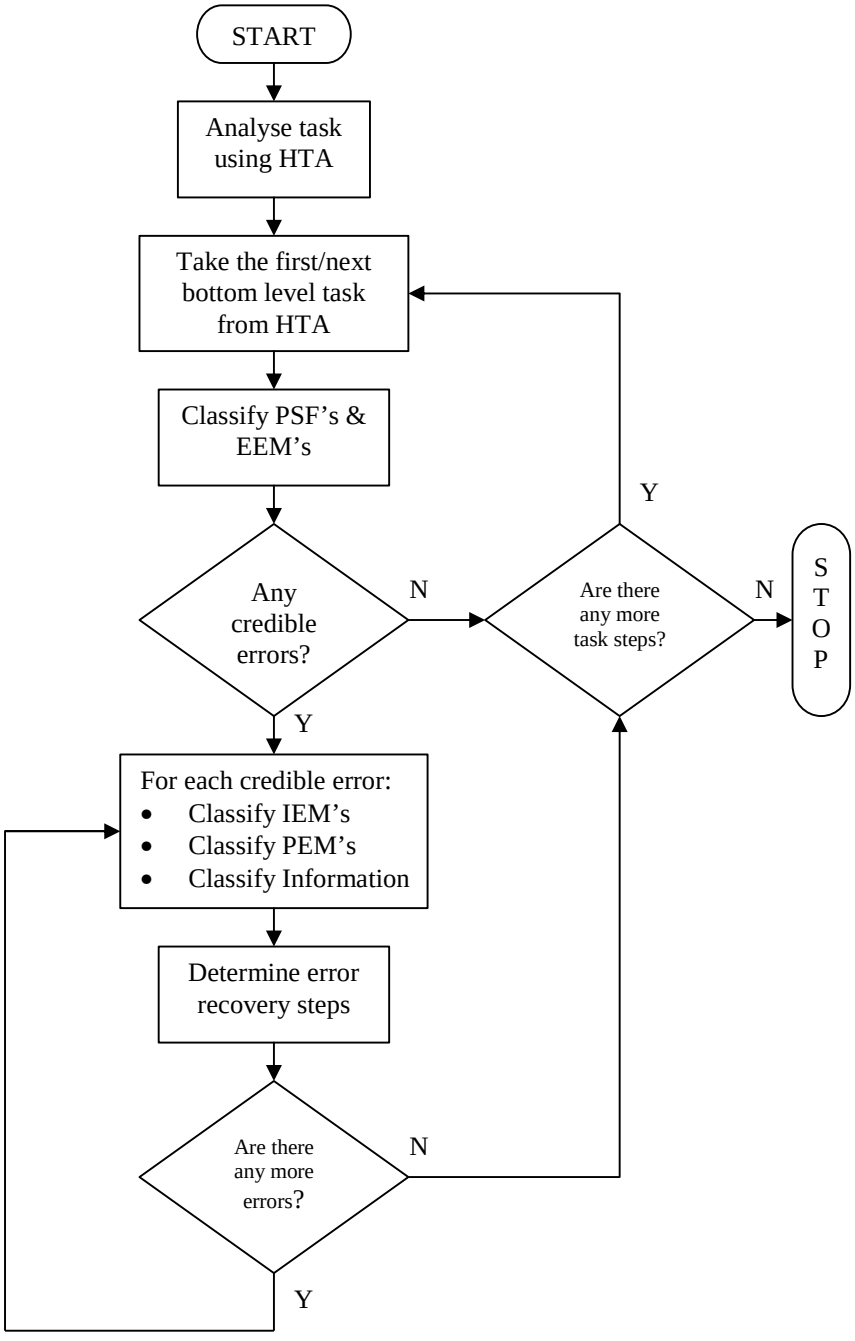
Figure 3. TRACEr's external error mode taxonomy

Step 3: IEM classification
For any credible errors, the analyst then determines which of the internal error modes (IEM's) are evident in the error. IEM's describe which cognitive function failed or could fail (Shorrock & Kirwan, 2002). Examples of TRACEr IEM's include *Late detection, misidentification, hearback error, forget previous actions, prospective memory failure, misrecall stored information and misprojection.*

Step 4: PEM classification
Next, the analyst has to determine the psychological cause or 'psychological error mechanism' (PEM) behind the error. Examples of TRACEr PEM's include *insufficient learning, expectation bias, false assumption, perceptual confusion, memory block, vigilance failure and distraction.*

Step 5: Error Recovery
Finally, once the error analyst has described the error and determined the EEM, IEM's and PEM's, error recovery steps for each error should be offered. This is based upon the analyst's subjective judgement.

Flowchart (Predictive TRACEr)



Procedure and advice (Retrospective Analysis)

Step 1: Analyse incident into 'error events'

Firstly, the analyst has to classify the task steps into error events i.e. the task steps in which an error was produced. This is based upon analyst judgement.

Step 2: Task Error Classification

The analyst then takes the first/next error from the error events list and classifies it into a task error from the task error taxonomy. The task error taxonomy contains thirteen categories describing controller errors. Task error categories include 'radar monitoring error', 'co-ordination error' and 'flight progress strip use error' (Shorrock and Kirwan, 2002).

Step 3: IEM Information Classification

Next the analyst has to determine the internal error mode (IEM) associated with the error. IEM's describe which cognitive function failed or could fail (Shorrock & Kirwan, 2002). Examples of TRACEr IEM's include *late detection, misidentification, hearback error, forget previous actions, prospective memory failure, misrecall stored information and misprojection*. When using TRACEr retrospectively, the analyst also has to use the *information* taxonomy to describe the 'subject matter' of the error i.e. what information did the controller misperceive? The information terms used are related directly to the IEM's in the IEM taxonomy. The information taxonomy is important as it forms the basis of error reduction within the TRACEr technique.

Step 4: PEM Classification

The analyst then has to determine the 'psychological cause' or psychological error mechanism (PEM) behind the error. Example PEM's used in the TRACEr technique include *Insufficient learning, expectation bias, false assumption, perceptual confusion, memory block, vigilance failure and distraction.*

Step 5: PSF Classification

Performance shaping factors are factors that influenced or have the potential to have influenced the operator's performance. The analyst has to use the PSF taxonomy to select any PSF's that were evident in the production of the error under analysis. TRACEr's PSF taxonomy contains both PSF categories and keywords. Examples of PSF's used in the TRACEr technique are shown below in figure 4.

PSF Category	**Example PSF keyword**
Traffic and Airspace	Traffic complexity
Pilot/controller communications	RT Workload
Procedures	Accuracy
Training and experience	Task familiarity
Workplace design, HMI and equipment factors	Radar display
Ambient environment	Noise
Personal factors	Alertness/fatigue
Social and team factors	Handover/takeover
Organisational factors	Conditions of work

Figure 4. Extract from TRACEr's PSF taxonomy

Step 6: Error detection and Error correction
Unique to the retrospective use of TRACEr, the error detection and correction stage provides the analyst with a set of error detection keywords. Four questions are used to prompt the analyst in the selection of error detection keywords (Source: Shorrock & Kirwan, 2002).

1. How did the controller become aware of the error? (e.g. action feedback, inner feedback, outcome feedback)
2. What was the feedback medium? (e.g. radio, radar display)
3. Did any factors, internal or external to the controller, improve or degrade the detection of the error?
4. What was the separation status at the time of error detection?

Once the analyst has classified the error detection, the error correction or reduction should also be classified. TRACEr uses the following questions to prompt the analyst in error correction/reduction classification (Source: Shorrock and Kirwan, 2002).

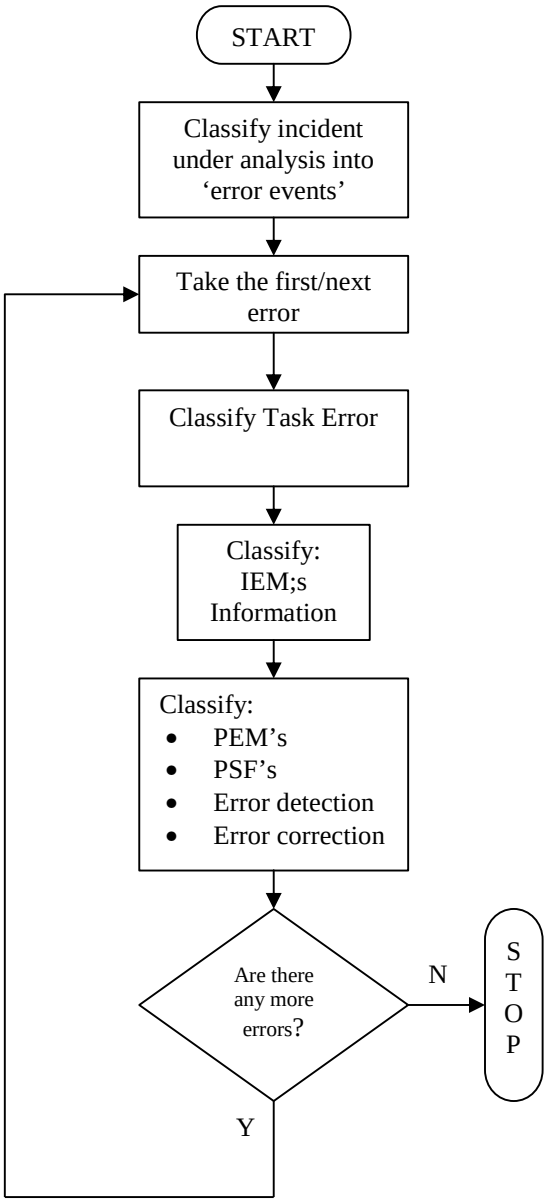
1. What did the controller do to correct the error? (e.g. reversal or direct correction, automated correction)
2. How did the controller correct the error? (e.g. turn or climb)
3. Did any factors, internal or external to the controller, improve or degrade the detection of the error?
4. What was the separation status at the time of the error correction?

Once the analyst has completes step 6, the next error should be analysed. Alternatively, if there are no more 'error events' then the analysis is finished.

Advantages

- TRACEr technique appears to be a very comprehensive approach to error prediction and error analysis, including IEM, PEM, EEM and PSF analysis
- TRACEr is based upon sound scientific theory, integrating Wickens (1992) model of information processing into its model of ATC.
- In a prototype study (Shorrock, 1997), a participant questionnaire highlighted comprehensiveness, structure, acceptability of results and usability as strong points of the technique (Shorrock and Kirwan, 2002).
- TRACEr has proved successful in analysing errors from AIRPROX reports and providing error reduction strategies.
- Used in the European human error in ATC (HERA) project.
- Developed specifically for ATC, based upon previous ATC incidents and interviews with ATC controllers.

Flowchart (Retrospective TRACEr)



Disadvantages

- The TRACEr technique appears unnecessarily over-complicated for what it actually is, a taxonomy based error analysis tool. A prototype study (Shorrock, 1997) highlighted a number of areas of confusion in participant use of the different categories (Shorrock and Kirwan, 2002).
- No validation evidence or studies using TRACEr.
- For complex tasks, analysis will become laborious and large
- Very high resource usage (time)**.** In a participant questionnaire used in the prototype study (Shorrock, 1997) resource usage (time and expertise) was the most commonly reported area of concern (Shorrock and Kirwan, 2002).

- Training time would be extremely high for such a technique.
- Extra work involved if HTA not already available
- Existing techniques using similar EEM taxonomies appear to be far simpler and much quicker (SHERPA, HET etc).

Example
For an example TRACEr analysis, the reader is referred to Shorrock & Kirwan (2000).

Related Methods
TRACEr is a taxonomic approach to HEI. A number of error taxonomy techniques exist, such as SHERPA, CREAM and HET. When applying TRACEr (both predictively and retrospectively) an initial HTA for the task/scenario under analysis is required.

Approximate training and application times
No data regarding training and application times for the TRACEr technique are presented in the literature. It is estimated that both the training and application times for TRACEr would be high.

Reliability and validity
There are no data available regarding the reliability and validity of the TRACEr technique. According to the authors (Shorrock and Kirwan, 2002) such a study is being planned. In a small study analysing error incidences from AIRPROX reports (Shorrock and Kirwan, 2002) it was reported, via participant questionnaire, that the TRACEr techniques strengths are its comprehensiveness, structure, acceptability of results and usability.

Tools needed
TRACEr analyses can be carried out using pen and paper. PEM, EEM, IEM, PSF taxonomy lists are also required. A HTA for the task under analysis is also required.

Bibliography
Isaac, A., Shorrick, S.T., Kirwan, B., (2002) Human Error in European air traffic management: The HERA project. *Reliability Engineering and System Safety*, Vol. 75 pp 257-272
Shorrock, S.T., Kirwan, B., (1999) The development of TRACEr: a technique for the retrospective analysis of cognitive errors in ATC. In: Harris, D. (Ed), *Engineering Psychology and Cognitive Ergonomics*, Vol. 3, Aldershot, UK, Ashgate
Shorrock, S.T., Kirwan, B., (2000) Development and application of a human error identification tool for air traffic control. *Applied Ergonomics*, Vol. 33 pp319-336

TAFEI - Task Analysis For Error Identification

Neville A. Stanton, Department of Design, Brunel University, Runnymede Campus, Egham, Surrey,TW20 0JZ, United Kingdom
Christopher Baber, School of Electronic, Electrical & Computing Engineering, The University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

Background and Applications

Task Analysis For Error Identification (TAFEI) is a method that enables people to predict errors with device use by modelling the interaction between the user and the device under analysis. It assumes that people use devices in a purposeful manner, such that the interaction may be described as a "cooperative endeavour", and it is by this process that problems arise. Furthermore, the technique makes the assumption that actions are constrained by the state of the product at any particular point in the interaction, and that the device offers information to the user about its functionality. Thus, the interaction between users and devices progresses through a sequence of states. At each state, the user selects the action most relevant to their goal, based on the System Image.

The foundation for the approach is based on general systems theory. This theory is potentially useful in addressing the interaction between sub-components in systems (i.e., the human and the device). It also assumes a hierarchical order of system components, i.e., all structures and functions are ordered by their relation to other structures and functions, and any particular object or event is comprised of lesser objects and events. Information regarding the status of the machine is received by the human part of the system through sensory and perceptual processes and converted to physical activity in the form of input to the machine. The input modifies the internal state of the machine and feedback is provided to the human in the form of output. Of particular interest here is the boundary between humans and machines, as this is where errors become apparent. We believe that it is essential for a method of error prediction to examine explicitly the nature of the interaction.

The theory draws upon the ideas of scripts and schema. We can imagine that a person approaching a ticket-vending machine might draw upon a 'vending machine' or a 'ticket kiosk' script when using a ticket machine. From one script, the user might expect the first action to be 'Insert Money', but from the other script, the user might expect the first action to be 'Select Item'. The success, or failure, of the interaction would depend on how closely they were able to determine a match between the script and the actual operation of the machine. The role of the comparator is vital in this interaction. If it detects differences from the expected states, then it is able to modify the routines. Failure to detect any differences is likely to result in errors. Following Bartlett's (1932) lead, the notion of schema is assumed to reflect a person's "…effort after meaning." (Bartlett, 1932), arising from the active processing (by the person) of a given stimulus. This active processing involves combining prior knowledge with information contained in the stimulus. While schema theory is not without its critics (see Brewer, 2000) for a review, the notion of an active processing of stimuli clearly has resonance with our proposal for rewritable routines. The reader might feel that there are similarities between the notion of rewritable routines and some of the research on mental models that was popular in the 1980s. Recent developments in the theory underpinning TAFEI by the authors have distinguished between global prototypical routines (i.e., a repertoire of stereotypical responses that allow people to

perform repetitive and mundane activities with little or no conscious effort) and local, state-specific, routines (i.e., responses that are developed only for a specific state of the system). The interesting part of the theory is the proposed relationship between global and local routines. It is our contention that these routines are analogous to global and local variables in computer programming code. In the same manner as a local variable in programming code, a local routine is overwritten (or rewritable in our terms) once the user have moved beyond the specific state for which it was developed. See Baber & Stanton (2002) for a more detailed discussion of the theory.

Examples of applications of TAFEI include prediction of errors in boiling kettles (Baber and Stanton, 1994; Stanton and Baber, 1998), comparison of word processing packages (Stanton and Baber, 1996; Baber and Stanton, 1999), withdrawing cash from automatic teller machines (Burford, 1993), medical applications (Baber and Stanton, 1999; Yamaoka and Baber, 2000), recording on tape-to-tape machines (Baber and Stanton, 1994), programming a menu on cookers (Crawford, Taylor and Po, 2000), programming video-cassette recorders (Baber and Stanton, 1994; Stanton and Baber, 1998), operating radio-cassette machines (Stanton and Young, 1999), recalling a phone number on mobile phones (Baber and Stanton, 2002), buying a rail ticket on the ticket machines on the London Underground (Baber and Stanton, 1996), and operating high-voltage switchgear in substations (Glendon and McKenna, 1995).

Domain of application
Public technology and product design.

Procedure and advice
Procedurally, TAFEI is comprised of three main stages. Firstly, Hierarchical Task Analysis (HTA – see Annett in this volume) is performed to model the human side of the interaction. Of course, one could employ any technique to describe human activity. However, HTA suits our purposes for the following reasons: i. it is related to Goals and Tasks; ii. it is directed at a specific goal; iii. it allows consideration of task sequences (through 'plans'). As will become apparent, TAFEI focuses on a sequence of tasks aimed at reaching a specific goal. Next, State-Space Diagrams (SSDs) are constructed to represent the behaviour of the artifact. Plans from the HTA are mapped onto the SSD to form the TAFEI diagram. Finally, a transition matrix is devised to display state transitions during device use. TAFEI aims to assist the design of artifacts by illustrating when a state transition is possible but undesirable (i.e., illegal). Making all illegal transitions impossible should facilitate the cooperative endeavour of device use.

For illustrative purposes of how to conduct the method, a simple, manually-operated, electric kettle is used in this example. The first step in a TAFEI analysis is to obtain an appropriate HTA for the device, as shown in figure 5. As TAFEI is best applied to scenario analyses, it is wise to consider just one specific goal, as described by the HTA (e.g., a specific, closed-loop task of interest) rather than the whole design. Once this goal has been selected, the analysis proceeds to constructing State-Space Diagrams (SSDs) for device operation.

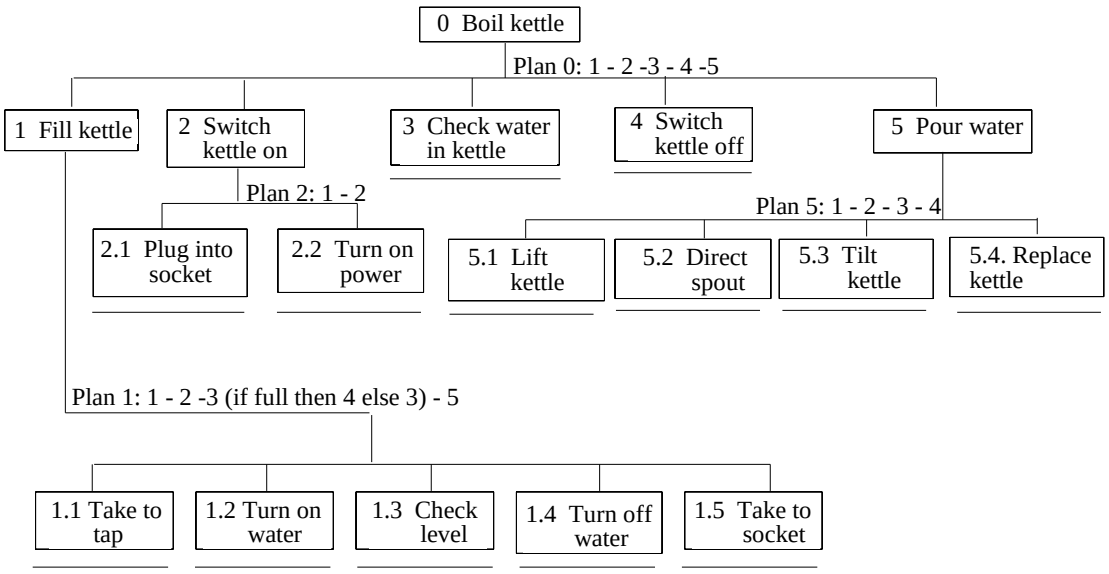


Figure 5. Hierarchical Task Analysis.

A SSD essentially consists of a series of states that the device passes from a starting state to the goal state. For each series of states, there will be a current state, and a set of possible exits to other states. At a basic level, the current state might be "off", with the exit condition "switch on" taking the device to the state "on". Thus, when the device is "off" it is 'waiting to…' an action (or set of actions) that will take it to the state "on". It is very important to have, on completing the SSD, an exhaustive set of states for the device under analysis. Numbered plans from the HTA are then mapped onto the SSD, indicating which human actions take the device from one state to another. Thus the plans are mapped onto the state transitions (if a transition is activated by the machine, this is also indicated on the SSD, using the letter 'M' on the TAFEI diagram). This results in a TAFEI diagram, as shown in figure three. Potential state-dependant hazards have also been identified.

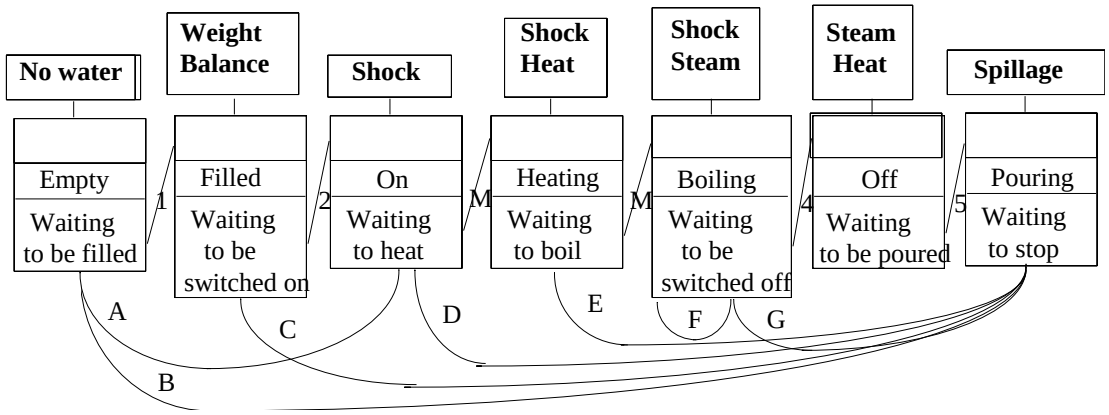


Figure 6. State-space TAFEI diagram

The most important part of the analysis from the point of view of improving usability is the transition matrix. All possible states are entered as headers on a matrix – see table 6. The cells represent state transitions (e.g., the cell at row 1, column 2 represents the transition between state 1 and state 2), and are then filled in one of

three ways. If a transition is deemed impossible (i.e., you simply cannot go from this state to that one), a "-" is entered into the cell. If a transition is deemed possible and desirable (i.e., it progresses the user towards the goal state - a correct action), this is a legal transition and "L" is entered into the cell. If, however, a transition is both possible but undesirable (a deviation from the intended path - an error), this is termed illegal and the cell is filled with an "I". The idea behind TAFEI is that usability may be improved by making all illegal transitions (errors) impossible, thereby limiting the user to only performing desirable actions. It is up to the analyst to conceive of design solutions to achieve this.

Table 6. Transition matrix

		TO STATE						
		Empty	Filled	On	Heating	Boiling	Off	Pouring
FROM STATE	Empty	---------	L (1)	I (A)	---------	---------	---------	I (B)
	Filled		---------	L (2)	---------	---------	---------	I (C)
	On			---------	L (M)	---------	---------	I (D)
	Heating					L (M)	---------	I (E)
	Boiling					I (F)	L (4)	I (G)
	Off							L (5)
	Pouring							

The states are normally numbered, but in this example the text description is used. The character "L" denotes all of the error-free transitions and the character "I" denotes all of the errors. Each error has an associated character (i.e., A to G), for the purposes of this example and so that it can be described in table four.

Table 7. Error descriptions and design solutions

Error	**Transition**	**Error description**	**Design solution**
A	1 to 3	Switch empty kettle on	Transparent kettle walls and/or link to water supply
B	1 to 7	Pour empty kettle	Transparent kettle walls and/or link to water supply
C	2 to 7	Pour cold water	Constant hot water or autoheat when kettle placed on base after filling
D	3 to 7	Pour kettle before boiled	Kettle status indicator showing water temperature
E	4 to 7	Pour kettle before boiled	Kettle status indicator showing water temperature
F	5 to 5	Fail to turn off boiling kettle	Auto cut-off switch when kettle boiling
G	5 to 7	Pour boiling water before turning kettle off	Auto cut-off switch when kettle boiling

Obviously the design solutions in table two are just illustrative and would need to be formally assessed for their feasibility and cost.

What TAFEI does best is enable the analysis to model the interaction between human action and system states. This can be used to identify potential errors and consider the task flow in a goal-oriented scenario. Potential conflicts and contradictions in task flow should come to light. For example, in a study of medical imaging equipment design, Baber & Stanton (1999) identified disruptions in task flow that made the device difficult to use. TAFEI enabled the design to be modified and led to the development of a better task flow. This process of analytical prototyping is key to the

use of TAFEI in designing new systems. Obviously, TAFEI can also be used to evaluate existing systems. There is a potential problem that the number of states that a device can be in could overwhelm the analyst. Our experience suggests that there are two possible approaches. First, only analyse goal-oriented task scenarios. The process is pointless without a goal and HTA can help focus the analysis. Second, the analysis can be nested at various levels in the task hierarchy, revealing more and more detail. This can make each level of analysis relatively self-contained and not overwhelming. The final piece of advice is to start with a small project and build up from that position.

Example

The following example of TAFEI was used to analyse the task of programming a video-cassette recorder. The task analysis, state-space diagrams and transition matrix are all presented. First of all the task analysis is performed to describe human activity, as shown in figure seven.

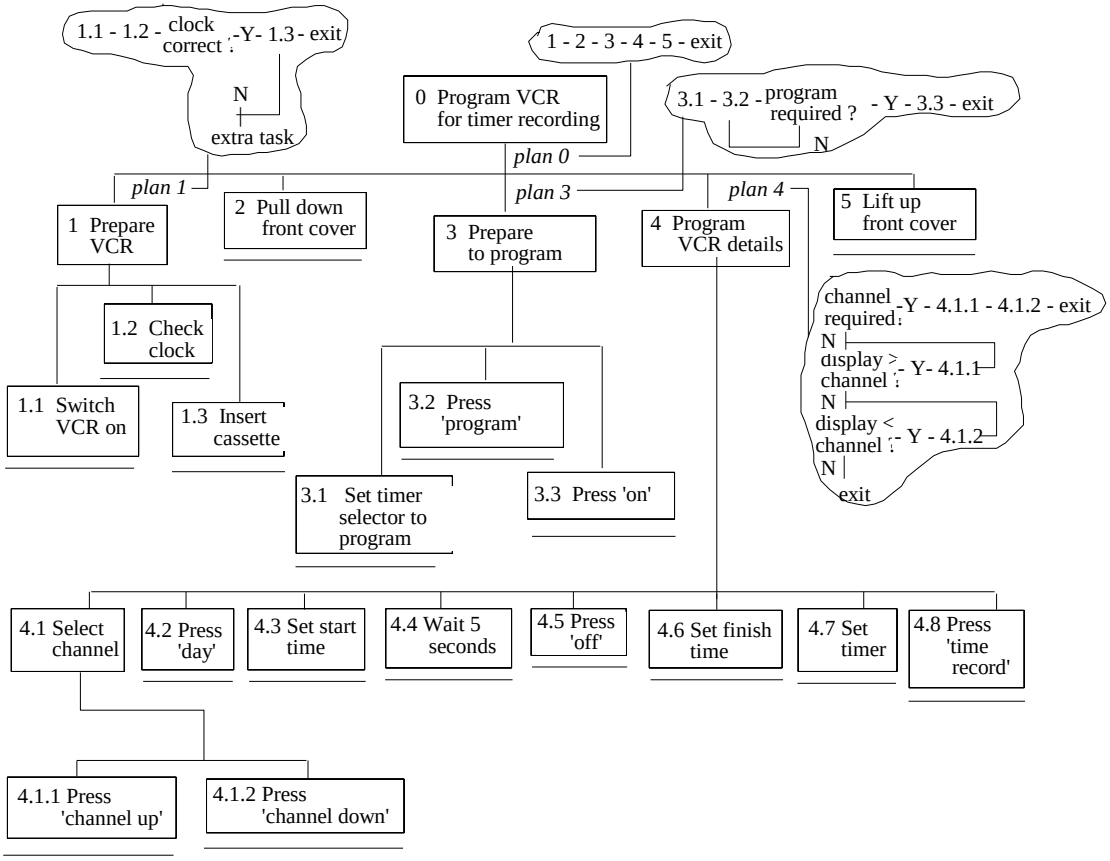


Figure 7. HTA of VCR programming task

Next, the state-space diagrams are drawn as shown in figure eight.

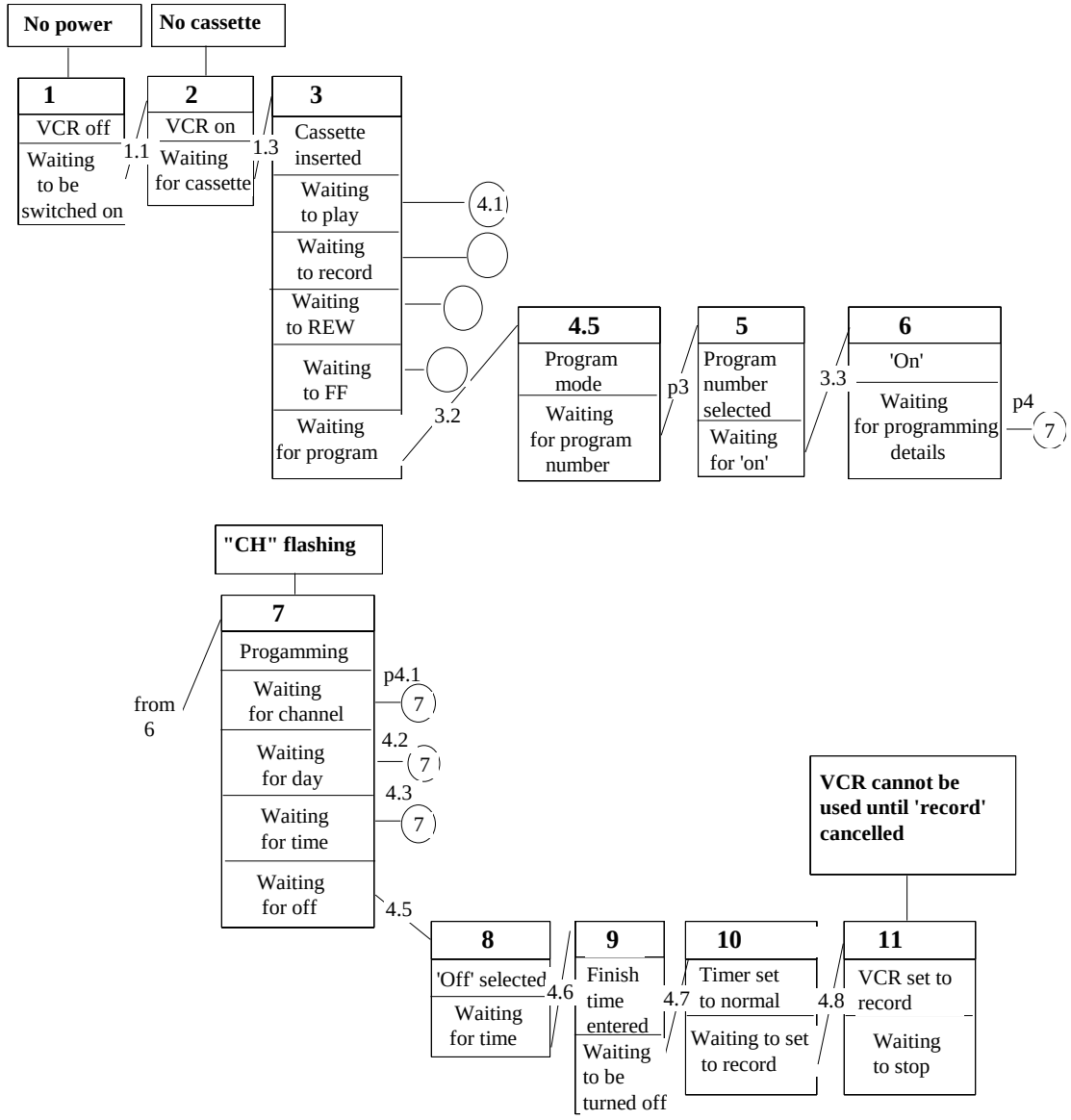


Figure 8. The TAFEI description.

From the TAFEI diagram, a transition matrix is compiled and each transition is scrutinised, as shown in table eight.

From state: \ To state:	1	2	3	4.5	5	6	7	8	9	10	11
1	-	L	I	-	-	-	-	-	-	-	-
2	L	-	L	I	-	-	-	-	-	-	-
3	L	-	-	L	-	-	-	-	-	-	-
4.5	I	-	-	-	L	I	-	-	-	-	-
5	I	-	-	-	-	L	I	I	-	-	-
6	I	-	-	-	-	-	L	I	-	-	-
7	I	-	-	-	-	-	[redacted]	L	-	-	-
8	I	-	-	-	-	-	-	-	L	-	-
9	I	-	-	-	-	-	-	-	-	L	-
10	I	-	-	-	-	-	-	-	-	-	L
11	-	-	-	-	-	-	-	-	-	-	-

Table 8. The transition matrix.

Thirteen of the transitions defined as ‘illegal’, these can be reduced to a subset of six basic error types:

A. Switch VCR off inadvertently.
B. Insert cassette into machine when switched off.
C. Programme without cassette inserted.
D. Fail to Select programme number.
E. Fail to wait for "on" light.
F. Fail to enter programming information.

In addition, one legal transition has been highlighted because it requires a recursive activity to be performed. These activities seem to be particularly prone to errors of omission. These predictions then serve as a basis for the designer to address the re-design of the VCR. A number of illegal transitions could be dealt with fairly easily by considering the use of modes in the operation of the device, such as switching off the VCR without stopping the tape and pressing play without inserting the tape.

Related methods

TAFEI is related to HTA for a description of human activity. Like SHERPA, it is used to predict human error with artefacts. Kirwan and colleagues recommend that multiple human error identification methods can be used to improve the predictive validity of the techniques. This is based on the premise that one method may identify an error that another one misses. Therefore using SHERPA and TAFEI may be better than using either alone. We have found that multiple analysts similarly improves performance of a method. This is based on the premise that one analyst may identify

an error that another one misses. Therefore using SHERPA or TAFEI with multiple analysts may perform better than one analyst with SHERPA or TAFEI.

Advantages

- Structured and thorough procedure.
- Sound theoretical underpinning.
- Flexible, generic, methodology.
- TAFEI can include error reduction proposals.
- TAFEI appears to be relatively simple to apply.
- "TAFEI represents a flexible, generic method for identifying human errors which can be used for the design of anything from kettles to computer systems." (Baber and Stanton, 1994)

Disadvantages

- Not a rapid technique, as HTA and SSD are prerequisites. Kirwan (1998) suggested that TAFEI is a resource intensive technique and that the transition matrix and State Space diagrams may rapidly become unwieldy for even moderately complex systems
- Requires some skill to perform effectively
- Limited to goal-directed behaviour
- TAFEI may be difficult to learn and also time consuming to train.
- It may also be difficult to acquire or construct the SSD's required for a TAFEI analysis. A recent study investigated the use of TAFEI for evaluating design induced pilot error and found that SSD's do not exist for Boeing and Airbus aircraft.

Approximate training and application times

Stanton & Young (1998, 1999) report that observational techniques are relatively quick to train and apply. For example, in their study of radio-cassette machines, training in the TAFEI method took approximately 3 hours. Application of the method by recently trained people took approximately 3 hours in the radio-cassette study to predict the errors.

Reliability and Validity

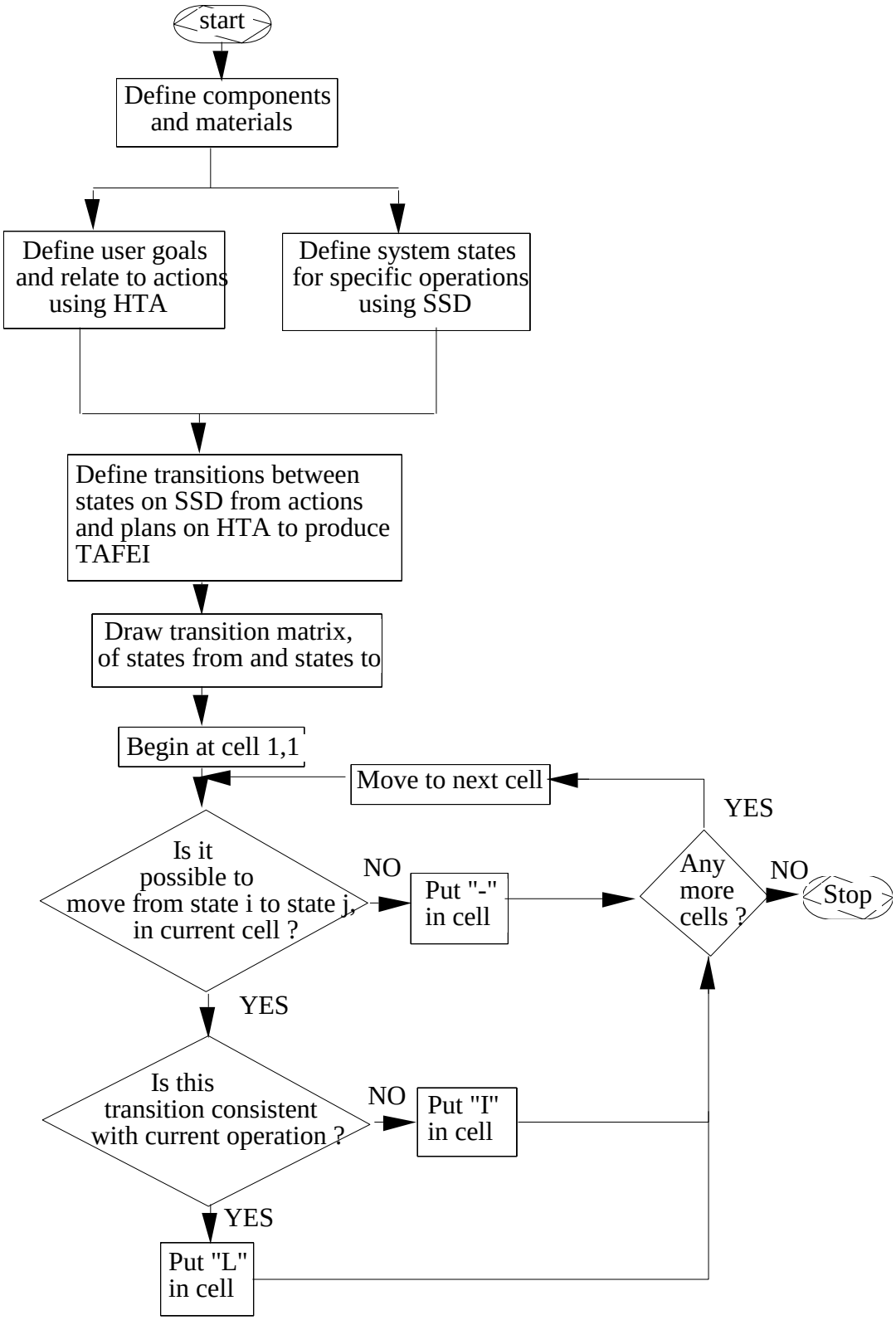
There are some studies that report on the reliability and validity of TAFEI for both expert and novice analysts. These data are reported in table nine.

Table 9. Reliability and validity data for TAFEI

	Novices*1	Experts*2
Reliability	r = 0.67	r = 0.9
Validity	SI = 0.79	SI = 0.9

Note:
*1, taken from Stanton & Baber (2002) Design Studies
*2, taken from Baber & Stanton (1996) Applied Ergonomics

Flowchart


start
Define components and materials
Define user goals and relate to actions using HTA
Define system states for specific operations using SSD
Define transitions between states on SSD from actions and plans on HTA to produce TAFEI
Draw transition matrix, of states from and states to
Begin at cell 1,1
Move to next cell
Is it possible to move from state i to state j, in current cell ?
NO
Put "-" in cell
Any more cells ?
NO
Stop
YES
YES
Is this transition consistent with current operation ?
NO
Put "I" in cell
YES
Put "L" in cell

Tools needed

TAFEI is a pen and paper based tool. There is currently no software available to undertake TAFEI, although there are software packages to support HTA.

Bibliography

Baber, C. & Stanton, N. A. (1994). Task analysis for error identification: a methodology for designing error-tolerant consumer products. Ergonomics, 37, 1923-1941.

Baber, C & Stanton, N. A. (1996) Human error identification techniques applied to public technology: predictions compared with observed use. Applied Ergonomics. 27 (2) 119-131.

Baber, C. & Stanton, N. A. (1999) Analytical prototyping. In: J. M. Noyes & M. Cook (eds) Interface Technology: the leading edge. Research Studies Press: Baldock.

Baber, C. and Stanton, N. A. (2002) Task Analysis For Error Identification: theory, method and validation. Theoretical Issues in Ergonomics Science 3 (2), 212-227.

Bartlett, F.C. (1932) Remembering: a study in experimental and social psychology, Cambridge: Cambridge University Press

Brewer, W.F. (2000) Bartlett's concept of the schema and its impact on theories of knowledge representation in contemporary cognitive psychology, In A. Saito (ed) Bartlett, Culture and Cognition, London: Psychology Press, 69-89

Burford, B. (1993) Designing Adaptive ATMs, Birmingham: University of Birmingham unpublished MSc Thesis

Crawford, J. O.; Taylor, C. & Po, N. L. W. (2001) A case study of on-screen prototypes and usability evaluation of electronic timers and food menu systems. International Journal of Human Computer Interaction 13 (2), 187-201.

Glendon, A.I. and McKenna, E.F. (1995) Human Safety and Risk Management. London: Chapman and Hall

Kirwan, B. (1994). A Guide to Practical Human Reliability Assessment. London: Taylor & Francis.

Stanton, N. A. (2002) Human error identification in human computer interaction. In: J. Jacko and A. Sears (eds) The Human Computer Interaction Handbook. (pp. 371-383) Mahwah, NJ: Lawrence Erlbaum Associates.

Stanton, N. A., & Baber, C. (1996). A systems approach to human error identification. Safety Science, 22, 215-228.

Stanton, N. A., & Baber, C. (1996). Task analysis for error identification: applying HEI to product design and evaluation. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland (eds.), Usability Evaluation in Industry (pp. 215-224). London: Taylor & Francis.

Stanton, N. A., & Baber, C. (1998). A systems analysis of consumer products. In N. A. Stanton (ed.), Human factors in consumer products (pp. 75-90). London: Taylor & Francis.

Stanton, N. A., & Baber, C. (2002) Error by design. Design Studies, 23 (4), 363-384.

Stanton, N. A. & Young, M. (1999) A Guide to Methodology in Ergonomics: Designing for Human Use. Taylor & Francis: London.

Yamaoka, T. and Baber, C. (2000) 3 point task analysis and human error estimation, Proceedings of the Human Interface Symposium 2000, Tokyo, Japan, 395-398

Human Error HAZOP (Hazard and Operability study)

Background and applications

The HAZOP (Hazard and Operability) study system analysis technique was first developed by ICI in the late 1960's. HAZOP was developed as a technique to investigate the safety or operability of a plant or operation and has been used extensively in the Nuclear Power and Chemical process industries. HAZOP (Kletz 1974) is a well-established engineering approach that was developed for use in process design audit and engineering risk assessment (Kirwan 1992a). The HAZOP type approach was developed as simply learning from past incidents became no longer acceptable in large-scale chemical plants (Swann and Preston 1995). Originally applied to engineering diagrams (Kirwan and Ainsworth 1992) the HAZOP technique involves the analyst applying guidewords, such as Not done, More than or Later than, to each step in a process in order to identify potential problems that may occur. Typically, HAZOP analyses are conducted on the final design of a system. Andow (1990) defines the HAZOP procedure as a disciplined procedure which generates questions systematically for consideration in an ordered but creative manner by a team of design and operation personnel carefully selected to consider all aspects of the system under review (Andow, 1990). When conducting a HAZOP type analysis, a HAZOP team is assembled, usually consisting of operators, design staff, human factors specialists and engineers. The HAZOP leader (who should be extensively experienced in HAZOP type analyses) guides the team through an investigation of the system design using the HAZOP 'deviation' guidewords. The HAZOP team consider guidewords for each step in a process to identify what may go wrong. The guidewords are proposed and the leader then asks the team to consider the problem in the following fashion (Swann and Preston, 1995):

- Which section of the plant is being considered?
- What is the deviation and what does it mean?
- How can it happen and what is the cause of the deviation?
- If it cannot happen, move onto the next deviation.
- If it can happen, are there any significant consequences?
- If there are not, move onto the next guideword.
- If there are any consequences, what features are included in the plant to deal with these consequences?
- If the HAZOP team believes that the consequences have not been adequately covered by the proposed design, then solutions and actions are considered.

Applying guide words like this in a systematic way ensures that all of the possible deviations are considered. The efficiency of the actual HAZOP analysis is largely dependent upon the HAZOP team.

There are a number of different variations of HAZOP style approaches, such as CHAZOP (Swann and Preston, 1995) and SCHAZOP (Kennedy and Kirwan, 1998). A more human factors orientated version emerged in the form of the Human Error HAZOP, aimed at dealing with human error issues (Kirwan and Ainsworth 1992). In the development of another HEI tool (PHECA) Whalley (1988) also created a set of human factors based guidewords, which are more applicable to human error. These Human Error guidewords are shown below. The error guidewords are applied to each bottom level task step in the HTA to determine any credible errors (i.e. those judged

by the subject matter expert to be possible). Once the analyst has recorded a description of the error, the consequences, cause and recovery path of the error are also recorded. Finally, the analyst then records any design improvements to remedy the error.

- Not Done
- Less Than
- More Than
- As Well As
- Other Than
- Repeated
- Sooner Than
- Later Than
- Mis-ordered
- Part Of

Domain of application
Nuclear Power and Chemical Process Industries.

Procedure and advice (Human Error HAZOP)
Step 1: Assembly of HAZOP team
The most important part of any HAZOP analysis is assembling the correct HAZOP team (Swann and Preston, 1995). The HAZOP team needs to possess the right combination of skills and experience in order to make the analysis efficient. The HAZOP team leader should be experienced in HAZOP type analysis so that the team can be guided effectively. For a human error HAZOP analysis of a nuclear petro-chemical plant, it is recommended that the team be comprised of the following personnel.

- HAZOP team leader
- Human Factors Specialist
- Human Reliability Analysis (HRA)/Human Error Identification (HEI) Specialist
- Project engineer
- Process engineer
- Operating team leader
- Control room operator(s)
- Data recorder

Step 2: Hierarchical Task Analysis (HTA)
Next, an exhaustive task description of system under analysis should be created, using Hierarchical Task Analysis. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Step 3: Guideword consideration
The HAZOP team takes the first/next bottom level task step from the HTA and

considers each of the associated HAZOP guidewords for the task step under analysis. This involves discussing whether the guideword could have any effect on the task step or not and also what type of error would result. If any of the guidewords are deemed credible by the HAZOP team, then they move onto step 4.

Step 4: Error description
For any credible guidewords, the HAZOP team should provide a description of the form that the resultant error would take e.g. operator fails to check current steam pressure setting. The error description should be clear and concise.

Step 5: Consequence analysis
Once the HAZOP team have described the potential error, its consequence should be determined. The consequence of the error should be described e.g. Operator fails to comprehend high steam pressure setting.

Step 6: Cause analysis
Next, the HAZOP team should determine the cause(s) of the potential error. The cause analysis is crucial to the remedy or error reduction part of the HAZOP analysis. Any causes should be recorded and described clearly.

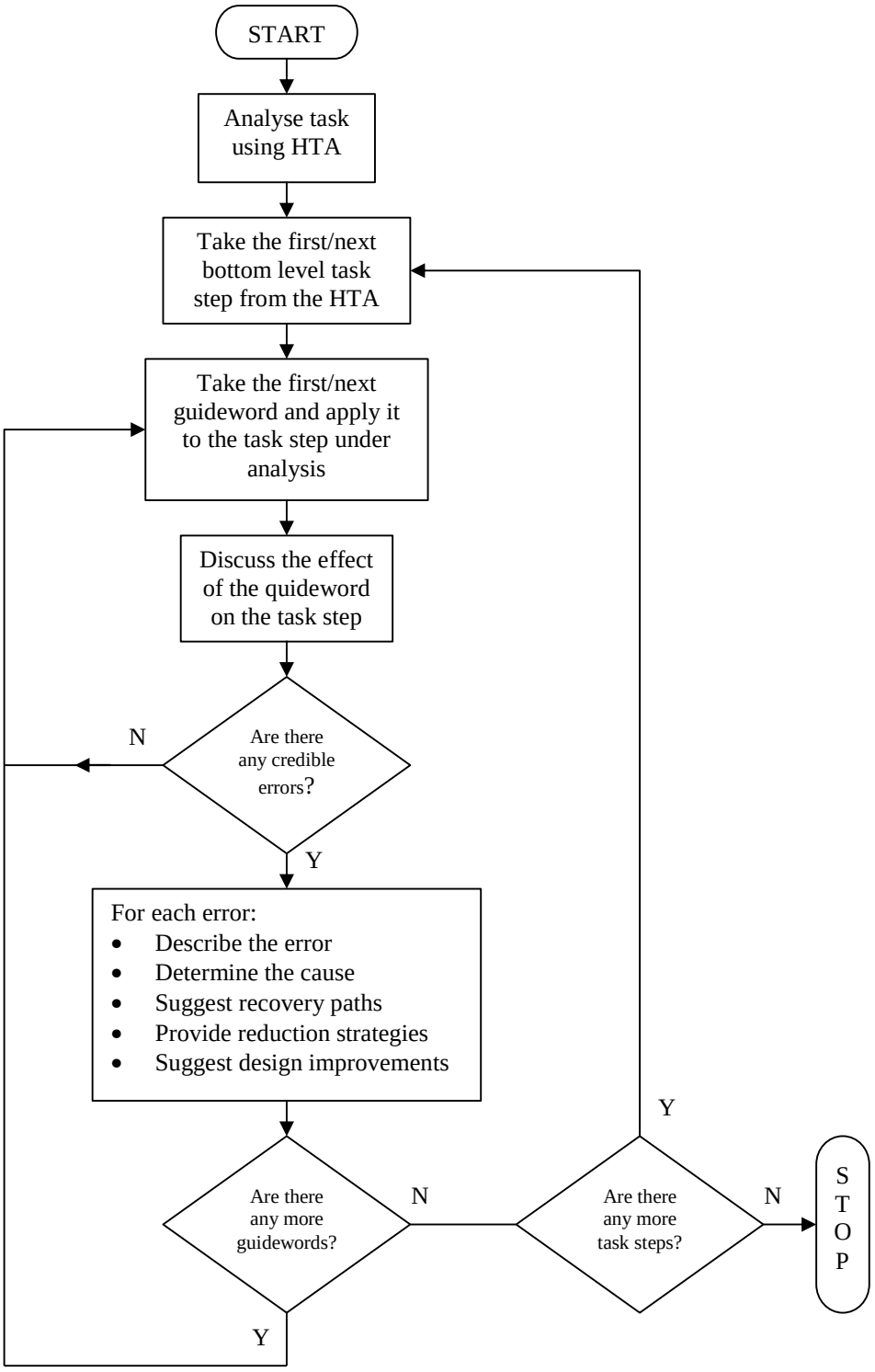
Step 7: Recovery Path analysis
In the recovery path analysis, any recovery paths that the operator can take after the described error has occurred to avoid the consequences are noted.

Step 8: Error Remedy
Finally, the HAZOP team propose any design or operational remedies that could reduce the chances of the error occurring. This is based upon subjective analyst judgement and domain expertise.

Advantages

- A correctly conducted HAZOP analysis has the potential to highlight all of the possible errors that could occur in the system.
- HAZOP has been used emphatically in many domains. HAZOP style techniques have received wide acceptance by both the process industries and the regulatory authorities (Andrews and Moss, 1993).
- "Two heads are better than one." Since a team of experts is used, the technique should be more comprehensive than other 'single analyst' techniques. This also removes the occurrence of 'far fetched' errors generated by single analyst techniques.
- "HAZOP can be readily extended to address human factors issues." (Kirwan and Ainsworth, 1992)
- Appears to be a very exhaustive technique.
- Easy to learn and use.
- Whalley's (1988) guidewords are generic, allowing the technique to be applied to a number of different domains.

Flowchart


START
Analyse task using HTA
Take the first/next bottom level task step from the HTA
Take the first/next guideword and apply it to the task step under analysis
Discuss the effect of the quideword on the task step
Are there any credible errors?
N
Y
For each error:
Describe the error
Determine the cause
Suggest recovery paths
Provide reduction strategies
Suggest design improvements
Are there any more guidewords?
N
Y
Are there any more task steps?
Y
N
STOP

Disadvantages

- The technique can be extremely time consuming. Typical HAZOP analyses can take up to several weeks to be completed.
- The technique requires a mixed team made up of operators, human factors specialists, designers, engineers etc. Building such a team and making sure they can all be brought together at the same time is often a difficult task.
- HAZOP analysis generates huge amounts of information that has to be recorded and analysed.
- Laborious.
- Disagreement within the HAZOP team may be a problem.
- The guidewords used are either limited or specific to nuclear petro-chemical industry.
- The human error HAZOP guidewords lack comprehensiveness (Salmon et al 2002)

Example

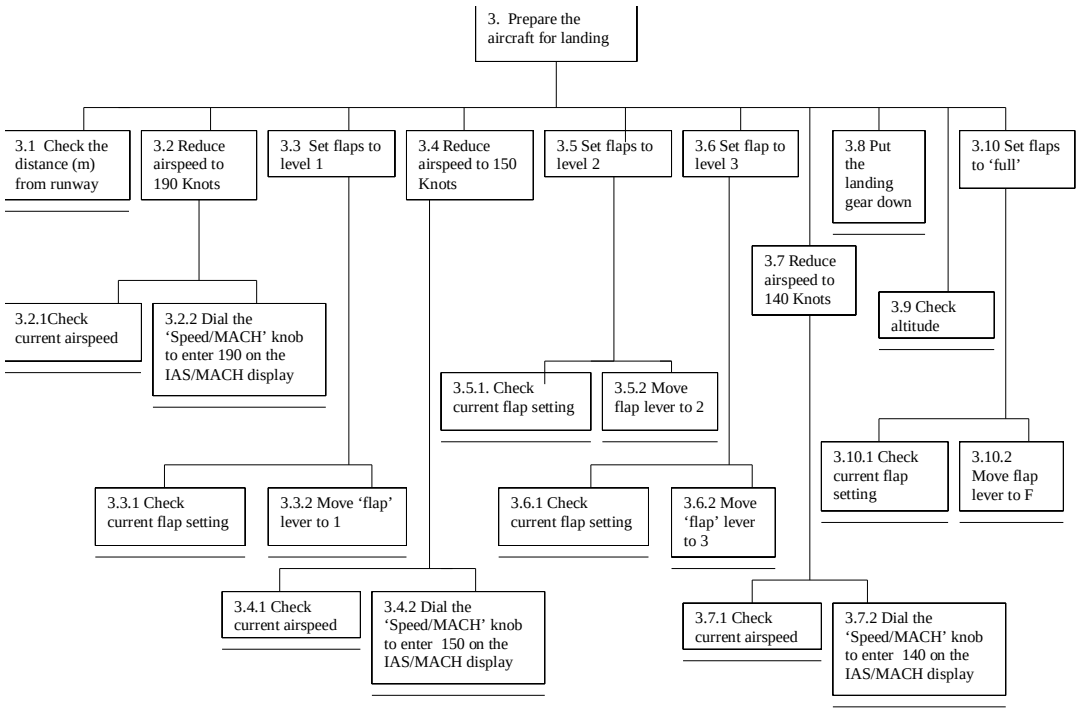


Figure 9. Extract of HTA of task 'Land A320 at New Orleans using the Auto-land system

Table 10. Extract of Human Error HAZOP analysis of task 'Land A320 at New Orleans using the Auto-land system

Task Step	Guide-word	Error	Consequence	Cause	Recovery path	Design Improvements
3.1 Check the distance from runway	Later than	Pilot checks the distance from the runway later than he should	Plane may be travelling to fast for that stage of the approach and also may have the wrong level of flap	Pilot inadequacy Pilot is preoccupied with another landing task	3.9	Auditory distance countdown inside 25N miles
3.2.1 Check current airspeed	Not done	Pilot fails to check current airspeed	Pilot changes airspeed wrongly i.e. may actually increase airspeed	Pilot is pre-occupied with other landing tasks	3.4.1	Auditory speed updates Bigger, more apparent speedo
	Mis-ordered	Pilot checks the current airspeed after he has altered the flaps	Plane may be travelling too fast for that level of flap or that leg of the approach	Pilot inadequacy Pilot is preoccupied with other landing tasks	3.4.1	Design flaps so each level can only be set within certain speed level windows
3.2.2 Dial the speed/mach knob to enter 190	Not done	Pilot fails to enter new airspeed	Plane may be travelling too fast for the approach	Pilot is pre-occupied with other landing tasks	3.4.2	Auditory reminder that the plane is travelling to fast e.g. overspeed display
	Less than	Pilot does not turn the Speed/Mach knob enough	The planes speed is not reduced enough and the plane may be travelling too fast for the approach	Poor control design Pilot inadequacy	3.4.2	One full turn for 1 knot Improved control feedback
	More than	Pilot turns the Speed/MACH knob too much	The planes speed is reduced too much and so the plane is travelling too slow for the approach	Poor control design Pilot inadequacy	3.4.2	Improved control feedback
	Sooner than	Pilot reduces the planes speed too early	The plane slows down too early	Pilot is preoccupied with other landing tasks Pilot inadequacy	3.4.2	Plane is travelling too slow auditory warning
	Other than	Pilot reduces the planes using the wrong knob e.g. HDG knob	Plane does not slow down to desired speed and takes on a heading of 190	Pilot is preoccupied with other landing tasks Pilot inadequacy	3.4.2	Clearer labelling of controls Overspeed auditory warning
3.3.1 Check the current flap setting	Not done	Pilot fails to check the current flap setting	The pilot does not comprehend the current flap setting	Pilot is preoccupied with other landing tasks Pilot inadequacy	3.4.2	Bigger/improved flap display/control Auditory flap setting reminders

Related methods

HAZOP type analyses are typically conducted on a HTA of the task under analysis. Engineering diagrams, flow-sheets, operating instructions and plant layouts are also required (Kirwan and Ainsworth, 1992)

Approximate training and application times

Swann and Preston (1995) report that studies on the duration of the HAZOP analysis process have been conducted, with the conclusion that a thorough HAZOP analysis carried out correctly would take over 5 years for a typical processing plant. This is clearly a worst-case scenario and impractical. More realistically, Swann and Preston (1995) suggest that ICI benchmarking shows that a typical HAZOP analysis would require about 40 meetings lasting approximately 3 hours each.

Reliability and Validity
The HAZOP type approach has been used emphatically over the last 4 decades in process control environments. However (Kennedy, 1997) reports that it has not been subjected to rigorous academic scrutiny (Kennedy and Kirwan, 1998). In a recent study (Stanton et al, 2003) reported that in a comparison of 4 HEI methods (HET, Human Error HAZOP, HEIST, SHERPA) when used to predict potential design induced pilot error, subjects using the human error HAZOP method achieved acceptable sensitivity in their error predictions (mean sensitivity index 0.62). Furthermore, only those subjects using the HET methodology performed better.

Tools needed
HAZOP analyses can be carried out using pen and paper. Engineering diagrams are also normally required. The human error taxonomy is also required for the human error HAZOP variation. A HTA for the task under analysis is also required.

Bibliography
Kennedy, R., & Kirwan, B. (1998) Development of a Hazard and Operability-based method for identifying safety management vulnerabilities in high risk systems, *Safety Science*, Vol. 30, Pages 249-274
Kirwan, B., Ainsworth, L. K. (1992) A guide to Task Analysis, Taylor and Francis, London, UK.
Kirwan, B. (1992) Human error identification in human reliability assessment. Part 1: Overview of approaches. *Applied Ergonomics* Vol. 23(5), 299 – 318
Stanton, N., Salmon, P., Young, M. S., (2003) UNPUBLISHED
Swann, C. D. and Preston, M. L. (1995) Twenty-five years of HAZOPs. *Journal of Loss Prevention in the Process Industries*, Vol. 8, Issue 6, 1995, Pages 349-353
Whalley (1988) Minimising the cause of human error. In B. Kirwan & L. K. Ainsworth (eds.) *A Guide to Task Analysis*. Taylor and Francis, London.

THEA – Technique for Human Error Assessment

Steven Pocock, University of York, Heslington, York, YO10 5DD, UK
Michael Harrison, University of York, Heslington, York, YO10 5DD, UK
Peter Wright, University of York, Heslington, York, YO10 5DD, UK
Paul Johnson, University of York, Heslington, York, YO10 5DD, UK

Background and applications

The Technique for Human Error Assessment (THEA) was developed primarily to aid designers/engineers in identifying potential problems between users and interfaces in the early design stages of systems design. The technique is a highly structured one that employs cognitive error analysis based upon Norman's (1988) model of action execution. The main aim of the development of THEA was to create a tool that could be used by non-human factors professionals. It is recommended that the technique should be used in the very early stages of systems design to identify any potential interface problems. Although THEA has its roots firmly in HRA methodology, it is suggested by the authors that the technique is more suggestive and also much easier to apply than typical HRA methods (Pocock et al 1997). Very similar to HEIST (Kirwan, 1994) THEA uses a series of questions in a checklist style approach based upon goals, plans, performing actions and perception/evaluation/interpretation. These questions were developed considering each stage of Norman's action execution model. THEA also utilises a scenario-based analysis, whereby the analyst exhaustively describes the scenario under analysis before any analysis is carried out. The scenario description gives the analyst a thorough description of the scenario under analysis, including information such as actions and any contextual factors, which may provide opportunity for an error to occur.

Domain of application

Generic.

Procedure and advice

Step 1: System description

Initially, a THEA analysis requires a formal description of the system and task or scenario under analysis. This system description should include details regarding the specification of the systems functionality and interface and also if and how it interacts with any other systems (Pocock, Harrison, Wright & Fields, 1997).

Step 2: Scenario description

Next, the analyst should provide a description of the type of scenario under analysis. The authors have developed a scenario template that assists the analyst in developing the scenario description. The scenario description template is shown in table 11.

Step 3: Task description

A description of the work that the operator or user would perform in the scenario is also required. This should describe goals, plans and intended actions.

Step 4: Goal decomposition

A HTA should be performed in order to give clarity and structure to the information presented in the scenario description. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve),

operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively redescribe the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Table 11. A template for describing scenarios (Source: Pocock, Harrison, Wright & Fields, 1997)

AGENTS
• The human agents involved and their organisations
• The roles played by the humans, together with their goals and responsibilities
RATIONALE
• Why is this scenario and interesting or useful one to have picked?
SITUATION AND ENVIRONMENT
• The physical situation in which the scenario takes place
• External and environmental triggers, problems and events that occur in this scenario
TASK CONTEXT
• What tasks are performed?
• Which procedures exist, and will they be followed as prescribed?
SYSTEM CONTEXT
• What devices and technology are involved?
• What usability problems might participants have?
• What effects can users have?
ACTION
• How are the tasks carried out in context?
• How do the activities overlap?
• Which goals do actions correspond to?
EXCEPTIONAL CIRCUMSTANCES
• How might the scenario evolve differently, either as a result of uncertainty in the environment or because of variations in agents, situation, design options, system and task context?
ASSUMPTIONS
• What, if any, assumptions have been made that will affect this scenario?

Step 5: Error Analysis

Next, the analyst has to identify and explain any human error that may arise during the operation of the system under analysis. THEA provides a structured questionnaire/checklist style approach in order to aid the analyst in identifying any possible errors. The analyst simply asks questions (from THEA) about the scenario under analysis in order to identify potentially problematic areas in the interaction between the operator and the system. The analyst should record the error, its causes and its consequences. Then questions are normally asked about each goal or task in the HTA, or alternatively, the analyst can select parts of the HTA where problems are anticipated. The THEA error analysis questions are comprised of four categories:

- Goals
- Plans
- Performing Actions
- Perception, Interpretation and evaluation

Examples of the THEA error analysis questions for each of the four categories are

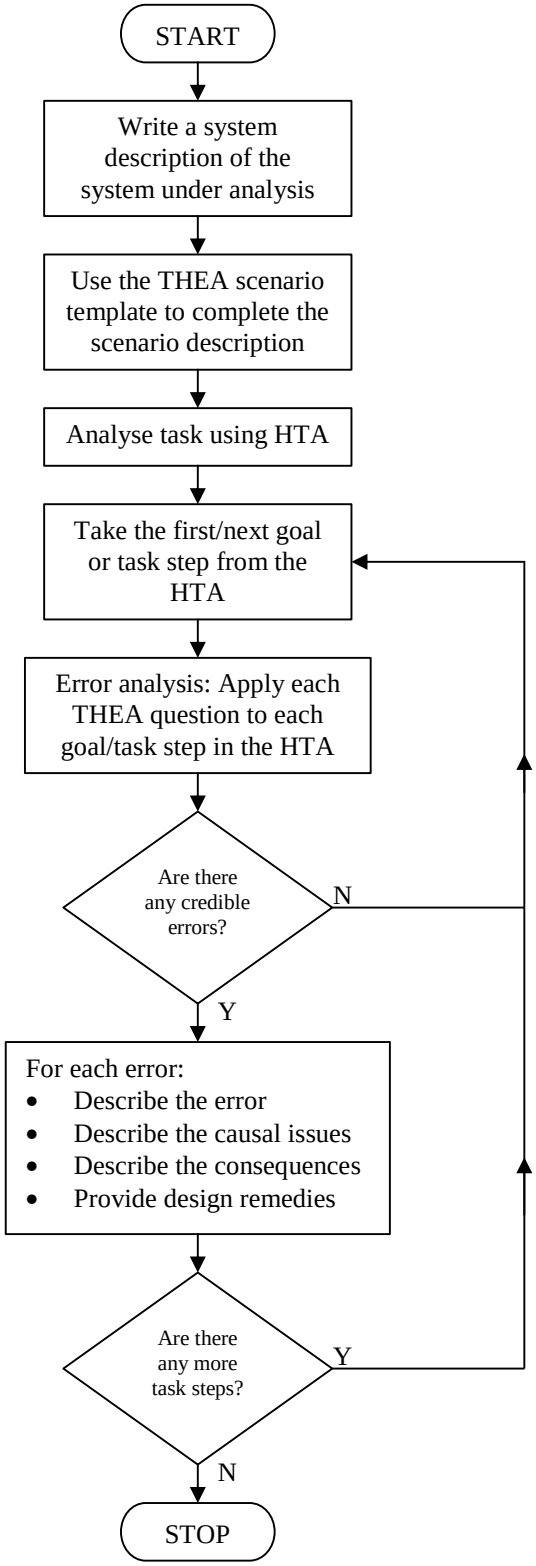
provided below.

Table 12. Example THEA error analysis questions

Questions	Consequences	Design Issues
Goals		
G1 – Are items triggered by stimuli in the interface, environment, or task?	If not, goals (and the tasks that achieve them) may be lost forgotten, or not activated, resulting in omission errors	Are triggers clear and meaningful? Does the user need to remember all of the goals
G2 – Does the user interface 'evoke' or 'suggest' goals?	If not, goals may not be activated, resulting in omission errors. If the interface does 'suggest' goals, they may not always be the right ones, resulting in the wrong goal being addressed	e.g. graphical display of flight plan shows pre-determined goals as well as current progress
Plans		
P1 - Can actions be selected in situ, or is pre-planning required?	If the correct action can only be taken by planning in advance, then the cognitive work may be harder. However, when possible, planning ahead often leads to less error-prone behaviour and fewer blind alleys	
P2 – Are there well practised and pre-determined plans	If a plan isn't well known or practised then it may be prone to being forgotten or remembered incorrectly. If plans aren't pre-determined, and must be constructed by the user, then their success depends heavily on the user possessing enough knowledge about their goals and the interface to construct a plan. If pre-determined plans do exist and are familiar, then thy might be followed inappropriately, not taking account of the peculiarities of the current context	
Performing actions		
A1 - Is there physical or mental difficulty in executing the actions?	Difficult, complex or fiddly actions are prone to being carried out incorrectly	
A2 – Are some actions made unavailable at certain times?		
Perception, Interpretation and evaluation		
I1 – Are changes in the system resulting from user action clearly perceivable?	If there is no feedback that an action has been taken, the user may repeat actions, with potentially undesirable effects	
I2 – Are the effects of user actions perceivable immediately?	If feedback is delayed, the user may become confused about the system state, potentially leading up to a supplemental (perhaps inappropriate) action being taken	

Step 6: Design Implications/recommendations
Once the analyst has identified an error, the final stage of the THEA analysis is to offer any design remedies that would eradicate the error identified. This is based on the subjective judgement of the analyst and the design issues section of the THEA questions, which prompts the analyst for design remedies.

Flowchart


START
Write a system description of the system under analysis
Use the THEA scenario template to complete the scenario description
Analyse task using HTA
Take the first/next goal or task step from the HTA
Error analysis: Apply each THEA question to each goal/task step in the HTA
Are there any credible errors?
N
Y
For each error:
Describe the error
Describe the causal issues
Describe the consequences
Provide design remedies
Are there any more task steps?
Y
N
STOP

Advantages

- THEA is a highly structured technique.
- The THEA technique can be used by non-human factors professionals.
- As it is recommended that THEA be used very early in the system life cycle, potential interface problems can be identified and eradicated very early in the design process.
- THEA error prompt questions are based on Norman's action execution model.
- THEA's error prompt questions aid the analyst in the identification of potential errors.
- THEA is more suggestive and easier to apply than typical HRA methods (Pocock, Harrison, Wright & Fields, 1997).
- Each error question has associated consequences and design issues to aid the analyst.
- THEA appears to be a very generic technique, allowing it to be applied to many domains, such as command and control.

Disadvantages

- Although error questions prompt the analyst for potential errors, THEA does not use any error modes and so the analyst may be unclear on the types of errors that may occur. HEIST (Kirwan, 1994) however, uses error prompt questions linked with an error mode taxonomy, which seems to be a much sounder approach.
- THEA is very resource intensive, particularly with respect to time taken to complete an analysis.
- Error consequences and design issues provided by THEA are very generic and limited.
- At the moment, there appears to be no validation evidence associated with THEA.
- HTA, task decomposition and scenario description create additional work for the analyst.
- For a technique that is supposed to be usable by non-human factors professionals, the terminology used in the error analysis questions section is confusing and hard to decipher. This could cause problems for non-human factors professionals.

Example (Source: Pocock, Harrison, Wright & Fields, 1997)
The following example is a THEA analysis of a video recorder programming task (Pocock, Harrison, Wright & Fields, 1997)

Table 13. Scenario details

SCENARIO NAME: Programming a video recorder to make a weekly recording
ROOT GOAL: Record a weekly TV programme
SCENARIO SUB-GOAL: Setting the recording date
ANALYST(S) NAME(S) & DATE:
AGENTS: A single user interfacing with a domestic video cassette recorder (VCR) via a remote control unit (RCU)
RATIONALE: The goal of programming this particular VCR is quite challenging. Successful programming is not certain
SITUATION & ENVIRONMENT: A domestic user wishes to make a recording of a television programme which occurs on a particular channel at the same time each week. The user is not very technologically aware and has not programmed this VCR previously. A reference handbook is not available, but there is no time pressure to set the machine – recording is not due to commence until tomorrow
TASK CONTEXT: The user must perform the correct tasks to set the VCR to record a television programme on three consecutive Monday evenings from 6pm-7pm on Channel 3. Today is Sunday
SYSTEM CONTEXT: The user has a RCU containing navigation keys used in conjunction with programming the VCR as well as normal VCR playback operation. The RCU has 4 scrolling buttons, indicating left, right, up, down. Other buttons relevant to programming are labelled OK and I.
ACTIONS: The user is required to enter a recording date into the VCR via the RCU using the buttons listed above. The actions appear in the order specified by the task decomposition.
EXCEPTIONAL CIRCUMSTANCES: None
ASSUMPTIONS: None

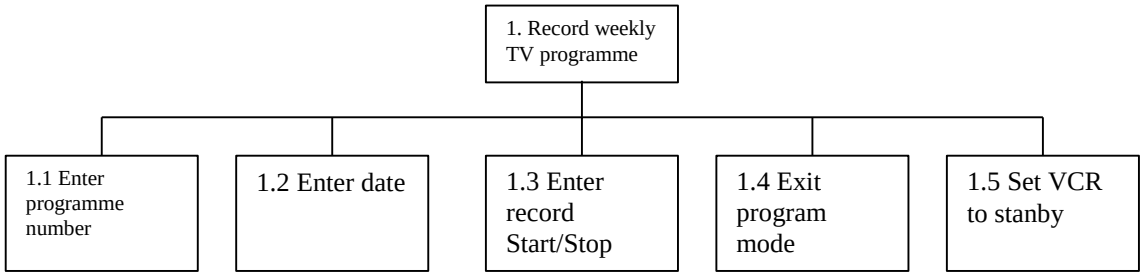


Figure 10. Video recorder HTA (adapted from Pocock, Harrison, Wright & Fields, 1997)

Table 14. Error Analysis Questionnaire (Source: Pocock, Harrison, Wright & Fields, 1997)

SCENARIO NAME: Programming a video recorder to make a weekly recording			
TASK BEING ANALYSED: Setting the recording date			
ANALYST(S) NAME(S) & DATE			
QUESTION	**CAUSAL ISSUES**	**CONSEQUENCES**	**DESIGN ISSUES**
GOALS, TRIGGERING, INITIATION			
G1 – Is the task triggered by stimuli in the interface, environment or the task itself?	Yes. (The presence of an 'enter date' prompt is likely to trigger the user to input the date at this point)		
G2 – Does the UI 'evoke' or 'suggest' goals	N/A. (The UI does not per se, strictly evoke or suggest the goal of entering the date)		
G3 – Do goals come into conflict>	There are no discernible goal conflicts		
G4 – Can the goal be satisfied without all its sub-goals being achieved?	NO. The associated sub-goal on this page of setting the DAILY/WEEKLY function may be overlooked. Once the date is entered, pressing the right cursor key on the RCU will enter the next 'ENTER HOUR' setting	Failure to set the DAILY/WEEKLY option. Once the ENTER HOUR screen is entered, the DAILY/WEEKLY option is no longer available	Suggest addition of an interlock so that the daily/weekly option cannot be bypassed
PLANS			
P1 – Can actions be selected in-situ, or is pre-planning required?	True. (Entering the date can be done 'on-the-fly'. No planning is required		
P2 – Are there well practised and pre-determined plans?	N/A. (A pre-determined plan, as such, does not exist, but the user should possess enough knowledge to know what to do at this step)		
P3 – Are there plans or actions that are similar? Are some used more often than others?	There are no similar or more frequently used plans or actions associated with this task.		
P4 – Is there feedback to allow the user to determine that the task is proceeding successfully towards the goal, and according to plan?	Yes. (As the user enters digits into the date field via the RCU, they are echoed back on screen	Task is proceeding satisfactorily towards the goal of setting the date, although the date being entered is not necessarily correct).	(See A1)
PERFORMING ACTIONS			
A1 – Is there physical or mental difficulty in performing the task?	Yes. The absence of any cues for how to enter the correct date format makes this task harder to perform	The user may try to enter the year or month instead of the day. Additionally, the user may try to add a single figure date, instead of preceding the digit with a zero.	Have an explanatory text box under the field or, better still, default today's date in the date field
A2 – Are some actions made unavailable at certain times?	No. (The only actions required of the user is to enter two digits into the blank field		

A3 – Is the correct action dependent on the current mode?	No. (The operator is operating in a single programming mode)		
A4 – Are additional actions required to make the right controls and information available at the right time?	Yes. The date field is presented blank. If the user does not know the date for recording (or today's date), the user must know to press the 'down' cursor key on the RCU to make today's date visible	The user may be unable to enter the date, or the date must be obtained from an external source. Also, if the user presses either the left or right cursor key, the 'enter date' screen is exited	1. Default current date into field 2. Prevent user from exiting 'enter date' screen before an entry is made (e.g. software lock-in)
PERCEPTION, INTERPRETATION & EVALUATION			
I1 – Are changes to the system resulting from user action clearly perceivable?	Yes. (Via on-screen changes to the date field)		
I2 – Are effects of such user actions perceivable immediately?	Yes. (Digit echoing of RCU key presses is immediate)		
I3 – Are changes to the system resulting from autonomous system actions clearly perceivable?	N/A. (The VCR performs no autonomous actions)		
I4 – Are the effects of such autonomous system actions perceivable immediately?	N/A		
I5 – Does the task involve monitoring, vigilance, or spells of continuous attention?	No. (There is no monitoring or continuous attention requirements on the user.		
I6 – Can the user determine relevant information about the state of the system from the total information provided?	NO. User cannot determine current date without knowing about the 'down' cursor key. Also, if date of recording is known, user may not know about the need to enter two digits.	If user doesn't know today's date, and only knows that, say, Wednesday, is when you want the recordings to commence, then the user is stuck	As A1
I7 – Is complex reasoning, calculation, or decision making involved?	No.		
I8 – If the user is interfacing with a moded system, is the correct interpretation dependent on the current mode?	N/A	It is not considered likely that the date field will be confused with another entry field e.g. hour	

Related methods
THEA is one of a number of HEI techniques. THEA is very similar to HEIST (Kirwan 1994) in that it uses error prompt questions to aid the analysis. A THEA analysis should be conducted on a HTA of the task under analysis.

Approximate training and application times
Although no training and application time is offered in the literature, it is apparent that the amount of training time would be minimal. The application time, however, would be high, especially for complex tasks.

Reliability and Validity
No data regarding reliability and validity are offered by the authors.

Tools needed
To conduct a THEA analysis, pen and paper is required. The analyst would also require functional diagrams of the system/interface under analysis and the THEA error analysis questions.

Bibliography
Pocock, S., Harrison, M., Wright, P., Johnoson, P. (2001) THEA: A technique for Human Error Assessment Early in Design, *unpublished work*
Pocock, S., Harrison, M., Wright, P., Fields, B. (2001) THEA – A Reference Guide. *Unpublished work*

HEIST – Human Error Identification in Systems Tool

Barry Kirwan, EUROCONTROL, Experimental Centre, BP15, F91222, Bretigny Sur Orge, France

Background and applications

The Human Error Identification in Systems Tool (HEIST) (Kirwan 1994) is a HEI technique that is based upon a series of tables containing questions or 'error identifier prompts' surrounding external error modes (EEM), performance shaping factors (PSF) and psychological error mechanisms (PEM). When using HEIST, the analyst identifies errors through applying a set of questions to all of the tasks involved in a scenario. The questions link EEM's (type of error) to relevant PSFs. All EEM's are then linked to PEM's (psychological error-mechanisms). Once this has been done, the recovery potential, consequences and error reduction mechanisms are noted in a tabular error-analysis format. This question and answer approach to error identification comes in the form of a table, which contains a code for each error-identifying question, the error identifier question, the external error mode (EEM), the identified cause (system cause or psychological error mechanism) and any error reduction guidelines offered. The HEIST tables and questions are based upon the Skill, Rule and Knowledge (SRK) framework (Rasmussen at al, 1981) i.e. Activation/Detection, Observation/Data collection, Identification of system state, Interpretation, Evaluation, Goal selection/Task definition, Procedure selection and Procedure execution. These error prompt questions are designed to prompt the analyst for potential errors. Each of the error identifying prompts are PSF based questions which are coded to indicate one of six PSFs. These performance shaping factors are Time (T), Interface (I), Training/Experience (E), Procedures (P), Task organisation (O), Task Complexity (C). The technique itself has similarities to a number of traditional HEI techniques such as SRK, SHERPA and HRMS (Kirwan 1994). Table 14 shows an extract of the HEIST table for *procedure execution.* There are eight HEIST tables in total. The analyst classifies the task step under analysis into one of the SRK behaviours and then applies the relevant table to the task step and determines whether any errors are credible or not. For each credible error, the analyst then records the system cause or PEM and error reduction guidelines (both of which are provided in the HEIST tables) and also the error consequence. Although it can be used as a stand-alone method, HEIST is also used as part of the HERA 'toolkit' methodology (Kirwan, 1998b) as a back up check for any errors identified. It is also suggested that the HEIST can be used by just one analyst and also that the analyst does not have to be an expert for the system under analysis (Kirwan, 1994).

Domain of application

Nuclear power and chemical process industries.

Procedure and advice

Step 1: Hierarchical Task Analysis (HTA)

The process begins with the analysis of work activities, using Hierarchical Task Analysis. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe the activity in greater degrees of detail. The analysis begins with an overall goal of

the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Table 14. Extract of Procedure Execution HEIST table (Source: Kirwan, 1994)

Code	**Error-identifier prompt**	**External error mode**	**System cause/PEM**	**Error reduction guidelines**
PET1	Could the operator fail to carry out the act in time?	Omission of action	Insufficient time available, inadequate time perception, crew co-ordination failure, manual variability, topographic misorientation	Training, team training and crew co-ordination trials, EOP's, ergonomics design of equipment
PET2	Could the operator carry out the task too early?	Action performed too early	Inadequate time perception, crew co-ordination failure	Training, perception cues, time-related displays, supervision
PEP1	Could the operator carry out the task inadequately?	Error of quality Wrong action Omission of action	Manual variability prompting, random fluctuation, misprompting, misperception, memory failure	Training, ergonomic design of equipment, ergonomic procedures, accurate and timely feedback, error-recovery potential, supervision
PEP2	Could the operator lose his/her place during procedure execution, or forget an item?	Omission of action Error of quality	Memory failure, interruption, vigilance failure, forget isolated act, misprompting, cue absent	Ergonomic procedures with built in checks, error recovery potential (error tolerant system design) , good system feedback, supervision and checking

Step 2: Task step classification
The analyst takes the first task step from the HTA and classifies it into one or more of the eight SRK behaviours (Activation/Detection, Observation/Data collection, Identification of system state, Interpretation, Evaluation, Goal selection/Task definition, Procedure selection and Procedure execution). For example, the task step 'Pilot dials in airspeed of 190 using the speed/MACH selector knob' would be classified as procedure execution. This part of the HEIST analysis is based entirely upon analyst subjective judgement.

Step 3: Error analysis
Next, the analyst should take the appropriate HIEST table and apply each of the error identifier prompts to the task step under analysis. Based upon subjective judgement, the analyst should determine whether or not any of the associated errors could occur during the task step under analysis. If the analyst deems an error to be credible, the error should be described and the EEM, system cause and PEM should be determined from the HEIST table.

Step 4: Error reduction analysis
For each credible error, the analyst should select the appropriate error reduction

guidelines from the HEIST table. Each HEIST error prompt has an associated set of error reduction guidelines. Whilst it is recommended that the analyst should use these, it is also possible for the analyst to propose their own design remedies.

Advantages

- As HEIST uses error identifier prompts based upon the SRK framework, the technique has the potential to be exhaustive.
- Error identifier prompts aid the analyst in error identification.
- Once a credible error has been identified, the HEIST tables provide the EEM's, PEM's and error reduction guidelines.

Disadvantages

- HEIST is very time consuming in its application.
- The need for an initial HTA creates further work for HEIST analysts.
- Although the HEIST tables provides error reduction guidelines, these are very generic and not really specific nor of any use e.g. ergonomic design of equipment and good system feedback.
- A HEIST analysis requires human factors/psychology professionals.
- No validation evidence is available for the HEIST.
- No evidence of the use of HEIST is available in the literature.
- Many of the error identifier prompts used by HEIST are repetitive.
- Salmon et al (2002) reported that HEIST performed poorly when used to predict potential design induced error on the flight task 'Land aircraft at New Orleans using the auto-land system'. Out of the four techniques HET, SHERPA, Human Error HAZOP and HIEST, subjects using HEIST performed the worst.

HEIST Example – Land A320 at New Orleans using the Autoland system
A HET analysis was conducted on the flight task 'Land A320 at New Orleans using the Autoland system.

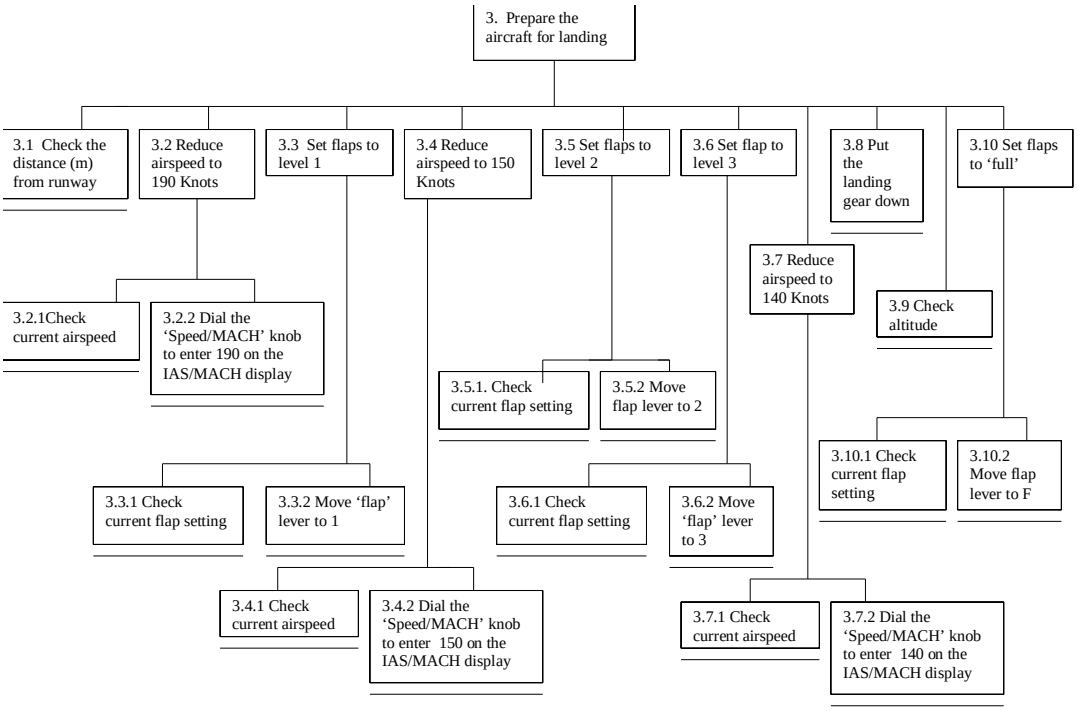


Figure 11. Extract of HTA 'Land at New Orleans using auto-land system' (Marshall et al (2003)

Table 15. Extract of HEIST analysis of the task 'Land at New Orleans using auto-land system (Salmon et al 2003).

Task step	**Error code**	**EEM**	**Description**	**PEM System cause**	**Consequence**	**Error reduction guidelines**
3.2.2	PEP3	Action on wrong object	Pilot alters the airspeed using the wrong knob e.g. heading knob	Topographic misorientation Mistakes alternatives Similarity matching	The airspeed is not altered and the heading will change to the value entered	Ergonomic design of controls and displays Training Clear labelling
3.2.2	PEP4	Wrong action	Pilot enters the wrong airspeed	Similarity matching Recognition failure Stereotype takeover Misperception Intrusion	Airspeed will change to the wrong airspeed	Training Ergonomic procedures with checking facilities Prompt system feedback

Related methods

A HEIST analysis should be conducted on a HTA of the task under analysis. The HEIST tables and error identifier prompts are also based upon the SRK framework approach. The use of error identifier prompts is similar to the approach used by THEA (Pocock et al 2001). HEIST is also used as a back-up check when using the HERA toolkit approach to HEI (Kirwan 1998b)

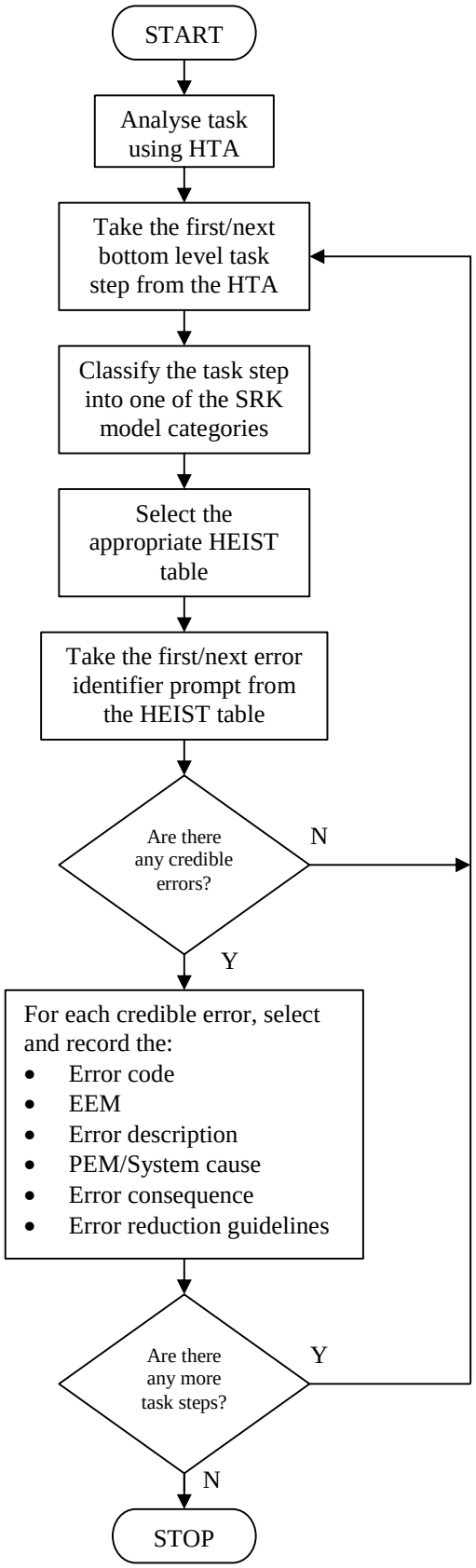
Approximate training and application times
Although no training and application time is offered in the literature, it is apparent that the amount of time in both cases would be high. When using HEIST to predict potential design induced pilot error, Marshall et al (2003) reported that the average training time for participants using the HEIST technique was 90 minutes. The average application time of HEIST in the same study was 110 minutes.

Reliability and Validity
The reliability and validity of the HEIST technique is questionable. Whilst no data regarding the reliability and validity are offered by the techniques authors, (Marshall et al 2003) report that subjects using HEIST achieved a mean sensitivity index of 0.62 at time 1 and 0.58 at time 2. This represents moderate reliability and validity ratings. In comparison to three other methods (SHERPA, HET and Human Error HAZOP) when used to predict design induced pilot error, the HEIST technique performed the worst (Salmon et al 2003).

Tools needed
To conduct a HEIST analysis, pen and paper is required. The analyst would also require functional diagrams of the system/interface under analysis and the eight HEIST tables containing the error identifier prompt questions.

Bibliography
Kirwan, B. (1994) A guide to Practical Human Reliability Assessment, Taylor and Francis, London.
Kirwan, B. (1998) "Human error identification techniques for risk assessment of high risk systems – Part 2: Towards a framework approach" *Applied Ergonomics,* **5**, 299 – 319
Karwowski, W. (1998) The Occupational Ergonomics Handbook, CRC Press, New York

Flowchart


START
Analyse task using HTA
Take the first/next bottom level task step from the HTA
Classify the task step into one of the SRK model categories
Select the appropriate HEIST table
Take the first/next error identifier prompt from the HEIST table
Are there any credible errors?
N
Y
For each credible error, select and record the:
Error code
EEM
Error description
PEM/System cause
Error consequence
Error reduction guidelines
Are there any more task steps?
Y
N
STOP

The Human Error and Recovery Assessment (HERA) framework

Barry Kirwan, EUROCONTROL, Experimental Centre, BP15, F91222, Bretigny Sur Orge, France

Background and applications

The HERA framework is a prototype multiple method or 'toolkit' approach to human error identification that was developed by Kirwan (1998a, 1998b) in response to a review of HEI methods, which suggested that no single HEI/HRA technique possessed all of the relevant components required for efficient HRA/HEI analysis. In conclusion to a review of thirty eight existing HRA/HEI techniques (Kirwan, 1998a), Kirwan (1998b) suggested that the best approach would be for practitioners to utilise a framework type approach to HEI, whereby a mixture of independent HRA/HEI tools would be used under one framework. Kirwan (1998b) suggested that one possible framework would be to use SHERPA, HAZOP, EOCA, Confusion matrix analyses, Fault symptom matrix analysis and the SRK approach together. In response to this conclusion, Kirwan (1998b) proposed the Human Error and Recovery Assessment (HERA) system, which was developed for the UK nuclear power and reprocessing industry. Whilst the technique has yet to be applied to a concrete system, it is offered in this review as a representation of the form that a HEI 'toolkit' approach may take.

Domain of application

Nuclear power and chemical process industries.

Procedure and advice

Step 1: Critical Task Identification

Before a HERA analysis is undertaken, the HERA team should determine how in-depth an analysis is required and also which tasks are to be analysed. Kirwan (1998b) suggests that the following factors should be taken into account:

- The nature of the plant being assessed and the cost of failure – Hazard potential of the plant under analysis.
- The operator's role – Criticality of the operator's role.
- The novelty of plant design –How new the plant is and how novel the control system used is.

A new plant that is classed as highly hazardous, with critical operator roles should require a very exhaustive HERA analysis. Alternatively, an older plant with no accident record and operator's with minor roles should require a scaled down, less exhaustive analysis. Furthermore, Kirwan (1998b) suggests that the HERA team should also consider the following logistical factors:

- System life cycle
- The extent to which the analysis is PSA driven
- Available resources

Once the depth of the analysis is decided upon, the HERA assessment team must then determine which operational stages to analyse e.g. normal operation, abnormal operation and emergency operation.

Step 2: Task Analysis

The next stage of the HERA analysis is to perform a task analysis for the scenarios chosen for analysis in question. Kirwan (1998b) recommends that two modules of task analysis are used in the HERA process. These are Initial Task Analysis (Kirwan,

1994) and HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992). Initial task analysis involves describing the scenario under analysis, including the following key aspects:

- Scenario starting condition
- The goal of the task
- Number and type of tasks involved
- Time available
- Personnel available
- Any adverse conditions
- Availability of equipment
- Availability of written procedures
- Training
- Frequency and severity of the event

Once the initial task analysis is completed, a HTA for the scenario under analysis should be completed.

Step 3: Error Analysis

The error analysis part of the HERA framework is made up of nine overlapping error identification modules.

a) Mission analysis – firstly, the HERA team must look at the scenario as a whole and determine whether there is scope for failure. The questions asked in the mission analysis are shown below.

- Could the task fail to be achieved in time?
- Could the task be omitted entirely?
- Could the wrong task be carried out?
- Could only part of the task be carried out unsuccessfully?
- Could the task be prevented or hampered by a latent or coincident failure?

The answer to at least one of these answers has to be yes for a HERA analysis to proceed.

b) Operations level analysis – The HERA team must identify the mode of failure.

c) Goals analysis - Goals analysis involves focussing on the goals identified in the HTA and determining if any goal related errors can occur. To do this, the HERA team use twelve goal analysis questions designed to highlight any potential 'goal errors'. An example of a goals analysis question used in HERA is, 'Could the operators have no goal, e.g. due to a flood of conflicting information; the sudden onset of an unanticipated situation; a rapidly evolving and worsening situation; or due to a disagreement or other decision making failure to develop a goal. The goal error taxonomy used in the HERA analysis is shown below.

- No goal
- Wrong goal
- Outside procedures
- Goal conflict
- Goal delayed
- Too many goals
- Goal inadequate

d) Plans analysis – Similar to the goals analysis, plans analysis involves focussing on the plans identified in the HTA to determine whether any plan related errors could

occur. The HERA team uses twelve plans analysis questions to identify any potential 'plan errors'. HERA plans analysis questions include, 'Could the operators fail to derive a plan, due to workload, or decision making failure', or, 'Could the plan not be understood or communicated to all parties'. The plan error taxonomy used in the HERA analysis is shown below.

- No plan
- Wrong plan
- Incomplete plan
- Plan communication failure
- Plan co-ordination failure
- Plan initiation failure
- Plan execution failure
- Plan sequence error
- Inadequate plan
- Plan termination failure

e) Error analysis – the HERA team uses an EEM taxonomy derived from SHERPA (Embrey, 1986) and THERP (Swain and Guttman, 1983) in order to identify potential errors. The EEM's are reviewed at each bottom level step in the HTA to identify any potential errors. This is based upon the subjective judgement of the HERA team.

The EEM taxonomy used in the HERA analysis is shown below.

Omission
- Omits entire task step
- Omits step in the task

Timing
- Action too late
- Action too early
- Accidental timing with other event
- Action too short
- Action too long

Sequence
- Action in the wrong sequence
- Action repeated
- Latent error prevents execution

Quality
- Action too much
- Action too little
- Action in the wrong direction
- Misalignment error
- Other quality or precision error

Selection error
- Right action on wrong object
- Wrong action on right object
- Wrong action on wrong object
- Substitution error

Information transmission error
- Information not communicated
- Wrong information communicated

Rule Violation

Other

Figure 12 – HERA EEM taxonomy

f) PSF based analysis – Explicit questions regarding environmental influences on performance are then applied to the task steps. This allows the HERA team to identify any errors caused by situational or environmental factors. There are seven PSF categories used in the HERA technique. These are time, interface, training and experience, procedures, organisation, stress and complexity. Each PSF question also has an associated EEM. An example of a HERA PSF question from each category is given below.

Time
- Is there more than enough time available? (**Too Late**)

Interface
- Is onset of the scenario clearly alarmed or cued, and is this alarm or cue compelling? (**Omission or detection failure**)

Training and experience
- Have operators been trained to deal with this task in the past twelve months? (**Omission, too late, too early**)

Procedures
- Are procedures required? (**Rule violation, wrong sequence, omission, quality error**)

Organisation
- Are there sufficient personnel to carry out the task and to check for errors? (**Action too late, wrong sequence, omission, error of quality**)

Stress
- Will the task be stressful, and are there significant consequences of task failure (**omission, error of quality, rule violation**)

Complexity
- Is the task complex or novel (**omission, substitution error, other**)

g) PEM based analysis

The analyst applies fourteen PEM questions in order to identify further errors. Similar to the PSF analysis, each PEM question has associated EEM's.

h) HEIST analysis

The HERA team should then perform a HEIST analysis for the task/system under analysis. HEIST is used to act as a 'back-up' check to ensure no potential errors are missed and to ensure comprehensiveness. HEIST is also used in order to provide error reduction guidelines.

i) Human Error HAZOP.

Finally, to ensure maximum comprehensiveness, a human error HAZOP style analysis should be performed.

Advantages
- The multi-method HERA framework ensures that it is highly exhaustive and comprehensive.
- Each of the questions surrounding the goals, PEM's , Plans and PSF analysis provide the HERA team with associated EEM's. This removes the problem of selecting the wrong error mode.
- The framework approach offers the analyst more than one chance to identify errors. This should ensure that no potential errors are missed.
- The HERA framework allows analysis teams to see the scenario from a number of different perspectives.
- HERA uses existing, proven HEI techniques, such as the human error HAZOP, THERP and SHEPRA techniques.

Disadvantages
- Such a framework approach would require a huge amount of time and resources to conduct an analysis.
- The technique could become very repetitive, with many errors being identified over and over again.

- Domain expertise would be required for a number of the modules.
- A HERA team would have to be constructed. Such a team requires a mixed group made up of operators, human factors specialists, designers, engineers etc. Building such a team and making sure they can all be brought together at the same time would be a difficult thing to do.
- Although the HERA technique is vast and contains a number of different modules, it is difficult to see how such an approach (using traditional EEM taxonomies) would perform better than far simpler and quicker approaches to HEI such as SHERPA and HET.
- The HERA framework seems too large and overcomplicated for what it actually offers.
- Due to the multitude of different techniques used, the training time for such an approach would be considerably high.

Example

HERA has yet to be applied. The following examples are extracts of a hypothetical analysis described by Kirwan (1992b). As the output is so large, only a small extract is shown below. For a more comprehensive example, the reader is referred to Kirwan (1992b).

Table 16. Extract of Mission analysis output (Source: Kirwan 1992b)

Identifier	Task step	Error identified	Consequence	Recovery	Comments
1. Fail to achieve in time	Goal 0: Restore power and cooling	Fail to achieve in time	Reactor core degradation	Grid re-connection	This is at the highest level of task-based failure description
2. Omit entire task	Goal 0: Restore power and cooling Goal A: Ensure reactor trip	Fail to restore power and cooling	Reactor core degradation Reactor core melt (ATWS)	Grid re-connection None	This is the anticipated transient without SCRAM (ATWS) scenario. It is not considered here but may be considered in another part of the risk assessment

Related methods
Any HERA analysis requires an initial task analysis and a HTA to be performed for the scenario and system under analysis. The HERA framework also uses the HEIST and Human Error HAZOP techniques as back up checks.

Approximate training times and application times
Although no training and application time is offered in the literature, it is apparent that the amount of time in both cases would be high, especially as analysts would have to be trained in all of the techniques within the HERA framework, such as initial task analysis, human error HAZOP, and HEIST.

Reliability and Validity
No data regarding reliability and validity are offered by the authors. The technique was proposed as an example of the form that such a technique would take. At the present time, the technique is yet to be applied.

Tools needed
The HERA technique comes in the form of a software package, although HERA analysis can be performed without using the software. This would require pen and paper and the goals, plans, PEM and PSF analysis questions. Functional diagrams for the system under analysis would also be required as a minimum.

Bibliography
Kirwan, B. (1996) Human Error Recovery and Assessment (HERA) Guide. Project IMC/GNSR/HF/5011, Industrial Ergonomics Group, School of Manufacturing Engineering, University of Birmingham, March.
Kirwan, B. (1998) Human error identification techniques for risk assessment of high-risk systems – Part 1: Review and evaluation of techniques. *Applied Ergonomics*, 29, pp157-177
Kirwan, B. (1998b) Human error identification techniques for risk assessment of high-risk systems – Part 2: Towards a framework approach. *Applied Ergonomics*, 5, pp299 – 319

SPEAR – System for Predictive Error Analysis and Reduction
Center for Chemical Process Safety (CCPS)

Background and applications
The System for Predictive Error Analysis (SPEAR) was developed by the Centre for Chemical Process Safety for use in the American chemical processing industry's HRA programme. SPEAR is a systematic approach to HEI that is very similar to other systematic HEI techniques, such as SHERPA. The main difference between SPEAR and SHERPA is that the SPEAR technique utilises performance-shaping factors (PSF) in order to identify any environmental or situational factors that may enhance the possibility of error. The SPEAR technique itself operates on the bottom level tasks (operations) of a HTA of the task under analysis. . Using subjective judgement, the analyst uses the SPEAR human error taxonomy to classify each task step into one of the five following behaviour types:

- **Action**
- **Retrieval**
- **Check**
- **Selection**
- **Transmission**

Each behaviour has an associated set of EEM's, such as *action incomplete, action omitted and right action on wrong object*. The analyst then uses the taxonomy and domain expertise to determine any credible error modes for the task in question. For each credible error (i.e. those judged by the analyst to be possible) the analyst should give a description of the form that the error would take, such as, 'pilot dials in wrong airspeed'. Next, the analyst has to determine how the operator can recover the error and also any consequences associated with the error. Finally, error reduction measures are proposed, under the categories of procedures, training and equipment.

Domain of application
Chemical process industries.

Procedure and advice
Step 1: Hierarchical Task Analysis (HTA)
The process begins with the analysis of work activities, using Hierarchical Task Analysis. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Step 2: PSF analysis
The analyst should take the first/next bottom level task step from the HTA and consider each of the PSF's for that task step. This allows the analyst to determine

whether the PSF's increase the possibility of error at any of the task steps. The PSF's used in the SPEAR technique can be found in Swain and Guttman (1983).

Step 3: Task Classification
Next, the analyst should classify the task step under analysis into one of the behaviour categories in the EEM taxonomy. Which EEM taxonomy is used is decided by the analyst but in this case the taxonomy shown in figure XX will be used. The analyst has to classify the task step into one of the behaviour categories; Action, Checking, Retrieval, Transmission, Selection and Plan.

Step 4: Error analysis
Taking the PSF's from step 2 into consideration, the analyst next considers each of the associated EEM's for the task step under analysis. Based upon the analyst's subjective judgement, any credible errors should be recorded and a description of the error should be noted.

Step 5: Consequence analysis
For each credible error, the analyst should record the associated consequence.

Step 6: Error reduction analysis
For each credible error, the analyst should offer any potential error remedies. The SPEAR technique uses three categories of error reduction guideline; Procedures, Training and Equipment. It is normally expected that a SPEAR analysis should offer one remedy for each of the three categories.

Advantages
- SPEAR provides a structured approach to HEI.
- Simple to learn and use.
- Unlike SHERPA, SPEAR also considers PSF's.
- Quicker than most HEI techniques.

Disadvantages
- HTA provides additional work for the analyst.
- Consistency of such techniques is questionable.
- Appears to be an almost exact replica of SHERPA.
- For large, complex tasks the analysis may become time consuming and unwieldy.

Related methods
Any SPEAR analysis requires an initial HTA to be performed for the task under analysis.

Approximate training times and application times
Since the technique is similar to the SHERPA technique, the training and application times specified would be the same.

Reliability and validity
No data regarding the reliability and validity of the SPEAR technique are available in the literature.

Tools needed

To conduct a SPEAR analysis, pen and paper is required. The analyst would also require functional diagrams of the system/interface under analysis and an appropriate EEM taxonomy, such as the SHERPA (Embrey, 1986) error mode taxonomy. A PSF taxonomy is also required, such as the one used in the THERP technique (Swain and Guttman, 1983)

Bibliography

Karwowski, W. (1998) The Occupational Ergonomics Handbook, CRC Press, New York.
Centre for Chemical Process Safety (CCPS). (1994) Guidelines for Preventing Human Error in Process Safety. New York: American Institute of Chemical Engineers.

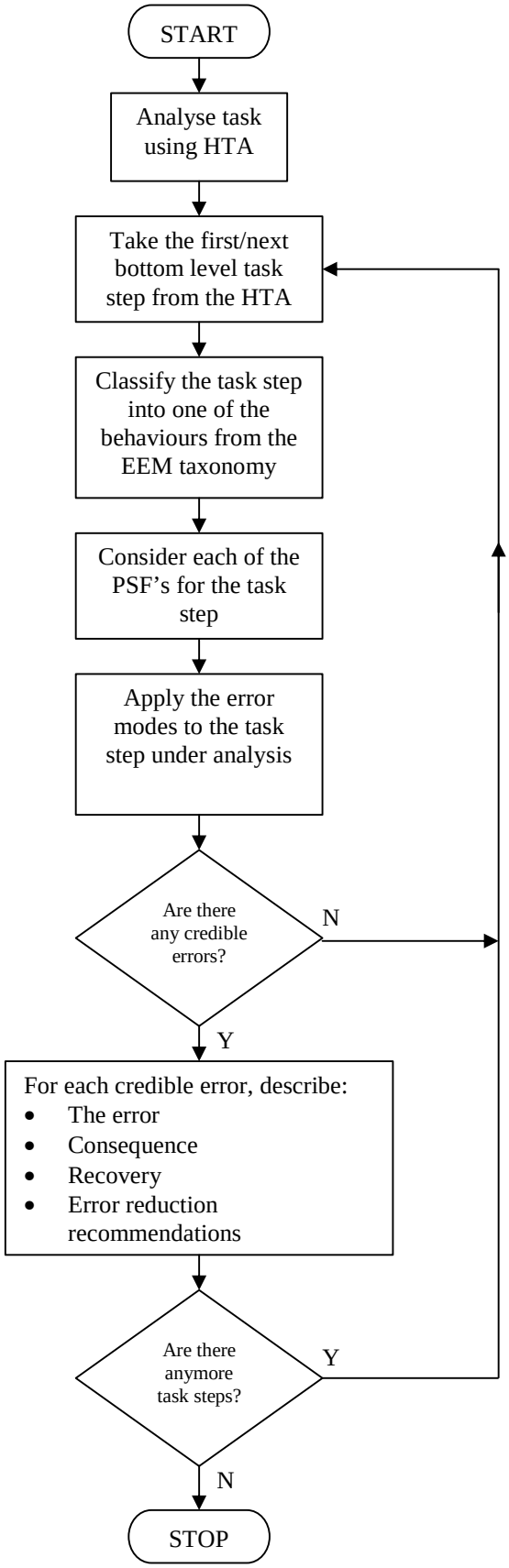
Example

The following example is an extract from a SPEAR analysis of a chlorine tanker-filling problem (CCPS, 1994 cited in Karwowski 1999).

Table 17. Example SPEAR output

Step	Error Type	Error Description	Recovery	Consequences	Error reduction recommendations		
					Procedures	Training	Equipment
2.3 Enter tanker target weight	Wrong information obtained (R2)	Wrong weight entered	On check	Alarm does not sound before tanker overfills	Independent validation of target weight	Ensure operator double checks entered date. Recording of values in checklist.	Automatic setting of weight alarms from unladen weight. Computerise logging system and build in checks on tanker reg. No. and unladen weight linked to warning system. Display differences
3.2.2 Check tanker while filling	Check omitted (C1)	Tanker not monitored while filling	On initial weight alarm	Alarm will alert the operator if correctly set. Equipment fault e.g. leaks not detected early and remedial action delayed	Provide secondary task involving other personnel. Supervisor periodically checks operation	Stress importance of regular checks for safety	Provide automatic log in procedure
3.2.3 Attend tanker during last 2-3 ton filling	Operation omitted (O8)	Operator fails to attend	On step 3.2.5	If alarm not detected within 10 minutes tanker will overfill	Ensure work schedule allows operator to do this without pressure	Illustrate consequences of not attending	Repeat alarm in secondary area. Automatic interlock to terminate loading if alarm not acknowledged. Visual indication of alarm.
3.2.5 Cancel final weight alarm	Operation omitted (O8)	Final weight alarm taken as initial weight alarm	No recovery	Tanker overfills	Note differences between the sound of the two alarms in checklist	Alert operators during training about differences in sounds of alarms	Use completely different tones for initial and final weight alarms
4.1.3 Close tanker valve	Operation omitted (O8)	Tanker valve not closed	4.2.1	Failure to close tanker valve would result in pressure not being detected during the pressure check in 4.2.1	Independent check on action. Use checklist	Ensure operator is aware of consequences of failure	Valve position indicator would reduce probability of error
4.2.1 Vent and purge lines	Operation omitted (O8)	Lines not fully purged	4.2.4	Failure of operator to detect pressure in lines could lead to leak when tanker connections broken	Procedure to indicate how to check if fully purged	Ensure training covers symptoms of pressure in line	Line pressure indicators at controls. Interlock device on line pressure
4.4.2 Secure locking nuts	Operation omitted (O8)	Locking nuts left unsecured	None	Failure to secure locking nuts could result in leakage during transportation	Use checklist	Stress safety implications of training	Locking nuts to give tactile feedback when secure

Flowchart


START
Analyse task using HTA
Take the first/next bottom level task step from the HTA
Classify the task step into one of the behaviours from the EEM taxonomy
Consider each of the PSF's for the task step
Apply the error modes to the task step under analysis
Are there any credible errors?
N
Y
For each credible error, describe:
The error
Consequence
Recovery
Error reduction recommendations
Are there anymore task steps?
Y
N
STOP

HEART - Human Error Assessment and Reduction Technique

J. C. Williams

Background and applications

HEART or the Human Error Assessment and Reduction technique (Williams, 1986) was designed primarily as a quick, simple to use and easily understood HEI technique. HEART is a highly procedural technique which attempts to quantify human error. The most significant aspect of the HEART technique is the fact that it aims only to deal with those errors that will have a gross effect on the system in question, in order to reduce the resource usage when applying the technique (Kirwan 1994). The method uses its own values of reliability and also 'factors of effect' for a number of error producing conditions (EPC). The HEART methodology has mainly been used in nuclear power plant assessments. The technique has been used in the UK for the Sizewell B risk assessment and also the risk assessments for UK Magnox and Advanced Gas-Cooled Reactor stations.

Domain of application

Nuclear power and chemical process industries.

Procedure and advice

Step 1: Hierarchical Task Analysis (HTA)

The process begins with the analysis of work activities, using techniques such as Hierarchical Task Analysis or tabular task analysis. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Step 2: The HEART screening process

The HEART technique uses a screening process, in the form of a set of guidelines that allow the analyst to identify the likely classes, sources and strengths of human error for the scenario under analysis (Kirwan, 1994).

Step 3: Task unreliability classification

The analyst must define the task under analysis in terms of its proposed nominal level of human unreliability. To do this, the analyst uses the HEART generic categories to classify the task to allow a human error probability to be assigned to it. For example, if the analysis was focussed upon an emergency situation on the flight deck, such as the one seen in the Sou city disaster, then the HEART analyst would classify this as A) Totally unfamiliar, performed at speed with no real idea of likely consequences. The probability associated with this would be 0.55. The HEART generic categories are shown in table 14.

Step 4: Identification of Error-Producing conditions
The next stage of the HEART is to identify any error producing conditions (EPC's) that would be applicable to the scenario/task under analysis. Again like the HEART generic categories, these EPC's have a critical effect on the HEP's produced. Table 19 shows the Error producing categories used in the HEART methodology.

Table 18 – HEART generic categories

Generic Task	**Proposed nominal human unreliability (5^{th} – 95^{th} percentile bounds)**
(A) Totally unfamiliar, performed at speed with no real idea of the likely consequences	0.55 (0.35 – 0.97)
(B) Shift or restore system to a new or original state on a single attempt without supervision or procedures	0.26 (0.14 – 0.42)
(C) Fairly simple task performed rapidly or given scant Attention	0.16 (0.12 – 0.28)
(D) Routine, highly practised, rapid task involving relatively Low level of skill	0.09 (0.06 – 0.13)
(E) Restore or shift a system to original or new state following procedures, with some checking	0.02 (0.007 – 0.045)
(F) Completely familiar, well designed, highly practised, routine task occurring several times per hour, performed at the highest possible standards by highly motivated, highly trained and experienced person, totally aware of the implications of failure, with time to correct potential error, but without the benefit of significant job aids	0.003 (0.0008 – 0.0009)
(G) Respond correctly to system command even when there is an augmented or automated supervisory system providing accurate interpretation of system stage	0.0004 (0.00008 – 0.009)
(H) Respond correctly to system command even when there is an augmented or automated supervisory system providing accurate interpretation of system stage	0.00002 (0.000006 – 0.009)

Step 5: Assessed proportion of effect
Once the analyst has identified any EPC's the next stage is to determine the assessed proportion of effect of each of the selected EPC's. This is a rating between 0 and 1 (0 = Low, 1 = High) and is based upon the analyst's subjective judgement.

Step 6: Remedial measures
Next the analyst has to determine whether there are any possible remedial measures that can be taken in order to reduce or stop the incidence of the identified error. Although the HEART technique does provide and some generic remedial measures, the analyst may be required to provide his own measures depending upon the nature of the error and the system under analysis. The remedial measures provided by the HEART methodology are generic and not system specific.

Step 7: Documentation stage
Throughout the HEART analysis, every detail should be recorded by the analyst. Once the analysis is complete, the HEART analysis should be converted into a presentable format.

Table 19. HEART EPC's (source – Kirwan, 1994)

Error producing condition (EPC)	**Maximum predicted Amount by which unreliability might change, going from good conditions to bad**
Unfamiliarity with a situation which is potentially important but which only occurs infrequently, or which is novel	**X17**
A shortage of time available for error detection and correction	**X11**
A low signal to noise ratio	**X10**
A means of suppressing or overriding information or features which is too easily accessible	**X9**
No means of conveying spatial and functional information to operators in a form which they can readily assimilate	**X8**
A mismatch between an operators model of the world and that imagined by a designer	**X8**
No obvious means of reversing an unintended action	**X8**
A channel capacity overload, particularly one caused by simultaneous presentation of non redundant information	**X6**
A need to unlearn a technique and apply one which requires the application of an opposing philosophy	**X6**
The need to transfer specific knowledge from task to task without loss	**X5.5**
Ambiguity in the required performance standards	**X5**
A mismatch between perceived and real risk	**X4**
Poor, ambiguous or ill-matched system feedback	**X4**
No clear, direct and timely confirmation of an intended action from the portion of the system over which control is exerted	**X4**
Operator inexperience	**X3**
An impoverished quality of information conveyed procedures and person-person interaction	**X3**
Little or no independent checking or testing of output	**X3**
A conflict between immediate and long term objectives	**X2.5**
No diversity of information input for veracity checks	**X2**
A mismatch between the educational achievement level of an individual and the requirements of the task	**X2**
An incentive to use other more dangerous procedures	**X2**
Little opportunity to exercise mind and body outside the immediate confines of the job	**X1.8**
Unreliable instrumentation	**X1.6**
A need for absolute judgements which are beyond the capabilities or experience of an operator	**X1.6**
Unclear allocation of function and responsibility	**X1.6**
No obvious way to keep track or progress during an activity	**X1.4**

Example

Table 16 shows an example of a HEART assessment output. (Source: Kirwan 1994)

Table 20 – HEART output

Type of Task - F		**Nominal Human Reliability – 0.003**	
Error Producing conditions	**Total HEART effect**	**Engineers POA**	**Assessed effect**
Inexperience	X3	0.4	((3 –1) x 0.4) + 1 = 1.8
Opp Technique	X6	1.0	((6 – 1) x 1.0) + 1 = 6.0
Risk Misperception	X4	0.8	((4 –1) x 0.8 + 1 = 3.4
Conflict of objectives	X2.5	0.8	((2.5 – 1) x 0.8) + 1 =2.2
Low Morale	X1.2	0.6	((1.2 – 1) x 0.6 + 1 = 1.12

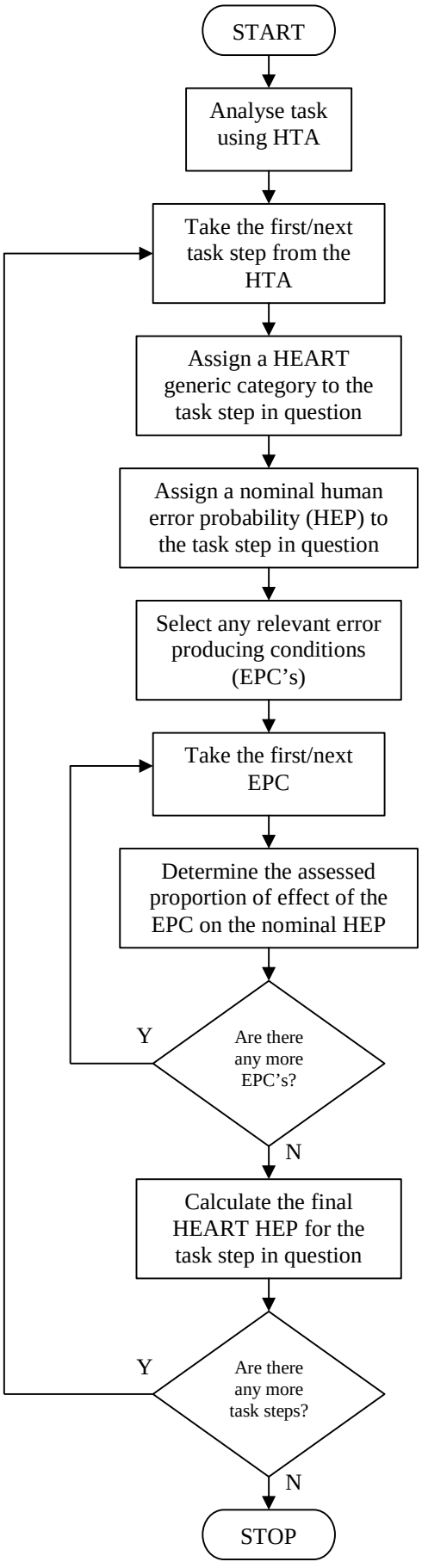
Assessed, nominal likelihood of failure;
= 0.003 x 1.8 x 6 x 3.4 x 2.2 x 1.12 = **0.27**

Thus, a HEP of 0.27 is calculated (just over 1 in 4). According to Kirwan (1994) this is a high predicted error probability and would warrant error reduction measures. In this instance, technique unlearning is the biggest contributory factor and so if error reduction were required, retraining or redesigning could be offered. Table 17 shows the remedial measures offered for each EPC in this example.

Table 18– Remedial measures (Source – Kirwan 1994)

Technique unlearning (x6)	The greatest possible care should be exercised when a number of new techniques are being considered that all set out to achieve the same outcome. They should not involve that adoption of opposing philosophies
Misperception of risk (x4)	It must not be assumed that the perceived level of risk, on the part of the user, is the same as the actual level. If necessary, a check should be made to ascertain where any mismatch might exist, and what its extent is
Objectives conflict (x2.5)	Objectives should be tested by management for mutual compatibility, and where potential conflicts are identified, these should either be resolved, so as to make them harmonious, or made prominent so that a comprehensive management-control programme can be created to reconcile such conflicts, as they arise, in a rational fashion
Inexperience (x3)	Personnel criteria should contain experience parameters specified in a way relevant to the task. Chances must not be taken for the sake of expediency
Low morale (x1.2)	Apart from the more obvious ways of attempting to secure high morale – by way of financial rewards, for example – other methods, involving participation, trust and mutual respect, often hold out at least as much promise. Building up morale is a painstaking process, which involves a little luck and great sensitivity

Flowchart


START
Analyse task using HTA
Take the first/next task step from the HTA
Assign a HEART generic category to the task step in question
Assign a nominal human error probability (HEP) to the task step in question
Select any relevant error producing conditions (EPC's)
Take the first/next EPC
Determine the assessed proportion of effect of the EPC on the nominal HEP
Are there any more EPC's?
Y
N
Calculate the final HEART HEP for the task step in question
Are there any more task steps?
Y
N
STOP

Advantages

- HEART appears to be quick and simple to use, involving little training.
- Each error-producing condition has a remedial measure associated with it.
- HEART gives the analyst a quantitative output.
- HEART uses fewer resources than other techniques such as SHERPA.
- Evidence of validity – Kirwan (1988, 1996, 1997), Waters (1989), Robinson (1981)

Disadvantages

- Doubts over the consistency of the technique remain. There is little in structure and in the task classification and assignment of error producing categories stages the analyst has no guidance. The result is that different analysts often use the technique differently. For example, for the assessed proportion of effect part of the HEART technique, Kirwan (1994) suggests that there is little published guidance available and that different analysts vary considerably in their approach.
- Although it has been involved in a number of validation studies, the HEART methodology still requires further validation.
- Neither dependence or EPC interaction is accounted for by HEART (Kirwan, 1994)
- HEART does not provide enough guidance for the analyst on a number of key aspects, such as task classification and also in determining the assessed proportion of effect.
- HEART is very subjective, reducing its reliability and consistency.
- The technique would require considerable development to be used in other domains, such as military operations.

Related Methods

Normally, a HEART analysis requires a task analysis description of the task or scenario under analysis. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is normally used. The HEART technique is a HRA technique, of which there are many, such as THERP (Swain & Guttman 1983) and JHEDI (Kirwan 1994).

Approximate training and application times

According to Kirwan (1994) the HEART technique is both quick to train and apply. The technique is certainly simple in its application and so the associated training and application time should be minimal.

Reliability and validity

Kirwan (1997) describes a validation of nine HRA techniques and reports that, of the nine techniques, HEART, THERP, APJ and JHEDI performed moderately well. A moderate level of validity for HEART was reported. In a second validation study (Kirwan 1997), HEART, THERP and JHEDI were validated. The highest precision rating associated with the HEART technique was 76.67%. Of 30 assessors, 23 displayed a significant correlation between their error estimates and the real HEP's. According to Kirwan (1997) the results demonstrate a level of empirical validity of the three techniques.

Tools needed

The HEART technique is a pen and paper tool. The associated HEART documentation is also required (HEART generic categories, HEART error producing conditions etc).

Bibliography

Kirwan, B. (1994) A guide to Practical Human Reliability Assessment. Taylor and Francis, London.

Kirwan, B. (1996) "The validation of three Human Reliability Quantification techniques – THERP, HEART and JHEDI: Part 1 – technique descriptions and validation issues" *Applied Ergonomics,* Vol **27**, 6, pp 359 – 373

Kirwan, B. (1997) "The validation of three Human Reliability Quantification techniques – THERP, HEART and JHEDI: Part 2 – Results of validation exercise" *Applied Ergonomics,* Vol **28**, 1, pp 17 – 25

Kirwan, B. (1997) "The validation of three Human Reliability Quantification techniques – THERP, HEART and JHEDI: Part 3 – Practical aspects of the usage of the techniques" *Applied Ergonomics,* Vol **28**, 1, pp 27 – 39

Williams, J. C. (1986) HEART – a proposed method for assessing and reducing human error. In 9th Advances in Reliability Technology Symposium, University of Bradford.

CREAM – The Cognitive Reliability and Error Analysis Method

Erik Hollnagel, Department of Computer and Information Science, University of Linkoping, LIU/IDA, S-581 83 Linkoping. erih@ida.liu.se

Background and applications

The Cognitive Reliability and Error Analysis Method (CREAM) (Hollnagel 1998) is a recently developed HEI/HRA method that was developed by the author in response to an analysis of existing HRA approaches. CREAM can be used both predictively, to predict potential human error, and retrospectively, to analyse and quantify error. CREAM. The CREAM technique consists of a method, a classification scheme and a model. According to Hollnagel (1998) CREAM enables the analyst to achieve the following:

1) Identify those parts of the work, tasks or actions that require or depend upon human cognition, and which therefore may be affected by variations in cognitive reliability.
2) Determine the conditions under which the reliability of cognition may be reduced, and where therefore the actions may constitute a source of risk.
3) Provide an appraisal of the consequences of human performance on system safety, which can be used in PRA/PSA.
4) Develop and specify modifications that improve these conditions, hence serve to increase the reliability of cognition and reduce the risk.

CREAM uses a model of cognition, the Contextual Control Model (COCOM). COCOM focuses on how actions are chosen and assumes that the degree of control that an operator has over his actions is variable and also that the degree of control an operator holds determines the reliability of his performance. The COCOM outlines four modes of control, *Scrambled control, Opportunistic control, Tactical control and Strategic control*. According to Hollnagel (1998) when the level of operator control rises, so does their performance reliability.

The CREAM technique uses a classification scheme consisting of a number of groups that describe the phenotypes (error modes) and genotypes (causes) of the erroneous actions. The CREAM classification scheme is used by the analyst to predict and describe how errors could potentially occur. The CREAM classification scheme allows the analyst to define the links between the causes and consequences of the error under analysis. Within the CREAM classification scheme there are three categories of causes (genotypes); *Individual, technological and organisational* causes. These genotype categories are then further expanded as follows:

1) Individual related genotypes – Specific cognitive functions, general person related functions (temporary) and general person related functions (permanent).
2) Technology related genotypes – Equipment, procedures, interface (temporary) and interface (permanent).
3) Organisation related genotypes – communication, organisation, training, ambient conditions, working conditions.

The CREAM technique uses a number of linked classification groups. The first classification group describes the CREAM error modes. Hollnagel (1998) suggests that the error modes denote the particular form in which an erroneous action can appear. The error modes used in the CREAM classification scheme are:

1) *Timing* – too early, too late, omission.
2) *Duration* – too long, too short.
3) *Sequence* – reversal, repetition, commission, intrusion.
4) *Object* – wrong action, wrong object.
5) *Force* – too much, too little.
6) *Direction* – Wrong direction.
7) *Distance* – too short, too far.
8) *Speed* – too fast, too slow.

These eight different error mode classification groups are then divided into the four sub-groups:

1) *Action at the wrong time* – includes the error modes timing and duration.
2) *Action of the wrong type* – includes the error modes force, distance, speed and direction.
3) *Action at the wrong object* – includes the error mode 'object'.
4) *Action in the wrong place* – includes the error mode 'sequence'.

The CREAM classification system is comprised of both phenotypes (error modes) and genotypes (causes of error). These phenotypes and genotypes are further divided into detailed classification groups, which are described in terms of general and specific consequents. The CREAM technique also uses a set of common performance conditions (CPC) that are used by the analyst to describe the context in the scenario/task under analysis. These are similar to PSF's used by other HEI/HRA techniques. Table 20 shows the CREAM common performance conditions.

Table 20 – Cream Common Performance Conditions

CPC Name	Level/Descriptors
Adequacy of organisation	The quality of the roles and responsibilities of team members, additional support, communication systems, safety management system, instructions and guidelines for externally orientated activities etc. **Very efficient/Efficient/Inefficient/Deficient**
Working Conditions	The nature of the physical working conditions such as ambient lighting, glare on screens, noise from alarms, task interruptions etc **Advantageous/Compatible/Incompatible**
Adequacy of MMI and operational support	The man machine interface in general, including the information available on control panels, computerised workstations, and operational support provided by specifically designed decision aids. **Supportive/Adequate/Tolerable/Inappropriate**
Availability of procedures/plans	Procedures and plans include operating and emergency procedures, familiar patterns of response heuristics, routines etc **Appropriate/Acceptable/Inappropriate**
Number of simultaneous goals	The number of tasks a person is required to pursue or attend to at the same time. **Fewer than capacity/Matching current capacity/More than capacity**
Available time	The time available to carry out the task **Adequate/Temporarily inadequate/Continuously inadequate**
Time of day (Circadian rhythm)	Time at which the task is carried out, inparticular whether or not the person is adjusted to the current time. **Day-time (adjusted)/Night time (unadjusted)**
Adequacy of training and experience	Level and quality of training provided to operators as familiarisation to new technology, refreshing old skills etc. Also refers to operational experience. **Adequate, high experience/Adequate, limited experience/Inadequate**
Crew collaboration quality	The quality of collaboration between the crew members, including the overlap between the official and unofficial structure, level of trust, and the general social climate among crew members. **Very efficient/Efficient/Inefficient/Deficient**

Domain of application
Although the technique was developed for the nuclear power industry, the author claims that it is a generic technique that can be applied in a number of domains involving the operation of complex, dynamic systems.

Procedure and advice (Prospective analysis)
Method/Procedure
Step 1: Task analysis
It is first important to analyse the situation or task. Hollnagel (1998) suggests that this should take the form of a HTA. It is also recommended that the analyst here should include considerations of the organisation and technical system, as well as looking at the operator and control tasks. If the system under analysis does not yet exist, then information from the design specifications can be used.

Step 2: Context description
The analyst should begin the analysis by firstly describing the context in which the scenario under analysis takes place. The CREAM CPC's are used to describe the scenario context.

Step 3: Specification of the initiating events
The analyst then needs to specify the initiating events that will be subject to the error predictions. Hollnagel (1998) suggests that PSA event trees can be used for this step. However, since a task analysis has already been conducted in step 1 of the procedure, it is recommended that this be used. The analyst(s) should specify the tasks or task steps that are to be subject to further analysis.

Step 4: Error Prediction
Using the CREAM, the analyst now has to describe how an initiating event could potentially into an error occurrence. To predict errors, the analyst should construct a modified consequent/antecedent matrix. The rows on the matrix show the possible consequents whilst the columns show the possible antecedents. The analyst starts by finding the classification group in the column headings that correspond to the initiating event (e.g. for missing information it would be communication). The next step is to find all the rows that have been marked for this column. Each row should point to a possible consequent, which in turn may be found amongst the possible antecedents. The author suggests that in this way, the prediction can continue in a straightforward way until there are no further paths left (Hollnagel 1998). Each error should be recorded along with the associated causes (antecedents) and consequences (consequents).

Step 5: Selection of task steps for quantification
Depending upon the analysis requirements, a quantitative analysis may be required. If so, the analyst should select the error cases that require quantification. It is recommended that if quantification is required, then all of the errors identified should be selected for quantification.

Step 6: Quantitative performance prediction
CREAM has a basic and extended method for quantification purposes. Since this review is based upon the predictive use of HEI techniques in the C4i design process,

which does not include error quantification. The reader is referred to Hollnagel (1998) for further information on the CREAM error quantification procedure.

Advantages

- CREAM has the potential to be extremely exhaustive.
- Context is considered when using CREAM.
- CREAM is a clear, structured and systematic approach to error identification/quantification.
- The same principles of the CREAM method can be used for both retrospective and predictive analyses.
- The method is not domain specific and the potential for use in different domains such as command and control is apparent.
- CREAM's classification scheme is detailed and exhaustive, even taking into account system and environmental (sociotechnical) causes of error.
- Section in Hollnagel (1998) on the links between consequents and antecedents is very useful.
- Can be used both qualitatively and quantitatively.

Disadvantages

- To the novice analyst, the method appears complicated and daunting.
- The exhaustiveness of the classification scheme serves to make the method larger and more resource intensive than other methods.
- CREAM has not been used extensively.
- It is apparent that the training and application time for the CREAM technique would be considerable.
- CREAM does not offer remedial measures i.e. ways to recover human erroneous actions are not given/considered.
- CREAM appears to be very complicated in its application.
- CREAM would presumably require analysts with knowledge of human factors and cognitive ergonomics.
- Application time would be high, even for very basic analyses.

Related methods

Hollnagel (1998) recommends that a task analysis such as HTA is carried out prior to a CREAM analysis. CREAM is a taxonomic approach to HEI. Other taxonomic approaches include SHERPA (Embrey 1986), HET (Marshall et al 2003) and TRACEr (Shorrock & Kirwan 2002).

Approximate training and application times

Although there is no data regarding training and application times in the literature, as the method appears large and quite complicated, it is predicted that the times will be high in both cases.

Reliability and Validity

Validation data for the CREAM technique is limited. Hollnagel, Kaarstad & Lee (1998) report a 68.6% match between errors predicted and actual error occurrences and outcomes when using the CREAM error taxonomy.

Tools needed

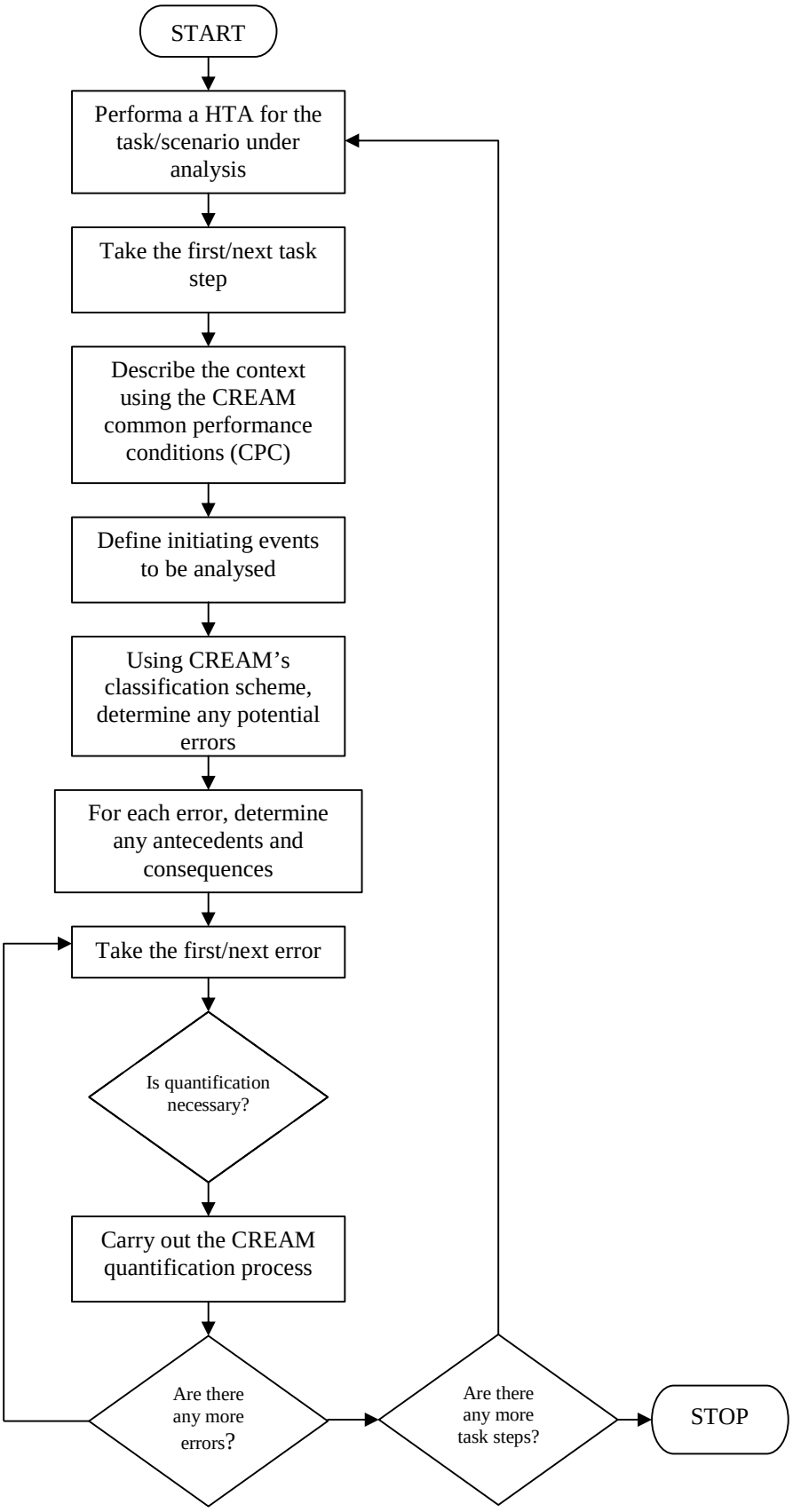
At its simplest, CREAM can be applied using simply pen and paper. A prototype software package has been developed to aid analysts (Hollnagel 1998).

Bibliography

Hollnagel, E. (1998) Cognitive Reliability and Error Analysis Method – CREAM, 1st Edition. Elsevier Science, Oxford, England

Kim, I. S. (2001) "Human reliability analysis in the man-machine interface design review" *Annals of Nuclear Energy*, **28**, 1069 - 1081

Flowchart – Prospective use



Charting Techniques

According to Kirwan & Ainsworth (1992) the first attempt to chart a work process was carried out by Gilbreth and Gilbreth in the 1920's. Since then, a number of charting and network techniques have been developed. The main aim of these techniques is to provide a graphical representation of a task, which is easier to understand than a typical text description (Kirwan & Ainsworth 1992). The charting of work processes is also a useful way of highlight essential task components and requirements. Charting techniques are used to depict graphically a task or process using standardised symbols. The output of charting techniques can be used to understand the different task steps involved a particular scenario, and also to highlight when each task step should occur and which technological aspect of the system interface is required. Charting techniques therefore represent both the human and system elements involved in the performance of a certain task or scenario (Kirwan & Ainsworth 1992). Charting techniques are particularly useful for representing team-based or distributed tasks, which are often exhibited in command and control systems. A process chart type analysis allows the specification of which tasks are conducted by which team member or technological component. A number of variations of charting techniques exist, including techniques used to represent operator decisions (DAD), and the causes of hardware and human failures (Fault tree analysis, Murphy diagrams). Typically used in the Nuclear Petro-chemical domain to understand, evaluate and represent the human and system aspects of a task, charting techniques have also been used in the analysis of operator tasks in other domains, including aviation, maritime, railway and air traffic control. Sanders & McCormick (1992) suggest that operational-sequence diagrams are developed during the design of complex systems in order develop a detailed understanding of the tasks involved in systems operation and that the process of developing the OSD may be more important than the actual outcome itself.

Process charts are probably the most simple form of charting technique, consisting of a single, vertical flow line which links up the sequence of activities that are performed in order to complete the task under analysis successfully. Operational Sequence Diagrams (OSD) are used to graphically describe the interaction between teams of operators and a system. The output of an OSD graphically depicts a task process, including the tasks performed and the interaction between operators over time, using standardised symbols. Event tree analysis is a task analysis technique that uses tree like diagrams to represent the various possible outcomes associated with operator tasks steps in a scenario. Decision Action Diagrams (DAD's) are used to depict the process of a scenario through a system in terms of the decisions required and actions to be performed by the operator in conducting the task or scenario under analysis. Fault trees are used to depict system failures and their causes. A fault tree is a tree like diagram, which defines the failure event and displays the possible causes in terms of hardware failure or human error (Kirwan & Ainsworth 1992). Murphy Diagrams (Pew et al 1981) are also used to graphically describe errors and their causes (proximal and distal).

The charting techniques reviewed in this document are shown below:

1. Process Charts
2. Operator event sequence diagrams
3. Event tree analysis
4. DAD – Decision Action Diagrams

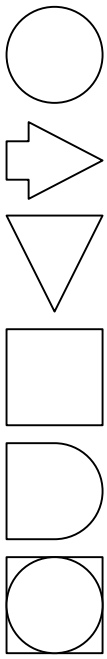
5. Fault tree analysis
6. Murphy Diagrams

It is hypothesised that a form of charting technique will be used in the design of the C4i system. More specifically, charting techniques will be used in order to represent operator and technological interaction in existing command and control situations. The resultant output will then be used to inform the design of the new C4i system, to highlight potential problems in existing command and control procedures, such as multiple task performance.

Process Charts

Background and applications

Process charts offer a systematic approach to describing tasks and provide a graphical representation of the task or scenario under analysis that is easy to follow and understand (Kirwan & Ainsworth 1992). Process charts are used to graphically represent separate steps or events that occur during the performance of a task or series of actions. Process charts were originally used to show the path of a product through its manufacturing process i.e. the construction of an automobile. Since the original use of process charts, however, there have been many variations in their use. It is suggested, for example, that process charts can be modified to refer to other entities, such as humans or information, as well as objects/products (Drury, 1990). Variations of the process chart methodology include Operation process charts, which show a chronological sequence of operations, inspections etc that are used in a process, and also the Triple resource chart, which has separate columns for the operator, the equipment used and also the material. In their simplest form, process charts consist of a single, vertical flow line which links up the sequence of activities that are performed in order to complete the task under analysis successfully. The main symbols used in a process chart reduced from 29 to 5 by the American Society of Mechanical Engineers in 1972 (Kirwan & Ainsworth 1992) and are shown below. These can be modified to make the analysis more appropriate for different applications.



Operation

Transportation

Storage

Inspection

Delay

Combined operations (e.g. inspection performed with an operation)

Once completed, a process chart analysis comes in the form of a single, top down flow line, which represents a sequence of task steps or activities. Time taken for each task step or activity can also be recorded as part of a process chart analysis.

Domain of application

Nuclear power and chemical process industries.

Procedure and advice
The symbols should be linked together in a vertical chart depicting the key stages in the task or process under analysis.

Step 1: Data collection
In order to construct a process chart, the analyst(s) must first obtain sufficient data regarding the scenario under analysis. It is recommended that the analyst(s) use various forms of data collection in this phase. Observational study should be used to observe the task (or similar types of task) under analysis. Interviews with personnel involved in the task (or similar tasks) should also be conducted. The type and amount of data collected in step 1 is dependent upon the analysis requirements. For example, if the output requires a cognitive component, techniques such as critical decision method and cognitive walkthrough can be used in step 1 to acquire the necessary data.

Step 2: Create task list
Firstly, the analyst should create a comprehensive list of the task steps involved in the scenario under analysis. These should then be put these into a chronological order. A HTA for the task or process under analysis may be useful here, as it provides the analyst with a thorough task description.

Step 3: Task step classification
Next, the analyst needs to classify each task step into one of the process chart behaviours; Operation, Transportation, Storage, Inspection, Delay or combined operation. Depending on the task under analysis, a new set of process chart symbols may need to be created. The analysts should take each task step or operation and determine, based on subjective judgement, which of the steps are operations. The analyst should then repeat this process for each of the process chart behaviour's.

Step 4: Create the process chart
Once all of the task steps/actions are sorted into operations, inspections etc, they should then be placed into the process chart. This involves linking each operation, transportation, storage, inspection, delay or combined operation in a vertical chart. Each task step should be placed in the order that they would occur when performing the task. Alongside the task steps symbol, another column should be placed, describing the task step fully.

Advantages
- Process charts are useful in that they show the logical structure of actions involved in a task.
- Process charts are simple to learn and construct.
- They have the potential to be applied to any domain.
- Process charts allow the analyst to observe how a task is undertaken.
- Process charts can also display task time information.
- Process charts can represent both operator and system tasks (Kirwan & Ainsworth, 1992).
- Process charts provide the analyst with a simple, graphical representation of the task or scenario under analysis.

Disadvantages
- For large tasks, a process chart may become large and unwieldy.

- When using process charts for complex, large tasks, chart construction will become very time consuming. Also, complex tasks require complex process charts.
- As process charts were originally developed to monitor a product being built, some of the symbols are irrelevant. An example of this would be using process charts in aviation. The symbols representing Transport and Storage would not be relevant. Modification of the symbols would have to occur for the method to be applied to domains such as aviation or command and control.
- Process charts do not take into account error, modelling only error free performance.
- Only a very limited amount of information can be represented in a process chart
- Process charts do not take into account cognitive processes.

Related methods

The process chart technique belongs to a family of charting or network techniques. Other techniques charting/networking techniques include input-output diagrams, functional flow diagrams, information flow diagrams, Murphy diagrams, critical path analysis, petri nets and signal flow graphs (Kirwan & Ainsworth 1992).

Approximate training and application times

The training time for such a technique should be low, representing the amount of time it takes for the analyst to become familiar with the process chart symbols.
Application time is dependent upon the size and complexity of the task under analysis. For small, simple tasks, the application time would be very low. For larger, more complex tasks, the application time would be high.

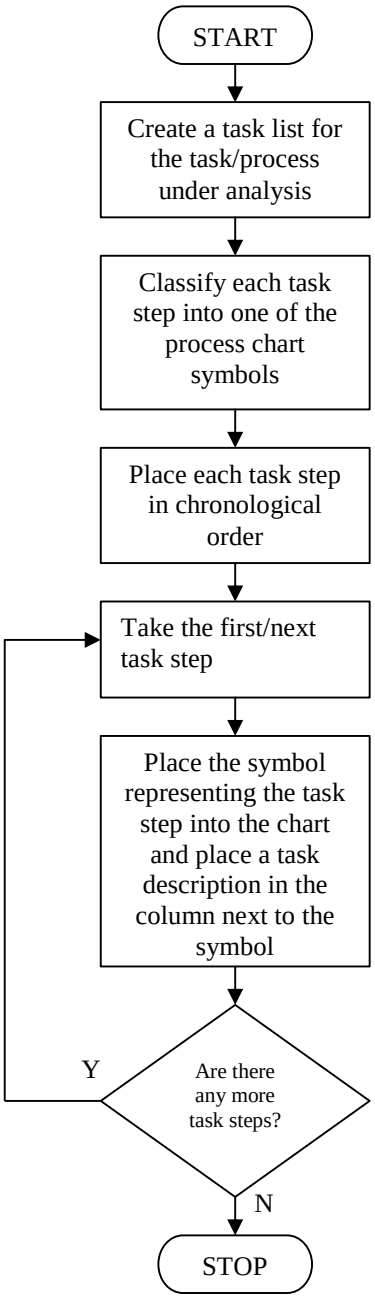
Reliability and Validity

No data regarding the reliability and validity of the technique are available in the literature.

Bibliography

Kirwan, B. & Ainsworth, L. K. (1992) A guide to Task Analysis, Taylor and Francis, London, UK.
Salvendy, G. (1997) Handbook of human factors and ergonomics, 2nd edition, John Wiley and Sons, Canada.
Drury, C. G. (1990) Methods for direct observation of performance, In Wilson, J. R. and Corlett, E. N (Eds) 'Evaluation of Human Work – A practical ergonomics methodology' Taylor and Francis, London.

Flowchart



Example

The following example is a process chart analysis for the landing task, 'land aircraft at New Orleans airport using the auto-land system' (Stanton et al 2003). A process chart analysis was conducted in order to assess the feasibility of using process chart type analysis in aviation. Process charts can also be used for the analysis of job or work processes involving teams of operator's. The second example is a process chart for a railroad operations task (adapted from Sanders & McCormick 1992).

1.1.1 Check the current speed brake setting
1.1.2 Move the speed brake lever to ‘full’ position
1.2.1 Check that the auto-pilot is in IAS mode
1.2.2 Check the current airspeed
1.2.3 Dial the speed/Mach knob to enter 210 on the IAS/MACH display
2.1 Check the localiser position on the HSI display
2.2.1 Adjust heading +
2.2.2 Adjust heading -
2.3 Check the glideslope indicator
2.4 Maintain current altitude
2.5 Press ‘APP’ button to engage the approach system
2.6.1 Check that the ‘APP’ light is on
2.6.2 Check that the ‘HDG’ light is on
2.6.3 Check that the ‘ALT’ light is off
3.1 Check the current distance from runway on the captains primary flight display
3.2.1 Check the current airspeed
3.2.2 Dial the speed/Mach knob to enter 190 on the IAS/MACH display
3.3.1 Check the current flap setting
3.3.2 Move the flap lever to setting ‘1’
3.4.1 Check the current airspeed
3.4.2 Dial the speed/Mach knob to enter 150 on the IAS/MACH display
3.5.1 Check the current flap setting
3.5.2 Move the flap lever to setting ‘2’
3.6.1 Check the current flap setting
3.6.2 Move the flap lever to setting ‘3’
3.7.1 Check the current airspeed
3.7.2 Dial the speed/Mach knob to enter 140 on the IAS/MACH display
3.8 Put the landing gear down
3.9 Check altitude
3.3.1 Check the current flap setting
3.3.2 Move the flap lever to ‘FULL’ setting

Figure 13. Task list for process chart example

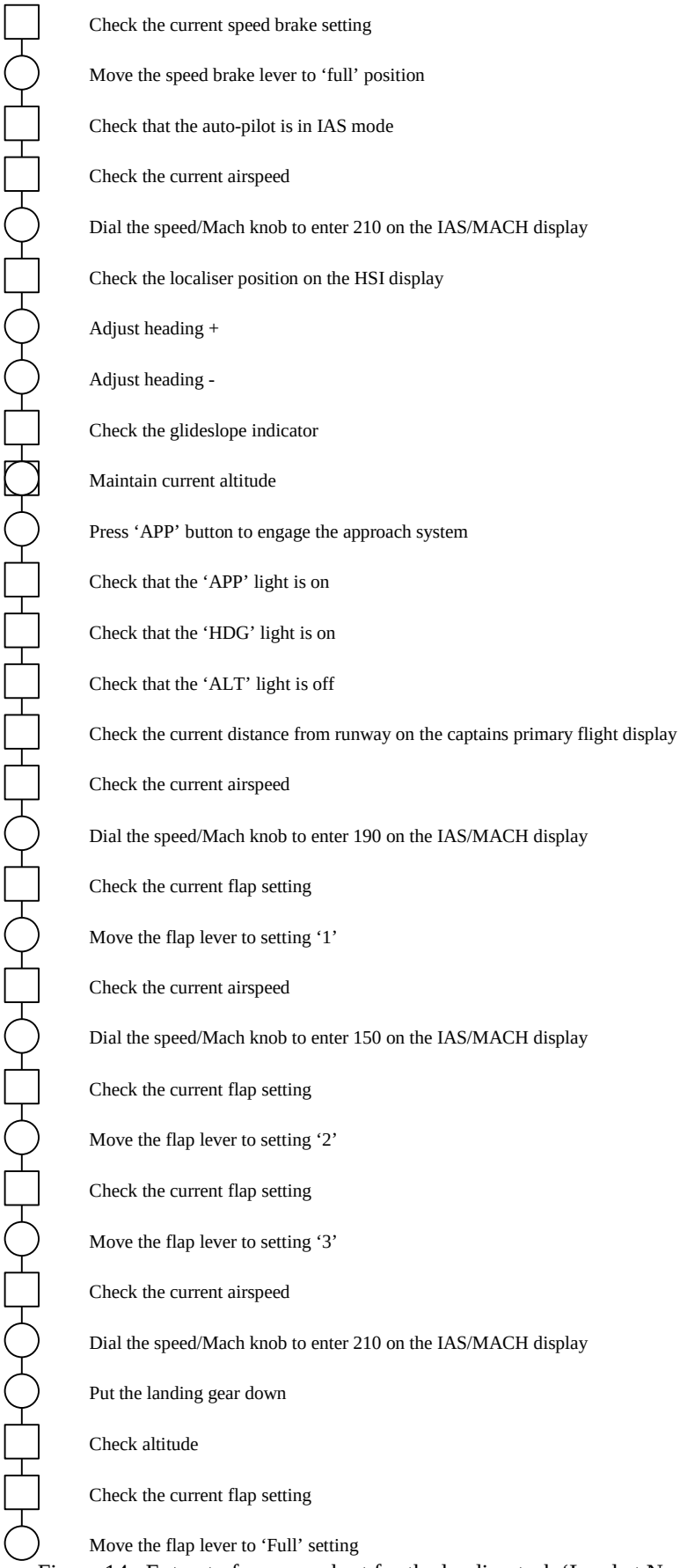
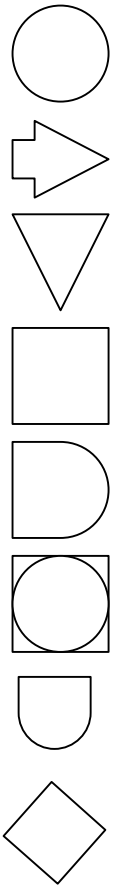


Figure 14. Extract of process chart for the landing task 'Land at New Orleans using the autoland system (Marshall et al 2003)

Operational Sequence Diagrams

Background and applications

Operational Sequence Diagrams (OSD) are used to graphically describe the interaction between teams of operators and a system. According to Kirwan and Ainsworth (1992), the original purpose of OSD analysis was to represent complex, multi-person tasks. The output of an OSD graphically depicts a task process, including the tasks performed and the interaction between operators over time, using standardised symbols. There are numerous forms of OSD's, ranging from a simple flow diagram representing task order, to more complex OSD which account for team interaction and communication, and often including a timeline of the scenario under analysis and potential sources of error. OSD's are typically used during the design of complex systems, such as nuclear petro-chemical processing plants. However, OSD's can also be constructed for existing systems and scenarios, in order to evaluate task structure. When constructing an OSD, a set of standardised symbols are typically used to represent operator actions and communications. These symbols are displayed below.



Operation

Transportation

Storage

Inspection

Delay

Combined operations (e.g. inspection performed with an operation)

Receipt – to receive information or objects

Decision

Domain of application

Nuclear power and chemical process industries.

Procedure and advice

Step 1: Data collection

In order to construct an OSD, the analyst(s) must first obtain sufficient data regarding the scenario under analysis. It is recommended that the analyst(s) use various forms of data collection in this phase. Observational study should be used to observe the task (or similar types of task) under analysis. Interviews with personnel involved in the task (or similar tasks) should also be conducted. The type and amount of data collected in step 1 is dependent upon the analysis requirements. For example, if the output requires a cognitive component, techniques such as critical decision method and cognitive walkthrough can be used in step 1 to acquire the necessary data.

Step 2: Conduct a task analysis

Once the data collection phase is completed, a detailed task analysis should be conducted for the scenario under analysis. The type of task analysis is determined by the analyst(s), and in some cases, a task list will suffice. However, it is recommended that a HTA is conducted. The task analysis should include the following:

- Operations or actions
- Transmission of information
- Receipt of information
- Operator decisions
- Storage of information or objects
- Delay's or periods of inactivity
- Inspections
- Transportations
- Timeline

Step 3: Convert task steps into OSD symbols

The next step in conducting an OSD analysis is to convert each task steps into an OSD symbol. The item should be classified and then converted into the relevant symbol.

Step 4: Construct the OSD diagram

Once each aspect of the task has been assigned a symbol, the OSD can be constructed. The OSD should include a timeline as the starting point, and each event in time should be entered into the diagram. The symbols involved in a particular task step should be linked by directional arrows.

Advantages

- OSD's display the task steps involved in a certain scenario. A number of task factors are included in the OSD analysis, such as actions, decisions, time and transmissions.
- OSD are useful for demonstrating the relationship between tasks, technology and team members
- OSD analysis seems to be very suited to analysing C4i type tasks or scenarios.
- OSD's can be used to analyse team-based tasks, including the interactions between team members.
- High face validity (Kirwan & Ainsworth 1992).
- The OSD output is extremely useful for task allocation and system design/analysis.

Disadvantages

- Constructing an OSD for large, complex tasks can be very difficult.
- For large, complex tasks, the technique is very time consuming to apply. Indeed, for very complex multi-agent scenarios it may become impossible to construct a coherent OSD.
- The initial data collection associated with OSD's is also very time consuming.
- OSD's can become cluttered and confusing (Kirwan & Ainsworth 1992).

Example

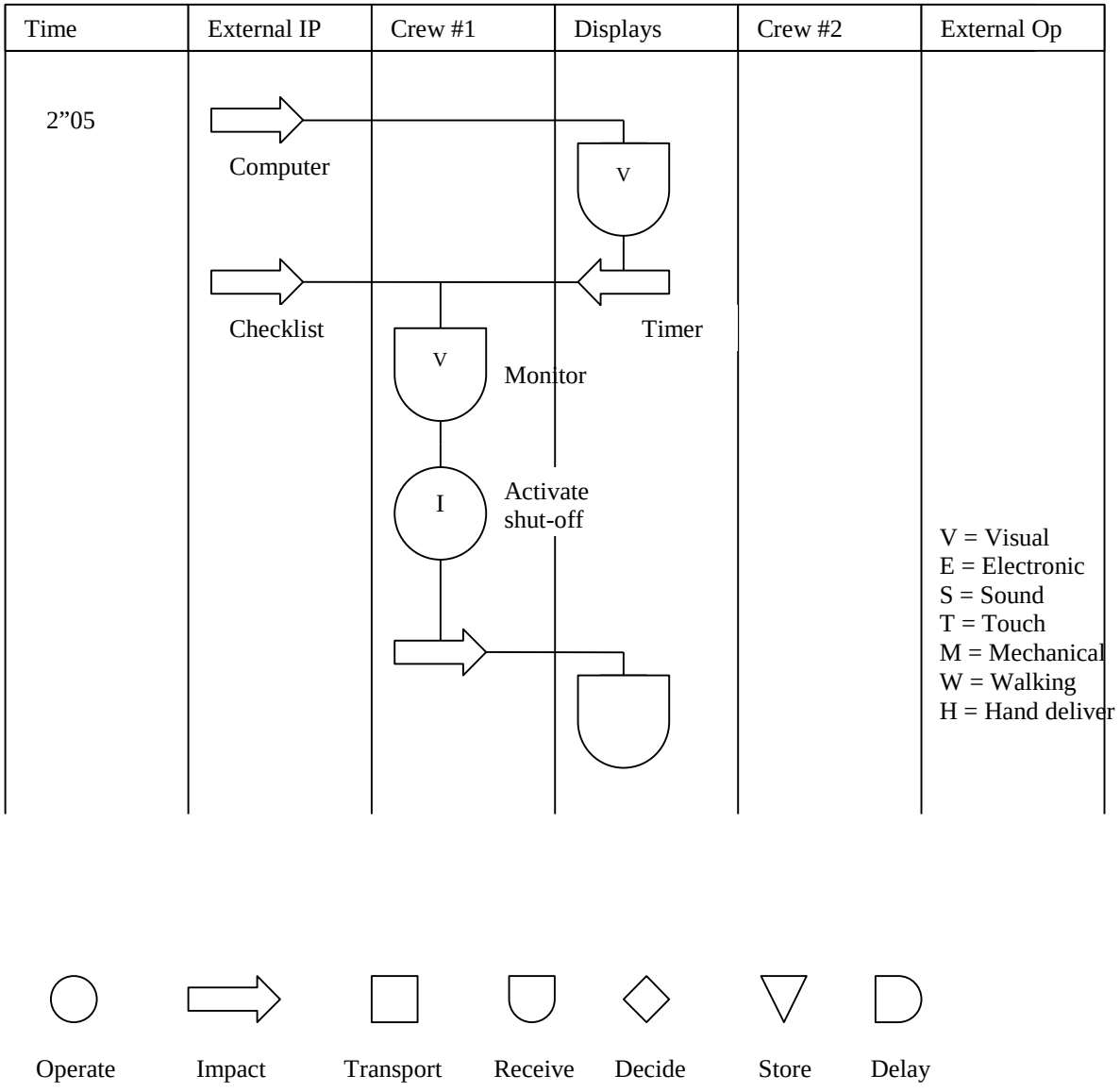


Figure 15. Example OSD

Related methods

Various types of OSD exist, including temporal operational sequence diagrams, partitioned operational sequence diagrams and spatial operational sequence diagrams (Kirwan & Ainsworth 1992). In the data collection phase, techniques such as observational study and interviews are typically used. Task analysis techniques such as HTA are also used during the construction of the OSD. Timeline analysis may also

be used in order to construct an appropriate timeline for the task or scenario under analysis.

Approximate training and application times
No data regarding the training and application time associated with the OSD technique are available in the literature. However, it is apparent that the training time for such a technique would be high. Similarly, the application time for the technique would be high, including the initial data collection phase of interviews and observational analysis.

Reliability and validity
According to Kirwan & Ainsworth, OSD techniques possess a high degree of face validity. Data regarding other aspects of the techniques validity and also reliability are not available.

Tools needed
When conducting an OSD analysis, pen and paper could be sufficient. However, to ensure that data collection is comprehensive, it is recommended that video or audio recording devices are used in conjunction with the pen and paper.

Bibliography
Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. Taylor and Francis, UK.
Sanders, M.S. & McCormick, E.J. (1993) Human Factors in Engineering and Design. McGraw-Hill Publications.

Event Tree analysis

Background and applications

Event tree analysis is a task analysis technique that uses tree like diagrams to represent the various possible outcomes associated with operator tasks steps in a scenario. Originally used in system reliability analysis (Kirwan & Ainsworth 1992), event tree analysis can also be applied to human operations to investigate possible actions and their consequences. Event tree output is normally made up of a tree like diagram consisting of nodes (representing task steps) and exit lines (representing the possible outcomes). Typically, success and failure outcomes are used, but for more complex analyses, multiple outcomes can be represented (Kirwan & Ainsworth 1992). Event tree analysis can be used to depict task sequences and their possible outcomes, to identify error potential within a system and to model team-based tasks. In the early stages of a system design, event tree analysis can be used to highlight potential error paths within a proposed system design, and can also be used to modify the design in terms of removing tasks which carry a multitude of associated task steps.

Domain of application

Nuclear power and chemical process industries.

Procedure and advice

Step 1: Define scenario(s) under analysis
Firstly, the scenario(s) under analysis should be clearly defined. Event tree analysis can be used to analyse either existing systems or system design concepts.

Step 2: Data collection phase
If the event tree analysis is concerned with an existing system, then data regarding the scenario under analysis should be collected. To do this, observational analysis, interviews and questionnaires are typically used. If the event tree analysis is based on a design concept, then storyboards can be used to depict the scenario(s) under analysis.

Step 3: Draw up task list
Once the scenario under analysis is defined clearly and sufficient data is collected, a comprehensive task list should be created. Each task step should be broken down to the operations level (as in HTA) and controls or interfaces used should also be noted. This initial task list should represent typical, error free performance of the task or scenario under analysis. It may be useful to consult with SME's during this process.

Step 4: Determine possible actions for each task step
Once the task list is created, the analyst should then describe every possible action associated with each task step in the task list. It may be useful to consult with SME's during this process. Every possible action associated with each task step should be recorded.

Step 5: Determine consequences associated with each possible action
Next, the analyst should take each action specified in step 4 and record the associated consequences.

Step 6: Construct event tree
Once steps 4 and 5 are complete, the analyst can begin to construct the event tree diagram. The event tree should depict all possible actions and their associated consequences.

Advantages
- Event tree analysis can be used to highlight a sequence of tasks steps and their associated consequences.
- Event tree analysis can be used to highlight error potential and error paths throughout a system.
- The technique can be used in the early design life cycle to highlight task steps that may become problematic (multiple associated response options) and also those task steps that have highly critical consequences.
- If used correctly, the technique could potentially depict anything that could possibly go wrong in a system.
- Event tree analysis is a relatively easy technique that requires little training.
- Event tree analysis has been used extensively in PSA/HRA.

Disadvantages
- For large, complex tasks, the event tree can become very large and complex.
- Can be time consuming in its application.
- Task steps are often not explained in the output.

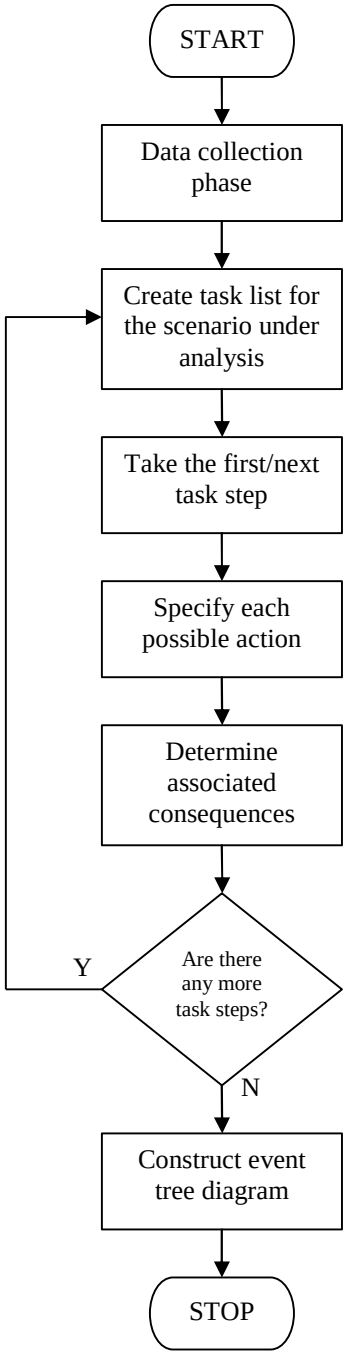
Related methods
According to Kirwan & Ainsworth (1992) there are a number of variations of the original event tree analysis technique, including operator action event tree analysis (OATS) (Hall et al 1982), human reliability analysis event tree analysis (HRAET) (Bell & Swain 1983). Event trees are also similar to fault tree analysis and operator sequence diagrams.

Reliability and validity
No data regarding the reliability and validity of the event tree technique are available.

Tools needed
An event tree can be conducted using pen and paper. If the event tree is based on an existing system, then observational analysis should be used, which requires video and audio recording equipment and a PC.

Bibliography
Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. Taylor and Francis, UK.

Flowchart


START
Data collection phase
Create task list for the scenario under analysis
Take the first/next task step
Specify each possible action
Determine associated consequences
Are there any more task steps?
Y
N
Construct event tree diagram
STOP

Example

Example A is an extract of an event tree that was constructed for the landing task, 'Land A320 at New Orleans using the autoland system' in order to investigate the use of event tree analysis for predicting design induced pilot error (Marshall et al 2003).

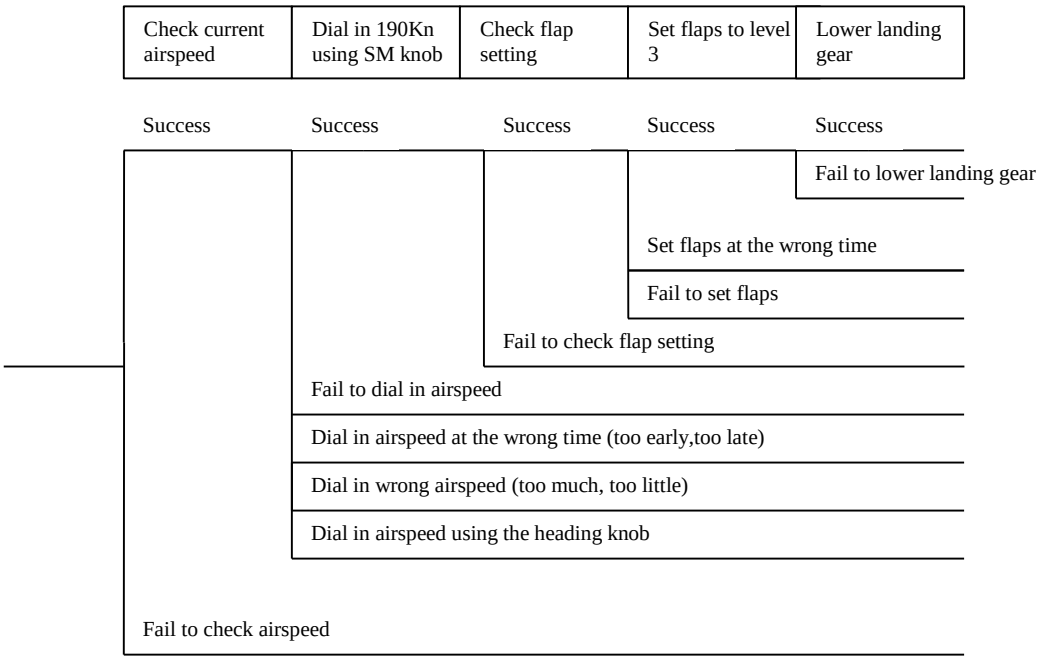


Figure 16. Extract of event tree diagram for the flight task 'Land at New Orleans using the autoland system' (Marshall et al 2003)

Decision Action Diagrams

Background and applications

Decision Action Diagrams (DAD's), also known as information flow diagrams (Kirwan & Ainsworth 1992) are used to depict the process of a scenario through a system in terms of the decisions required and actions to be performed by the operator in conducting the task or scenario under analysis. Decisions are represented by diamonds and each decision option available to the system operator is represented by exit lines. In their simplest form, the decision options are usually 'Yes' or 'No', however depending upon the complexity of the task and system, multiple options can be represented. The DAD output diagram should display all of the possible outcomes at each task step in a process. DAD analysis can be used to evaluate existing systems or to inform the design of system's and task processes. DAD's could potentially be used to depict the decisions and actions exhibited in command and control scenarios.

Domain of application

Nuclear power and chemical process industries.

Procedure and advice

Step 1: Data collection

In order to construct a DAD, the analyst(s) must first obtain sufficient data regarding the scenario under analysis. It is recommended that the analyst(s) use various forms of data collection in this phase. Observational study should be used to observe the task (or similar types of task) under analysis. Interviews with personnel involved in the task (or similar tasks) should also be conducted. The type and amount of data collected in step 1 is dependent upon the analysis requirements. For example, if the output requires a cognitive component, techniques such as critical decision method and cognitive walkthrough can be used in step 1 to acquire the necessary data.

Step 2: Conduct a task analysis

Once the data collection phase is completed, a detailed task analysis should be conducted for the scenario under analysis. The type of task analysis is determined by the analyst(s), and in some cases, a task list will suffice. However, it is recommended that when constructing a DAD, a HTA for the scenario under analysis is conducted.

Step 3: Construct DAD

Once the task or scenario under analysis is fully understood, the DAD can be constructed. This process should begin with the first decision available to the operator of the system. Each possible outcome or action associated with the decision should be represented with an exit line from the decision diamond. Each resultant action and outcome for each of the possible decision exit lines should then be specified. This process should be repeated for each task step until all of the possible decision outcomes for each task have been exhausted.

Advantages

- A DAD can be used to depict the possible options that an operator faces at each task step. This can be used to inform the design of the system or process i.e. task steps that have multiple options associated with them can be redesigned.
- DAD's are relatively easy to construct and require little training.
- DAD's could potentially be used for error prediction purposes.

Disadvantages

- In their current form, DAD's do not cater for the cognitive component of task decisions.
- It would be very difficult to model parallel activity using DAD's.
- DAD's do not cater for processes involving teams. Constructing a team DAD would appear to be extremely difficult.
- It appears that a HTA for the task or scenario under analysis would be sufficient. A DAD output is very similar to the plans depicted in a HTA.
- For large, complex tasks, the DAD would be difficult and time consuming to construct.
- The initial data collection phase involved in the DAD procedure adds a considerable amount of time to the analysis.
- Reliability and validity data for the technique is sparse.

Related methods

DAD's are also known as information flow charts (Kirwan & Ainsworth 1992). The DAD technique is related to other process chart techniques such as operation sequence diagrams and also task analysis techniques such as HTA. When conducting a DAD type analysis, a number of data collection techniques are used, such as observational analysis and interviews. A task analysis (e.g. HTA) of the task/scenario under analysis may also be required.

Approximate training and application times

No data regarding the training and application times associated with DAD's are available in the literature. It is hypothesised that the training time for DAD's would be minimal or low. The application time associated with the DAD technique is dependent upon the task and system under analysis. For complex scenarios with multiple options available to the operator involved, the application time would be high. For more simple procedural tasks, the application time would be very low. The data collection phase of the DAD procedure would require considerable time, particularly when observational analysis is used.

Reliability and validity

No data regarding the reliability and validity of the DAD technique are available.

Tools needed

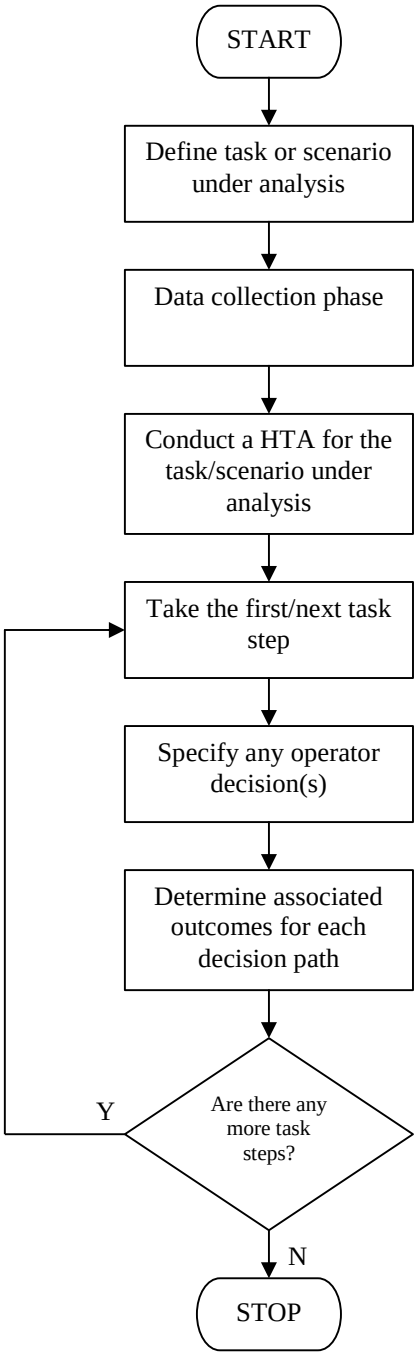
Once the initial data collection is complete, the DAD technique can be conducted using pen and paper. The tools required for the data collection phase are dependent upon the techniques used. Typically, observation is used, which would require video and audio recording equipment and a PC.

Bibliography

Kirwan, B. (1994) A Guide to Practical Human Reliability Assessment. UK, Taylor and Francis.

Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. UK, Taylor and Francis.

Flowchart


START
Define task or scenario under analysis
Data collection phase
Conduct a HTA for the task/scenario under analysis
Take the first/next task step
Specify any operator decision(s)
Determine associated outcomes for each decision path
Are there any more task steps?
Y
N
STOP

Example

The following example is a DAD taken from Kirwan & Ainsworth (1992).

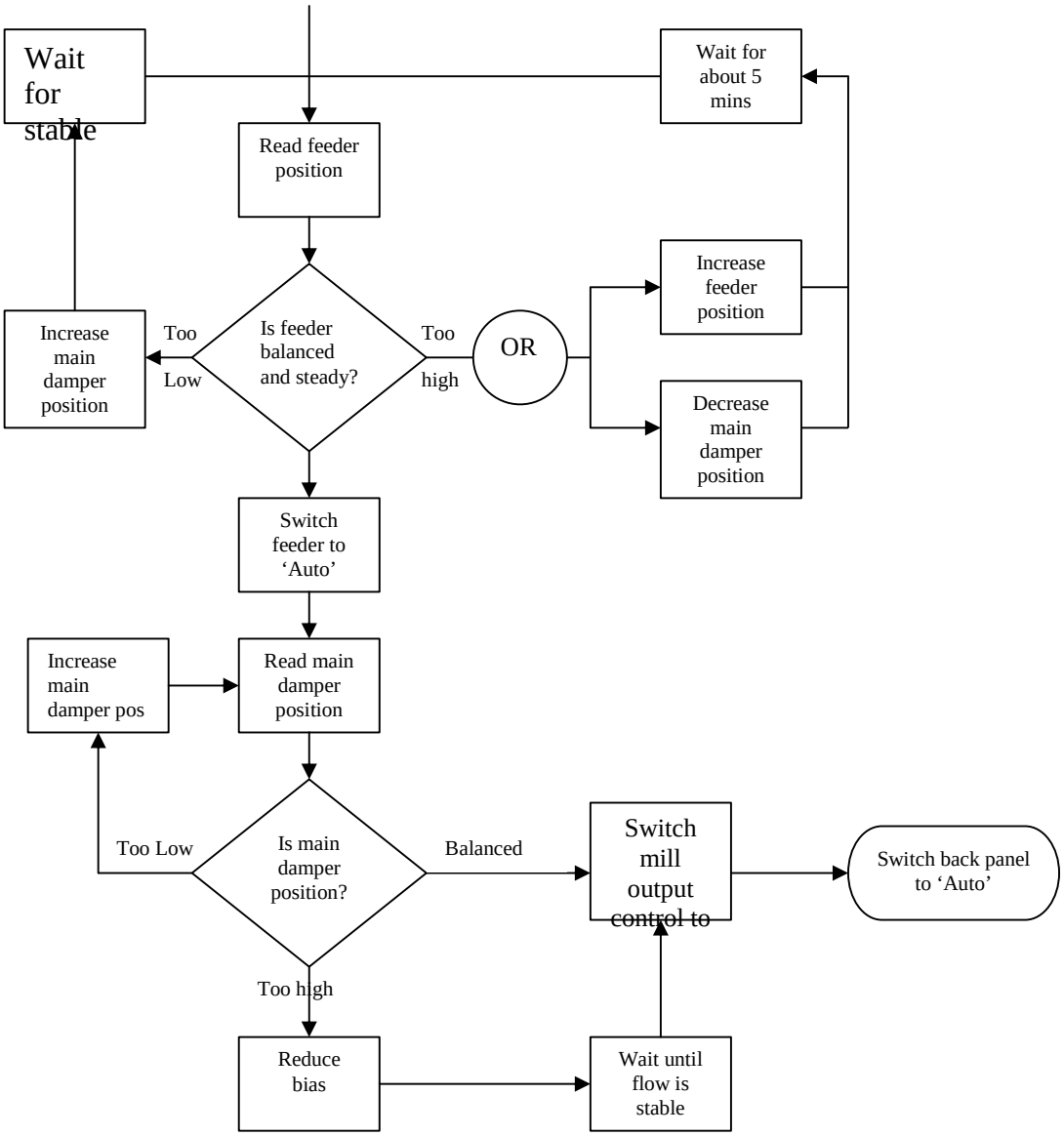


Figure 17. Decision-Action Diagram. Adapted from Kirwan & Ainsworth (1992).

Fault Trees

Background and application

Fault trees are used to depict system failures and their causes. A fault tree is a tree like diagram, which defines the failure event and displays the possible causes in terms of hardware failure or human error (Kirwan & Ainsworth 1992). Fault tree analysis was originally developed for the analysis of complex systems in the aerospace and defence industries (Kirwan & Ainsworth 1992) and they are now used extensively in probabilistic safety assessment (PSA). Although typically used to evaluate events retrospectively, fault trees can be used at any stage in the design process to predict failure events and their causes. The fault tree can be used to show the type of failure event and its various causes. Typically, the failure event or top event (Kirwan and Ainsworth 1992) is placed at the top of the fault tree, and the contributing events are placed below. The fault tree is held together by AND and OR gates, which link contributing events together. An AND gate is used when more than one event causes a failure i.e. contributing factors are involved. The events placed directly underneath an AND gate must occur together for the failure event above to occur. An OR gate is used when the failure event could be caused by more than one contributory event in isolation, but not together. The event above the OR gate may occur if any one of the events below the OR gate occurs. A fault tree analysis could be used in the design of a system in order to contribute to the eradication of potential failure causes.

Domain of application

Nuclear power and chemical process industries.

Procedure and advice

Step 1: Define failure event

The failure or event under analysis should be defined first. This may be an actual event that has occurred or an imaginary event. This event is the top event in the fault tree. If using the technique to analyse a failure event in an existing system, then the failure event under analysis makes up the top event. However, if the technique is being used to predict how failure events could occur in the system design concept, then failure events or scenarios should be offered by the design team.

Step 2: Determine causes of failure event

Once the failure event has been defined, the causes of the event need to be determined. The nature of the causes analysed is dependent upon the focus of the analysis. Typically, human error and hardware failures are considered (Kirwan & Ainswoth 1992).

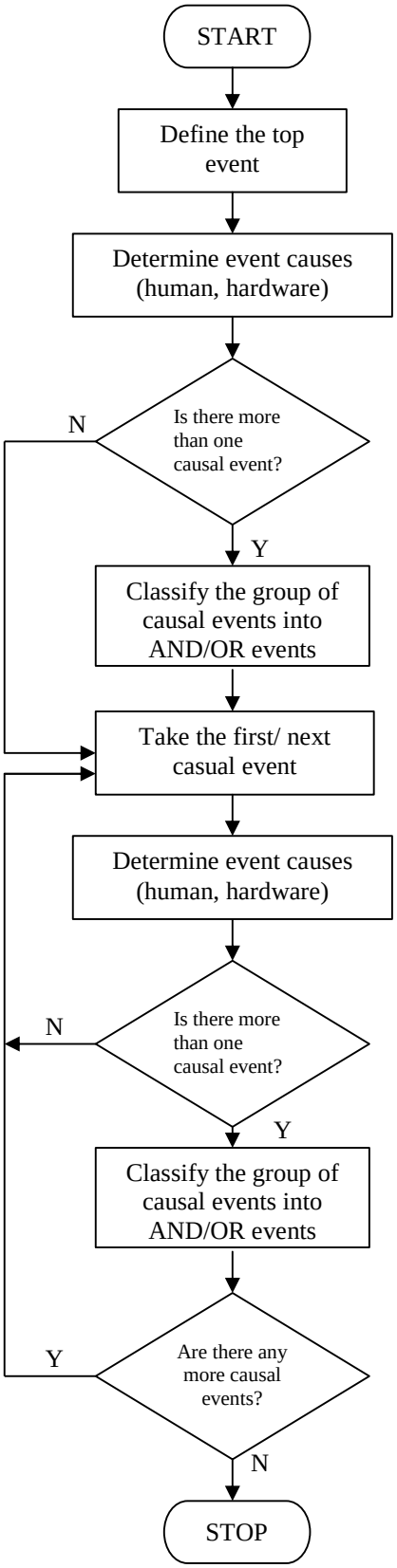
Step 3: AND/OR classification

Once the cause(s) of the failure event are determined, they should be classified into AND or OR causes. If the two or more cause events contribute to the failure event, they are classified as OR events. If two or more cause events are responsible for the failure even when they occur separately, then they are classified as OR events.

Steps 2 and 3 should be repeated until each of the initial causal events and associated causes are investigated and described fully.

Step 4: Construct Fault tree diagram
Once all events and their causes have been defined fully, they should be put into the fault tree diagram. The fault tree should begin with the main failure or top event at the top of the diagram with it's associated causes linked underneath as AND/OR events. Then, the causes of these events should be linked underneath as AND/OR events. The diagram should continue until all events and causes are exhausted fully.

Flowchart



Advantages

- Fault trees are useful in that they define possible failure events and their causes. This is especially useful when looking at failure events with multiple causes.
- Fault tree type analysis has been used extensively in PSA.
- Although most commonly used in the analysis of nuclear power plant events, the technique is generic and can be applied in any domain.
- Fault trees can be used to highlight potential weak points in a system design concept (Kirwan & Ainsworth 1992).
- The technique could be particularly useful in modelling team-based errors, where a failure event is caused by multiple events distributed across a team of personnel.
- Fault tree analysis has the potential to be used during the design process in order to remove potential failures associated with a system design.

Disadvantages

- When used to depict failures in large, complex systems, fault tree analysis can be very difficult and time consuming to apply. The fault tree itself can also quickly become large and complicated.
- To utilise the technique quantitatively, a high level of training may be required (Kirwan & Ainsworth 1992).

Related methods

The fault tree technique is often used with event tree analysis (Kirwan & Ainsworth 1992).

Approximate training and application times

No data regarding the training and application times associated with fault tree analysis are available in the literature. It is hypothesised that the training time for fault trees would be minimal or low. The application time associated with the fault tree technique is dependent upon the task and system under analysis. For complex failure scenarios, the application time would be high. For more simple failure events, the application time would be very low.

Reliability and validity

No data regarding the reliability and validity of the DAD technique are available

Tools needed

Fault tree analysis can be conducted using pen and paper. If the analysis were based upon an existing system, an observational analysis of the failure event under analysis would be useful. This would require video and audio recording equipment.

Bibliography

Kirwan, B. (1994) A Guide to Practical Human Reliability Assessment. UK, Taylor and Francis.

Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. UK, Taylor and Francis.

Example

The following example is taken from Kirwan (1994)

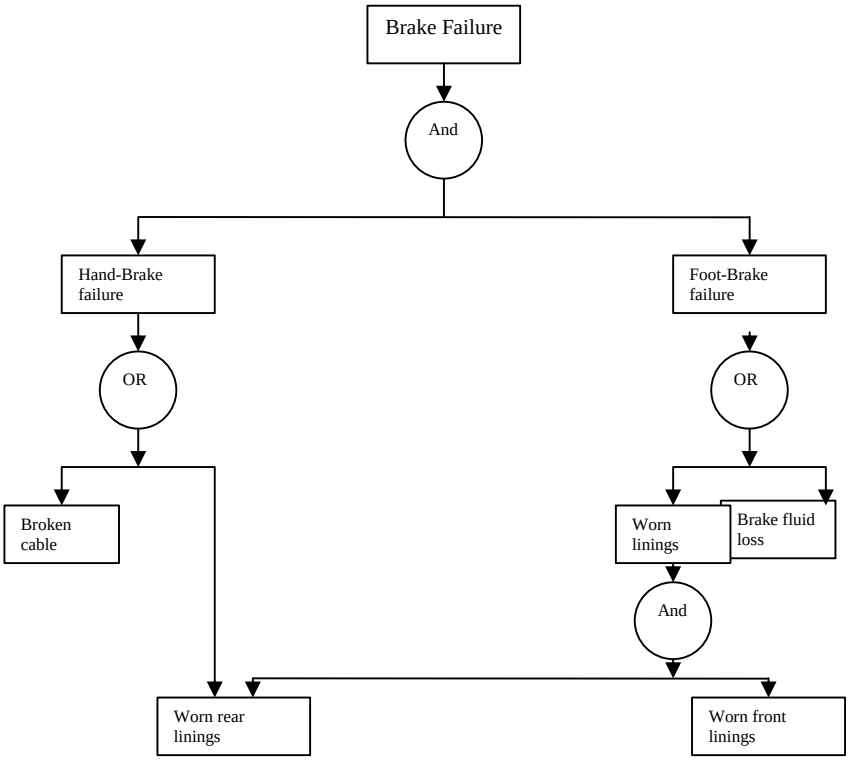


Figure 18. Fault tree for brake failure scenario

.

Murphy Diagrams

Background and applications

Murphy diagrams (Pew et al, 1981) were first developed as part of a study commissioned by the Electronic Power Research Institute in the USA and were originally used for the retrospective examination of errors in process control rooms. Murphy diagrams are based on the notion that "if anything can go wrong, it will go wrong" (Kirwan & Ainsworth 1992). The technique is very similar to fault tree analysis in that errors of failures are analysed in terms of their potential causes. Although originally used for the retrospective analysis of error events whereby the analyst conducts eight Murphy diagrams for the error under analysis, there is no reason why the technique could not be used to predict potential error events associated task steps in a scenario. Each task step is classified into one of eight the eight decision making process classifications below:

- Activation/Detection
- Observation and data collection
- Identification of system state
- Interpretation of situation
- Task definition/selection of goal state
- Evaluation of alternative strategies
- Procedure selection
- Procedure execution

The Murphy diagram begins with the top event being split into success and failure nodes. Obviously, the success event requires no further analysis, and so the analyst should describe the failure event. Next the analyst takes the 'failure' outcome and defines the sources of the error that have an immediate effect. These are called the proximal sources of error. The analyst then takes each proximal error source and breaks it down further so that the causes of the proximal error sources are defined. These proximal error causes are termed the distal causes. For example, if the failure was 'Procedure incorrectly executed', the proximal sources could be 'wrong switches chosen', 'switches incorrectly operated' or 'switches not operated'. The distal sources for 'wrong switches chosen' could then be further broken down into 'deficiencies in placement of switches', 'inherent confusability in switch design' or 'training deficiency' (Kirwan & Ainsworth 1992). The Murphy diagram technique could be used to highlight error causes and consequences in a design concept. More importantly, perhaps, the technique appears to have the potential to be used in the analysis of team-based operations, highlighting distributed task requirements and distributed error causes.

Domain of application

Nuclear power and chemical process industries.

Procedure and advice

The following procedure is intended to act as a set of guidelines when using the technique for the prediction of error events and their causes.

Step 1: Define task/scenario under analysis
The first step in a Murphy Diagram analysis is to define the task or scenario under analysis.

Step 2: Data collection
If the analyst(s) possess insufficient data regarding the scenario under analysis, then data regarding similar scenarios in similar systems should be collected. Techniques used for the data collection would include direct observation and interviews.

Step 3: Define error events
Once sufficient data regarding the scenario under analysis is collected, the analysis begins with the definition of the first error. The analyst(s) should define the error clearly.

Step 4: Classify error activity into decision making category
Once the error under analysis is described, the activity leading up to the error should be classified into one of the eight decision making process categories.

Step 5: Determine error consequence and causes
Once the error is described and classified, the analysis begins. The analyst(s) should determine the consequences of the error event and also determine possible consequences associated with the error. The error causes should be explored fully, with proximal and distal sources described.

Step 6: Construct Murphy Diagram
Once the consequences, proximal and distal sources have been explored fully, the Murphy diagram for the error in question should be constructed.

Step 7: Propose design remedies
For the purpose of error prediction in the design of systems, it is recommended that the Murphy diagram is extended to include an error or design remedy column. The analyst(s) should use this column to propose design remedies for the identified errors, based upon the causes identified.

Advantages

- Easy technique to use and learn, requiring little training.
- Murphy diagrams present a useful way for the analyst to identify a number of different possible causes for a specific error.
- High documentability.
- Each task step failure is exhaustively described, including proximal and distal sources.
- The technique has the potential to be applied to team-based tasks, depicting teamwork and failures with multiple team-based causes.
- Murphy diagrams have the potential to use little resources (low cost, time spent etc).
- Although developed for the retrospective analysis of error, there appears to be no reason why it cannot be used predictively.

Disadvantages

- Its use as a predictive tool is uncertain – "While it is easy to use for the analysis of predictions, its predictive utility as an HEI tool is uncertain, again because there is little in the way of published literature on such applications." (Kirwan, 1994)
- Could become large and unwieldy for large, complex tasks.

- There is little guidance for the analyst.
- Consistency of the method can be questioned.
- Design remedies are based entirely upon the analyst's subjective judgement.
- It would be difficult to model time on a Murphy diagram.

Example

There is no evidence of the technique being used in the way proposed above in the literature. Consequently, a mock-up analysis of how the technique could be used in the design of command and control systems is provided.

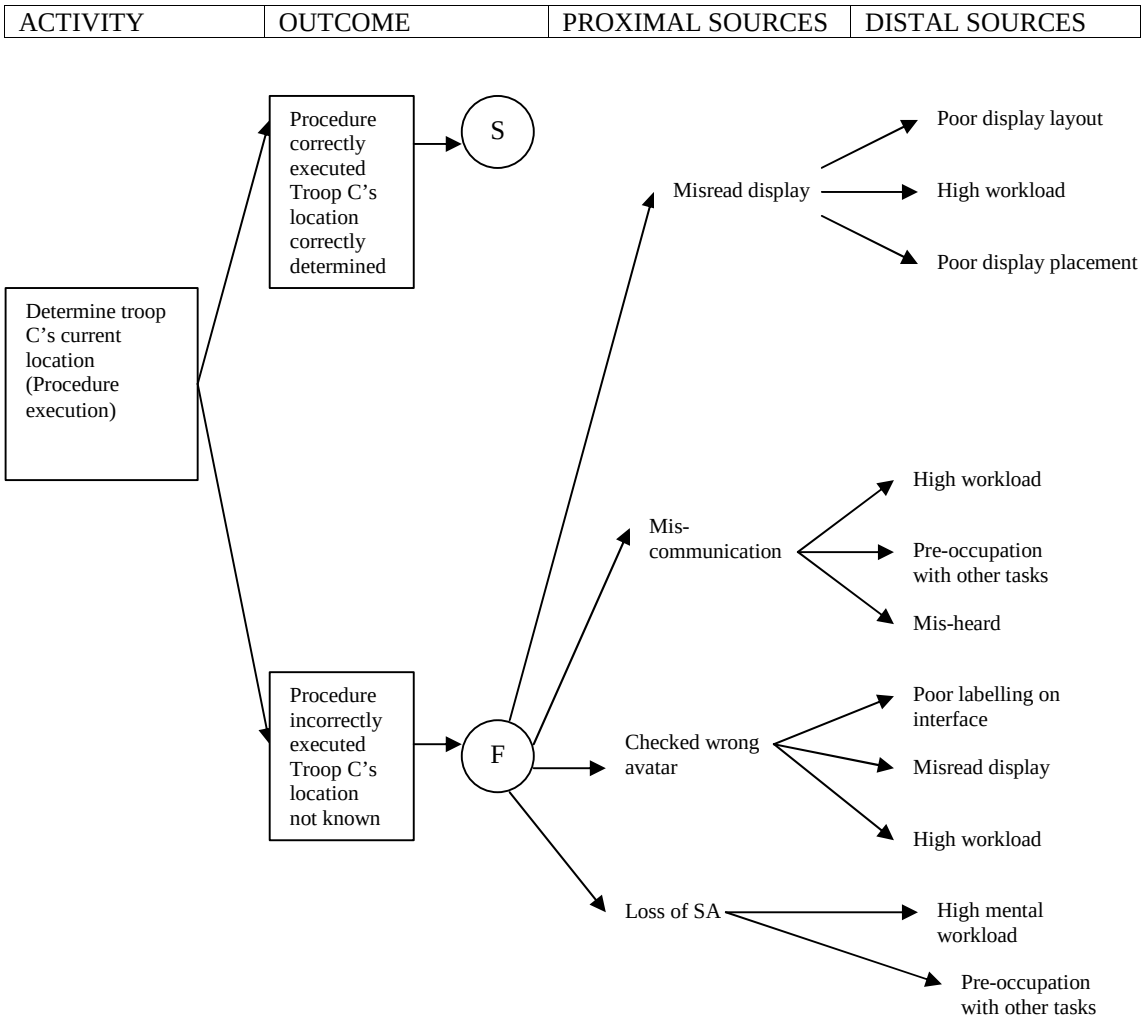


Figure 19. Mock Murphy Diagram

Related methods

Murphy diagrams are very similar to fault tree analysis in that they depict failure events and their causes. The Murphy diagram technique is also similar in its output to operator sequence diagrams.

Approximate training and application times
The training time for the technique would be minimal. The application time would depend upon the task or scenario under analysis. For error incidences with multiple causes and consequences, the application time would be high.

Reliability and validity
No data regarding the reliability and validity of Murphy diagrams are available in the literature.

Tools needed
The technique can be conducted using pen and paper. A PC is normally used to construct the Murphy diagram.

Bibliography
Kirwan, B. Ainsworth, L. K (1992) A guide to Task Analysis, Taylor and Francis, London, UK.
Kirwan, B. (1992a) Human error identification in human reliability assessment. Part 1: Overview of approaches. *Applied Ergonomics* Vol. 23(5), 299 – 318
Kirwan, B. (1992b) Human error identification in human reliability assessment. Part 2: detailed comparison of techniques. Applied Ergonomics, 23, 371-381.

Task Analysis techniques

Probably the most commonly used group of techniques by the HF practitioner, task analysis techniques are used to understand and represent human and system performance in a particular task or scenario under analysis. According to Diaper & Stanton (2003) there are, or at least have been, over 100 task analysis techniques described in the literature. Task analysis involves identifying tasks, collecting task data, analysing the data so that tasks are understood, and then producing a documented representation of the analysed tasks (Stanton 2003). Typical Task analysis techniques break down tasks or scenarios into the required individual task steps, in terms of the required human-machine and human-human interactions. According to Kirwan & Ainsworth (1992) task analysis can be defined as the study of what an operator (or team of operators) is required to do, in terms of actions and cognitive processes, to achieve system goals. A number of different variations of task analysis techniques exist, including hierarchical task analysis (HTA), tabular task analysis (TTA), verbal protocol analysis (VPA), critical path analysis (CPA) and goals, operators, methods and selection rules (GOMS).

The use of task analysis techniques is widespread, with applications in a wide range of domains, including military operations, aviation (Marshall et al 2003), air traffic control, driving (Walker 2001), public technology, product design and nuclear petro-chemical domains. According to Annett (In Press) a survey of defence task analysis studies demonstrated its use in system procurement, manpower analysis, interface design, operability assessment and training specification. According to Diaper (2003) task analysis is potentially the most powerful technique available to HCI practitioners, and it has potential application at each stage in system design and development. Stanton (2003) also suggests that task analysis is the central method for the design and analysis of system performance, involved in everything from design concept to system development and operation. Stanton (2003) also highlights the role of task analysis in task allocation, procedure design, training design and interface design.

Hierarchical task analysis (HTA) involves breaking down the task under analysis into a hierarchy of goals, operations and plans. Tasks are broken down into hierarchical set of tasks, sub tasks and plans. Critical path analysis (CPA) is a project management tool that is used to calculate the combination of tasks that will most affect the time taken to complete a job. GOMS (Card, Moran & Newell 1983) attempts to define the user's goals, decompose these goals into sub-goals and demonstrate how the goals are achieved through user interaction. Verbal protocol analysis (VPA) is used to derive the processes, cognitive and physical, that an individual uses to perform a task. VPA involves creating a written transcript of operator behaviour as they perform the task under analysis. Task decomposition (Kirwan & Ainsworth 1992) can be used to create a detailed task description regarding a particular task. Specific categories are used to exhaustively describe the task under analysis, such as actions, goals, controls, error potential and time constraints.

Whilst its use is ongoing and widespread, the concept of task analysis has also evolved, with task analysis techniques now considering the cognitive aspects of work (CTA, CDM), and work distributed across teams and systems (TTA, CUD). Cognitive task analysis techniques, such as the critical decision method (CDM) (Klein 2003), and applied cognitive task analysis (ACTA) (Millitello & Hutton 2003) use

probe interview techniques in order to analyse, understand and represent the unobservable cognitive processes associated with tasks or work. Team task analysis techniques attempt to describe the process of work across teams or distributed systems. Annett (In press) reports the use of HTA for analysing an anti-submarine warfare team task (Annett et al 2000).

Task analysis techniques can be used during the design of systems or to evaluate existing systems and processes. The usefulness of task analysis techniques is heightened by the fact that most HF techniques require some sort of task analysis output as their input, such as SHERPA (Embrey 1986), HET (Marshall et al 2003) and TAFEI (Baber & Stanton 1996).

The task analysis techniques reviewed in this document are shown below.

1. HTA - Hierarchical Task Analysis
2. CPA – Critical Path Analysis
3. GOMS – Goals, Operators, Methods and Selection Rules
4. Task Decomposition
5. VPA – Verbal Protocol Analysis

Task analysis will be used throughout the design lifecycle of the C4i system for a number of purposes, such as representing and understanding existing C4 systems and processes, task allocation, task or process design and the evaluation of design concepts.

HTA - Hierarchical Task Analysis

John Annett, Department of Psychology, Warwick University, Coventry CV4 7AL

Background and Applications

HTA was developed at the University of Hull in response to the need to analyse complex tasks, such as those found in the chemical processing and power generation industries (Annett, Duncan, Stammers & Gray, 1971). The training of process control operators was a matter of concern since the 'time and motion' style methods of task analysis, developed for routine repetitive manual operations used in manufacturing industry did little justice to skills in modern 'automated' industries involving less physical activity combined with a high degree of cognitive skill and knowledge on the part of the operator.

HTA has been widely used in a number of domains, including the process control and power generation industries, military applications (Kirwan & Ainsworth, 1992; Ainsworth & Marshall, 1998/2000), aviation (Marshall et al 2003). Annett (2003) also reports that HTA has been adapted for use in many human factors applications including training (Shepherd, 2002), design (Lim & Long, 1994), error and risk analysis (Baber & Stanton, 1994;) and the identification and assessment of team skills (Annett, Cunningham & Mathias-Jones, 2000).

HTA involves breaking down the task under analysis into a hierarchy of goals, operations and plans. Tasks are broken down into hierarchical set of tasks, sub tasks and plans. The goals, operations and plans categories used in HTA are described below.

- Goals – The unobservable task goals associated with the task in question.
- Operations – The observable behaviours or activities that the operator has to perform in order to accomplish the goal of the task in question.
- Plans – The unobservable decisions and planning made on behalf of the operator.

Domain of application

Generic.

Procedure and advice

Step 1: Determine the overall goal of the task

The overall task goal of the task under analysis should first be specified at the top of the hierarchy i.e. Land Boeing 737 at New Orleans Airport using the ‘Auto-land system’ or ‘boil kettle’.

Step 2: Determine task sub-goals

The next step of the HTA is to break the overall goal down into four or five meaningful sub-goals, which together make up the overall goal. In a HTA analysis of a Ford in-car radio (Stanton & Young 1999) the task, “listen to in car entertainment”, was broken down into the following sub-goals:

- Check unit status,
- Press on/off button,
- Listen to the radio,
- Listen to cassette,
- Adjust audio preferences

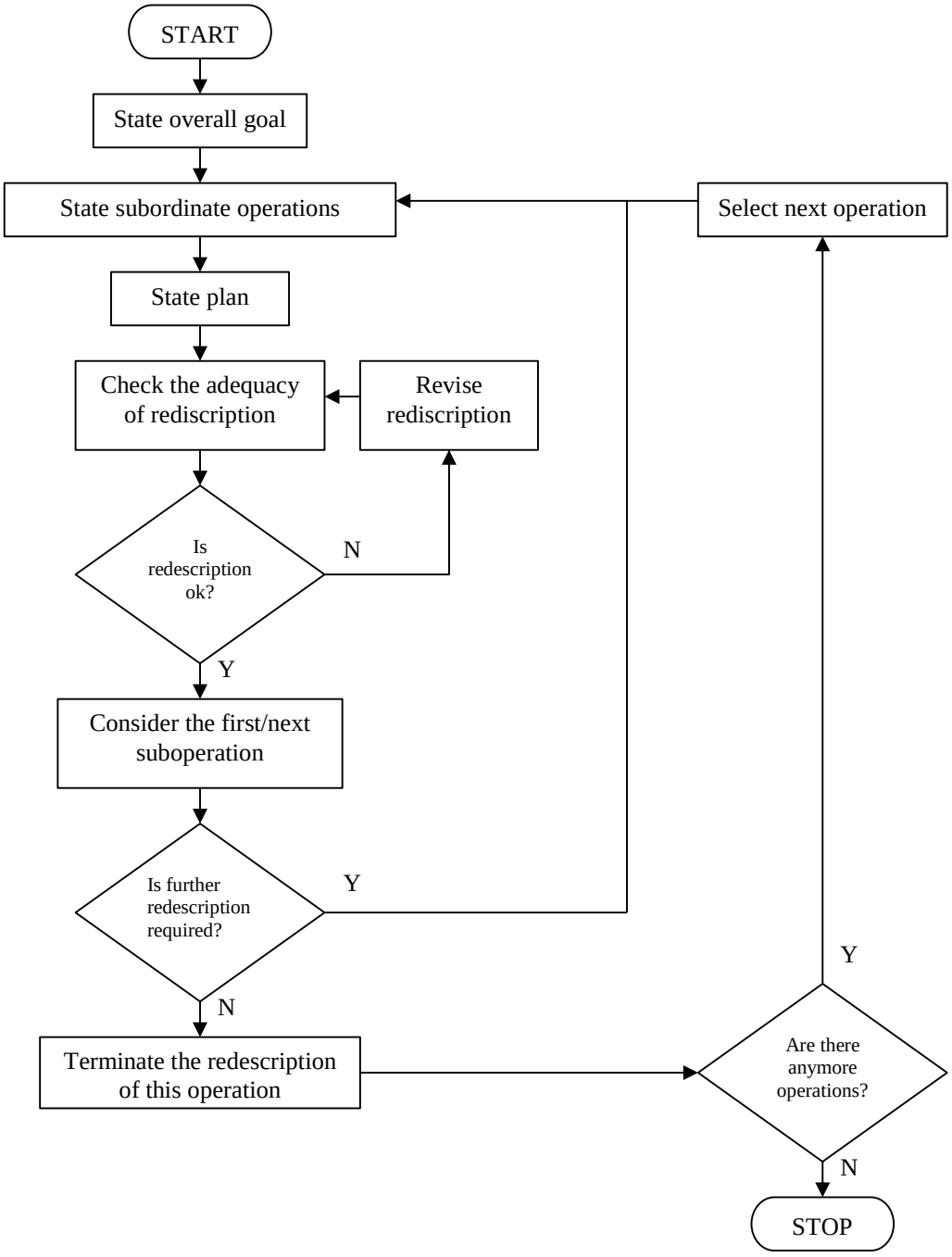
Step 3: Sub-goal decomposition
The sub-goals identified in step two should then be broken down into further sub goals and operations, according to the task. This process should go on until an appropriate sub-goal is reached. The bottom level of any branch in a HTA will always be an operation. Whilst everything above an operation specifies goals, operations actually say what needs to be done. Thus operations are actions to be made by the operator. Underneath the sub-goals, the analyst basically enters what needs to be done to achieve the sub-goal.

Step 4: Plans analysis
Once all of the sub-goals have been fully described, the plans need to be added. Plans dictate how the goals are achieved. A simple plan would say, Do 1, then 2, and then 3. Once the plan is completed, the operator returns to the super-ordinate level. Plans do not have to be linear and can come in any form such as Do 1, Or 2 And 3. Once the goals, sub-goals, operations and plans are exhausted, a complete diagram made up of these four aspects of the task makes up an HTA. If required, this can be tabulated.

Advantages
- HTA is a technique that is both easy to learn and easy to implement.
- HTA is the starting point for numerous human factors techniques.
- Quick to use in most instances.
- Comprehensive technique covers all sub-tasks of the task in question.
- HTA has been used extensively in a wide range of contexts.
- Conducting an HTA gives the user a great insight into the task under analysis.
- HTA is an excellent technique to use when requiring a task description for further analysis. If performed correctly, the HTA should depict everything that needs to be done in order to complete the task in question.
- As a generic method HTA is adaptable to a wide range of purposes.
- Tasks can be analysed to any required level of detail, depending on the purpose.
- When used correctly HTA provides an exhaustive analysis of the problem addressed.

Disadvantages
- Provides mainly descriptive information rather than analytical information.
- HTA contains little that can be used directly to provide design solutions.
- HTA does not cater for the cognitive components of a task.
- Can be time consuming for the more complex and larger tasks.
- Requires handling by an analyst well trained in a variety of methods of data collection and in relevant human factors principles.
- Requires time in proportion to the complexity of the task and the depth of the analysis.

Flowchart


START
State overall goal
State subordinate operations
State plan
Check the adequacy of rediscription
Revise rediscription
Is redescription ok?
N
Y
Consider the first/next suboperation
Is further redescription required?
Y
N
Terminate the redescription of this operation
Select next operation
Y
Are there anymore operations?
N
STOP

Example
The following example is a HTA of the task 'boil kettle'. This is typically the starting point in the training process of HTA.

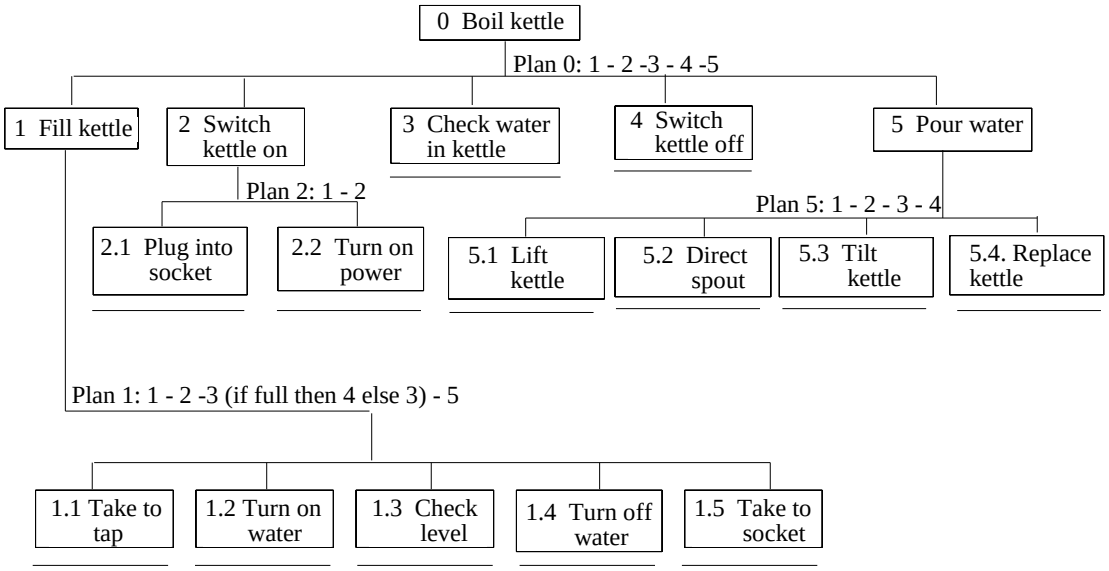


Figure 20. HTA of the task 'boil kettle'

Related Methods
HTA is widely used in HF and often forms the first step in a number of analyses, such as HEI, HRA and mental workload assessment. Annett (2003) reports that HTA has been used in a number of applications, for example as the first step in the TAFEI method for hazard and risk assessment (Baber & Stanton, 1994), in SHERPA for predicting human error (Baber & Stanton, 1996), in MUSE usability assessment (Lim & Long, 1994), the SGT method for specification of information requirements (Ormerod, Richardson & Shepherd, 1998/2000), and the TAKD method for the capture of task knowledge requirements in HCI (Johnson, Diaper & Long, 1984).

Approximate Training and Application Times
According to Annett (2003), a study by Patrick, Gregov and Halliday (2000) gave students a few hours training with not entirely satisfactory results on the analysis of a very simple task, although performance improved with further training. A survey by Ainsworth & Marshall (1998/2000) found that the more experienced practitioners produced more complete and acceptable analyses. Stanton & Young (1999) report that the training and application time for HTA is substantial. The application time associated with HTA is dependent upon the size and complexity of the task under analysis. For large, complex tasks, the application time for HTA would be high.

Reliability and Validity
According to Annett (2003), the reliability and validity of HTA is not easily assessed. Stanton & Young (1999) report that, in a comparison of twelve HF techniques, HTA achieved an acceptable level of validity and a poor level of reliability.

Tools needed.
HTA can be carried out using only pencil and paper.

Bibliography

Ainsworth, L. & Marshall, E. (1998) Issues of quality and practicality in task analysis: preliminary results from two surveys. Ergonomics 41(11), 1604-1617. Reprinted in J. Annett & N..A. Stanton (2000) op.cit. Pp. 79-89.

Annett, J., Duncan, K.D., Stammers, R.B. & Gray, M. (1971) Task Analysis. London: HMSO.

Annett, J., Cunningham, D.J., & Mathias-Jones, P. (2000). A method for measuring team skills. Ergonomics, 43(8), 1076-1094.

Annett, J. & Stanton, N.A. (Eds) (2000) Task Analysis. London: Taylor & Francis.

Baber, C. & Stanton, N.A. (1994) Task analysis for error identification. Ergonomics 37, 1923-1941.

Baber, C. & Stanton, N.A. (1996) Human error identification techniques applied to public technology: predictions compared with observed use. Applied Ergonomics 27, 119-131.

Johnson, P., Diaper, D. & Long, J. 1984. Tasks, skills and knowledge: Task analysis for knowledge-based descriptions. In B. Shackel (Ed.) Interact '84 - First IFIP Conference on Human-Computer Interaction. Amsterdam: Elsevier. Pp. 23-27.

Kirwan, B. & Ainsworth, L. (1992) A Guide to Task Analysis. London: Taylor & Francis.

Lim, K.Y. & Long, J. (1994) The MUSE Method for Usability Engineering. Cambridge: Cambridge University Press.

Ormerod, T.C., Richardson, J. & Shepherd, A. (1998) Enhancing the usability of a task analysis method: A notation and environment for requirements. Ergonomics 41(11), 1642-1663. Reprinted in Annett & Stanton (2000) op.cit. Pp. 114-135.

Patrick, J., Gregov, A. & Halliday, P. (2000) Analysing and training task analysis. Instructional Science, 28(4), 51-79.

Shepherd, A. (2002). Hierarchical Task Analysis. London: Taylor & Francis.

CPA - Critical Path Analysis for Multimodal Activity

Chris Baber, School of Electronic, Electrical & Computing Engineering, University of Birmingham, Edgbaston, Birmingham. B15 2TT, UK

Background and Applications

The idea of using time as the basis for predicting human activity has its roots in the early Twentieth Century; specifically in the "Scientific Management" of Fredrick Taylor (although the idea of breaking work into constituent parts and timing these parts can be traced to the Industrial Revolution in the Eighteenth Century). The basic idea of such approaches was to simplify work and then seek ways of making the work as efficient as possible, i.e., to reduce the time taken for each task-step and, as a consequence, to reduce the overall time for the activity. Obviously, such an approach is not without problems. For example, Taylor faced Presidential Select Committee hearings in the USA when workers rioted or went on strike in response to the imposition of his methods. At a more basic level, there is no clear evidence that there is 'one best way' to perform a sequence of tasks, and people often are adept in employing several ways. Thus, while the timing of task-steps can be seen as fairly straightforward, the combination of the task-steps into meaningful wholes is problematic.

In recent years, human-computer interaction has sought techniques that will allow 'modelling' of the interaction between user and computer in order to determine whether a proposed design will be worth developing. One such set of techniques involves breaking activity into discrete tasks and then defining times for these tasks. Combining the tasks into sequences would then result in a prediction of overall time for the sequence. This is basically the approach that the Keystroke-Level Model (see 'Related Methods' section).

Researchers have been investigating approaches that will allow them to combine discrete tasks in more flexible ways. One such approach draws on critical path analysis (CPA), which is a project management tool that is used to calculate the combination of tasks that will most affect the time taken to complete a job (see Harrison, 1997 or Lockyer and Gordon, 1994 for more detailed descriptions of CPA as a project management technique). Any change in the tasks on the 'critical path' will change the overall job completion time (and changes in tasks off the critical path, within limits, can be accommodated without problem). In the version presented in this chapter, the critical path is defined both in terms of time, so that a task will need to be completed before a subsequent task can begin, and modality, so that two tasks sharing the same modality must be performed in series.

One of the earliest studies that employed critical path analysis in HCI was reported by Gray et al. (1993) and Lawrence et al (1995). In this study, a telephone company wanted to re-equip its exchanges with new computer equipment. Critical path analysis was used to investigate the relationship between computer use and other activities in call handling. It was shown that computer use did not lie on the critical path, so investment in such equipment would not have improved performance.

Domain of application

HCI.

Procedure and advice

Step 1: Define tasks. This could take the form of a task analysis, or could be a simple decomposition of the activity into constituent tasks. Thus, the Activity of 'Accessing an automated teller machine' might consist of the following task steps: 1. Retrieve card from wallet, 2. Insert card into ATM, 3. Recall PIN, 4. Wait for screen to change, 5. Read prompt, 6. Type in digit of PIN, 7. Listen for confirmatory beep, 8. Repeat steps 6 and 7 for all digits in PIN, 9. Wait for screen to change.

Step 2: Define the tasks in terms of input and output sensory modality: Manual (left or right hand), Visual, Auditory, Cognitive, Speech. There will also be times associated with various system responses. Table 20 relates task step to modality. The table might require a degree of judgement from the analyst, e.g., some task steps might require more than modality or might not easily fit into the scheme. However, taking the dominant modality usually seems to work.

Table 20. Relating task step to modality

Task step	Manual-L	Manual-R	Speech	Auditory	Visual	Cognitive	System
Retrieve card	X	X					
Insert card		X					
Recall PIN						X	
Screen change							X
Read prompt					X		
Type digit		X					
Listen for beep				X			
Screen change							X

Step 3: Construct a chart showing the task sequence and the dependencies between tasks. As mentioned above, dependency is defined in terms of time, i.e., a specific task needs to be completed before another task can commence, and modality, i.e., two tasks in the same modality must occur in series. Figure 21 shows a chart for the worked example. The example takes the task sequence up to the first digit being entered, for reasons of space (the other four digits will need to be entered, with the user pausing for the 'beep' prior to the next digit, and the final screen change will occur for the sequence to be completed). In this diagram, an action-on-arrow approach is used. This means that each node is linked by an action, which takes a definable length of time. The nodes are numbered, and also have spaces to insert earliest start time and latest finish time (see step 5).

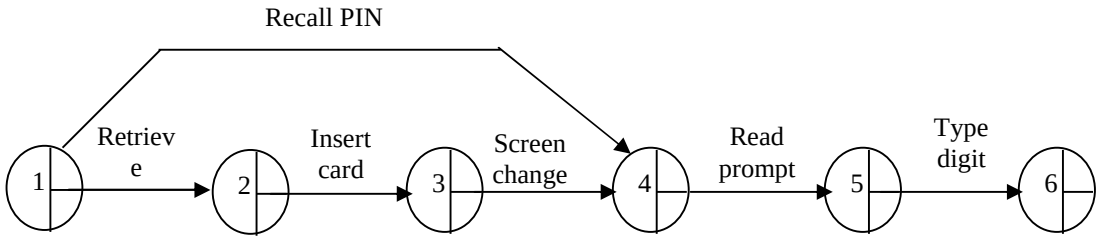


Figure 21. Initial part of CPA chart

Step 4: Assign times to the tasks.
Table 21 provides a set of times for the example. Appendix A provides a larger set of data. The diagram shown in figure 21 can be redrawn in the form of a table, which helps in the following steps (see table 21).

Step 5: Calculate forward pass. Begin at the first node of figure 21 and assign an earliest start time of 0. The finish time for task from this node will be 0 + the duration of the task step; in this case, 'retrieve card' takes 500ms, so the earliest finish time will be 500ms. Enter these values onto table 20, and move to the next node. The earliest finish time of one task becomes the earliest start time (EST) for the next task. A simple rule is to calculate Es on the forward pass. When more than one task feed into a node, take the highest time. Repeat the steps until you reach the last node.

Table 21. Critical path calculation table – forward pass

Task step	Duration	Earliest start	Latest start	Earliest finish	Latest finish	Float
Retrieve card	500ms	0		500		
Insert card	350ms	500		850		
Recall PIN	780ms	0		780		
Screen change	250ms	850		1100		
Read prompt	350ms	1100		1450		
Type digit	180ms	1450		1630		
Wait for beep	100ms	1630		1730		

Step 6: Calculate backward pass. Begin at the last node and assign a latest finish time (in this case, the time will equal the earliest finish time). To produce the latest start time, subtract the task duration from the latest finish time. The time on the connection becomes the latest finish time (LFT) for that task. When more than one task feed into a node, take the lowest time. Repeat the steps until you reach the first node.

Table 22. Critical path calculation table

Task step	Duration	Earliest start	Latest start	Earliest finish	Latest finish	Float
Retrieve card	500ms	0	0	500	500	0
Insert card	350ms	500	500	850	850	0
Recall PIN	780ms	0	320	780	1100	320
Screen change	250ms	850	850	1100	1100	0
Read prompt	350ms	1100	1100	1450	1450	0
Type digit	180ms	1450	1450	1630	1630	0
Wait for beep	100ms	1630	1630	1730	1730	0

Step 7: Calculate critical path. The critical path consists of all nodes that have zero difference between EST and LFT. In this example, the task step on 'recall PIN' has none-zero float, which means that it can be started up to 320ms into the other tasks without having an impact on total task performance. It is possible to perform the calculations using commercial software, such as MicroSoft Project (although this works in terms of days, hours, and months rather than milliseconds or seconds, so can produce some misleading calculations unless you set all of the parameters appropriately). Alternatively, you can perform the calculations using MicroSoft Excel (see Appendix B).

Advantages

- CPA allows the analyst gain a better understanding of the task via splitting the task into the activities that need to be carried out in order to ensure successful task completion.
- CPA allows the consideration of parallel unit task activity (Baber and Mellor, 2001), KLM does not.
- CPA gives predicted performance task times for the full task and also for each task step.
- CPA determines a logical, temporal description of the task in question.
- CPA does not require a great deal of training
- Structured and comprehensive procedure
- Can accommodate parallelism in user performance
- Provides reasonable fit with observed data
- Olson and Olson (1990) suggest that CPA can be used to address the shortcomings of KLM.

Disadvantages

- Can be tedious and time consuming for complex tasks
- CPA only models error free performance and cannot deal with unpredictable events such as the ones seen in man-machine interactions.
- Modality can be difficult to define
- Can only be used for activities that can be described in terms of performance times
- Times not available for all actions
- Can be overly reductionistic, particularly for tasks that are mainly cognitive in nature.

Related methods

The earliest, and most influential, model of transaction time was the Keystroke Level Model (Card et al., 1983). The Keystroke Level Model (KLM) sought to decompose human activity into unit-tasks and to assign standard times to each of these unit-tasks. Transaction time was calculated by summing all standard-times. KLM represents a particular approach to HCI, which can be thought of as reducing humans to engineering systems, i.e., with 'standardised', predictable actions, which can be assigned standard times. KLM has proven to be effective at predicting transaction time, within acceptable limits of tolerance, e.g., usually within 20% of the mean time observed from human performance (Card et al., 1983; Olson and Olson, 1990). However, there are a number of criticisms that have been levelled at KLM, including the following:

- KLM assumes 'expert' performance, where the definition of an 'expert' is a person who uses the most efficient strategy to perform a sequence of unit-tasks and who works as fast as possible without error;
- KLM ignores flexibility in human activity;
- KLM ignores other unit-task activity or variation in performance;
- KLM assumes that unit-tasks are combined in series, i.e., that performance is serial and that there is no parallel activity.

The first criticism has been the subject of much discussion; experts are users with a wide repertoire of methods and techniques for achieving the same goal, rather than people programmed with a single efficient procedure. Thus, a technique that reduces performance to a simple, linear description will obviously miss the variability and subtlety of human performance. Furthermore, non-expert users will typically exhibit a wide variety of activity, and the notion that this activity can be reduced to 'one-best way' is questionable.

The main response to the second criticism is that the approach seeks to produce 'engineering approximations' of human performance, rather than a detailed description (Card et al., 1983). As such, the approach can be considered as a means of making task analysis 'dynamic' (in the sense that times can be applied to unit-tasks in order to predict the likely performance time of a sequence of such unit-tasks). This shifts the debate from the utility of KLM per se and onto the inherent reductionism of task analysis techniques. Recent discussions of human-computer interaction have tended to focus on the broad range of issues associated with the context of HCI, and have argued against descriptions which focus too narrowly on one-person using one-computer. It is proposed that a requirement of user modelling techniques ought to be that they can adequately reflect that range of activities that a user performs, giving the context of work. Consequently, KLM might be too narrowly focused on one-user performing one-task using one-computer (following one-best way of working), and alternative methods should be developed too rectify these problems.

The third criticism has been the subject of less debate, although there have been attempts to capture performance variation. Researchers have examined how systems respond to definable variability in performance. For example, speech recognition systems can be defined by their recognition accuracy, and it is important to know how variation in recognition accuracy can influence system efficiency. Rudnicky and Hauptmann (1991) have used Markov models to describe HCI, working from the assumption that dialogues progress through a sequence of states, and that each state can be described by its duration. By varying state transition parameters, it is possible to accommodate variation in recognition accuracy of speech recognizers. Ainsworth (1988) employs a slightly different technique to the same end. His work models the impact of error correction and degradation of recognition accuracy on transaction time. We have used unit-task-network models (specifically MicroSaint) to investigate error correction and the effects of constraint on speech-based interaction with computers (Hone and Baber, 1999). Examination of the issues surrounding the combination of unit-times for prediction of human performance raises questions concerning the scheduling of unit-tasks and the coordination of activity. It also leads to concerns over how unit-tasks might be performed in parallel (which relates to the fourth criticism).

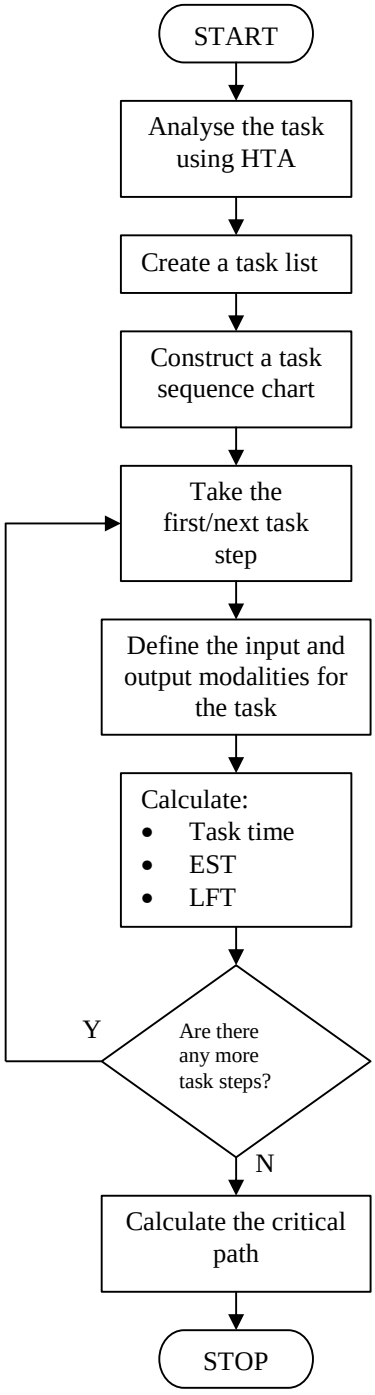
Approximate training and application times

Although no data regarding the training and application time of CPA is available, it is suggested that the training time would be low, and that the application time would also be low, although this is dependent upon the task under analysis. For complex, larger tasks, the application time would be high.

Reliability and Validity

Baber and Mellor (2001) compared predictions using critical path analysis with the results obtained from user trials, and found that the 'fit' between observed and predicted values had an error of less than 20%. This suggests that the approach can provide robust and useful approximations of human performance.

Flowchart



Tools needed
CPA can be conducted using pen and paper.

Bibliography
Ainsworth, W. (1988) Optimization of string length for spoken digit input with error Correction. International Journal of Man Machine Studies 28 573-581
Baber, C. and Mellor, B.A. (2001) Modelling multimodal human-computer interaction using critical path analysis, International Journal of Human Computer Studies 54 613-636
Card, S.K., Moran, T.P. and Newell, A. (1983) *The Psychology of Human-Computer Interaction* Hillsdale, NJ: LEA
Gray, W.D., John, B.E. and Atwood, M.E. (1993) Project Ernestine: validating a GOMS analysis for predicting and explaining real-world performance *Human-Computer Interaction 8* 237-309
Harrison, A. (1997) *A Survival Guide to Critical Path Analysis* London: Butterworth-Heinemann
Hone, K.S. and Baber, C. (1999) Modelling the effect of constraint on speech-based human computer interaction *International Journal of Human Computer Studies 50* 85-105
ISO 9241 (1998) *Ergonomics of office work with VDTs-guidance on usability*, Geneva: International Standards Office
ISO 13407 (1999) *Human-centred design processes for interactive systems*, Geneva: International Standards Office
ISO 9126 (2000) *Software engineering - product quality*, Geneva: International Standards Office
Lawrence, D., Atwood, M.E., Dews, S. and Turner, T. (1995) Social interaction in the use and design of a workstation: two contexts of interaction In P.J. Thomas (ed) *The Social and Interactional Dimensions of Human-Computer Interaction* Cambridge: Cambridge University Press 240-260
Lockyer, K. and Gordon, J. (1991) *Critical Path Analysis and Other Project Network Techniques* London: Pitman
Olson, J.R. and Olson, G.M. (1990) The growth of cognitive modelling in human-computer interaction since GOMS *Human-Computer Interaction 3* 309-350

GOMS – Goals, Operators, Methods and Selection Rules
Card, Moran & Newell (1983)

Background and applications
The GOMS technique is part of a family of HCI orientated techniques that is used to provide a description of human performance in terms of the user's goals, operators, methods and selection rules. GOMS attempts to define the user's goals, decompose these goals into sub-goals and demonstrate how the goals are achieved through user interaction. GOMS can be used to provide a description of how a user performs a task, to predict performance times and to predict human learning. Whilst the GOMS techniques are most commonly used for the evaluation of existing designs or systems, it is also feasible that they could be used to inform the design process, particularly to determine the impact of a design on the user. Within the GOMS family, there are four techniques. The four GOMS techniques are described below:

- NGOMSL
- KLM
- CMN-GOMS
- CPM-GOMS

The GOMS techniques are based upon the assumption that the user's interaction with a computer is similar to solving problems. Problems are broken down into sub-problems, and these sub-problems are broken down further. Four basic components of human interaction are used within the GOMS technique. These are defined below:
1) *Goals* – The goal represents exactly what the user wishes to achieve through the interaction. The goals are decomposed until an appropriate stopping point is achieved.
2) *Operators* – The operators are the motor or cognitive actions that the user performs during the interaction. The goals are achieved through performing the operators.
3) *Methods* – The methods describe the user's procedures for accomplishing the goals in terms of operators and sub-goals. Often there are more than one set of methods available to the user.
4) *Selection Rules* – When there is more than one method for achieving a goal available to a user, selection rules highlight which of the available methods should be used.

Domain of application
HCI.

Procedure and advice
Step 1: Define the user's top-level goals
Firstly, the analyst should describe the user's top-level goals. Kieras (2003) suggests that the top-level goals should be described at a very high level. This ensures that any methods are not left out of the analysis.

Step 2: Goal decomposition
Once the top-level goal or set of goals has been specified, the next step is to break down the top-level goal into a set of sub-goals. According to Kieras (2003) the analyst should always assume that each top-level goal is achieved through the performance of a series of smaller steps.

Step 3: Describe operators
Operators are actions executed by the user to achieve a goal or sub-goal. In the next stage of the GOMS analysis, each goal/sub goal should be considered and high level operators described. Each high level operator should be replaced with another goal/method set until the analysis is broken down to the level desired by the analyst (Kieras 2003).

Step 4: Describe methods
Methods describe the procedures or set of procedures used to achieve the goal (Kirwan and Ainsworth 1992). In this stage of the GOMS analysis, the analyst should describe each set of methods that the user could use to achieve the task. Often there are a number of different methods available to the user, and the analyst is encouraged to include all possible methods.

Step 5: Describe selection rules
If there is more than one method of achieving a goal, then the analyst should determine selection rules for the goal. Selection rules predict which of the available methods will be used by the user to achieve the goal.

Advantages
- GOMS can be used to provide a hierarchical description of task activity.
- The methods part of a GOMS analysis allows the analyst to describe a number of different potential task routes.
- GOMS analysis can aid designers in choosing between systems, as performance and learning times can be specified.

Disadvantages
- GOMS is a difficult technique to apply. Far simpler task analysis techniques are available.
- Time consuming.
- Appears to be restricted to HCI. As it was developed specifically for use in HCI, most of the language is HCI orientated.
- A high level of training and practice would be required.
- GOMS does not deal with error occurrence.
- GOMS analysis is limited as it only models error-free, expert performance.
- Context is not taken into consideration.
- The GOMS methods remain largely unvalidated outside of HCI.

Example
The following example is taken from Card, Moran & Newell (1983).

GOAL: EDIT-MANUSCRIPT
. GOAL: EDIT-UNIT-TASK *repeat until no more unit tasks*
. . GOAL: ACQUIRE-UNIT-TASK
. . . GET-NEXT-PAGE *if at end of manuscript*
. . . GET-NEXT-TASK
. . GOAL: EXECUTE-UNIT-TASK
. . . GOAL: LOCATE-LINE
. . . .(Select: USE-QS-METHOD
: USE-LF-METHOD
. . . GOAL: MODIFY-TEXT
. . . . (Select: USE-S-COMMAND
: USE-M-COMMAND
. . . . VERIFY-EDIT

Related methods
There are four main techniques within the GOMS family. These are NGOMSL, KLM, CMN-GOMS and CPM-GOMS.

Approximate training and application times
For non-HCI experienced practitioners, it is hypothesised that the training time would be medium to high. The application time associated with the GOMS technique is dependent upon the size and complexity of the task under analysis. For large, complex tasks involving many operators and methods, the application time for GOMS would be very high. However, for small, simplistic tasks the application time would be minimal.

Reliability and validity
The use of GOMS in HCI has been validated extensively. According to Salvendy (1997) Card et al (1983) reported that for a text-editing task, the GOMS technique predicted the user's methods 80-90% of the time and also the user's operators 80-90% of the time. The validation of the GOMS technique in applications outside of the HCI domain is limited.

Tools needed
GOMS can be conducted using pen and paper. The system/programme or device under analysis is required.

Bibliography
Card, S. K., Moran, T. P.; Newell, A. (1983) The Psychology of Human Computer Interaction. Lawrence Erlbaum Associates.
Kieras, D. (2003) GOMS Models for Task Analysis. In D. Diaper & N. Stanton (Eds) *The Handbook of Task Analysis for Human-Computer Interaction*. Pp 83-117. Lawrence Erlbaum Associates.
Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. Taylor and Francis, UK.

Salvendy, G. (1997) Handbook of Human Factors and Ergonomics. 2nd Edition. Canada, John Wiley and Sons.

VPA - Verbal Protocol Analysis
Various

Background and applications
Verbal protocol analysis (VPA) is used to make 'valid inferences' from the content of discourse (Weber 1990). In other words, VPA is used to derive the processes, cognitive and physical, that an individual uses to perform a task. VPA involves creating a written transcript of operator behaviour as they perform the task under analysis. The transcript is based upon the operator 'thinking aloud'. VPA has been used extensively within human factors as a means of gaining an insight into the cognitive aspects of complex behaviours. Walker (In Press) reports VPA's use in areas as diverse as steel melting (Bainbridge 1974), Internet usability (Hess 1999) and driving (Walker, Stanton & Young 2001).

Domain of application
Generic.

Procedure and advice
There are no 'set' rules as such for conducting a verbal protocol analysis. The following procedure is an adaptation of the procedure recommended by Walker (In Press).

Step 1: Define scenario under analysis
Firstly, the scenario to be analysed should be determined. A HTA is often used at this stage, in order to specify which tasks are to be analysed. In a study by Walker, Stanton & Young (2001) participants were required to drive a vehicle around a pre-determined test route. In the analysis of control room operations, analysing a set of representative scenarios may be useful.

Step 2: Instruct/Train the participant
Once the scenario is set, the participant should be briefed regarding what is required of them during the analysis. What they should report verbally is clarified here. Walker (In Press) suggests that most importantly, the participant should be informed that they should continue talking even when what they are saying does not seem to make much sense. A small demonstration should also be given to the participant at this stage. A practice run may also be undertaken, although this is not always necessary.

Step 3: Begin scenario and record data
The participant should begin to perform the scenario under analysis. The whole scenario should be audio recorded by the analyst. It is also recommended that a video recording is made.

Step 4: Verbalisation of transcript
Once collected, the data should be transcribed into a written form. An excel spreadsheet is normally used. This aspect of VPA is particularly time consuming and laborious.

Step 5: Encode verbalisations
The verbal transcript (written form) should then be categorised or coded. Depending upon the requirements of the analysis, the data is coded into one of the following five categories; words, word senses, phrases, sentences or themes. The encoding scheme chosen should then be encoded according to a rationale determined by the aims of the research being undertaken. Walker (In Press) suggests that this should involve attempting to ground the encoding scheme according to some established theory or approach, such as mental workload or situation awareness. The analyst should also develop a set of written instructions for the encoding scheme. These instructions should be strictly adhered to and constantly referred to during the encoding process (Walker In Press). Once the encoding type, framework and instructions are completed, the analyst should proceed to encode the data. Various computer software packages are available to aid the analyst with this process, such as General Enquirer, TextQuest and Wordstation.

Step 6: Devise other data columns
Once the encoding is complete, the analyst should devise any 'other' data columns. This allows the analyst to note any mitigating circumstances that may have affected the verbal transcript.

Step 7: Establish Inter and Intra-rater reliability
Reliability of the encoding scheme then has to be established (Walker In Press). In VPA, reliability is established through reproducibility i.e. independent raters need to encode previously analyses.

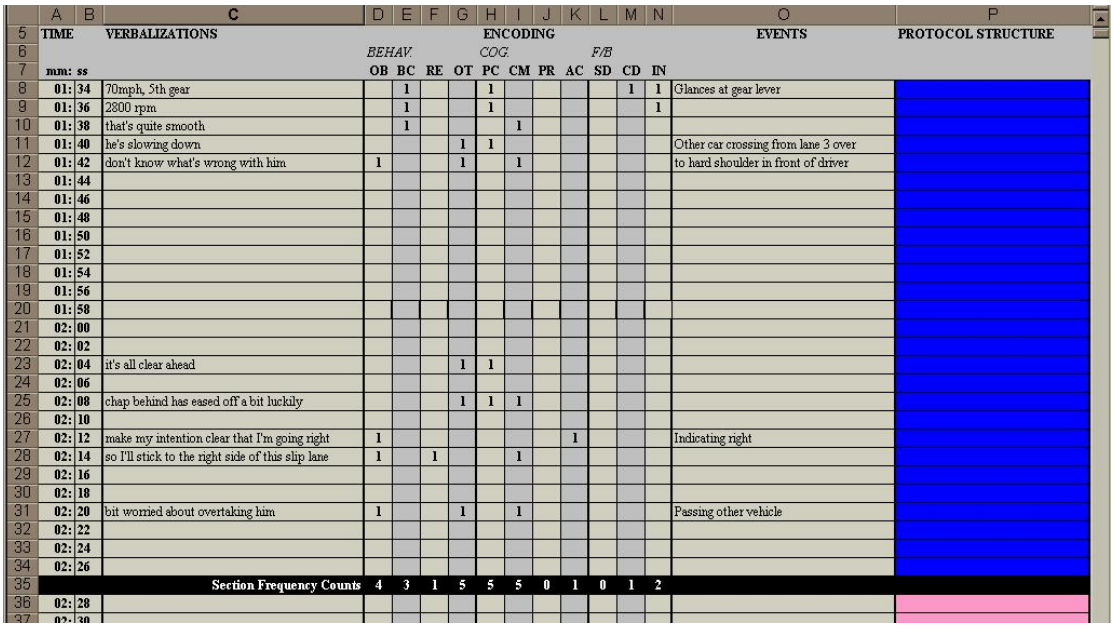
Step 8: Perform pilot study
The protocol analysis procedure should now be tested within the context of a small pilot study. This will demonstrate whether the verbal data collected is useful, whether the encoding system works, and whether inter and intra rater reliability are satisfactory. Any problems highlighted through the pilot study should be refined before the analyst conducts the VPA for real.

Step 9: Analyse structure of encoding
Finally, the analyst can analyse the results from the VPA. During any VPA analysis the responses given in each encoding category require summing, and this is achieved simply by adding up the frequency of occurrence noted in each category. Walker (In Press) suggests for a more fine grained analysis, the structure of encodings can be analysed contingent upon events that have been noted in the 'other data' column(s) of the worksheet, or in light of other data that have been collected simultaneously.

Example
The following example is a VPA taken from Walker (In Press).
Example of Protocol Analysis Recording, Transcription and Encoding Procedure for an On-Road Driving Study

Figure 22. Digital Audio/Video Recording of Protocol Analysis Scenario

This digital video image (figure 21) is taken from the study reported by Walker, Stanton, and Young (2001) and shows how the Protocol Analysis was performed with normal drivers. The driver in Figure 21 is providing a concurrent verbal protocol whilst being simultaneously videoed. The driver's verbalisations and other data gained from the visual scene are transcribed into the transcription sheet in Figure 22.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
5	TIME		VERBALIZATIONS	ENCODING											EVENTS	PROTOCOL STRUCTURE
6				BEHAV.				COG.				F/B				
7	mm:	ss		OB	BC	RE	OT	PC	CM	PR	AC	SD	CD	IN		
8	01:	34	70mph, 5th gear		1			1					1	1	Glances at gear lever	
9	01:	36	2800 rpm		1			1						1		
10	01:	38	that's quite smooth		1				1							
11	01:	40	he's slowing down				1	1							Other car crossing from lane 3 over	
12	01:	42	don't know what's wrong with him	1			1		1						to hard shoulder in front of driver	
13	01:	44														
14	01:	46														
15	01:	48														
16	01:	50														
17	01:	52														
18	01:	54														
19	01:	56														
20	01:	58														
21	02:	00														
22	02:	02														
23	02:	04	it's all clear ahead				1	1								
24	02:	06														
25	02:	08	chap behind has eased off a bit luckily				1	1	1							
26	02:	10														
27	02:	12	make my intention clear that I'm going right	1							1				Indicating right	
28	02:	14	so I'll stick to the right side of this slip lane	1		1			1							
29	02:	16														
30	02:	18														
31	02:	20	bit worried about overtaking him	1			1		1						Passing other vehicle	
32	02:	22														
33	02:	24														
34	02:	26														
35			Section Frequency Counts	4	3	1	5	5	5	0	1	0	1	2		
36	02:	28														

Figure 23. Transcription and Encoding Sheet

Figure 22 illustrates the 2-second incremental time index, the actual verbalisations provided by the driver's verbal commentary, the encoding categories, the events column and the protocol structure.

In this study three encoding groups were defined: behaviour, cognitive processes, and feedback. The behaviour group defined the verbalisations as referring to the driver's own behaviour (OB), behaviour of the vehicle (BC), behaviour of the road environment (RE), and behaviour of other traffic (OT). The cognitive processes group was sub divided into perception (PC), comprehension (CM), projection (PR), and action execution (AC). The feedback category offered an opportunity for vehicle feedback to be further categorised according to whether it referred to system or control dynamics (SD or CD), or vehicle instruments (IN). The cognitive processes and feedback encoding categories were couched in relevant theories in order to establish a conceptual framework. The events column was for noting road events from the simultaneous video log, and the protocol structure was colour coded according to the road type being travelled upon. In this case the shade corresponds to a motorway, and would permit further analysis of the structure of encoding contingent upon road type. The section frequency counts simply sum the frequency of encoding for each category for that particular road section.

Advantages

- Verbal protocol analysis provides a rich data source.
- Protocol analysis is particularly effective when used to analyse sequences of activities.
- Verbalisations can provide a genuine insight into cognitive processes.
- Domain experts can provide excellent verbal data.
- Verbal protocol analysis has been used extensively in a wide variety of domains.
- Simple to conduct with the right equipment.

Disadvantages

- Data analysis (encoding) can become extremely laborious and time consuming.
- Verbal Protocol Analysis is a very time consuming method to apply (data collection and data analysis).
- It is difficult to verbalise cognitive behaviour. Researchers have been cautioned in the past for relying on verbal protocol data (Militello & Hutton 2000).
- Verbal commentary can sometimes serve to change the nature of the task.
- Complex tasks involving high demand can often lead to a reduced quantity of verbalisations (Walker In Press).
- Strict procedure is often not adhered to fully.
- VPA is prone to bias on the participant's behalf.

Related methods

Verbal protocol analysis is related to observational techniques such as walkthroughs and direct observation. Task analysis techniques such as HTA are often used in constructing the scenario under analysis.

Approximate training and application times

Although the technique is very easy to train, VPA can be very time consuming in its application. Walker (In Press) suggests that if transcribed and encoded by hand, 20 minutes of verbal transcript data at around 130 words per minute can take between 6 to 8 hours to transcribe and encode.

Reliability and validity

Walker (In Press) suggests that the reliability of the technique is reassuringly good. For example, Walker, Stanton and Young (2001) used two independent raters and established inter-rater reliability at Rho=0.9 for rater 1 and Rho=0.7 for rater 2. Intra rater reliability during the same study was also high, being in the region of Rho=0.95.

Tools needed

A VPA can be conducted using pen and paper, a digital audio recording device and a video recorder if required. The device/system under analysis is also required. In analysing the data obtained in the VPA, Microsoft Excel is normally required, although this can be done using pen and paper. A number of software packages can also be used by the analyst, including Observer, General Enquirer, TextQuest and Wordstation.

Bibliography

Walker, G. H (In Press) Verbal Protocol Analysis. In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.

Walker, G.H., Stanton, N.A., & Young, M.S. (2001). An on-road investigation of vehicle feedback and its role in driver cognition: Implications for cognitive ergonomics. *International Journal of Cognitive Ergonomics*, 5(4), 421-444

Weber, R.P. (1990). Basic Content Analysis. Sage Publications, London.

Wilson, J.R., & Corlett, N.E. (1995). Evaluation of Human Work: A Practical Ergonomics Methodology. Taylor and Francis, London.

Task Decomposition
Barry Kirwan
L. K. Ainsworth

Background and applications
Kirwan and Ainsworth (1992) present an overview of a task decomposition methodology (also known as tabular task analysis) that can be used to gather a detailed task description regarding a particular task. Task decomposition begins with a task description, such as a HTA describing how each step of the task under analysis is performed. The analyst then gathers further information about specific aspects of each task step (such as time taken, controls used, cues initiating each action etc). The information for each of the task steps can then be presented using a set of sub-headings. This allows the relevant information for each task step to be decomposed into a series of statements regarding the task (Kirwan and Ainsworth 1992). The categories used to decompose the task steps should be chosen by the analyst based on the requirements of the analysis. There are numerous decomposition categories that can be used and new categories can be developed if required by the analysis. According to Kirwan and Ainsworth (1992), Miller (1953) was the first practitioner to use the task decomposition technique. Miller (1953) suggested that each task step should be decomposed around the following categories:

- Description
- Subtask
- Cues initiating action
- Controls used
- Decisions
- Typical errors
- Response
- Criterion of acceptable performance
- Feedback

This set of decomposition categories appears dated and inadequate for an analysis of command and control systems. It is recommended that the analyst should develop a set of specific categories for the system under analysis. The task decomposition technique can be used at any stage in the design process, either in the early design stages to provide a detailed task analysis and determine which aspects of the task require further system design inputs or to evaluate existing operational systems or devices.

Domain of application
Generic.

Procedure and advice
Step 1: Hierarchical task analysis
The first step in a task decomposition analysis involves creating an initial task description of the task under analysis. For this purpose it is recommended that HTA is used. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe

the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Step 2: Create task descriptions
Once an initial HTA for the task under analysis has been conducted, the analyst should create a set of clear task descriptions for each of the different task steps. These descriptions can be derived from the HTA. The task description should give the analyst enough information to determine exactly what has to be done to complete each task element. The detail of the task descriptions should determined by the requirements of the analysis.

Step 3: Choose decomposition categories
Once a sufficient description of each task step is created, the analyst should choose the appropriate decomposition categories. Kirwan and Ainsworth (1992) suggest that there are 3 types of decomposition categories – descriptive, organisation-specific and modelling. Table 22 contains a taxonomy of descriptive decomposition categories that have been used in various studies (Kirwan and Ainsworth, 1992).

Table 23. Table showing decomposition categories (Source: Kirwan & Ainsworth 1992)

Description of task	Task difficulty
Description	Task criticality
Type of activity/behaviour	Amount of attention required
Task/action verb	**Performance on the task**
Function/purpose	Performance
Sequence of activity	Time taken
Requirements for undertaking task	Required speed
Initiating cue/event	Required accuracy
Information	Criterion of response adequacy
Skills/training required	**Other activities**
Personnel requirements/manning	Subtasks
Hardware features	Communications
Location	Co-ordination requirements
Controls used	Concurrent tasks
Displays used	**Outputs from the task**
Critical values	Output
Job aids required	Feedback
Nature of the task	**Consequences/Problems**
Actions required	Likely/typical errors
Decisions required	Errors made/problems
Responses required	Error consequences
Complexity/Task complexity	Adverse conditions/hazards

Step 4: Information collection
Once the decomposition categories have been chosen, the analyst should create an information collection form for each decomposition category. The analyst should then work through each of these forms, recording task descriptions and gathering the additional information required for each of the decomposition headings. To gather this information, Kirwan and Ainsworth (1992) suggest that there are many possible

methods to use, including observation, system documentation, procedures, training manuals and discussions with system personnel and designers. VPA and walkthrough analysis can also be used.

Step 5: Construct task decomposition
The analyst should then put the collected data into a task decomposition. The table will be made up of all of the decomposition categories chosen for the analysis. The detail included in the table is also determined by the scope of the analysis.

Advantages
- Through choosing which decomposition categories to use, the analyst can determine the direction of the analysis.
- Flexible technique, allowing any factors associated with the task to be assessed.
- A task decomposition analysis has the potential to provide a very comprehensive analysis of a particular task.
- The structure of the method ensures that all issues of interest are considered and evaluated for each of the task steps (Kirwan and Ainsworth, 1992).
- The method is entirely generic and can be used in any domain.
- Task decomposition provides a much more detailed description of tasks than traditional task analysis techniques do.
- As the analyst has control over the decomposition categories used, potentially any aspect of a task can be evaluated. In particular, the technique could be adapted to assess the cognitive components associated with tasks (goals, decisions, SA).
- Potentially extremely exhaustive, if the correct decomposition categories are used.

Disadvantages
- As the task decomposition is potentially so exhaustive, it is a very time consuming technique to apply and analyse. The HTA only serves to add to the high application time. Furthermore, obtaining information about the tasks (observation, interview etc) creates even more work for the analyst.
- Task decomposition can be laborious to perform, involving observations, interviews etc.
- The development of decomposition categories would also add further time costs. For use in command and control military environments, it is apparent that a set of categories would have to be developed.

Example
A task decomposition was performed on the landing task, “Land at New Orleans using the autoland system”. An extract of the analysis is shown below. Data collection included the following:
- Walkthrough of the flight task.
- Questionnaire administered to A320 pilots.
- Consultation with training manuals.
- Performing the flight task in aircraft simulator
- Interview with A320 pilot.

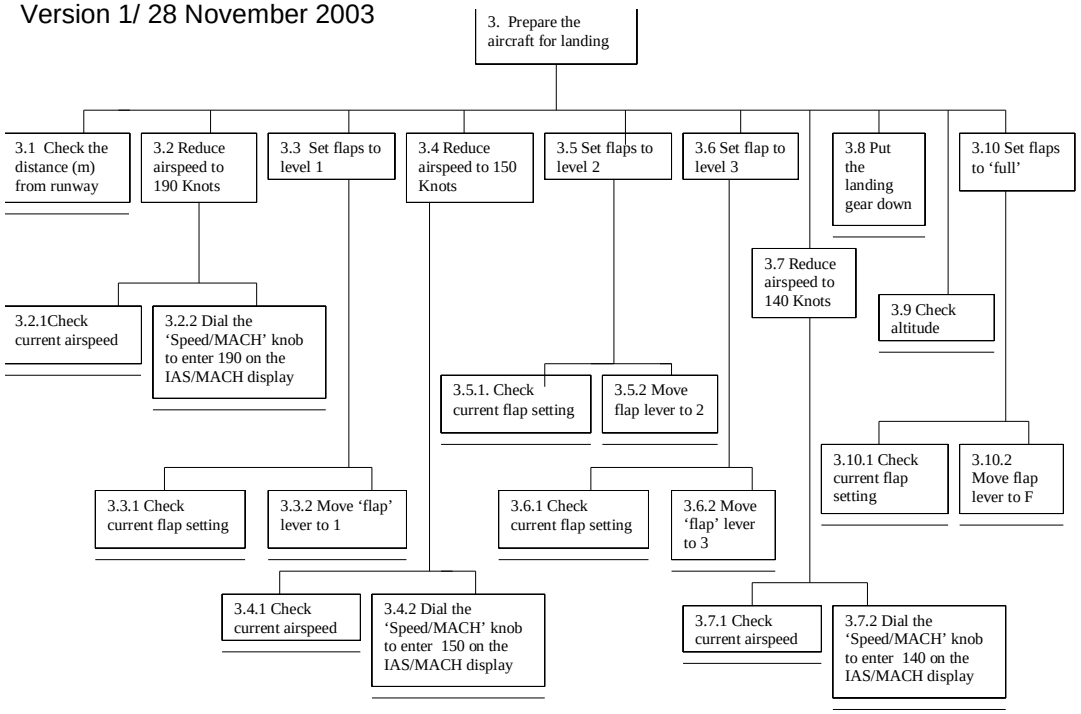
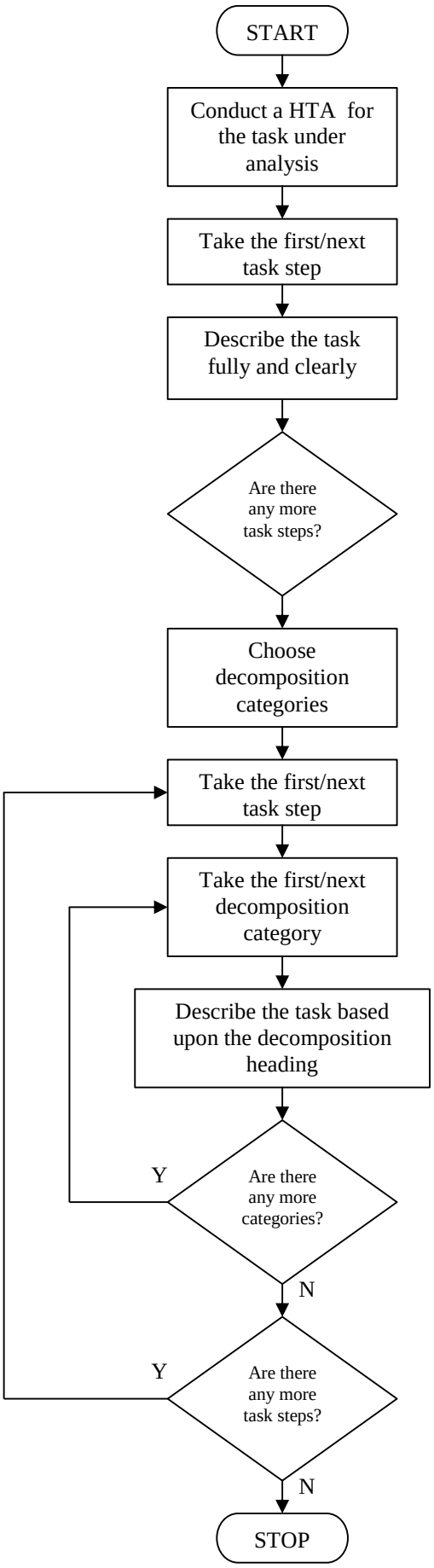


Figure 24. Extract of HTA 'Land at New Orleans using auto-land system'

Table 24. Extract of task decomposition analysis for flight task 'Land at New Orleans using the autoland system'

Task description 3.2.2 Dial the speed/MACH knob to enter 190 knots on the IAS/MACH display	**Complexity** Medium. The task involves a number of checks in quick succession and also the use of the Speed/MACH knob, which is very similar to the HDG/Track knob.
Initiating cue/event Check that the distance from the runway is 15 miles	**Difficulty** Low
Displays used Captains Primary Flight display IAS/MACH window (Flight control unit) Captains navigation display	**Criticality** High. The task is performed in order to reduce the aircrafts speed so that the descent and approach can begin.
Controls used IAS/MACH Knob	**Feedback provided** Speed/MACH window displays current airspeed value. CPFD displays airspeed.
Actions required Check distance from runway on CPFD Dial in 190 using the IAS/MACH display Check IAS/MACH window for speed value	**Probable errors** a) Using the wrong knob i.e. the HDG/Track knob b) Failing to check the distance from runway c) Failing to check current airspeed d) Dialling in the wrong speed value e) Fail to enter new airspeed
Decisions required Is distance from runway 15 miles or under? Is airspeed over/under 190knots? Have you dialled in the correct airspeed (190Knots)? Has the aircraft slowed down to 190knots?	**Error consequences** a) Aircraft will change heading to 190 b) Aircraft may be too close or too far way from the runway c) Aircraft travelling at the wrong airspeed d) Aircraft may be travelling to fast for the approach

Flowchart


START
Conduct a HTA for the task under analysis
Take the first/next task step
Describe the task fully and clearly
Are there any more task steps?
Choose decomposition categories
Take the first/next task step
Take the first/next decomposition category
Describe the task based upon the decomposition heading
Are there any more categories?
Y
N
Are there any more task steps?
Y
N
STOP

Related Methods
The task decomposition technique relies on a number of separate methods for its input. The initial task description required is normally provided by a HTA for the task under analysis. Data collection for the task decomposition analysis can involve any number of ergonomics methods. Normally, observational techniques, interviews, walkthrough and questionnaire type analyses are used in a task decomposition analysis. Task decomposition is primarily a task analysis technique.

Approximate training and application times
As a number of techniques are used within a task decomposition analysis, the training time associated with the technique is high. Not only would an inexperienced practitioner require training in the task decomposition technique itself (which incidentally would be minimal), but they would also require training in HTA and any techniques that would be used in the data collection part of the analysis. Also, due to the exhaustive nature of a task decomposition analysis, the associated application time is also very high. Kirwan and Ainsworth (1992) suggest that task decomposition can be a lengthy process and that its main disadvantage is the huge amount of time associated with collecting the required information.

Reliability and validity
At present, no data regarding the reliability and validity of the technique is offered in the literature. It is apparent that such a technique may suffer from reliability problems, in terms of eliciting the same data during different analysis of similar systems.

Tools needed
The tools needed for a task decomposition analysis are determined by the scope of the analysis and the techniques used for the data collection process. Task decomposition is primarily a pen and paper technique. For the data collection process, visual and audio recording equipment would be required. The system under analysis is required in some form, either in mock up, prototype or operational form.

Bibliography
Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. Taylor and Francis, UK.

Cognitive Task Analysis techniques

Operators of complex dynamic systems face an increasing demand upon their cognitive skills and resources. As system complexity increases, operators require training in specific cognitive skills and processes in order to keep up. System designers require an analysis of the cognitive skills and demands associated with the operation of the system under design in order to propose design concepts, allocate tasks, develop training procedures and to evaluate operator competence. Traditional task analysis techniques such as HTA only cater for the observable actions exhibited by system operators. As a result, a number of techniques have been developed in order to aid the HF practitioner in evaluating and describing the cognitive processes involved in system operation. Cognitive task analysis (CTA) techniques are used to describe and represent the unobservable cognitive aspects of task performance. Militello & Hutton (2000) suggest that CTA techniques focus upon describing and representing the cognitive elements that underlie goal generation, decision-making and judgements. CTA techniques are used to describe the mental processes used by system operators in completing a task or set of tasks. According to Chipman, Schraagen & Shalin (2000), CTA is an extension of traditional task analysis techniques used to describe the knowledge, thought processes and goal structures underlying observable task performance. CTA output is often used to inform the design of systems, processes and training procedures. Typical CTA techniques use observational, interview and questionnaire techniques in order to elicit specific data regarding the mental processes used by system operators. The use of CTA techniques is widespread, with applications in a number of domains, including firefighting (Militello & Hutton 2000), aviation (O'Hare et al 2000), nuclear power plant operation, emergency services (O'Hare et al 2000), air traffic control, military operations and even white-water rafting (O'Hare et al 2000).

Flanagan (1954) first probed the decisions and actions taken by pilots in near accidents using the critical incident technique. Klein (1989) proposed the critical decision method (CDM), which is a development of the critical incident technique, and uses cognitive probes to analyse decision-making during non-routine incidents (Klein 1989). Probes such as 'What were your specific goals at the various decision points?' 'What features were you looking for when you formulated your decision?' and 'How did you know that you needed to make the decision?' are used to analyse operator decisions during non-routine events. Applied Cognitive Task Analysis (ACTA) (Millitello & Hutton 2000) is a toolkit of interview techniques that can be used to elicit information regarding cognitive demands associated with the task or scenario under analysis. The ACTA framework can be used to determine the cognitive skills and demands associated with a particular task or scenario. The cognitive walkthrough technique (Polson et al 1992) focuses upon the usability of an interface, in particular the ease of learning associated with the interface. Based upon traditional design walkthrough techniques and a theory of exploratory learning (Polson and Lewis), the cognitive walkthrough technique consists of a set of criteria that the analyst must evaluate each task and the interface under analysis against. These criteria focus on the cognitive processes required to perform the task (Polson et al 1992).

The main problem associated with the use of cognitive task analysis techniques is the considerable amount of resources required. Using techniques such as interviews and observations, CTA techniques require considerable time and effort to conduct.

Access to SME's is also required, as is great skill on the analyst's behalf. CTA techniques are also criticised for their reliance upon the recall of events or incidents from the past. Klein (2003) suggests that methods that analyse retrospective incidents are associated with concerns of data reliability, due to evidence of memory degradation.

The following CTA techniques are reviewed in this document.

1. ACTA – Applied Cognitive Task Analysis
2. CDM – Critical Decision Method
3. Cognitive Walkthrough
4. Critical Incident Technique

CTA techniques will be employed during the design process of the C4i system. CTA will be used to describe the current mental processes required in existing command and control systems. This will then inform the design of the C4i system in highlighting problems with the existing systems, contributing to task allocation and training and also specifying the requirements of the new system. CTA techniques will also be used to analyse C4i design concepts and prototypes.

ACTA – Applied Cognitive Task Analysis

Laura G. Millitello & Robert J. B. Hutton, Klein Associates Inc., 582 E. Dayton-Yellow Springs Road, Fairborn, Ohio 43524, USA.

Background and applications

Applied Cognitive Task Analysis (ACTA) is a toolkit of interview techniques that can be used to elicit information regarding cognitive demands associated with the task or scenario under analysis. The techniques within the ACTA framework can be used to determine the cognitive skills and demands associated with a particular task or scenario. The output of an ACTA is typically used to aid system design. The ACTA technique or framework is made up of three interview techniques designed to allow the analyst to elicit relevant information from operators. Originally used in the fire fighting domain, ACTA was developed as part of a Navy Personnel Research and Development Center funded project as a solution to the inaccessibility and difficulty associated with using existing cognitive task analysis type methods (Militello & Hutton 2000). The overall goal of the project was to develop and evaluate techniques that would allow system designers to extract the critical cognitive elements of a particular task. The ACTA approach is designed to be used by system designers and no training in cognitive psychology is required (Militello & Hutton 2000). The ACTA procedure consist of the following components:

1. Task diagram interview
   The task diagram interview is used to give the analyst an overview of the task under analysis. The task diagram interview also allows the analyst to identify any cognitive aspects of the task that require further analysis.

2. Knowledge audit
   During the knowledge audit part of ACTA, the analyst determines the expertise required for each part of the task. The analyst probes subject matter experts (SME's) for specific examples.

3. Simulation Interview
   The simulation interview allows the analyst to probe specific cognitive aspects of the task based upon a specific scenario.

4. Cognitive demands table
   The cognitive demands table is used to group and sort the data.

Domain of application

Generic.

Procedure and advice

Step 1: Task Diagram Interview

Firstly, the analyst should conduct the task diagram interview with the relevant SME. The task diagram interview is used to provide the analyst with a clearer picture of the task under analysis and also to aid the analyst in highlighting the various cognitive elements associated with the task. According to Militello & Hutton (2000) the SME should first be asked to decompose the task into relevant task steps. Militello & Hutton (2000) recommend that the analyst should use questions such as, 'Think about what you do when you (perform the task under analysis). Can you break this task down into less than six, but more than three steps?' This process gives a verbal

protocol type analysis, with the SME verbalising the task steps. Once the task is broken down into a number of separate task steps, the SME should then be asked to identify which of the task steps require cognitive skills. Militello & Hutton (2000) define cognitive skills as judgements, assessments, problem solving and thinking skills. Once the task diagram interview is complete, the analyst should possess a very broad overview of the task under analysis, including the associated cognitive requirements.

Step 2: Knowledge audit
Next, the analyst should proceed with the knowledge audit interview. This allows the analyst to identify instances during the task under analysis where expertise is used. The knowledge audit interview is based upon the following knowledge categories that are linked to expertise (Millitello & Hutton 2000):

- Diagnosing and Predicting
- Situation Awareness
- Perceptual skills
- Developing and knowing when to apply tricks of the trade
- Improvising
- Meta-cognition
- Recognising anomalies
- Compensating for equipment limitations

The analyst should use the ACTA knowledge audit probes in order to elicit the appropriate responses. Once a probe has been administered, the analyst should then query the SME for specific examples of critical cues and decision-making strategies. Potential errors should then be discussed. The list of knowledge audit probes is shown below (Source: Militello & Hutton 2000).

Basic Probes

- **Past and Future**. Experts can figure out how a situation developed, and they can think into the future to see where the situation is going. Among other things, this can allow experts to head off problems before they develop.
  Is there a time when you walked into the middle of a situation and knew exactly how things got there and where they were headed?
- **Big Picture**. Novices may only see bits and pieces. Experts are able to quickly build an understanding of the whole situation – the big picture view. This allows the expert to think about how different elements fit together and affect each other.
  Can you give me an example of what is important about the big picture for this task? What are the major elements you have to know and keep track of?
- **Noticing**. Experts are able to detect cues and see meaningful patterns that less experienced personnel may miss altogether.
  Have you had experiences where part of a situation just 'popped' out at you; where you noticed things going on that others didn't catch? What is an example?
- **Job Smarts**. Experts learn how to combine procedures and work the task in the most efficient way possible. They don't cut corners, but they don't waste time and resources either.
  When you do this task, are there ways of working smart or accomplishing more with less – that you have found especially useful?

- **Opportunities/Improvising**. Experts are comfortable improvising – seeing what will work in this particular situation; they are able to shift directions to take advantage of opportunities.
  Can you think of an example when you have improvised in this task or noticed an opportunity to do something better?
- **Self-Monitoring**. Experts are aware of their performance; they check how they are doing and make adjustments. Experts notice when their performance is not what it should be (this could be due to stress, fatigue, high workload etc)
  Can you think of a time when you realised that you would need to change the way you were performing in order to get the job done?

Optional Probes

- **Anomalies**. Novices don't know what is typical, so they have a hard time identifying what is atypical. Experts can quickly spot unusual events and detect deviations. And, they are able to notice when something that ought to happen, doesn't.
  Can you describe an instance when you spotted a deviation from the norm, or knew something was amiss?
- **Equipment difficulties**. Equipment can sometimes mislead. Novices usually believe whatever the equipment tells them; they don't know when to be sceptical.
  Have their been times when the equipment pointed in one direction but your own judgement told you to do something else? Or when you had to rely on experience to avoid being led astray by the equipment?

Step 3: Simulation Interview

Next, is the simulation interview, which allows the analyst to understand the cognitive processes involved in the task under analysis. The SME is presented with a scenario. Once the scenario is completed, the analyst should prompt the SME to recall any major events, including decisions and judgements. Each event or task step in the scenario should be probed for situation awareness, actions, critical cues, potential errors and surrounding events. Militello & Hutton (2000) present a set of simulation interview probes, shown below.

For each major event, elicit the following information

- As the (job you are investigating) in this scenario, what actions, if any, would you take at this point in time?
- What do you think is going on here? What is your assessment of the situation at this point in time?
- What pieces of information led you to this situation assessment and these actions?
- What errors would an inexperienced person be likely to make in this situation?

Any information elicited here should be recorded in a simulation interview table. An example simulation interview table is shown in Table 24.

Table 25. Example simulation interview table (Source: Militello & Hutton 2000)

Events	Actions	Assessment	Critical Cues	Potential errors
On scene arrival	Account for people (names) Ask neighbours Must knock on or knock down to make sure people aren't there	Its a cold night, need to find place for people who have been evacuated	Night time Cold > 15° Dead space Add on floor Poor materials, metal girders Common attic in whole building	Not keeping track of people (could be looking for people who are not there)
Initial attack	Watch for signs of building collapse If signs of building collapse, evacuate and throw water on it from outside	Faulty construction, building may collapse	Signs of building collapse include: What walls are doing: cracking What floors are doing: groaning What metal girders are doing: clicking, popping Cable in old buildings hold walls together	Ventilating the attic, this draws the fire up and spreads it through the pipes and electrical system

Step 4: Cognitive demands table

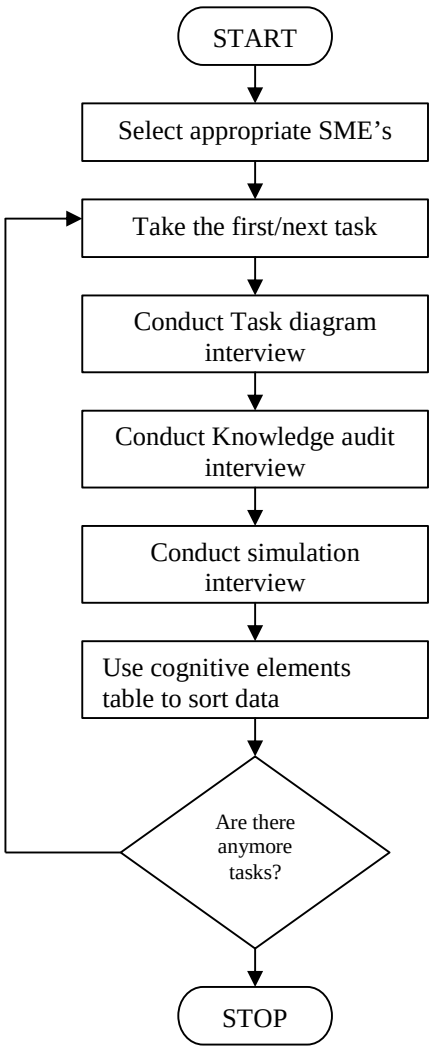
Once the knowledge audit and simulation interview are completed, it is recommended (Militello & Hutton 2000) that a cognitive demands table is used to sort and analyse the collected data. This table is used to help the analyst focus on the most important aspects of the collected data. The analyst should prepare the cognitive demands table based upon the goals of the particular project that they are applying ACTA to. An example of a cognitive demands table is shown in table 25 (Militello & Hutton 2000).

Table 25. Example cognitive demands table (Militello & Hutton 2000).

Difficult cognitive element	**Why difficult?**	**Common errors**	**Cues and strategies used**
Knowing where to search after an explosion	Novices may not be trained in dealing with explosions. Other training suggests you should start at the source and work outward	Novice would be likely to start at the source of the explosion. Starting at the source is a rule of thumb for most other kinds of incidents	Start where you are most likely to find victims, keeping in mind safety considerations Refer to material data sheets to determine where dangerous chemicals are likely to be Consider the type of structure and where victims are likely to be Consider the likelihood of further explosions. Keep in mind the safety of your crew
Finding victims in a burning building	There are lots of distracting noises. If you are nervous or tired, your own breathing makes it hard to hear anything else	Novices sometimes don't recognise their own breathing sounds; they mistakenly think they hear a victim breathing	Both you and your partner stop, hold your breath and listen Listen for crying, victims talking to themselves, victims knocking things over etc

Once the ACTA analysis is complete, the analyst has a set of data that can be used to inform either the design of the systems or the design of the training procedures.

Flowchart



Advantages

- Analysts using the technique do not require training in cognitive psychology.
- Requires fewer resources than traditional cognitive task analysis techniques (Militello & Hutton 2000).
- Militello & Hutton (2000) reported that in a usability questionnaire focussing on the use of the ACTA techniques, ratings were very positive. The data indicated that participants found the ACTA techniques easy to use and flexible, and that the output of the interviews was clear and the knowledge representations to be useful.
- Probes and questions are provided for the analyst, facilitating relevant data extraction.

Disadvantages

- The quality of data is very much dependent upon the skill of the analyst.
- The consistency of such a technique is questionable.

- The technique would appear to be time consuming in its application. In a validation study (Militello & Hutton 2000) participants using the ACTA techniques were given 3 hours to perform the interviews and 4 hours to analyse the data.
- The training time for the ACTA techniques is also quite high. Militello & Hutton (2000) gave participants an initial 2 hour workshop introducing cognitive task analysis and then a 6 hour workshop on the ACTA techniques.
- The analysis of the data appears to be a laborious process.
- As with most cognitive task analysis techniques, ACTA requires further validation. At the moment there is little in the way of validation studies associated with the ACTA techniques.
- The quality of the data obtained depends both on the SME's used and the analyst applying the techniques. Militello & Hutton (2000) suggest that some people are better interviewers than others and also that some SME's are more useful than others.

Related methods

Each of the techniques used within the ACTA toolkit is an interview type approach. According to the authors (Militello & Hutton 2000), the technique is a streamlined version of existing cognitive task analysis techniques. The ACTA techniques also require SME's to walkthrough the task in their head, which is an approach very similar to that of walkthrough or cognitive walkthrough type analysis. The interview techniques used in the ACTA technique provide an output that is very similar to that of VPA.

Approximate training and application times

In a validation study (Militello & Hutton 2000), participants were given 8 hours of training, consisting of a 2 hour introduction to cognitive task analysis and a 6 hour workshop on the ACTA techniques. This represents a medium to high training time for the ACTA techniques. In the same study, the total application times for each participant was 7 hours, consisting of 3 hours applying the interviews and 4 hours analysing the data. This represents a moderate application time for the ACTA techniques.

Reliability and Validity

Militello & Hutton (2000) suggest that there are no well established metrics that exist in order to establish the reliability and validity of cognitive task analysis techniques. However, a number of attempts were made to establish the reliability and validity of the ACTA techniques. In terms of validity, three questions were addressed:

1) Does the information gathered address cognitive issues?
2) Does the information gathered deal with experience based knowledge as opposed to classroom based knowledge?
3) Do the instructional materials generated contain accurate information that is important for novices to learn?

Each item in the cognitive demands table was examined for its cognitive content. It was found that 93% of the items were related to cognitive issues. To establish the level of experience based knowledge elicited, participants were asked to subjectively rate the proportion of information that only highly experienced SME's would know. In the fire fighting study, the average was 95% and in the EW study, the average was 90%. The importance of the instructional materials generated was validated via

domain experts rating the importance and accuracy of the data elicited. The findings indicated that the instructional materials generated in the study contained important information for novices (70% fire fighting, 95% EW). The reliability of the ACTA techniques was assessed by determining whether the participants using the techniques generated similar information. It was established that participants using the ACTA techniques were able to consistently elicit relevant cognitive information.

Tools needed

ACTA is a pencil and paper tool. The analyst should also possess the knowledge audit and simulation interview probes. A tape recorder or Dictaphone may also be useful to aid the recording and analysis of the data.

Bibliography

Militello, L. G. & Militello, J. B. (2000) Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. In J. Annett & N. S Stanton (Eds) *Task Analysis,* pp 90-113. UK, Taylor & Francis

Cognitive Walkthrough

Peter G. Polson, Clayton Lewis, Cathleen Wharton, John Rieman, Institute of Cognitive Science, Department of Psychology and the Department of Computer Science, University of Colorado, Boulder, CO 80309-0345, USA

Background and applications

The cognitive walkthrough technique is a methodology for evaluating the usability of user interfaces. The main driver behind the techniques development was the goal to develop and test a theoretically based design methodology that could be used in actual design and development situations (Polson et al 1992). The main criticism of existing walkthrough techniques suggests that they are actually unusable in actual design situations (Polson et al 1992). The technique is designed for use early in the design process of a user interface, however the technique could also be used on existing user interfaces as an evaluation tool. Based upon traditional design walkthrough techniques and a theory of exploratory learning (Polson and Lewis), the technique focuses upon the usability of an interface, in particular the ease of learning associated with the interface. The cognitive walkthrough technique consists of a set of criteria that the analyst must evaluate each task and the interface under analysis against. These criteria focus on the cognitive processes required to perform the task (Polson et al 1992). Although originally developed for use in software engineering, it is apparent that the technique could be used to evaluate an interface in any domain.

The cognitive walkthrough process involves the analyst 'walking' through each user/operator action involved in a task step. The analyst then considers each criteria and the effect the interface has upon the user's goals and actions. The criteria used in the cognitive walkthrough technique are shown below: (Source: Polson et al 1992). Each task step or action is analysed separately using this criteria.

1. Goal structure for a step

1.1 Correct goals. What are the appropriate goals for this point in the interaction? Describe as for initial goals.

1.2 Mismatch with likely goals. What percentage of users will not have these goals, based on the analysis at the end of the previous step. Based on that analysis, will all users have the goal at this point, or may some users dropped it or failed to form it. Also check the analysis at the end of the previous step to see if there are any unwanted goals, not appropriate for this step that will be formed or retained by some users. (% 0 25 50 75 100)

2. Choosing and executing the action

Correct action at this step.....

2.1 Availability. Is it obvious that the correct action is a possible choice here? If not, what percentage of users might miss it? (% 0 25 50 75 100)

2.2 Label. What label or description is associated with the correct action?

2.3 Link of label to action. If there is a label or description associated with the correct action, is it obvious, and is it clearly linked with this action? If not, what percentage of users might have trouble? (% 0 25 50 75 100)

2.4 Link of label to goal. If there is a label or description associated with the correct action, is it obvious, and is it clearly linked with this action? If not, what percentage of users might have trouble? (% 0 25 50 75 100)

2.5 No label. If there is no label associated with the correct action, how will users relate this action to a current goal? What percentage might have trouble doing so? (% 0 25 50 75 100)
2.6 Wrong choices. Are there other actions that might seem appropriate to some current goal? If so, what are they, and what percentage of users might choose one of these? (% 0 25 50 75 100)
2.7 Time out. If there is a time out in the interface at this step does it allow time for the user to select the appropriate action? How many users might have trouble? (% 0 25 50 75 100)
2.8 Hard to do. Is there anything physically tricky about executing the action? If so, what percentage of users will have trouble? (% 0 25 50 75 100)

3. Modification of goal structure. Assume the correct action has been taken. What is the systems response?
3.1 Quit or backup. Will users see that they have made progress towards some current goal? What will indicate this to them? What percentage of users will not see progress and try to quit or backup? (% 0 25 50 75 100)
3.2 Accomplished goals. List all current goals that have been accomplished. Is it obvious from the system response that each has been accomplished? If not, indicate for each how many users will not realise it is complete.
3.3 Incomplete goals that look accomplished. Are there any current goals that have not been accomplished, but might appear to have based upon the system response? What might indicate this? List any such goals and the percentage of users who will think that they have actually been accomplished.
3.4 "And-then" structures. Is there an "and-then" structure, and does one of its sub-goals appear to be complete? If the sub-goal is similar to the supergoal, estimate how many users may prematurely terminate the "and-then" structure.
3.5 New goals in response to prompts. Does the system response contain a prompt or cue that suggests any new goal or goals? If so, describe the goals. If the prompt is unclear, indicate the percentage of users who will not form these goals.
3.6 Other new goals. Are there any other new goals that users will form given their current goals, the state of the interface, and their background knowledge? Why? If so, describe the goals, and indicate how many users will form them. NOTE these goals may or may not be appropriate, so forming them may be bad or good.

Domain of application
Generic.

Procedure and advice (adapted from Polson et al 1992)
The cognitive walkthrough procedure is made up of two phases, the preparation phase and the evaluation phase. The preparation phase involves selecting the set of tasks to analyse and determining the task sequence. The evaluation phase involves the analysis of the interaction between the user and the interface, using the criteria outlined above.

Step 1: Select tasks to be analysed
Firstly, the analyst should select the set of tasks that are to be analysed. To thoroughly examine the interface in question, an exhaustive set of tasks should be used. However, if time is limited, then the analyst should try to select a set of tasks that involve all aspects of the interface.

Step 2: Create task descriptions
Each task selected by the analyst must be described fully from the point of the user.

Step 3: Determine the correct sequence of actions
For each of the selected tasks, the appropriate sequence of actions required to complete the task must be specified. A HTA of the task would be useful for this part of the cognitive walkthrough analysis.

Step 4: Identify user population
Next, the analyst should determine the potential users of the interface under analysis. A list of user groups should be created.

Step 5: Describe the user's initial goals
The final part of phase one of a cognitive walkthrough analysis is to determine and record the user's initial goals. The analyst should record what goals the user has at the start of the task. This is based upon the analyst's subjective judgement.

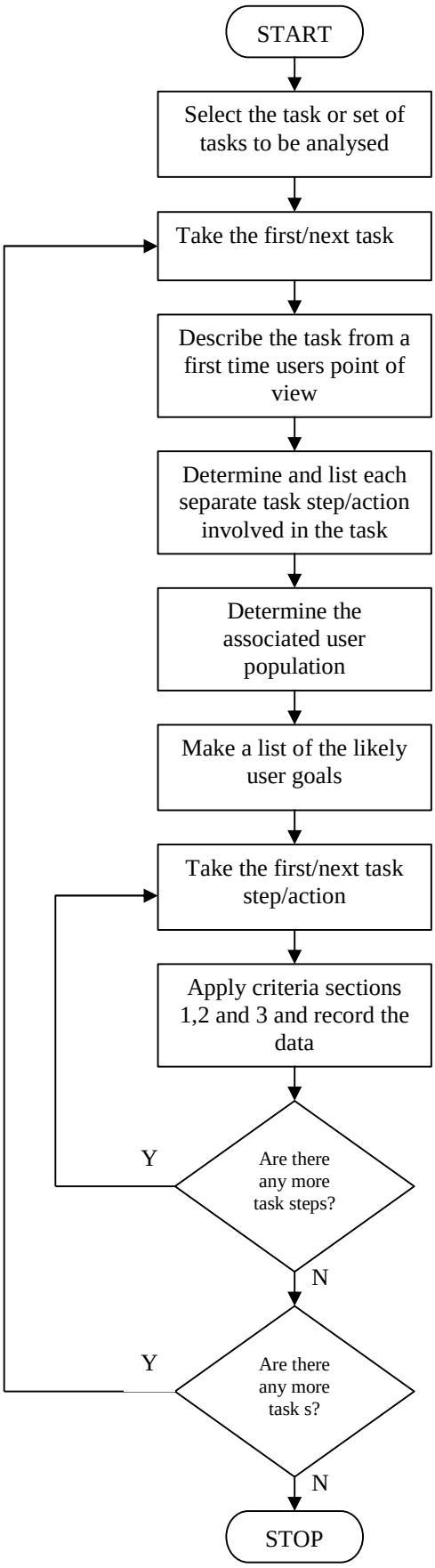
Step 6: Analyse the interaction between user and interface
Phase 2, the evaluation phase, involves analysing the interaction between the user and the interface under analysis. Here, the analyst should 'walk' through each task, applying the criteria outlined above as they go along. The cognitive walkthrough evaluation concentrates on 3 aspects of the user interface interaction:
1) Relationship between the required goals and the goals that the user actually have
2) The problems in selecting and executing an action
3) Changing goals due to action execution and system response

The analyst should record the results for each task step. This can be done via video, audio or pen and paper techniques.

Advantages
- The cognitive walkthrough technique presents a structured approach to highlighting the design flaws of an interface.
- Can be used very early in the design cycle of an interface.
- Designed to be used by non-cognitive psychology professionals.
- The cognitive walkthrough technique is based upon sound underpinning theory, including Norman's model of action execution.
- Easy to learn and apply.
- The output from a cognitive walkthrough analysis appears to be very useful.

Disadvantages
- The cognitive walkthrough technique is limited to cater only for ease of learning of an interface.
- Requires validation.
- May be time consuming for more complex tasks.
- Recorded data would require in depth analysis in order to be useful.
- A large part of the analysis is based upon analyst skill. For example, the percentage estimates used with the walkthrough criteria require a 'best guess'.
- Cognitive walkthrough requires access to the personnel involved in the task(s) under analysis.

Flowchart


START
Select the task or set of tasks to be analysed
Take the first/next task
Describe the task from a first time users point of view
Determine and list each separate task step/action involved in the task
Determine the associated user population
Make a list of the likely user goals
Take the first/next task step/action
Apply criteria sections 1,2 and 3 and record the data
Are there any more task steps?
Y
N
Are there any more task s?
Y
N
STOP

Example
The following example is an extract of a cognitive walkthrough analysis of a phone system task presented in Polson et al (1992).

Task – Forward all my calls to 492-1234

Task list
1. Pick up the handset
2. Press ##7
3. Hang up the handset
4. Pick up the handset
5. Press **7
6. Press 1234
7. Hang up the handset

Goals:
75% of users will have FORWARD ALL CALLS TO 492 1234 (Goal)
PICK UP HANDSET (Sub-goal)
and then SPECIFY FORWARDING (Sub-goal)

25% of users will have FORWARD ALL CALLS TO 492 1234
PICK UP HANDSET
and then CLEAR FORWARDING
and then SPECIFY FORWARDING

Analysis of ACTION 1: Pick up the handset
Correct goals
FORWARD ALL CALLS TO 492 1234
PICK UP HANDSET
and then CLEAR FORWARDING
and then SPECIFY FORWARDING

75% of the users would therefore be expected to have a goal mismatch at this step, due to the required clear forwarding sub-goal that is required but not formed (Polson et al 1992).

Related methods
The cognitive walkthrough technique is a development of the traditional design walkthrough methods (Polson et al 1992). HTA or tabular task analysis could also be used when applying cognitive walkthrough technique in order to provide a description of the sequence of actions.

Approximate training and application times
No data regarding the training and application time for the technique are offered by the authors. It is estimated that the training time for the technique would be quite high. It is also estimated that the application time for the technique would be high, particularly for large, complex tasks.

Reliability and validity
Lewis et al (1990) reported that in a cognitive walkthrough analysis of four answering machine interfaces about half of the actual observed errors were identified. More critically, the false alarm rate (errors predicted in the cognitive walkthrough analysis but not observed) was extremely high, at almost 75%. In a study on voicemail directory, Polson et al (1992) reported that half of all observed errors were picked up in the cognitive walkthrough analysis. It is apparent that the cognitive walkthrough

technique requires further validation in terms of the reliability and validity of the technique.

Tools needed

The cognitive walkthrough technique can be conducted using pen and paper. The analyst would also require the walkthrough criteria sections 1, 2 and 3 and the cognitive walkthrough start up sheet. For larger analyses, the analyst may wish to record the process using video or audio recording equipment. The device/interface under analysis is also required.

Bibliography

Lewis, C., Polson, P., Wharton, C. & Rieman, J. (1990) Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces. In Proceedings of CHI'90 conference on Human Factors in Computer Systems, pp 235-241, New York: Association for Computer Machinery

Polson, P. G., Lewis, C., Riemant, J. & Wharton, C. (1992) Cognitive walkthroughs: a method for theory based evaluation of user interfaces. *International Journal of Man-Machine Studies*, 36 pp741-773.

Wharton, C., Rieman, J., Lewis, C., Polson, P. The Cognitive Walkthrough Method: A Practitioners Guide

Critical Decision Method
Gary Klein, Klein Associates, 1750 Commerce Center Boulevard, North Fairborn, OH 45324-6362

Background and applications
The Critical Decision Method is a semi-structured interview technique that uses a set of cognitive probes in order to elicit information regarding expert decision-making. According to the authors, the technique can serve to provide knowledge engineering for expert system development, identify training requirements, generate training materials and evaluate the task performance impact of expert systems (Klein, Calderwood & MacGregor 1989). The technique is a development of the Critical Incident Technique (Flanagan 1954) and was developed in order to study naturalistic decision-making strategies of experienced personnel. CDM has been applied to personnel in a number of domains involving complex and dynamic systems, including fire fighting, military and paramedics (Klien, Calderwood & MacGregor 1989).

Domain of application
Generic.

Procedure and advice (adapted from Klein, Calderwood & MacGregor 1989)
When conducting a CDM analysis, it is recommended that a pair of analyst's are used. Klein & Armstrong (In Press) suggests that when using only one analyst, data may be missed or not recorded. The CDM analysis process should be recorded using a video recording device or an audio recording device.

Step 1: Select the Incident to be analysed
The first part of a CDM analysis is to select the incident that is to be analysed. Depending upon the purpose of the analysis, the type of incident may already be selected. CDM normally focuses on non-routine incidents, such as emergency scenario's, or highly challenging incidents. If the type of incident is not already known, the CDM analysts may select the incident via interview with system personnel, probing the interviewee for recent high risk, highly challenging, emergency situations. The interviewee involved in the CDM analysis should be the primary decision maker in the chosen incident.

Step 2: Gather and record account of the incident
Next the interviewee should be asked to provide a description of the incident in question, from its starting point (i.e. alarm sounding) to its end point (i.e. when the incident was classed as 'under control'.

Step 3: Construct Incident Timeline
The next step in the CDM analysis is to construct an accurate timeline of the incident under analysis. The aim of this is to give the analysts a clear picture of the incident and its associated events, including when each event occurred and what the duration of each event was. According to Klein, Calderwood & MacGregor (1989) the events included in the timeline should encompass any physical events, such as alarms sounding, and also 'mental' events, such as the thoughts and perceptions of the interviewee during the incident. The construction of the incident timeline serves to increase the analyst's knowledge and awareness of the incident whilst simultaneously focussing the interviewee's attention on each event involved in the incident.

Step 4: Identify Decision Points
Whilst constructing the timeline, the analysts should select specific decisions of interest for further analysis. Each selected decision should then be probed or analysed further. Klein, Calderwood & MacGregor (1989) suggest that decision points where other courses of action were available to the operator should be probed further.

Step 5: Probe selected decision points
Each decision point selected in step 4 should be analysed further using a set of specific probes. The probes used are dependent upon the aims of the analysis and the domain in which the incident is embedded. Klein, Calderwood & MacGregor (1989) summarise the probes that have been used in CDM's in the past.

Probe Type	**Probe Content**
Cues	What were you seeing, hearing, smelling.....?
Knowledge	What information did you use in making this decision, and how was it obtained?
Analogues	Were you reminded of any previous experience?
Goals	What were your specific goals at this time?
Options	What other courses of action were considered by or available to you?
Basis	How was this option selected/other options rejected? What role was being followed?
Experience	What specific training or experience was necessary or helpful in making this decision?
Aiding	If the decision was not the best, what training, knowledge or information could have helped?
Time Pressure	How much time pressure was involved in making this decision? (offer scale here)
Situation Assessment	Imagine that you were asked to describe the situation to a relief officer at this point, how would you summarise the situation?
Hypotheticals	If a key feature of the situation had been different, what difference would it have made in your decision?

A set of revised CDM probes were developed by O'Hare et al (2000).

Goal Specification	What were your specific goals at the various decision points?
Cue Identification	What features were you looking for when you formulated your decision? How did you that you needed to make the decision? How did you know when to make the decision?
Expectancy	Were you expecting to make this sort of decision during the course of the event? Describe how this affected your decision making process.
Conceptual	Are there any situations in which your decision would have turned out differently? Describe the nature of these situations and the characteristics that would have changed the outcome of your decision.
Influence of uncertainty	At any stage, were you uncertain about either the reliability of the relevance of the information that you had available? At any stage, were you uncertain about the appropriateness of the decision?
Information integration	What was the most important piece of information that you used to formulate the decision?
Situation Awareness	What information did you have available to you at the time of the decision?
Situation Assessment	Did you use all of the information available to you when formulating the decision? Was there any additional information that you might have used to assist in the formulation of the decision?
Options	Were there any other alternatives available to you other than the decision you made?

Decision blocking - stress	Was there any stage during the decision making process in which you found it difficult to process and integrate the information available? Describe precisely the nature of the situation
Basis of choice	Do you think that you could develop a rule, based on your experience, which could assist another person to make the same decision successfully? Why/Why not?
Analogy/generalisation	Were you at any time, reminded of previous experiences in which a similar decision was made? Were you at any time, reminded of previous experiences in which a different decision was made?

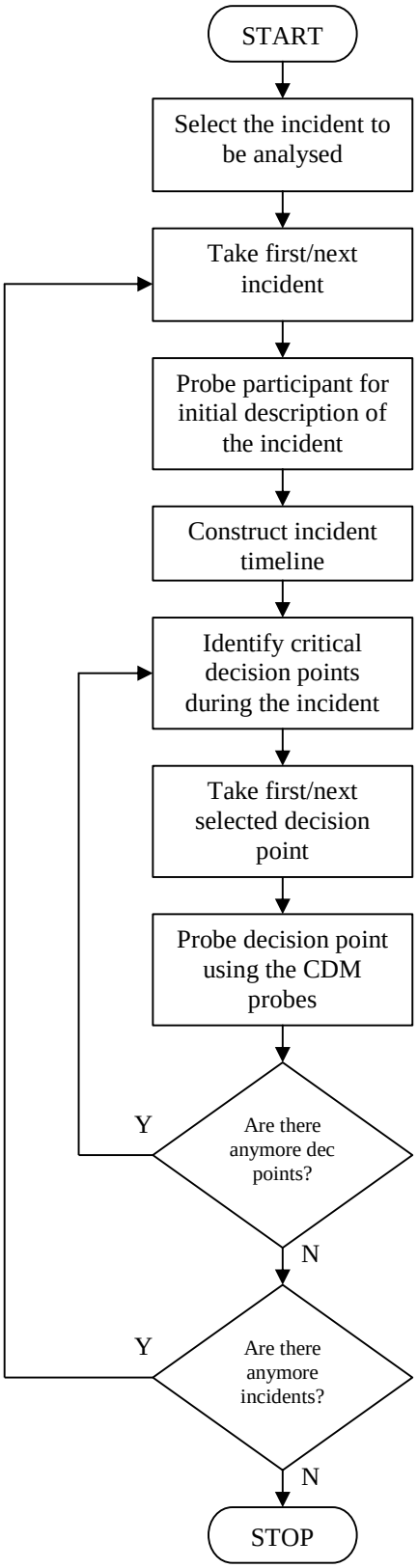
Advantages

- The CDM can be used to elicit specific information regarding decision making in complex systems.
- The technique requires relatively little effort to apply.
- The incidents which the technique concentrates on have already occurred, removing the need for costly, time consuming to construct event simulations.
- Once familiar with the technique, CDM is easy to apply
- Has been used extensively in a number of domains and has the potential to be used anywhere.
- Real life incidents are analysed using the CDM, ensuring a more comprehensive, realistic analysis than simulation techniques.
- The cognitive probes used in the CDM have been used for a number of years and are efficient at capturing the decision making process (Klein & Armstrong In Press).

Disadvantages

- The reliability of such a technique is questionable. Klein & Armstrong (In Press) suggests that methods that analyse retrospective incidents are associated with concerns of data reliability, due to evidence of memory degradation.
- CDM will never be an exact description of an incident.
- The CDM is a resource intensive technique. The data analysis part is especially time consuming.
- A high level of expertise and training is required in order to use the CDM to its maximum effect (Klein & Armstrong In Press).
- The CDM requires a team (minimum of 2) of interviewer's for each interviewee.
- The CDM relies upon interviewee verbal reports in order to reconstruct incidents. How far a verbal report accurately represents the cognitive processes of the decision maker is questionable. Facts could be easily misrepresented by interviewee's. Certainly, glorification of events would be one worry associated with this sort of analysis.
- After the fact data collection has a number of concerns associated with it. Such as degradation, correlation with performance etc

Flowchart


START
Select the incident to be analysed
Take first/next incident
Probe participant for initial description of the incident
Construct incident timeline
Identify critical decision points during the incident
Take first/next selected decision point
Probe decision point using the CDM probes
Are there anymore dec points?
Y
N
Are there anymore incidents?
Y
N
STOP

Example
O'Hare et al (2000) report the use of the CDM to analyse expert white water rafting guides. Seventeen raft guides with varying degrees of experience were interviewed using the CDM. The participants were asked to describe any an incident in which they were required to make a critical decision or series of critical decisions (O'Hare et al 2000). The CDM analysis produced seventeen non-routine critical incidents including a total of 52 decision points. The most common critical incident elicited involved the retrieval of clients who had fell into the water. According to O'Hare et al (2000) it was also found that expert raft guides considered no more than two action options when making a decision in a critical situation. In comparison, trip leaders (less experience) either developed a single course of action used this approach or they developed up to five courses of action, considering each one until the most appropriate course of action became evident. In conclusion O'Hare et al (2000) reported that expert guides were able to retrieve the most appropriate option without comparing multiple action options whilst less experienced trip leaders use a mixture of analytical and intuitive decision styles and novice guides act upon their original course of action specification.

Related Methods
The CDM is an extension of the original Critical Incident Technique (Flanagan 1954), which involved identifying factors contributing to success or failure in a particular scenario. The CDM is also closely related to other cognitive task analysis (CTA) techniques, in that it uses probes to elicit data regarding task performance from participants. Other CTA techniques include ACTA and cognitive walkthrough analysis.

Approximate training and application times
Klein & Armstrong (In Press) report that the training time associated with the CDM would be high. In terms of application, the normal application time for CDM is around 2 hours (Klein, Calderwood & MacGregor 1989). The data analysis part of the CDM would, however, add considerable time to the overall analysis. For this reason, it is suggested that the CDM application time, including data collection and data analysis, would be considerably high.

Reliability and validity
The reliability of the CDM is questionable. It is apparent that such an approach may elicit different data from similar incidents when applied by different analysts on separate participants. Klein & Armstrong (In Press) suggests that there are concerns associated with the reliability of the CDM due to evidence of memory degradation.

Tools needed
When conducting a CDM analysis, pen and paper could be sufficient. However, to ensure that data collection is comprehensive, it is recommended that video or audio recording equipment is used. A set of 'cognitive' probes are also required. The type of probes used are dependent upon the focus of the analysis.

Bibliography
Flanagan, J. C. (1954). The Critical Incident Technique. *Psychological Bulletin*, 51, 327-358.

Klein, G. & Armstrong, A. A. (In Press) Critical Decision Method. In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.
Klein, G. A., Calderwood, R., & MacGregor, D. (1989) Critical Decision Method for Eliciting Knowledge. *IEEE Transactions on Systems*, Man and Cybernetics, 19(3), 462-472
O'Hare, D., Wiggins, M., Williams, A., & Wong, W. (2000) Cognitive task analyses for decision centred design and training. In J. Annett & N. Stanton (eds) *Task Analysis*, pp 170-190. UK, Taylor and Francis

Critical Incident Technique

Background and applications

Critical incident technique (CIT) (Flanagan 1954) is an interview technique that is used to collect specific data regarding incidents or events and associated operator decisions and actions made. The technique was first used to analyse aircraft incidents that almost led to accidents and has since been used extensively and developed in the form of CDM (Klein 2003). CIT involves using interview techniques to facilitate operator recall of critical events or incidents, including what actions and decisions made by themselves and colleagues and why they made them. Although the technique is typically used to analyse incidents involving existing systems, it is offered here as a way of analysing events in similar systems to that of the system being designed. CIT can be used to highlight vulnerable system features or poorly designed system features and processes. The CIT probes used by Flanagan (1954) are shown below. It is recommended that new probes be developed when using the technique as these may be dated and over-simplistic.

- Describe what led up to the situation
- Exactly what did the person do or not do that was especially effective or ineffective
- What was the outcome or result of this action?
- Why was this action effective or what more effective action might have been expected?

Domain of application

Aviation.

Procedure and advice

Step 1: Select the Incident to be analysed

The first part of a CIT analysis is to select the incident or group of incidents that are to be analysed. Depending upon the purpose of the analysis, the type of incident may already be selected. CIT normally focuses on non-routine incidents, such as emergency scenario's, or highly challenging incidents. If the type of incident is not already known, the CIT analysts may select the incident via interview with system personnel, probing the interviewee for recent high risk, highly challenging, emergency situations. The interviewee involved in the CDM analysis should be the primary decision maker in the chosen incident. CIT can also be conducted on groups of operators.

Step 2: Gather and record account of the incident

Next the interviewee(s) should be asked to provide a description of the incident in question, from its starting point (i.e. alarm sounding) to its end point (i.e. when the incident was classed as 'under control'.

Step 3: Construct Incident Timeline

The next step in the CIT analysis is to construct an accurate timeline of the incident under analysis. The aim of this is to give the analysts a clear picture of the incident and its associated events, including when each event occurred and what the duration of each event was. According to Klien, Calderwood & MacGregor (1989) the events included in the timeline should encompass any physical events, such as alarms sounding, and also 'mental' events, such as the thoughts and perceptions of the

interviewee during the incident. The construction of the incident timeline serves to increase the analyst's knowledge and awareness of the incident whilst simultaneously focussing the interviewee's attention on each event involved in the incident.

Step 4: Select required incident aspects
Once the analyst has an accurate description of the incident, the next step is to select specific incident points that are to be analysed further. The points selected are dependent upon the nature and focus of the analysis. For example, if the analysis is focussing upon team communication, then aspects of the incident involving team communication should be selected.

Step 5: Probe selected incident points
Each incident aspect selected in step 4 should be analysed further using a set of specific probes. The probes used are dependent upon the aims of the analysis and the domain in which the incident is embedded. The analyst should develop specific probes before the analysis begins. In an analysis of team communication, the analyst would use probes such as 'Why did you communicate with team member B at this point?', 'How did you communicate with team member B', 'Was there any mis-communication at this point' etc.

Advantages

- The CIT can be used to elicit specific information regarding decision making in complex systems.
- The technique requires relatively little effort to apply.
- The incidents which the technique concentrates on have already occurred, removing the need for costly, time consuming to construct event simulations.
- CIT is easy to apply
- Has been used extensively in a number of domains and has the potential to be used anywhere.
- Real life incidents are analysed using the CIT, ensuring a more comprehensive, realistic analysis than simulation techniques.
- CIT is a very flexible technique.
- Cost effective.
- High face validity (Kirwan & Ainsworth 1992).

Disadvantages

- The reliability of such a technique is questionable. Klien (2003) suggests that methods that analyse retrospective incidents are associated with concerns of data reliability, due to evidence of memory degradation.
- A high level of expertise in interview techniques is required.
- After the fact data collection has a number of concerns associated with it. Such as degradation, correlation with performance etc.
- Relies upon the accurate recall of events.
- Operators may not wish to recall events or incidents in which there performance is under scrutiny.
- Analyst(s) may struggle to obtain accurate descriptions of past events.

Related methods
CIT was the first interview type technique focussing upon past events or incidents. A number of techniques have been developed as a result of the CIT, such as the critical decision method (Klein 2003). CIT is an interview technique that is also similar to walkthrough type techniques.

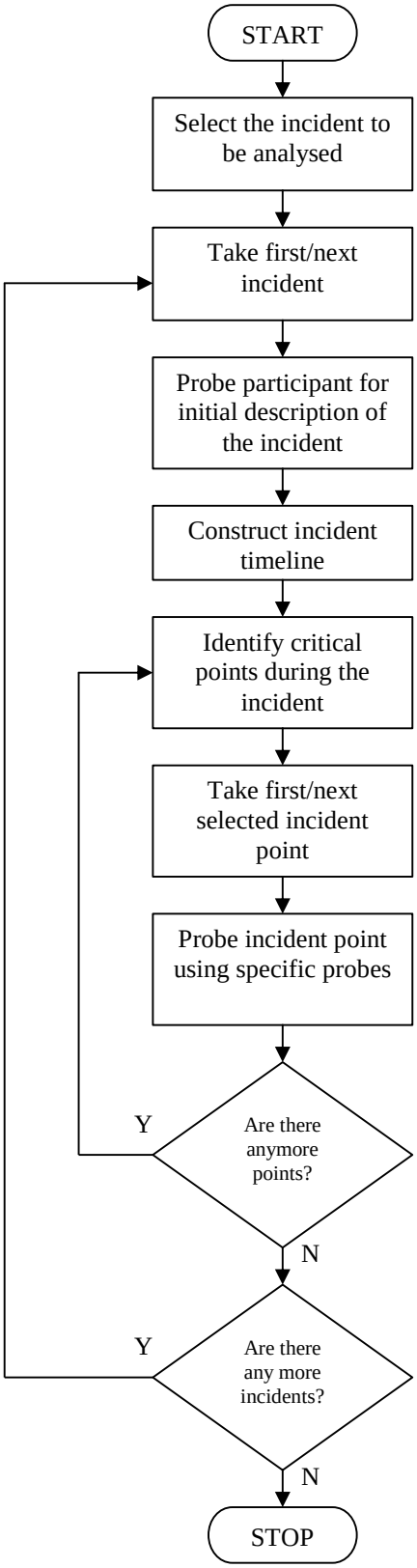
Approximate training and application times
Provided the analyst is experienced in interview techniques, the training time for CIT is minimal. However, for analysts with no interview experience, the training time would be high. Application time for the CIT is typically low, although for complex incidents involving multiple agents, the application time could increase considerably.

Reliability and validity
The reliability of the CIT is questionable. It is apparent that such an approach may elicit different data from similar incidents when applied by different analysts on separate participants. Klien (2003) suggests that there are concerns associated with the reliability of the CDM (similar technique) due to evidence of memory degradation. Also, recalled events may be correlated with performance and also subject to bias.

Tools needed
CIT can be conducted using pen and paper. It is recommended however, that the analysis is recorded using video and audio recording equipment.

Bibliography
Flanagan, J. C. (1954). The Critical Incident Technique. *Psychological Bulletin*, 51, 327-358.
Klein, G. & Armstrong, A. A. (2003) Critical Decision Method. In Stanton et al (Eds) *Handbook of Human Factors and Ergonomics methods*. UK, Taylor and Francis.
Klein, G. A., Calderwood, R., & MacGregor, D. (1989) Critical Decision Method for Eliciting Knowledge. *IEEE Transactions on Systems*, Man and Cybernetics, 19(3), 462-472

Flowchart


START
Select the incident to be analysed
Take first/next incident
Probe participant for initial description of the incident
Construct incident timeline
Identify critical points during the incident
Take first/next selected incident point
Probe incident point using specific probes
Are there anymore points?
Y
N
Are there any more incidents?
Y
N
STOP

Team Task Analysis techniques

A more recent theme in the task analysis arena is team task analysis. According to Savoie (1998) (cited by Salas In Press) the use of teams has risen dramatically with reports of 'team presence' by workers rising from 5% in 1980 to 50% in the mid 1990s. However, the increased use of teams has been accompanied by a swift realisation that team performance is extremely complex to understand and often flawed. Salas (In Press) suggests that whilst there are a number of advantages associated with the use of teams, there are also a number of disadvantages. This has led to an increased focus upon team performance in complex, dynamic systems. As a result, a number of team task analysis (TTA) techniques have emerged. TTA techniques are used to describe team performance in terms of requirements (knowledge, skills and attitudes) and the tasks that require either teamwork or individual performance (Burke 2003). According to Baker, Salas and Bowers (1998) TTA refers to the analysis of a team tasks and also the assessment of a teams teamwork requirements (Knowledge, skills and abilities). TTA output are typically used to develop team training procedures, evaluate team performance, and to identify operational and teamwork skills required within teams (Burke 2003). According to Salas (In Press) optimising team performance and effectiveness involves understanding a number of components surrounding the use of teams, such a communication and task requirements, team environments and team objectives. The team task analysis techniques reviewed in this document attempt to analyse such components.

Comms Usage Diagram (CUD) (Watts & Monk 2000) is a team task analysis technique that is used to describe collaborative activity between teams of personnel situated in different geographical locations. The output of CUD describes how and why communications between a team occur, which technology is involved in the communication, and the advantages and disadvantages of the technology used. Social Network Analysis (SNA) (Driskell & Mullen In Press) is a technique used to analyse and represent the relationships existing between teams of personnel or social groups. The analysis of these relationships can be used to demonstrate the different types of relationships, the importance and the number of relationships within a team. Groupware Task Analysis (GTA) (Van Welie & Van Der Veer 2003) is a team task analysis technique that is used during the design process to study and evaluate group or team activities in order to highlight design requirements for similar team systems. The technique involves describing the existing system or process (Task model 1) and then specifying design concepts in task model 2. HTA (T) (Annett 2000) is a development of hierarchical task analysis that caters for team-based tasks. Team Task Analysis (TTA) is a task analysis technique that provides a description of tasks distributed across a team and the requirements associated with the tasks in terms of operator knowledge, skills, and abilities. TTA aims to analyse team-based scenarios by gathering data regarding teamwork (individuals interacting or co-ordinating tasks that are important to the teams goals) and taskwork (individuals performing individual tasks).

TTA techniques are crucial in the design of C4i systems. Command and control systems employ teams of personnel with various interacting roles. Tasks are performed both individually and as a team, and interaction between team members is frequent. According to Swezey et al (2000), military tasks are characterised by dynamic changes of information and resources among groups and team members, and

by co-ordination of task activities. TTA techniques will be employed to analyse team performance and requirements in existing command and control systems. During the design lifecycle of the C4i system, TTA techniques will be used to inform task allocation and determine team requirements.

The TTA techniques reviewed in this document are shown below:

1. Comms Usage Diagrams
2. Social Network Analysis
3. Groupware Task Analysis
4. Team Task Analysis
5. HTA (T)

CUD - Comms Usage Diagram

Leon Watts, Department of Psychology, University of York, York, Y01 5DD, UK
Andrew Monk, Department of Psychology, University of York, York, Y01 5DD, UK

Background and applications

Comms Usage Diagram (CUD) (Watts & Monk 2000) is a task analysis technique that is used to describe collaborative activity between teams of personnel situated in different geographical locations. The output of CUD describes how and why communications between a team occur, which technology is involved in the communication, and the advantages and disadvantages of the technology used. The CUD technique was originally developed and applied in telecommunications, whereby the technique was used to analyse 'telemedical consultation' (Watts & Monk 2000), involving a medical practitioner offering advice regarding a medical ailment from a different location to the advice seeker. In conducting a CUD type analysis, data is typically collected via observational study, talk through type analysis and interviews (Watts and Monk 2000) and then collaborative activity is described in the CUD output table. According to Watts and Monk (2000) an analysis of collaborative activity should take into account the following factors:

1) What are the primary activities that constitute the work in question?
2) Which of these primary activities are interactions between agents (distinguished from interactions with equipment)?
3) Who else may participate (i.e. who has access to the ongoing work)?
4) The contemporaneity of agents' activities (from which the potential for opportunistic interaction might be determined).
5) The space where the activities are taking place.
6) How accessibility to primary activities is made available, through the resources that provide relevant information about them and the resources that broker interactions between the primary agents once initiated.

Domain of application

Medical telecommunications.

Procedure and advice

There is no set procedure offered by the authors for the CUD technique. The following procedure is intended to act as a set of guidelines for conducting a CUD analysis.

Step 1: Data collection

The first phase of a CUD analysis is to collect specific data regarding the task or scenario under analysis. Watts & Monk (2000) recommend that interviews, observations and task talk-through should be used to collect the data. Specific data regarding the personnel involved, activity, task steps, communication between personnel, technology used and geographical location should be collected.

Step 2: Complete Initial Comms report

Following the data collection phase, the raw data obtained should be put into a report form. According to Watts & Monk (2000) the report should include the location of the technology used, the purpose of the technology, the advantages and disadvantages of using such technology, and graphical account of a typical consultation session.

The report should then be made available to all personnel involved for evaluation and reiteration purposes.

Step 3: Construct CUD output table
The graphical account developed in step 3 forms the basis for the CUD output. The CUD output contains a description of the task activity at each geographical location and the collaboration between personnel at each location. Arrows should then be used to represent the communications between personnel at different locations. For example, if person A at site A communicates with person B at site B, the two should be linked with a two-way arrow. Column three of the CUD output table specifies the technology used in the communication and column 4 lists any good points, problems, flaws, advantages and disadvantages observed when using the particular technology during the communication.

Step 4: Construct participant-percept matrix
For instances where personnel are communicating with each other at the same geographical location (co-present) it is assumed that they can see and hear each other (Watts & Monk 2000). If environmental conditions may obstruct communication at the same site, the participant percept matrix is constructed in order to represent the awareness between participants. The percept matrix is explained further in the example section.

Advantages

- CUD offers a thorough description of collaborative activity, including the order of activity, the personnel involved, the technology used and it's associated advantages and disadvantages.
- A CUD output could be very useful in highlighting communication problems, their causes and potential solutions.
- CUD type analysis seems to be very suited to analysing command and control scenarios.
- It appears that the CUD technique could be modified in order to make it more comprehensive. In particular, a timeline and error occurrence could be incorporated into the CUD output table.
- Although the CUD technique was developed and originally used in telecommunications, it is a generic technique and could potentially be applied in any domain involving communication or collaboration.

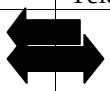
Disadvantages

- Neither time nor error occurrence are catered for by the CUD technique in its current form.
- The initial data collection phase of the CUD technique is very time consuming and labour intensive, including interviews, observational analysis and talk-through analysis.
- No validity or reliability data are available for the technique.
- Application of the CUD technique appears to be limited.
- A team of analysts would be required to conduct a CUD analysis.

Example

The following example is an extract of a CUD analysis that was conducted to assess the suitability of the use of videophones in medical collaborations (Watts & Monk 2000).

Table 26. Example CUD output (adapted from Watts & Monk 2000)

Peterhead Treatment Room	Aberdeen Royal Infirmary Teleradiology Workstation	Comms Resource	Effects of communication medium used
GP discusses X-Ray (N, P, R, Rd) 	C discusses X-Ray	Videophone – handsfree Videophone – picture Teleradiography	+ For all: Freedom to hear and attempt to speak at will - For all: Sound subject to false switching and delay - For GP and C: confidentiality lost
Nurse re-scans X-Ray (GP, P, R, Rd)	Consultant requests better X-Ray image	Image scanner + Teleradiography	+ For GP: learns how to diagnose a new kind of borderline case + For Radiographer: Learns more about radiology + For All: Fast turn-around of expert X-Ray interpretation

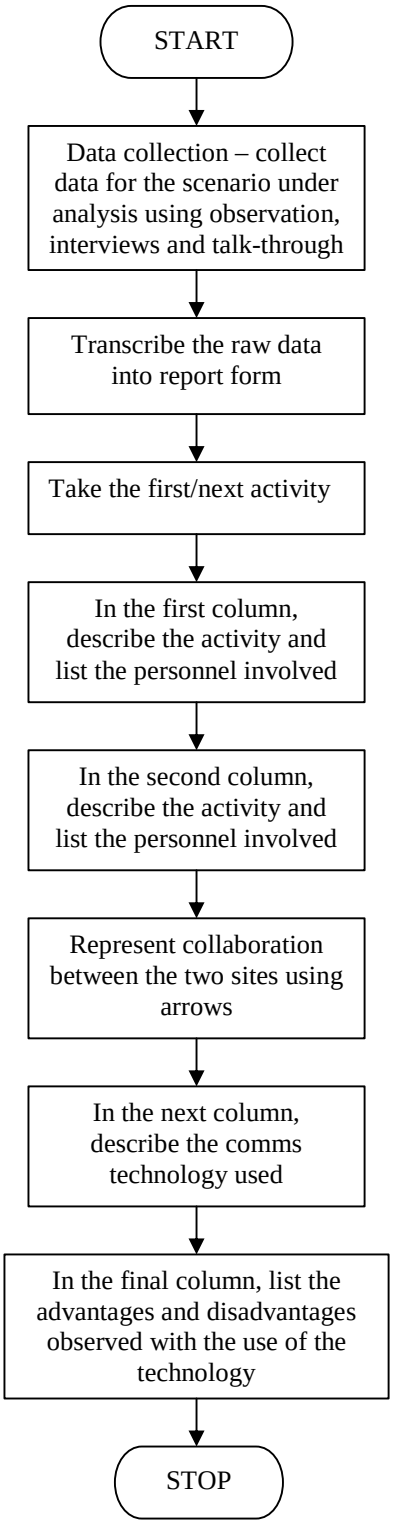
Key:
C = Consultant
P = Patient
R = Relative
Rd = Radiographer
N = Nurse

Table 27 – Participant-percept matrix from site A. Consultants and GP using handsets, GP in front of camera

Percept	GP	Patient (P)	Nurse (N)	Consultant (C)
Hear GP voice	C+	C+	C+	E+
Hear P voice	C+	C+	C+	
Hear N voice	C+	C+	C+	
Hear C voice	E+			C+
See GP face		C+	C?	E+
See P face	C+		C?	E+
See N face		C?		E?
See C face	E+	E+	E+	
See X-Ray	E+	E?	E?	E+
See P's problem	C+	C+	C?	

Key: C+ = Copresent, can hear or see. C? = copresent, can sometimes hear or sometimes see. E+ = Electronic, can hear or see. E? = Electronic, can sometimes hear or sometimes see. Empty cells indicate that the percept is not available to the participant.

Flowchart


START
Data collection – collect data for the scenario under analysis using observation, interviews and talk-through
Transcribe the raw data into report form
Take the first/next activity
In the first column, describe the activity and list the personnel involved
In the second column, describe the activity and list the personnel involved
Represent collaboration between the two sites using arrows
In the next column, describe the comms technology used
In the final column, list the advantages and disadvantages observed with the use of the technology
STOP

Related methods
During the data collection phase, a number of different techniques are used, such as observational analysis, interviews and talk-through type analysis. The CUD technique itself is predominantly a team task analysis technique that focuses upon collaboration or communication.

Approximate training and application times
Whilst no data regarding training and application times for the technique are available, it is apparent that the training time would be low, assuming that the practitioner was already proficient in data collection techniques such as interviews and observational analysis. The application time of the technique, although dependent upon the scenario under analysis, would be high, due to the initial data collection phase.

Reliability and validity
No data regarding the reliability and validity of the technique are available.

Tools needed
A CUD analysis would require the tools associated with the data collection techniques used by the analyst(s). Visual and audio recording equipment would typically be used to record the scenario under analysis and any interviews, and a PC and observer software used to analyse the data. For the CUD comms table, pen and paper are used.

Bibliography
Watts, L. A., & Monk, A. F. (2000) Reasoning about tasks, activities and technology to support collaboration. In J. Annett & N. Stanton (Eds) Task Analysis. UK, Taylor and Francis.

Social Network Analysis

Background and applications

Social Network Analysis (SNA) is a technique used to analyse and represent the relationships existing between teams of personnel or social groups. A social network is a set or team of actors (such as members of a military infantry unit) that possess relationships with one another (Driskell & Mullen In Press). The analysis of these relationships can be used to demonstrate the different types of relationships, the importance and the number of relationships within a specified group. According to Driskell and Mullen (In Press), SNA utilises mathematical and graphical procedures to represent relationships within a group. SNA output typically provides a graphical depiction and a mathematical analysis of the relationships exhibited within the group under analysis. For the mathematical analysis part of SNA, Driskell & Mullen (In Press) recommend that the concept of centrality is rated. Centrality is divided into three components;

1) Degree – represents the number of positions in the group that are in direct contact with the position in question.
2) Betweenness – the number of times a position falls between pairs of positions in the group.
3) Closeness – the extent to which the position in question is close to the other positions in the group.

Each component should be rated between 0 and 1 (0 = Low centrality, 1 = High centrality).

Domain of application

Generic.

Procedure and advice (adapted from Driskell & Mullen (In Press))

Step 1: Define network or group

The first step in a SNA involves defining the network or group of networks that are to be analysed. For example, when analysing command and control networks, a number of different control room networks could be considered, such as military, police, ambulance, railway and air traffic control rooms.

Step 2: Define scenarios

Typically, a SNA requires that the network is analysed in a specific scenario. Once the type of network under analysis has been defined, the scenario within which they will be analysed should be defined. For a thorough analysis, a number of different scenarios should be analysed.

Step 3: Define set of relationships

Once the type of network or group and scenario has been defined, it is then useful to define the relationships within the network that are to be analysed. A number of relationships are typically considered, including roles (e.g. footsoldier, commander), interaction between network members (e.g. information communication, task collaboration) and environmental relationships (e.g. location, proximity). The relationships considered in a SNA are dependent upon the focus and scope of the analysis. For example, in analysing relationships within command and control networks, the relationships considered would include roles, communication, location,

task performance and task collaboration. Whilst it is recommended that the relationships under analysis are defined before any data collection occurs, it should also be stressed that the set of defined relationships are not rigid, in that any novel relationships undefined but exhibited by the network during the data collection phase can also be added to the analysis.

Step 4: Data collection
Once the network and the relationships to be analysed are defined clearly, the data collection phase can begin. The data collection phase involves the collection of specific data on the relationship variables specified during the required scenarios. Typical human factors data collection techniques should be used in this process, such as observational analysis, interviews and questionnaires. Data can be collected either in real world settings or in scenario simulations.

Step 5: Measure/Analyse relationships
The relationships observed should then be analysed or measured. According to Driskell & Mullen (2003) centrality should be measured via assessing three concepts;

1) Degree – number of positions in the network in direct contact with a given position.
2) Betweenness – The number of times a position falls between pairs of other positions within the network under analysis.
3) Closeness – the extent to which a position is close to the other positions within the network under analysis.

Each concept should be rated between 0 and 1 (0 = Low centrality, 1 = High centrality).

Step 6: Construct social network graph/matrix
Typically, social networks are represented in graphs or matrices.

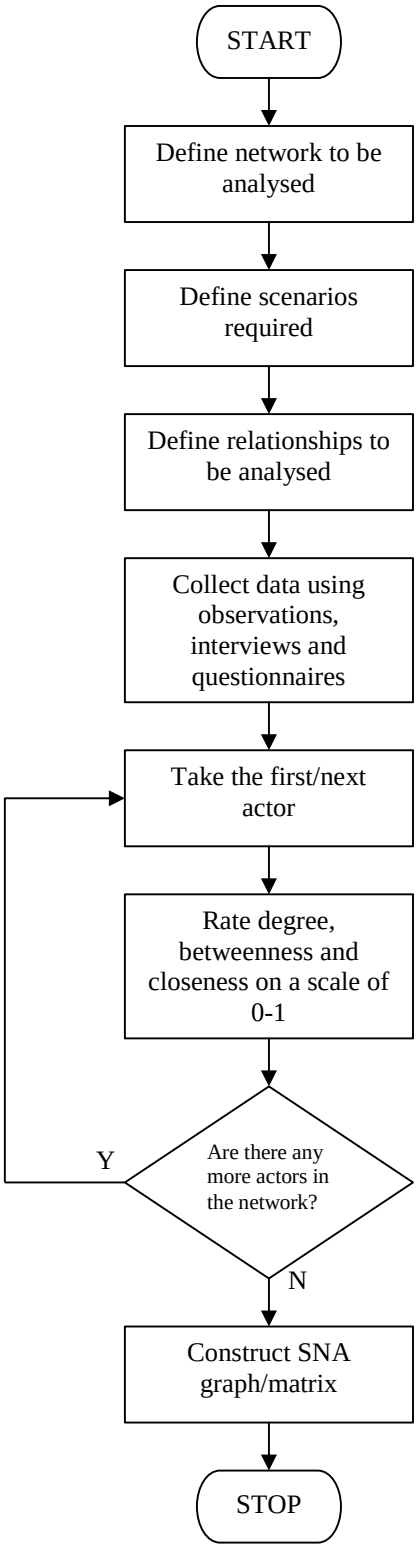
Advantages

- SNA could be used to highlight the importance of positions within a network or group. Conversely, those positions that appear to be of little importance to the network could also be classified.
- SNA analyses the importance of relationships between operators in a specified network.
- SNA seems to be suited to analysing the importance of relationships in control room networks.
- SNA is a generic technique that has the potential to be applied in any domain.

Disadvantages

- For complex networks, it would be difficult to conduct a SNA.
- The data collection phase involved in a SNA is resource intensive.
- SNA would require more training than other team task analysis techniques.
- The SNA would be prone to the various flaws associated with observational analysis, interviews and questionnaires.
- SNA is time consuming in its application.

Flowchart


START
Define network to be analysed
Define scenarios required
Define relationships to be analysed
Collect data using observations, interviews and questionnaires
Take the first/next actor
Rate degree, betweenness and closeness on a scale of 0-1
Are there any more actors in the network?
Y
N
Construct SNA graph/matrix
STOP

Example
The following example is taken from Driskell & Mullen (In Press).

Table 28 – SNA matrice

Matrices for two social networks.												
Network A							Network B					
	A	B	C	D	E			A	B	C	D	E
A	-	0	1	0	0		A	-	1	1	0	1
B	0	-	1	0	0		B	1	-	1	0	0
C	1	1	-	1	1		C	1	1	-	1	1
D	0	0	1	-	0		D	0	0	1	-	1
E	0	0	1	0	-		E	1	0	1	1	-

Figure 25 - Five-person networks, with indices of centrality.

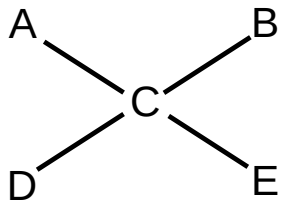
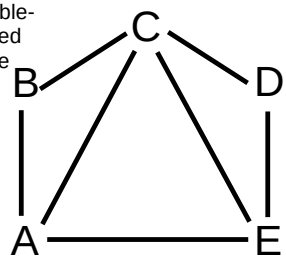
Network A	Position	Degree	Betweenness	Closeness
Wheel 	A	.25	.00	.57
	B	.25	.00	.57
	C	1.00	1.00	1.00
	D	.25	.00	.57
	E	.25	.00	.57
	Network A	1.00	1.00	1.00
Network B				
Double-barred circle 	A	.75	.08	.80
	B	.50	.00	.67
	C	1.00	.33	1.00
	D	.50	.00	.67
	E	.75	.08	.80
	Network B	.50	.29	.62

Figure 24 – SNA output (Driskell & Mullen In Press)

Related methods
In terms of rating relationships in networks the SNA technique appears to be unique. In the data collection phase, techniques such as observational study, interviews and questionnaires are typically used.

Approximate training and application times
Although no data regarding the training and application times associated with SNA are available in the literature, it is apparent that it would be high in both cases. SNA appears to be complex in its application and so training and application time would be high. The data collection involved in a SNA would also add further time cost to the techniques overall application time. Of course, the actual application time would be dependent upon the complexity of the network(s) under analysis. A SNA of complex networks involving a high number of actors with numerous relationships would require a great deal of time.

Reliability and validity
No data regarding the reliability and validity of the SNA technique are available.

Tools needed
The SNA can be conducted using pen and paper, once the data collection phase is complete. The tools required during the data collection phase for a SNA would be dependent upon the type of data collection techniques used. Observational analysis, interviews and questionnaires would normally require visual and audio recording equipment (video cameras, minidisc recorder, PC). Driskell & Mullen (In Press) recommend that the UCINET and STRUCTURE software packages are used during a SNA.

Bibliography
Driskell, J. E. & Mullen, B. (In Press) Social Network Analysis. In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.

Groupware Task Analysis

Martijn van Welie & Gerrit C. Van Der Veer, Department of Computer Science, Vrije University, The Netherlands

Background and applications

Groupware Task Analysis (GTA) is a team task analysis technique that is used to study and evaluate group or team activities in order to inform the design and analysis of similar team systems. GTA comprises a conceptual framework focussing upon the relevant aspects that require consideration when designing systems or processes for teams or organisation. The technique entails the description of two task models.

1. Task model 1 – Task model 1 is essentially a description of the situation at the current time in the system that is being designed. This is developed in order to enhance the design teams understanding of the current work situation. In the design of C4i systems, Task Model 1 would include a description of the command and control systems that are currently used.
2. Task model 2 – Task model 2 involves re-designing the current system or situation outlined in task model 1. This should include technological solutions to problems highlighted in task model 1 and also technological answers to requirements specified (Van Welie & Van Der Veer 2003). Task model 2 should represent a model of the future task world when the new design is implemented.

According to (Van Welie & Van Der Veer 2003), task models should consist of the following components.

- Agents – refers to the personnel involved in the system under analysis, including teams and individuals. Agents should be described in terms of their goals, roles (which tasks the agent is allocated), organisation (relationship between agents and roles) and characteristics (agent experience, skills etc)
- Work – The task or tasks under analysis should be described, including unit and basic task specification (Card, Moran & Newell 1983). It is recommended that a HTA is used for this aspect of task model 1. Events (triggering conditions for tasks) should also be described.
- Situation – the situation description should include a description of the environment and any objects in the environment.

The techniques used when conducting a GTA are determined by the available resources. For guidelines on which techniques to employ the reader is referred to Van Welie & Van Der Veer (2003). Once the two task models are completed, the design of the new system can begin, including specification of functionality and also the way in which the system is presented to the user (Van Welie & Van Der Veer 2003). According to the authors, the task model can be used to answer the following design questions (Van Welie & Van Der Veer 2003).

- What are the critical tasks?
- How frequently are those tasks performed?
- Are they always performed by the same user?
- Which types of user are there?
- Which roles do they have?
- Which tasks belong to which roles?
- Which tasks should be possible to undo?
- Which tasks have effects that cannot be undone?

- Which errors can be expected?
- What are the error consequences for users?
- How can prevention be effective?

Domain of application
Generic.

Procedure and advice
Step 1: Define system under analysis
The first step in a GTA is to define the system(s) under analysis. For example, in the design of C4i systems, existing command and control systems would be analysed, including railway, air traffic control, security and gas network command and control systems.

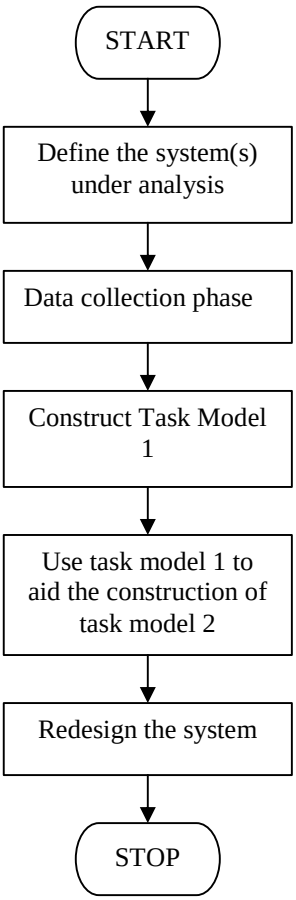
Step 2: Data collection phase
Before task model 1 can be constructed, specific data regarding the existing systems under analysis should be collected. Traditional technique should be used during this process, including observational analysis, interviews and questionnaires. The data collected should be as comprehensive as possible, including information regarding the task (specific task steps, procedures, interfaces used etc), the personnel (roles, experience, skills etc) and the environment.

Step 3: Construct task model 1
Once sufficient data regarding the system or type of system under analysis has been collected, task model 1 should be constructed. Task model 1 should completely describe the situation as it currently stands, including the agents, work and situation categories outlined above.

Step 4: Construct task model 2
The next stage of the GTA is to construct task model 2. Task model 2 involves re-designing the current system or situation outlined in task model 1. The procedure used for constructing task model 2 is determined by the design teams, but may include focus groups, scenarios and brainstorming sessions.

Step 5: Redesign the system
Once task model 2 has been constructed, the system re-design should begin. Obviously, this procedure is dependent upon the system under analysis and the design team involved. The reader is referred to Van Welie & Van Der Veer (2003) for guidelines.

Flowchart



Advantages

- GTA output provides a detailed description of the system requirements and highlights specific issues that need to be addressed in the new design.
- Task model 2 can potentially highlight the technologies required and their availability.
- GTA provides the design team with a detailed understanding of the current situation and problems.
- GTA seems to be suited to the analysis of existing command and control systems.

Disadvantages

- GTA appears to be extremely resource intensive and time consuming in its application.
- Limited evidence of use in the literature.
- The technique provides limited guidance for its application.
- A large team of analysts would be required in order to conduct a GTA analysis.

Example

For an example GTA, the reader is referred to Van Welie & Van Der Veer (2003).

Related methods
GTA analysis is team task analysis technique and so is related to CUD, SNA and team task analysis. When using GTA, a number of different techniques can be employed, including observation, interviews, surveys, questionnaires and HTA.

Approximate training and application times
It estimated that the training and application times for the GTA technique would be very high.

Reliability and Validity
There are no data regarding the reliability and validity of the GTA technique available in the literature.

Tools needed
Once the initial data collection phase is complete, GTA can be conducted using pen and paper. The data collection phase would require video and audio recording devices and a PC.

Bibliography
Van Welie, M., & Van Der Veer, G. (2003) Groupware Task Analysis. In E. Hollnagel (Ed) Handbook of Cognitive Task Design. Pp 447 – 477. Lawrence Erlbaum Associates Inc.

Team Task Analysis

C. Burke, Institute for Simulation and Training, University of Central Florida, 3280 Progress Dr. Orlando, FL 32826-0544, USA

Background and applications

Team Task Analysis (TTA) is a task analysis technique that provides a description of tasks distributed across a team and the requirements associated with the tasks in terms of operator knowledge, skills, and abilities. According to Baker, Salas and Bowers (1998) TTA refers to the analysis of a team tasks and also the assessment of a teams teamwork requirements (Knowledge, skills and abilities) and that TTA forms the foundation for all team resource management functions. The recent increase in the use of teams and a renewed focus upon team training and team performance measures has been accompanied by a renewed research emphasis upon TTA techniques (Baker & Salas, 1996; Bowers, Baker, & Salas, 1994; Bowers, Morgan, Salas, & Prince, 1993, Campion, Medsker, & Higgs, 1993; Campion, Papper, & Medsker; 1996). Typically, TTA is used to inform team task design, team training procedures and team performance measurement. TTA aims to analyse team-based scenarios by gathering data regarding teamwork and taskwork.

- Teamwork – individuals interacting or co-ordinating tasks that are important to the teams goals (Baker and Salas,XXXX).
- Taskwork – individuals performing individual tasks.

According to Burke (2003), the TTA procedure has not yet been widely adopted by organisations, with the exception of the US military and aviation communities. Never the less, TTA appears to be a very useful procedure for eliciting data regarding operating skills and team co-ordination. Although a set procedure for TTA does not exist, Burke (2003) attempted to integrate the existing TTA literature into a set of guidelines for conducting a TTA.

Domain of application

Generic.

Procedure and advice (Adapted from Burke 2003)

Step 1: Conduct requirements analysis

Firstly, a requirements analysis should be conducted. This involves clearly defining the task scenario to be analysed, including describing all duties involved and also conditions under which the task is to be performed. Burke (2003) also suggests that when conducting the requirements analysis, the methods of data collection to be used during the TTA should be determined. Typical TTA data collection methods are observational techniques, interviews, questionnaires, interviews and surveys. The requirements analysis should also involve determining the participants that will be involved in the data collection process, including occupation and number.

Step 2: Task identification

Next, the tasks involved in the scenario under analysis should be identified and listed. A HTA could potentially be used for this step. Burke (2003) recommends that interviews with SME's, observation and source documents should be used to identify the full set of tasks. Once each individual task step is identified, a task statement should be written (for each task step), including the following information:

- Task name
- Task goals

- What the individual has to do to perform the task
- How the individual performs the task
- Which devices, controls, interfaces are involved in the task
- Why the task is required

Step 3: Identify teamwork taxonomy
Once all of the tasks involved in the scenario under analysis have been identified and described fully, a teamwork taxonomy should be identified (Burke 2003). The aim of this is to determine which of the tasks involved in the scenario are taskwork (individual) and which are teamwork (team). According to Burke (2003) several teamwork taxonomies exist in the literature.

Step 4: Conduct a co-ordination analysis
Once the teamwork taxonomy is defined, a co-ordination analysis should be conducted. The aim of this is to identify which of the identified tasks require the team to co-ordinate their activities (Burke 2003) to perform the task i.e. which of the tasks require teamwork. Burke (2003) suggests that surveys should be used for this process, however a number of techniques can be used, such as questionnaires and interviews with SME's.

Step 5: Determine relevant taskwork and teamwork tasks
At this stage of the TTA, the analyst should have a list of all the tasks involved in the scenario under analysis, and a list of taskwork and teamwork tasks. The next step of the TTA is to determine the relevance of each of the tasks. Burke suggests that likert scale questionnaire is used for this step and that the following task factors should be rated:

- Importance to train
- Task frequency
- Task difficulty
- Difficulty of learning
- Importance to job

A standardised set of task indices is yet to be developed (Burke 2003). It is recommended that the task indices used should be developed based upon the overall aims and objectives of the TTA.

Step 6: Translation of tasks into KSAO's
Next, the knowledge, skills, abilities and attitudes (KSAO) for each of the relevant task steps should be determined. Normally, interviews or questionnaires are used to elicit the required information from SME's.

Step 7: Link KSAO's to team tasks
The final step in the TTA is to link the KSAO's identified in step 6 to the individual tasks. Burke (2003) suggests that is most often achieved through the use of surveys completed by SME's. According to Burke (2003), the SME is asked if the KSAO for the task is helpful or irrelevant.

Advantages

- TTA goes further than individual task analysis techniques by specifying the knowledge, skills and abilities required to complete each task step.

- The output from TTA can be used in the development of team training procedures and in team job design.
- The TTA output specifically states which tasks are team based and which tasks are individually performed. This is extremely useful when designing new systems.
- TTA can be used to address team task performance issues.
- TTA provides a systematic view of the tasks that make up the scenario under analysis.
- TTA could be used in the identification of team-based errors.

Disadvantages

- TTA is a hugely time consuming technique to conduct.
- SME's and domain experts are required throughout the procedure. The acquisition of SME's can sometimes prove very difficult.
- There is no rigid procedure for the TTA technique. As a result, reliability is questionable.
- Great skill is required on behalf of the analyst in order to elicit the required information throughout the TTA procedure.

Related methods

There are a number of different approaches to team task analysis, such as TTRAM, CUD and SNA. TTA also utilises a number of human factors data collection techniques, such as interviews, questionnaires and surveys.

Approximate training and application times

Due to the methods infancy, there are limited estimates for the training and application times associated with the TTA technique. It is estimated that it would be high for both training and application. Certainly the use of interviews, questionnaires and surveys during the technique ensure high application and analysis times. Burke (2003) estimates

Tools needed

The tools required for conducting a TTA are dependent upon the methodologies used during the procedure. TTA can be conducted using pen and paper, and a visual or audio recording device. A PC with a word processing package such as Microsoft Word is normally used to transcribe and sort the data.

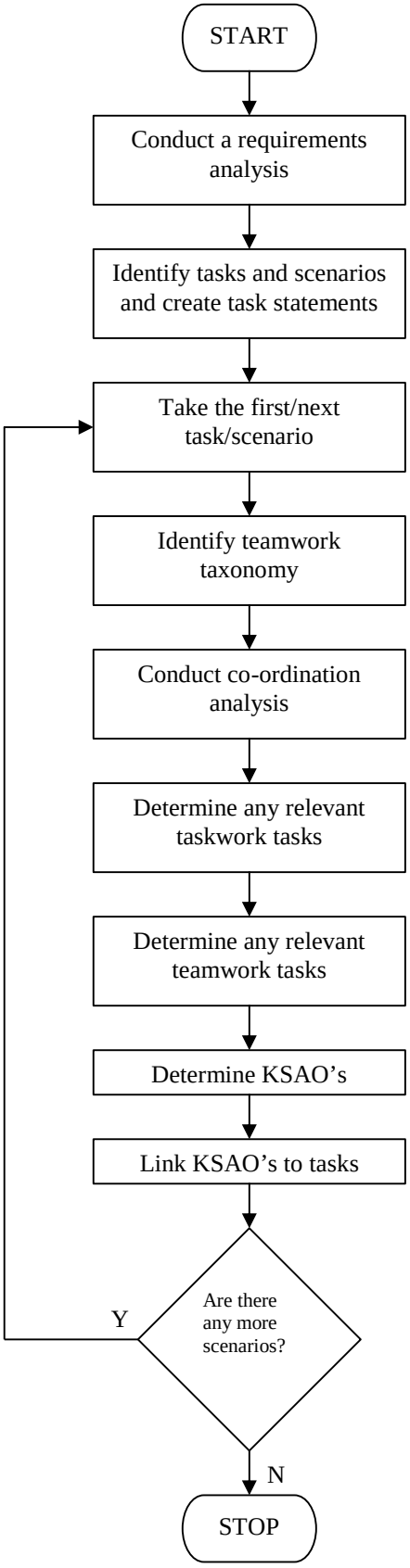
Bibliography

Bowers, C. A., Morgan, B. B., Salas, E., & Prince, C. (1993) Assessment of co-ordination demand for aircrew co-ordination training. *Military Psychology*, 5(2), 95-112.

Bowers, C. A., Baker, D. P., & Salas, E. (1994). Measuring the importance of teamwork: The reliability and validity of job/task analysis indices for team training design. *Military Psychology*, 6(4), 205-214

Burke, C. S. (2003) Team Task Analysis. In N. Stanton, Hedge, Hendrick , K. Brookhuis , E. Salas (Eds) *Handbook of Human Factors and Ergonomics Methods.* UK, Taylor and Francis

Flowchart


START
Conduct a requirements analysis
Identify tasks and scenarios and create task statements
Take the first/next task/scenario
Identify teamwork taxonomy
Conduct co-ordination analysis
Determine any relevant taskwork tasks
Determine any relevant teamwork tasks
Determine KSAO's
Link KSAO's to tasks
Are there any more scenarios?
Y
N
STOP

HTA (T)

John Annett, Department of Psychology, Warwick University, Coventry CV4 7AL

Background and Applications

HTA involves breaking down the task under analysis into a hierarchy of goals, operations and plans. Tasks are broken down into hierarchical set of tasks, sub tasks and plans. The goals, operations and plans categories used in HTA are described below.

- Goals – The unobservable task goals associated with the task in question.
- Operations – The observable behaviours or activities that the operator has to perform in order to accomplish the goal of the task in question.
- Plans – The unobservable decisions and planning made on behalf of the operator.

A more recent variation of HTA that caters for the task analysis of team-based tasks is described by Annett (In Press).

Domain of application

Generic.

Procedure and advice

The reader is referred to the HTA procedure and advice section on page XX

Example

The following example is taken from an analysis of anti-submarine warfare teams (Annett In Press). According to the author, the purpose of the analysis was to identify and measure team skills critical to successful anti-submarine warfare.

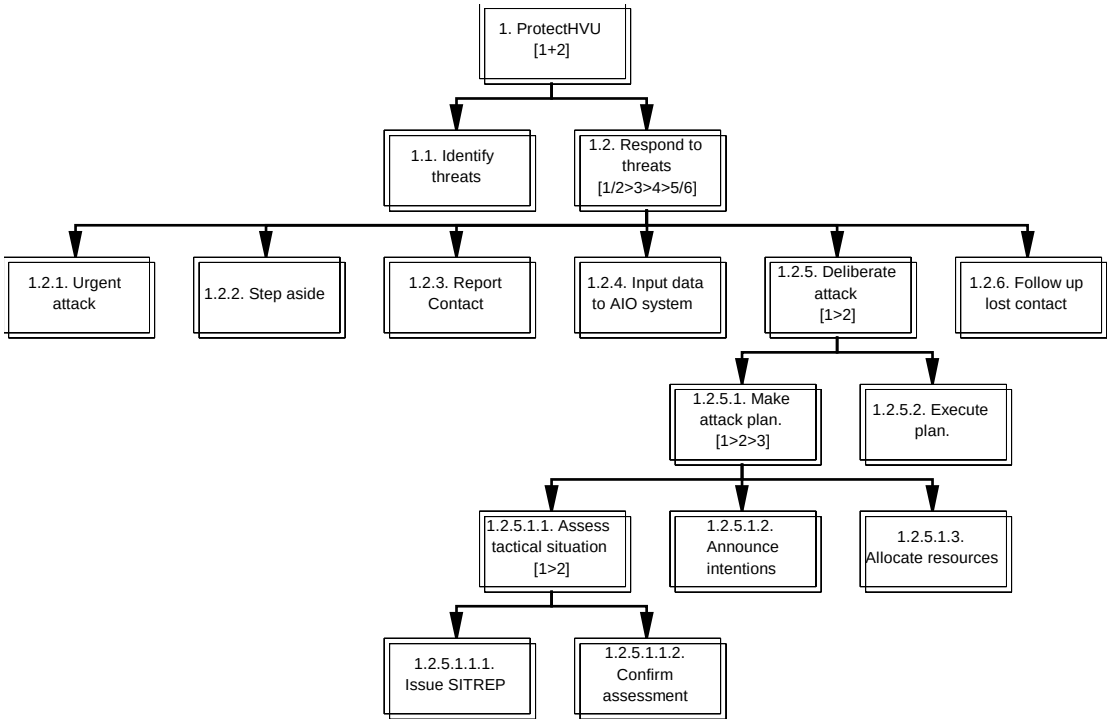
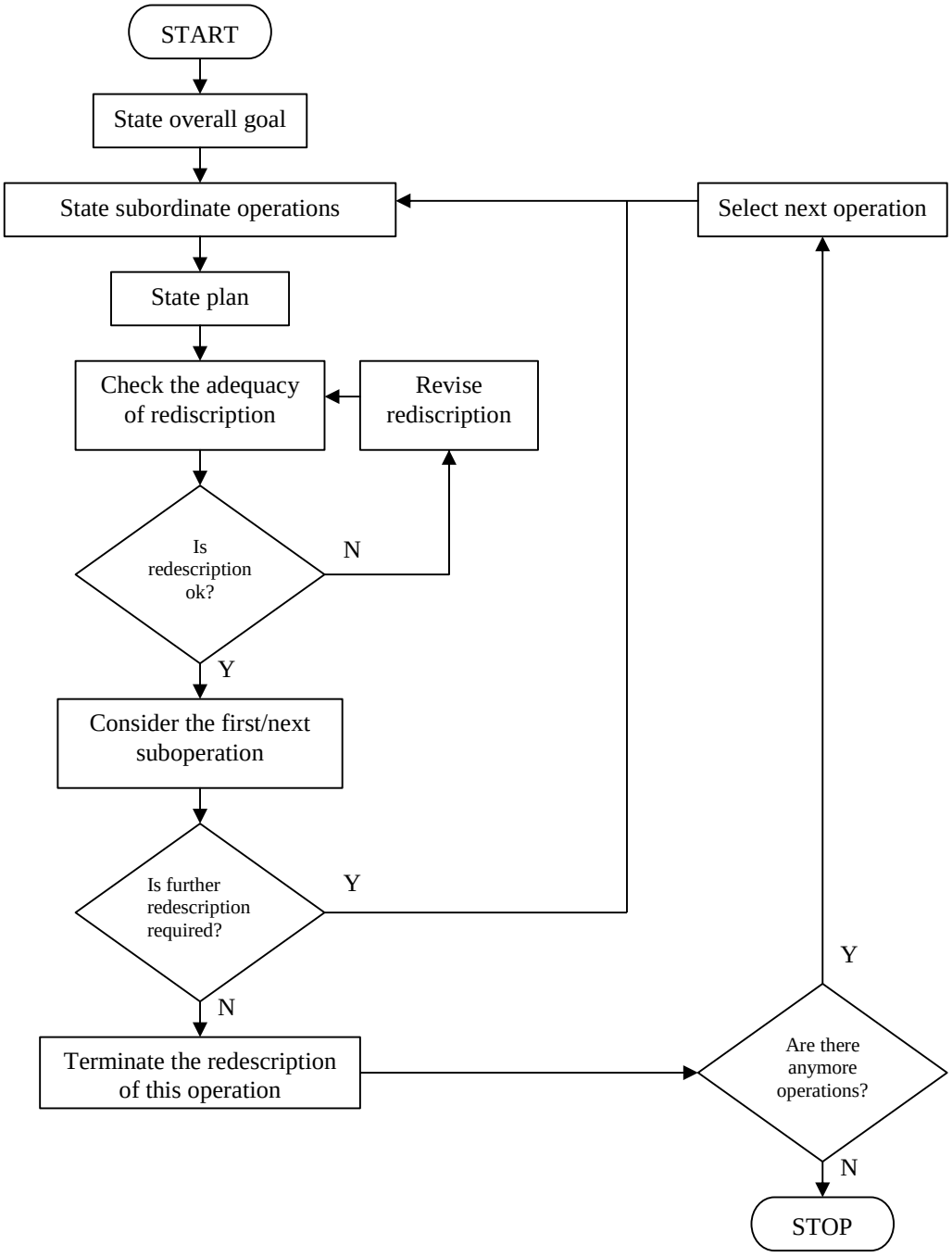


Figure 26. Extract from an analysis of an Anti-submarine Warfare tea Task (Source: Annett In Press).

Table 29. Tabular form of selected ASW team operations. (Source: Annett In Press)

1. Protect Highly Valued Unit (HVU) [1+2]	Goal: Ensure safe & timely arrival of HVU Teamwork: PWO in unit gaining initial contact with threat assumes tactical command and follows standard operating procedures in this role. Plan: Continues to monitor threats[1] whilst responding to identified threat. Criterion measure: Safe & timely arrival of HVU
1.2. Respond to threats. [1/2>3>4>5/6]	Goal: Respond to threat according to classification. Teamwork: PWO selects response based on information provided by other team members. Plan: If threat is immediate (e.g. torpedo) go to urgent attach [1.2.1.] else execute 2,3,4 and 5 or 6. Criterion measure: Appropriate response with minimal delay.
1.2.5. Deliberate attack. [1>2]	Goal: Get weapon in water within 6 minutes. Teamwork: See further breakdown below. Plan: Make attack plan then execute. Criterion measure: Time elapsed since classification and/or previous attack.
1.2.5.1. Make attack plan. [1>2>3]	Goals: Plan understood and accepted by team. Teamwork: Information regarding tactical situation and resources available from team members to PWO. Plan: Assess tactical situation; announce intentions; allocate resources. Criterion measure: Accurate information provided.
1.2.5.1.1. Assess tactical situation. [1>2]	Goal: Arrive at correct assessment of tactical situation. Teamwork: PWO must gather all relevant information by up-to-date status reports from own team and sensors and other friendly forces. Plan: Issue SITREP then confirm assessment. Criterion measures: Correct assessment; time to make assessment.
1.2.5.1.1.1. Issue SITREP	Goal: To ensure whole team is aware of threat situation and to provide an opportunity for other team members to check any omissions or errors in tactical appreciation. Teamwork: PWO issues situation report (SITREP) at appropriate time; all team members check against information they hold. Criterion measure: All team members have accurate tactical information.
1.2.5.1.1.2. Confirm tactical assessment	Goal: Construct an accurate assessment of the threat and of resources available to meet it. Teamwork: Final responsibility lies with the PWO but information provided by and discussion with other team members essential to identify and resolve any inconsistencies. Criterion measure: Accurate assessment in light of information and resources available.

Flowchart


START
State overall goal
State subordinate operations
State plan
Check the adequacy of rediscription
Revise rediscription
Is redescription ok?
N
Y
Consider the first/next suboperation
Is further redescription required?
Y
N
Terminate the redescription of this operation
Select next operation
Y
Are there anymore operations?
N
STOP

Related Methods
HTA is widely used in HF and often forms the first step in a number of analyses, such as HEI, HRA and mental workload assessment. Annett (In Press) reports that HTA has been used in a number of applications, for example as the first step in the TAFEI method for hazard and risk assessment (Baber & Stanton, 1994), in SHERPA for predicting human error (Baber & Stanton, 1996), in MUSE usability assessment (Lim & Long, 1994), the SGT method for specification of information requirements (Ormerod, Richardson & Shepherd, 1998/2000), and the TAKD method for the capture of task knowledge requirements in HCI (Johnson, Diaper & Long, 1984).

Approximate Training and Application Times
According to Annett (2003), a study by Patrick, Gregov and Halliday (2000) gave students a few hours training with not entirely satisfactory results on the analysis of a very simple task, although performance improved with further training. A survey by Ainsworth & Marshall (1998/2000) found that the more experienced practitioners produced more complete and acceptable analyses. Stanton & Young (1999) report that the training and application time for HTA is substantial. The application time associated with HTA is dependent upon the size and complexity of the task under analysis. For large, complex tasks, the application time for HTA would be high.

Reliability and Validity
There are no data regarding the reliability and validity of HTA used for team task analysis purposes available in the literature.

Tools needed.
HTA can be carried out using only pencil and paper.

Bibliography
Annett, J. (In Press) Hierarchical Task Analysis (HTA). In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.

Design techniques

Design techniques are a general classification used for the purposes of this review. A design technique in this is case merely implies that the technique is one that is used by designers during the early design process. This includes interface design techniques (such as Link analysis checklists and heuristics), and group design techniques (such as focus groups and design scenarios). The techniques reviewed in this document represent those HF techniques that are typically used during the early design process of systems.

Link analysis is an interface analysis and design technique that records and represents the nature, frequency and importance of links between elements of a systems interface. Used to improve interface design, link analysis defines links (hand or eye movements) between elements of the interface under analysis. The interface is then re-designed based upon these links, with the most often linked elements of the interface relocated to increase their proximity to one another.

Layout analysis is another interface analysis technique that is used to offer a redesign of the interface under analysis. Layout analysis involves arranging the interface components into functional groupings, and then organising these groups by importance of use, sequence of use and frequency of use. The layout analysis output offers a redesign based upon the user's model of the task.

Task centred system design (TCSD) is a quick and easy approach to evaluating system design involving the identification of the potential users and tasks associated with the design concept and evaluating the design using design scenario's and a walkthrough type analysis. The technique offers a redesign of the interface or system design under analysis as its output. A typical TCSD involves gathering data from an existing design and redesigning the system using design scenarios and system task walkthrough's.

A Focus group is a group interview approach that involves using a group of SME's to discuss a particular design concept or prototype. Focus groups are extremely flexible and can be used for almost any purpose.

Scenario based design involves the use of imaginary scenarios to communicate or evaluate design concepts. A set of scenarios depicting the future use of the design concept are proposed and performed, and the design concept is evaluated. Scenarios typically use how, why and what if questions to evaluate and modify a design concept.

The checklist style approach is a very simple approach whereby the analyst checks the product or system design against a pre-defined set of criteria in order to evaluate the design. Conducting a checklist analysis is a matter of simply inspecting the device against each point on the chosen checklist. .

Heuristic type analysis is one of the simplest design techniques available, involving simply obtaining analyst(s) subjective opinions on a design concept or product. In conducting a heuristic analysis, an analyst or team of analysts should interact with the design under analysis and make observations regarding the usability, quality, and error potential of the design.

Walkthrough analysis is a very simple procedure used by designers whereby experienced system operators perform a walkthrough or demonstration of a task or set of tasks using the system under analysis. Walkthroughs are typically used early in the design process to envisage how a design would work and also to evaluate and modify the design concept. A walkthrough involves an operator walking through a scenario performing the actions that would occur and describing the functions of controls and displays used.

The design techniques reviewed in this document are shown below:

1. Link Analysis
2. Layout Analysis
3. TCSD – Task-Centred System Design
4. Focus groups
5. Scenario analysis
6. Checklists
7. Heuristic analysis
8. Walkthrough analysis

Link Analysis

Background and applications

Link analysis is an interface evaluation technique that is used to determine 'links' in a system between interface components and operations and to determine the nature, frequency and importance of these links. Links are defined as movements of attentional gaze or position between parts of the system, or communication with other system elements. For example, if an operator performing a task is required to press button A and then button B in sequence to accomplish the task goal, a link between button's A and B is defined. Link analysis uses spatial diagrams as its output, with each link represented by a straight line between the 'linked' interface elements. Specifically aimed at aiding the design of interfaces and systems, link analyses most obvious use is in the area of workspace-layout optimisation (Stanton & Young 1999) i.e. the placement of controls and displays according first to their importance, then to their frequency of use, then to their function within the system and finally to their sequence of use (Grandjean 1988). Link analysis was originally aimed at process control rooms (Stanton and Young, 1999) but it can be applied to any system where the user exhibits hand or eye movements. Link analysis can be used to analyse the layout of panel displays in any domain, such as driving, control rooms, aviation, air traffic control etc. Link analysis has the potential to be used in the design and evaluation of any interface or system and to date has been used in a wide variety of domains, including the design and evaluation of In-Car Radios (Stanton and Young 1999). When conducting a link analysis, establishing the links between system/interface components is normally achieved through a walkthrough or observation analysis. Link analysis can be used to analyse either hand or eye movements. The output of a link analysis is normally a link diagram and also a link table (both depict the same information). The link diagram and table can be used to suggest revised layouts of the components for the device, based on the premise that links should be minimised in length, particularly if they are important or frequently used.

Domain of application

Generic.

Procedure and advice

Link analysis can be used to analyse either hand or eye movements.

Step 1: Task analysis/list
Initially, a task analysis or task list for the task under analysis should be constructed. A representative set of tasks with the device in question should be listed.

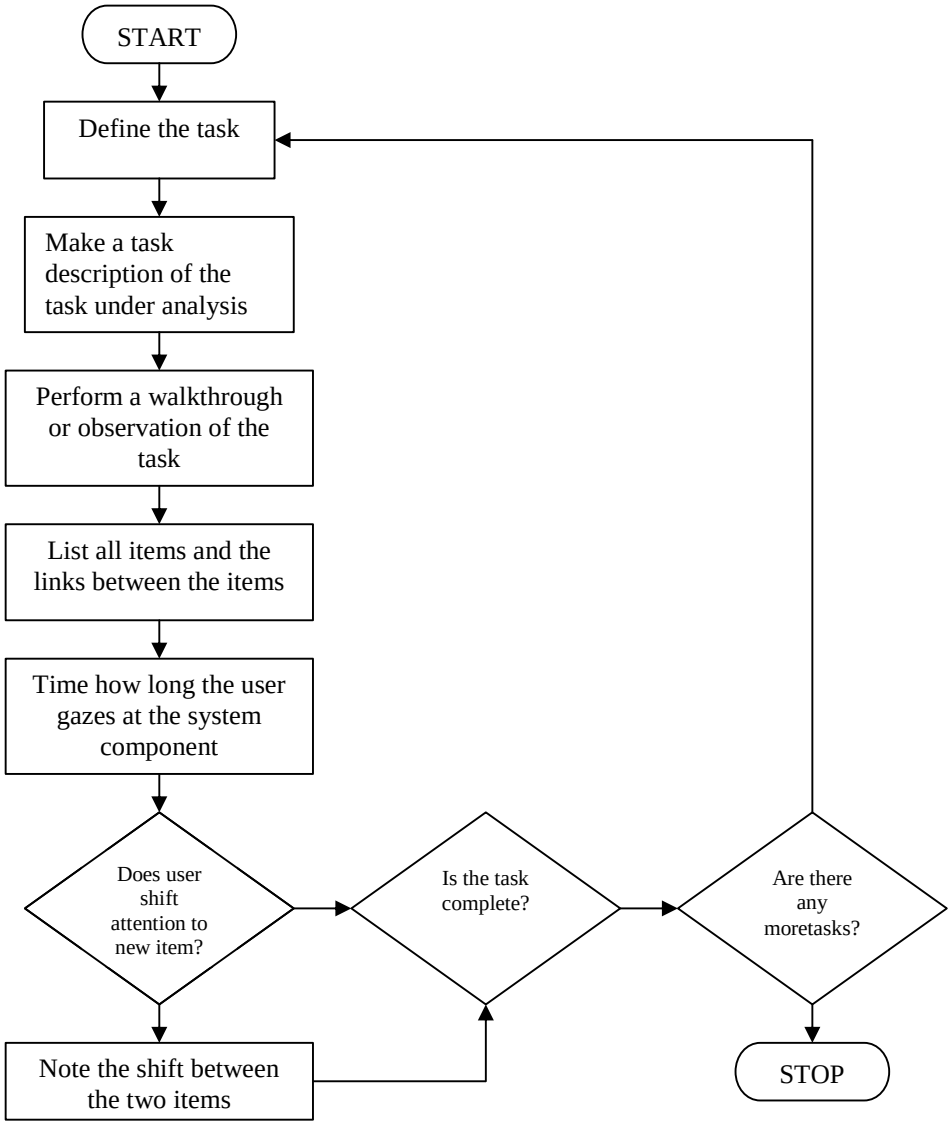
Step 2: Data collection
The analyst should then collect data on the task(s) under analysis. This normally includes performing a walkthrough and an observational analysis of the task. The analyst should record which components are linked by hand/eye movements and how many times these links occur during the task.

Step 3: Link diagram
The analyst should then create a schematic layout of the device/system/interface under analysis. The recorded links should then be added to the diagram in the form of lines joining the linked elements or components. One line represents one link.

Step 4: Link table
Next, the analyst should then complete a link table, which displays the same information as the link diagram, only in a tabular format. Components take positions at the heads of the rows and columns and the numbers of links are entered in the appropriate cells.

Step 5: Redesign
Although not compulsory as part of a link analysis, a redesign for the interface under analysis is normally offered. The redesign is based upon reducing the length between the linked interface components; particularly the most important and frequently used linked components.

Flowchart



Advantages

- Link analysis is a very straightforward technique that requires very little training.
- Link analysis is a quick technique that offers an immediately useful output.
- Link analysis output helps to generate design improvements.
- Link analysis has been used extensively in the past in a number of domains.
- Link analysis output prompts logical redesign of system interfaces.
- Link analysis can be used throughout the design process to evaluate and modify design concepts.

Disadvantages

- A link analysis requires preliminary data collection such as observation, walkthrough type analysis and a HTA.
- Link analysis only considers the basic physical relationship between the user and the system. Cognitive processes and error mechanisms are not accounted for.
- Link analysis output is not easily quantifiable.

Example

The following example is a link analysis performed on the SHARP RG-F832E In-Car Radio (Stanton & Young (1999).

Task List

1. Switch unit on
2. Adjust Volume
3. Adjust Bass
4. Adjust Treble
5. Adjust Balance
6. Choose new Pre-set
7. Use Seek, then Store station
8. Use Manual search, then store station
9. Insert Cassette
10. Autoreverse, then Fast Forward
11. Eject cassette and switch off

Table 30. Table showing Ford In-Car Radio components and functions (Stanton & Young 1999)

A = On/Off/Volume/Balance/Fadar	H = Tape Eject Button
B = Treble Bass	I = Cassestte Compartment
C = Station Preset Buttons	J = Fast Wind/Programme Buttons
D = FM Mono Stereo Button	K = Tuning Up/Down Buttons
E = DX-Local Button	L = Tuning Scan/Seek Buttons
F = Band Selector Button	M = Tuning Scan/Seek Buttons
G = ASPM/Preset Memory Scan Button	

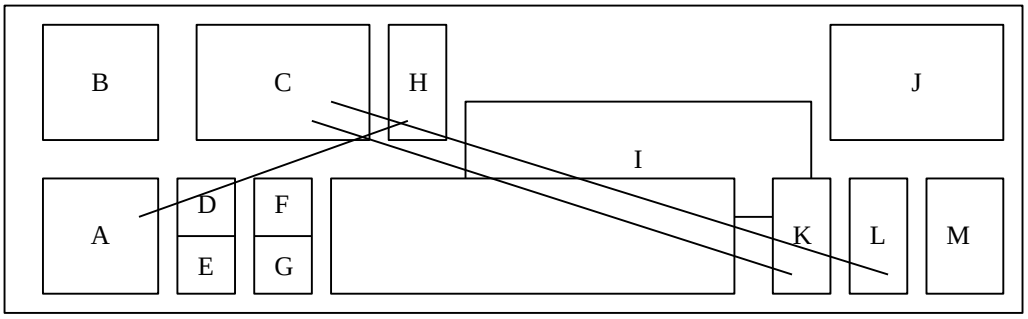


Figure 27. Link diagram for Ford In-Car Radio (Stanton & Young 1999)

Table 31. Link table for Ford In-Car Radio (Stanton & Young 1999)

A	**X**												
B		**X**											
C			**X**										
D				**X**									
E					**X**								
F						**X**							
G							**X**						
H	**1**							**X**					
I									**X**				
J										**X**			
K		**1**									**X**		
L		**1**										**X**	
M													X
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**

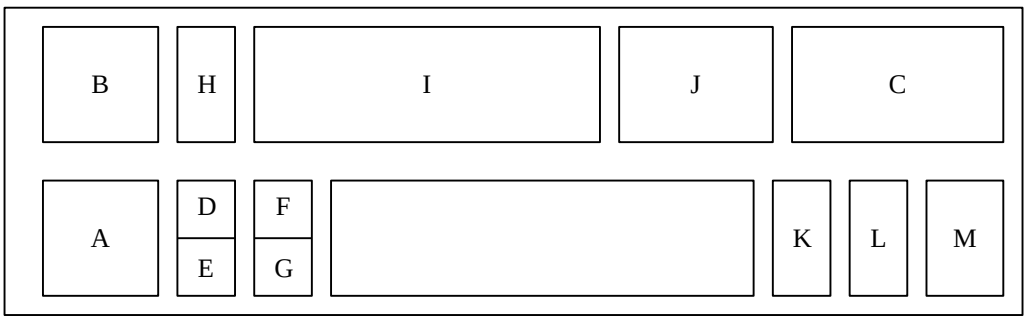


Figure 28. Revised design for Ford In-Car Radio (Stanton & Young 1999)

Related methods
A link analysis normally requires an initial task description to be created for the task under analysis, such as a HTA. Also, an observation or walkthrough analysis should be performed in order to establish the links between components in the system. Stanton & Young (1999) also suggest that it is helpful to be in the possession of a HTA for the device under analysis.

Approximate training and application times
Stanton & Young (1999) report that the link analysis technique is relatively fast to train and practice and also that execution time is moderate compared to a number of other techniques, including SHERPA, layout analysis, repertory grids, checklists and TAFEI.

Reliability and Validity
Stanton & Young (1999) reported that Link analysis performed particularly well on measures of intra-rater reliability and predictive validity. They also reported, however, that the technique is let down by poor inter-rater reliability.

Tools needed
When conducting a link analysis the analyst should have the device under analysis, pen and paper, and a stopwatch. For the observation part of the analysis, a video recording device is required. An eye tracker device can also be used to record fixations during the task performance.

Bibliography
Drury, C. G. (1990) Methods for direct observation of performance, In Wilson, J. And Corlett, E. N. Evaluation of Human Work: A practical Ergonomics Methodology, 2nd Edition, London, Taylor and Francis, 45 –68
Kirwan, B. (1994) A guide to practical human reliability Assessment, London, Taylor and Francis
Stanton, N. A, and Young, M. S. (1999) A guide to methodology in ergonomics: Designing for human use, London, Taylor and Francis.

Layout Analysis

Background and applications

Layout analysis is similar to link analysis in that it is based on spatial diagrams of the product and its output directly addresses interface design. Layout analysis simply analyses an existing design and suggests improvements to the interface arrangements based on functional grouping. The theory behind layout analysis is that the interface should mirror the users structure of the task and the conception of the interface as a task map greatly facilitates design (Easterby, 1984). Layout analysis begins by simply arranging all of the components of the interface into functional groupings. These groups are then organised by importance of use, sequence of use and frequency of use. The components within each functional group are then re-organised, once more this is done according to importance, sequence and frequency of use. The components within a functional group will then stay in that group throughout the analysis and they cannot move anywhere else in the re-organisation stage. At the end of the process, the analyst has redesigned the device in accordance with the users model of the task based upon importance, sequence and frequency of use.

Domain of application

Generic.

Procedure and advice

Step 1: Schematic diagram
First, the analyst should create a schematic diagram for the device under analysis. This diagram should have each interface element clearly labelled.

Step 2: Arrange interface components into functional groupings
The analyst should then arrange the components of the interface into functional groupings. For example, the interface components of a Ford In-Car Radio were arranged into the functional groups radio and cassette (Stanton & Young 1999). This is based entirely upon the analyst's subjective judgement.

Step 3: Arrange functional groupings into importance of use
Next, the analyst should arrange the functional groupings into importance of use. The analyst may want to make the most important functional group the most readily available on the interface. Again this is based entirely on the analyst's subjective judgement.

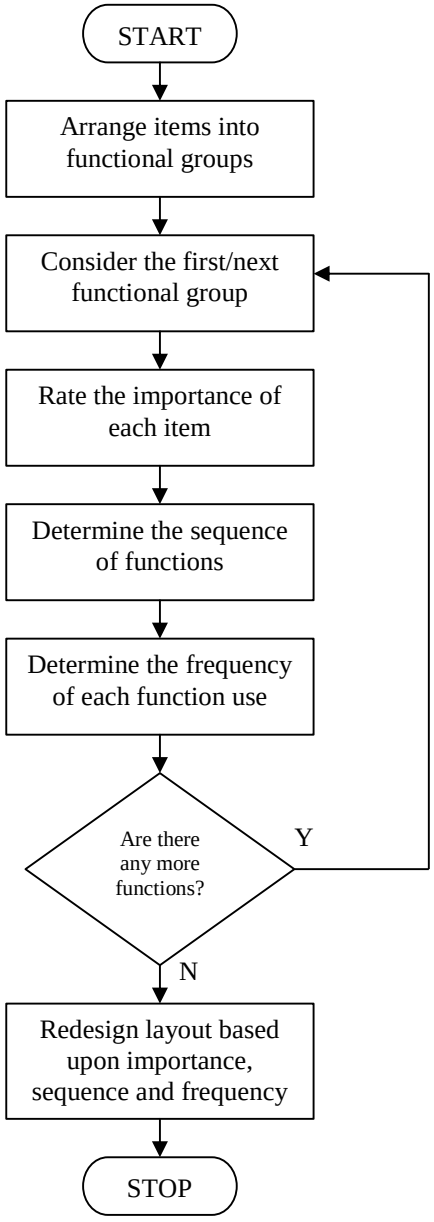
Step 4: Arrange functional groupings into sequence of use
The analyst should then repeat step 3, only this time arranging the functional groupings into sequence of use.

Step 5: Arrange functional groupings into frequency of use
The analyst should then repeat step 3, only this time arranging the functional groupings into frequency of use. At the end of the process, the analyst has redesigned the device according to the user's model of the task.

Step 6: Redesign the interface
Once the functional groups have been organised into importance, sequence and frequency of use, the interface should be redesigned. The analyst should base the

interface redesign upon the three categories (importance, sequence, frequency of use). For example, the analyst may wish to make the most important and frequently used aspect of the interface the most readily available.

Flowchart



Advantages

- Layout analysis is very easy to implement and also simple to use.
- Low resource usage.
- Layout analysis requires very little training.
- Layout analysis is a very quick technique to perform, offering an immediately useful output.
- Can be applied to paper diagrams of the device/interface under analysis.

- The output provided by the technique is immediately useful, offering a redesign of the interface under analysis based upon importance, sequence and frequency of use.

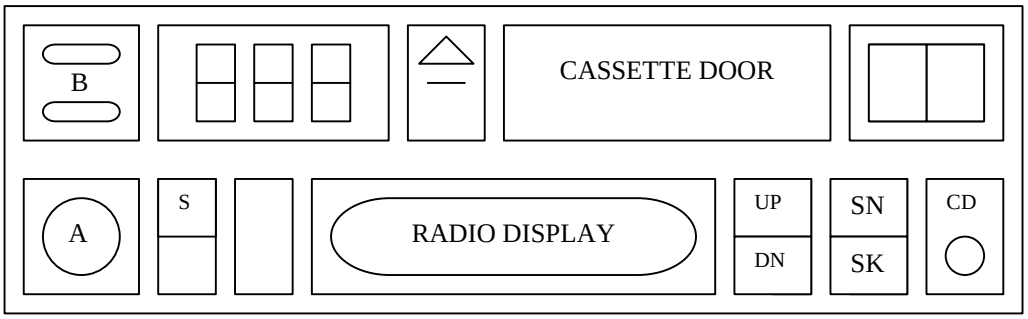
Disadvantages

- Poor reliability/Validity (Stanton & Young 1999).
- The output of the technique is very limited i.e. it only caters for layout. Errors and Task times are ignored.
- Literature regarding layout analysis is extremely sparse.
- If an initial HTA is required, application time can rise dramatically.
- Conducting a layout analysis for complex interfaces may be very difficult and time consuming.

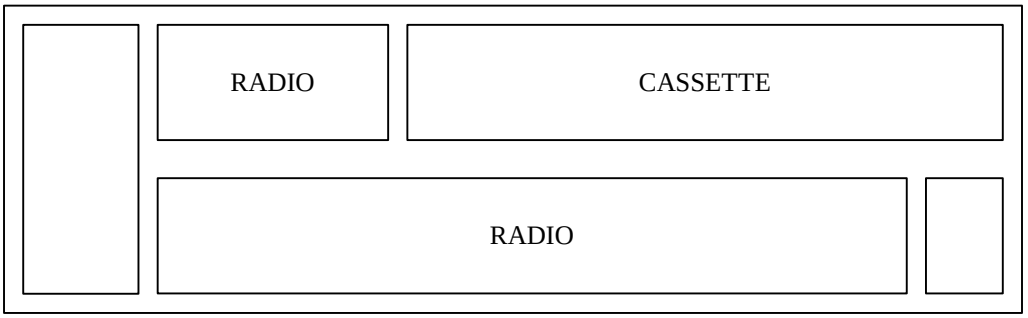
Example

The following layout analysis was conducted on a SHARP RG-F832E In-Car Radio (Stanton & Young 1999)

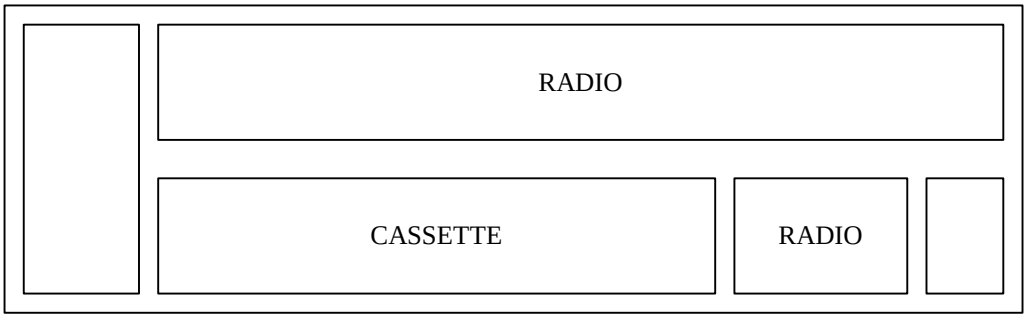
Initial design



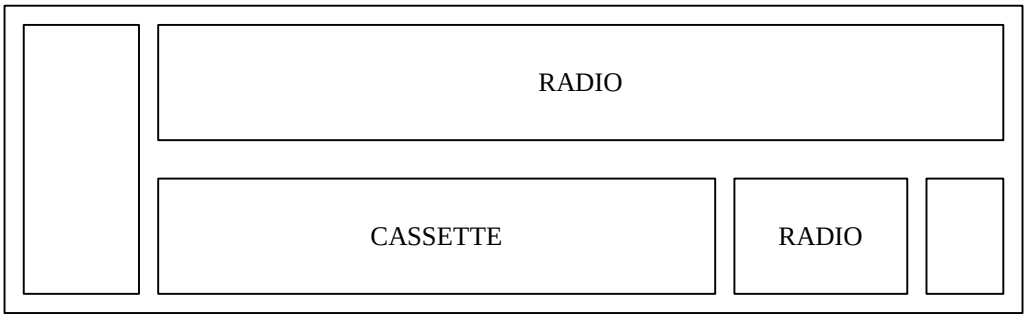
Functional groupings



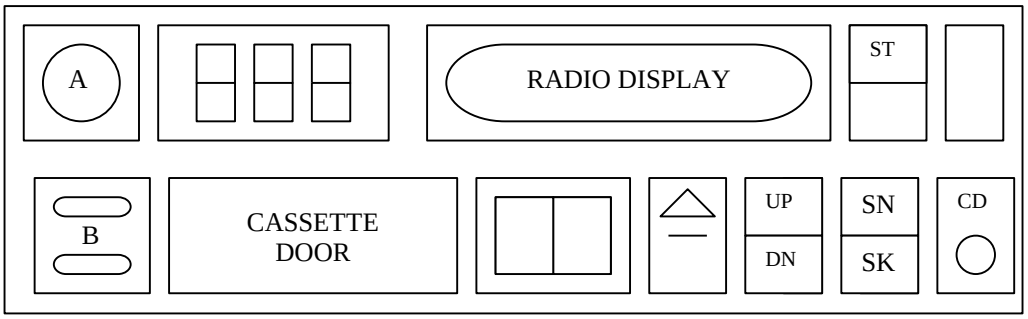
Importance of use



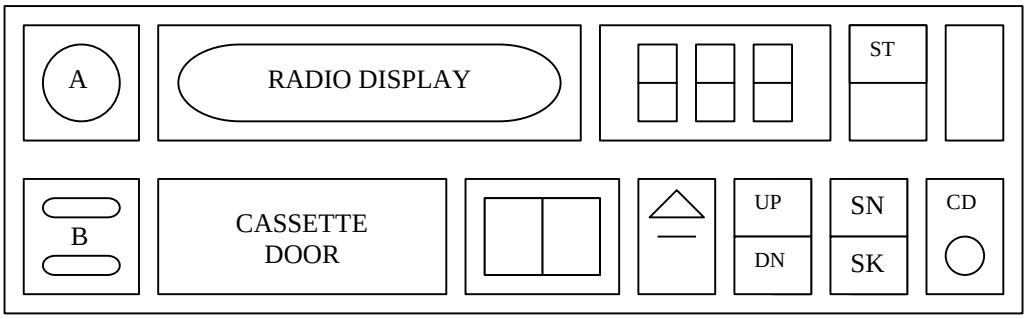
Sequence of use



Within functional groupings



Revised design by importance, frequency and sequence of use



Related methods
Layout analysis is very similar to link analysis in its approach to interface design.

Approximate training and application times
Stanton & Young (1999) report that little training is required for layout analysis and that it is amongst the quickest of twelve techniques to apply. If an initial HTA is required, the application time would rise considerably.

Reliability and validity
Stanton & Young (1999) report poor statistics for intra-rater reliability and predictive validity for layout analysis.

Tools needed
Layout analysis is a pen and paper tool.

Bibliography

Stanton, N. A. And Young, M. S (1999) A guide to methodology in Ergonomics, Taylor and Francis, London.
Easterby, R. (1984) Tasks, processes and display design. In R. Easterby & H. Zwaga (Eds.), *Information Design* (pp. 19-36). Chichester: Whiley.

TCSD - Task-Centred System Design

Background and applications

Task-Centred System Design is a simple, low cost and resource efficient approach to evaluating system design concepts involving the identification of the potential users and the tasks associated with the design concept and evaluating the design using design scenario's and a walkthrough type analysis. The techniques main appeal lies in its quick and easy application and the immediate usefulness of its output. The technique offers a redesign of the interface or system design under analysis as its output. TCSD is both easy to learn and apply. Greenberg (2003) divides the TCSD procedure into 4 main phases;

1) Identification phase – Involves specifying potential system users and example tasks.
2) User-centred requirements analysis – involves determining which user groups and which tasks will be catered for by the design.
3) Design through Scenario's – involves the assessment and modification of the design concept through use of design scenario's or storybooks.
4) Evaluation – involves the evaluation of the design concept via walkthrough type analysis.

A typical TCSD involves gathering data from an existing design and redesigning the system using design scenarios and system task walkthrough's.

Domain of application

Generic.

Procedure and advice

Step 1: Identification of potential users

The first step in a TCSD analysis is to identify the potential end users of the design under analysis. Specific user groups should be described. Observation and interviews are normally used to gather this data. The analyst should produce a representative list of user groups.

Step 2: Specification of example tasks

Once the specific user groups have been defined, a representative set of tasks for the system under analysis should be defined. This data is also collected through observation and interviews. The data for steps 1 and 2 are normally collected at the same time i.e. observing different users performing different tasks. Once the set of representative tasks is defined fully, each individual task should be given a task description. Greenberg (2003) suggests that each task description should adhere to five rules:

1) Description should describe what the user wants to do but not how they will do it.
2) Description should be very specific.
3) Description should describe a complete job.
4) Description should identify the users and reflect their interests.
5) When put together as a set of task descriptions, a wide range of users and task types should be described.

Once the list of tasks is complete, they should be checked and verified by the system end users. Task descriptions that are incomplete should be rewritten.

Step 3: Determine system users
The next step forms the first part of phase 2, the user-centred requirements analysis. Typically, system design cannot cater for all possible users. Step 3 involves determining which users or user groups the proposed design will cater for. Greenberg (2003) suggests that user's should be put into typical user types or groups. Greenberg also suggests that the different user types or groups should be categorised as *absolutely must include, should include if possible and exclud*e. For example, for a military command and control system design concept, the user groups falling into the absolutely must include group would be Gold command users, silver command users and bronze command personnel (foot soldiers, infantrymen).

Step 4: Determine system tasks
The next task in the TCSD process involves clearly specifying which tasks the system design will cater for. Similar criteria to that used in step 3 (absolutely must include, should include if possible and exclude) are used (with the addition of a 'could include' category) to categorise each task described in step 2.

Step 5: Generate design scenarios
Once step's 1 to 4 are complete, the analyst(s) should have a set of clearly defined end users and a set of tasks that the design will cater for. The actual design of the system can now begin. To do this, the TCSD informs the design process via the use of design scenario's or storybooks. A number of different design scenarios should be created, each one exploring how the design could cope with the scenario under analysis. Whilst no guidelines are offered regarding which scenario's and how many, it is recommended that a scenario involving each of the 'absolutely must include', 'should include if possible' and 'could include' tasks identified in step 4 should be created.

Step 6: Evaluate and modify design concept using scenario
Once a set of design scenario's have been specified, they should be used to continually evaluate and modify the design concept. Each scenario should be taken individually and applied to the system design, with team members questioning the efficiency of the design with respect to the events that unfold during each scenario. This is a continuous process, with each design scenario effectively testing the design concept. This process should continue until the team are happy with the system design.

Step 8: Perform task walkthrough
Once all of the scenarios have been applied to the design and the design team are happy with the end design concept, the design is tested further and more thoroughly using a walkthrough analysis. Depending upon resources available (time, money) SME's or members of the design team can be used. However, walkthroughs using SME's or system operators would produce more valid results. Essentially, the walk-through involves role-playing, putting oneself in the mind and context of the user (Greenberg 2003). Lewis and Reiman (1993) propose the following procedure for performing task-centred walk-throughs.

- Select one of the task scenarios
- For each of the users/actions in the task:
    - Can you build a believable story that motivates the user's actions?

- Can you rely on the user's expected knowledge and training about the system?
- If you cannot, you have located a problem in the interface.
- Note the problem and any comments or solutions that come to mind.

Once a problem is identified, assume it has been repaired.
Go to the next step in the task.
Once all of the scenarios have been subjected to a walk-through, the end design should be complete.

Advantages

- TCSD is a simplistic technique to use that immediately informs system design.
- Design modifications occur naturally throughout the analysis.
- Considers the end users and the set of tasks that the design is required to support.
- The use of design scenario's allow the design to be evaluated as it would be used.
- Correctly assembled TCSD teams can be very powerful.
- The design concept is evaluated and modified as a result of a TCSD analysis.
- Not as resource intensive as other techniques.

Disadvantages

- Validity and reliability of the technique is questionable.
- The use of such a simplistic technique in the design of a miltary command and control may be questioned.
- Whilst the techniques simplicity is the main advantage associated with its use, this leads to criticisms regarding depth of the analysis.
- Although TCSD is not as resource intensive as other techniques, it is still a time consuming technique to apply.
- Assembling the TCSD team may prove difficult. For example, a TCSD analysis for the design of a military command and control system would require numerous specialists (human factors, military, design, system operators etc). Getting such a team together in one place at one time could prove very difficult.
- TCSD generates huge amounts of data.

Example

The following example is adapted from a TCSD analysis of a catalogue based department store (Greenberg 2003). As the end output of TCSD is typically very large, only extracts of the analysis are shown below. The example is based upon the evaluation and redesign of an in-store computer ordering system. For a more detailed example, the reader is referred to Greenberg (2003).

Table 32. User types

Customers	**Sales Clerks**
First time V's Repeat customers	Experienced and trained
Computer knowledgeable V's Computer niave	New staff member; has passed introductory training session
Typists V's Non typists	
Willing to use the computer V's Unwilling	
People with disabilities who may have trouble with fine motor control	

Flowchart

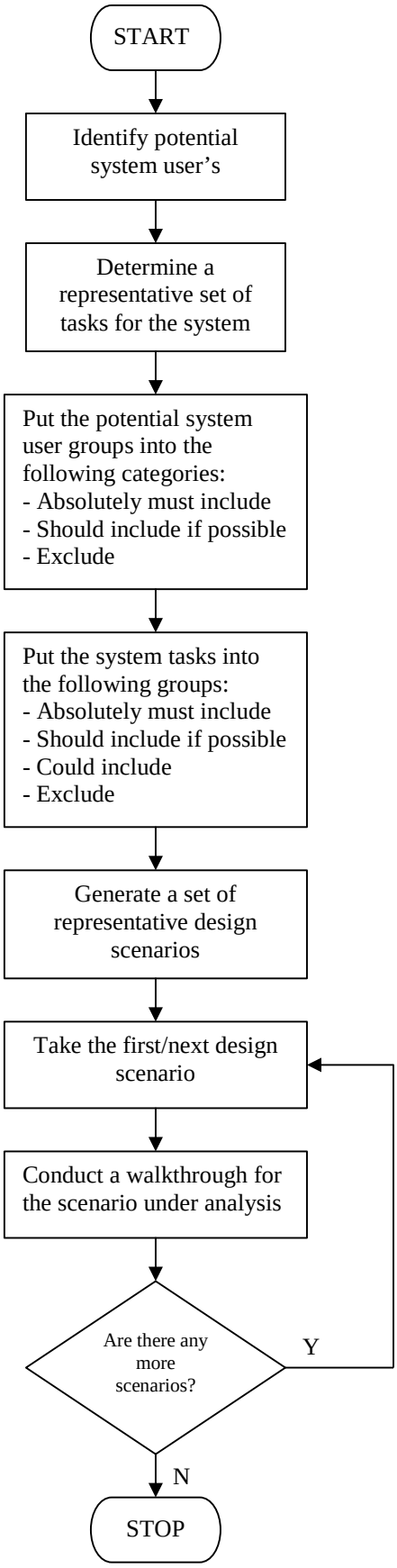

START
Identify potential system user's
Determine a representative set of tasks for the system
Put the potential system user groups into the following categories:
- Absolutely must include
- Should include if possible
- Exclude
Put the system tasks into the following groups:
- Absolutely must include
- Should include if possible
- Could include
- Exclude
Generate a set of representative design scenarios
Take the first/next design scenario
Conduct a walkthrough for the scenario under analysis
Are there any more scenarios?
Y
N
STOP

Table 33. Tasks to be catered for by the end design

Choosing merchandise	**Pay by**	**Reviewing cost**	**Merchandise pickup**
One item	Cash	Individual item cost	Immediate
Multiple items	Credit or debit card	Total costs	Delivery
Modifying the selected list of items	Invoice	Comparison shopping	

Table 34. Example TCSD walk-through

Task step	Knowledge? Believable? Motivated?	Comments/Solutions?
a. Enters store	Okay	Finding paper catalogues is not a problem in the current store
b. Looks for catalogue	Okay if paper catalogue is used, but what if the catalogue is online	However, we were not told if the paper catalogue would still be used or if the catalogue would be made available on-line Note – ask cheap shop about this. If they are developing an electronic catalogue, we will have to consider how our interface will work with it. For now, we assume that only a paper catalogue is used.
c. Finds red JPG stroller in catalogue	Okay	The current paper catalogue has proven itself repeatedly as an effective way for customers to brows cheap shop merchandise and to locate products.
d. Looks for computer	Modest problem	As a first time customer, Fred does not know that he needs to order through the computer. Unfortunately, we do not know how the store plans to tell customers that they should use the computer. Is there is a computer next to every catalogue or are there a limited number of computers on separate counters? Are there signs telling Fred what to do? Note: Ask cheap shop about the store layout and possible signage Possible solution: Instead of screen 1, a startup screen can clearly indicate what the computer is for (e.g., “Order your items here” in large letters.

Related methods

In conducting a TCSD analysis, a number of different human factors techniques can be utilised. Observational techniques and interviews are typically used to collect data regarding the system users and the type of tasks that the system caters for. Design scenarios and Walkthrough analysis are also used to evaluate the design concept. Greenberg (2003) suggests that to make a TCSD analysis more comprehensive, heuristic type analysis is often used.

Approximate training and application times

The training time for the TCSD technique would be minimal. The application time, including observations, interviews, the generation of scenarios and the application of walk-through type analysis would be high.

Reliability and validity
The reliability of the TCSD technique is questionable. Greenberg (2003) suggests that it is not a precise technique and that task or user groups are likely to be overlooked. Indeed, it is apparent that when used by different analysts, the technique may offer strikingly different results. The validity of such a technique is a hard thing to define

Tools needed
TCSD can be conducted using pen and paper. However, for the observational analysis, it is recommended that visual and/or audio recording devices are used.

Bibliography
Lewis, C., & Reiman, J. (1993). Task centred user interface design: A practical introduction. Boulder, CO: University of Colorado. Shareware book available from ftp.cs.colorado.edu/pub/cs/distribs/clewis/HCI-Design-Book
Greenberg, S. (2003) Working through Task-Centred System Design. In D. Diaper & N. Stanton (Eds) The Handbook of Task Analysis for Human Computer Interaction. Lawrence Erlbaum Associates Inc.

Design Scenario Analysis
Various

Background and applications
Design scenarios are a storybook style approach used to help designers and design teams propose, evaluate and modify design concepts. According to Go & Carroll (2003) a scenario is a description that contains actors, assumptions about the environment, goals and objectives, sequences of actions and events. Scenario analyses are used throughout the design cycle to depict new system designs in future contexts. Scenario analyses typically come in the form of sketch storyboards depicting a proposed future operation of the device/system being designed. At its most basic level, a scenario type analysis involves proposing a design concept and querying the design using who, what, when, why and how type questions (Go & Carroll 2003). Once a scenario is created, design ideas and changes can be added to the storyboard and the design is modified as a result. Scenarios are also used to communicate design concepts to other organisations or design teams. One of the main reasons for using scenario analysis is that it is much cheaper to sketch and act out a future scenario than it is to develop a simulation of one. Scenario type analysis are a powerful design tool that have been applied to the design process in a number of different domains, such as HCI, requirements engineering, object oriented design, systems design and strategic planning (Go & Carroll 2003). The appeal of scenarios lies in the techniques flexibility, whereby the focus and nature of the analysis is based entirely upon the analyst(s) requirements, and the direction of the analysis is entirely up to the analysis team.

Domain of application
Generic.

Procedure and advice
There are no set rules for scenario type analysis. A rough guide proposed by the author is presented below.

Step 1: Determine representative set of scenarios
The first step in a scenario analysis is to develop and describe a representative set of scenarios for the system under analysis. Each scenario should be described fully, including the scenario aims, objectives and activities as well as any input devices, displays or interfaces used in the scenario. The personnel involved, the context within which the scenario may take, individual goals, actions and possible outcomes should also be stipulated. A scenario description table should be constructed at this point, containing all of the relevant information regarding the scenario, such as goals, objectives, task steps, input devices, output devices etc.

Step 2: Scenario Observation
Scenarios are normally based upon an observation of similar scenarios to the scenario under analysis. The analyst(s) should record and observe the scenario under analysis. If the system or design concept does not yet exist, the scenario should be 'made-up' from scratch using techniques such as group brainstorming. Any novel scenarios observed or elicited that were not expressed in step 1 should also be added to the scenario description table. Interviews and questionnaires may also be used to elicit information regarding potential scenarios.

Step 3: Act out the scenario
The analyst or team of analysts should then create the scenario in the form of a storyboard. The scenario should be based upon the system being designed, with future contexts and situations being added to the scenario as the analysis progresses. Team members should offer intervention, proposing different contexts and events, such as 'what would happen if' and 'how would the operator cope if...'. This allows the scenario team to evaluate every possibility that occur with the design concept. Problem scenarios are particularly useful for evaluating a design concept. This part of the scenario analysis is the most crucial and should involve maximum experimentation with the proposed design concept. All assumptions and resultant design modifications should be recorded. The process should continue until the design team is satisfied that all possible scenarios have been exhausted and the end design is complete.

Advantages
- Scenario analyses offer a quick and easy method of seeing the design concept working in future contexts. This can help highlight any design flaws and future problems associated with the initial design.
- Scenario analysis is a very flexible technique.
- Scenario type analyses promote broad thinking.
- Scenario analyses can provide a format for communicating design concepts and issues between designers and design teams.
- Quick, cheap and easy technique to apply.
- Scenario type analysis can also be used to develop operator mental models.
- Scenario analysis output is immediately useful, giving a sketch drawing of the design in action and also highlighting any problems that may be encountered.
- Any number of scenario's can be evaluated, ranging from 'normal' to 'worse case' scenario's.

Disadvantages
- Scenarios are not very precise and many potential scenarios may be missed or left out by the analysis team.
- Could be time consuming for large scenarios.
- To reap the full benefit of a scenario analysis, a multi-discipline team needs to be put together. This could prove quite difficult to achieve.

Example
For an example scenario analysis, the reader is referred to Go & Carroll (2003)

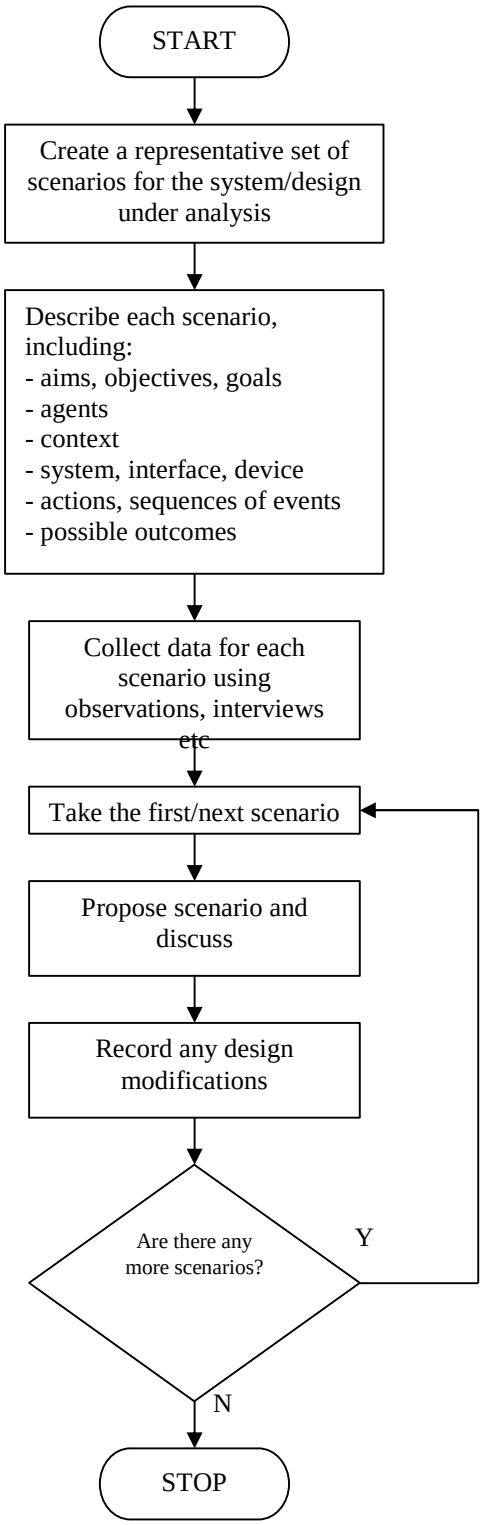
Related methods
Scenario analysis involves the collection of data using techniques such as observations, interviews and questionnaires. Scenario techniques are also similar to role-play techniques, which are also used by designers to visualise potential product use.

Approximate training and application times
The technique is simple to use and so training time is estimated to be very low. Application time can vary, as there are no set end points to a scenario and new scenario's can be added to existing ones at any point. The size of the scenario also has an effect upon the length of the analysis.

Reliability and validity

The reliability of the technique is questionable. Scenario teams may fail to capture all of the potential future scenarios of a design in a scenario analysis. Similarly, the technique may produce inconsistent results for the same design, when applied by different teams.

Flowchart



Tools needed

Scenarios are typically conducted using pen and paper. For the data collection part of scenario analysis, it is recommended that visual and/or audio recording equipment is used.

Bibliography

Go, K., & Carroll, J., M. (2003) Scenario-Based Task Analysis. In D. Diaper & N. Stanton (Eds) *The Handbook of Task Analysis for Human-Computer Interaction, London*, Lawrence Erlbaum Associates.

Verplank, B., Fulton, J., Black, A., Moggridge, B. (1993) Observation and Invention – Use of scenarios in interaction design. Tutorial notes for InterCHI 93. Amsterdam

Checklists
Various

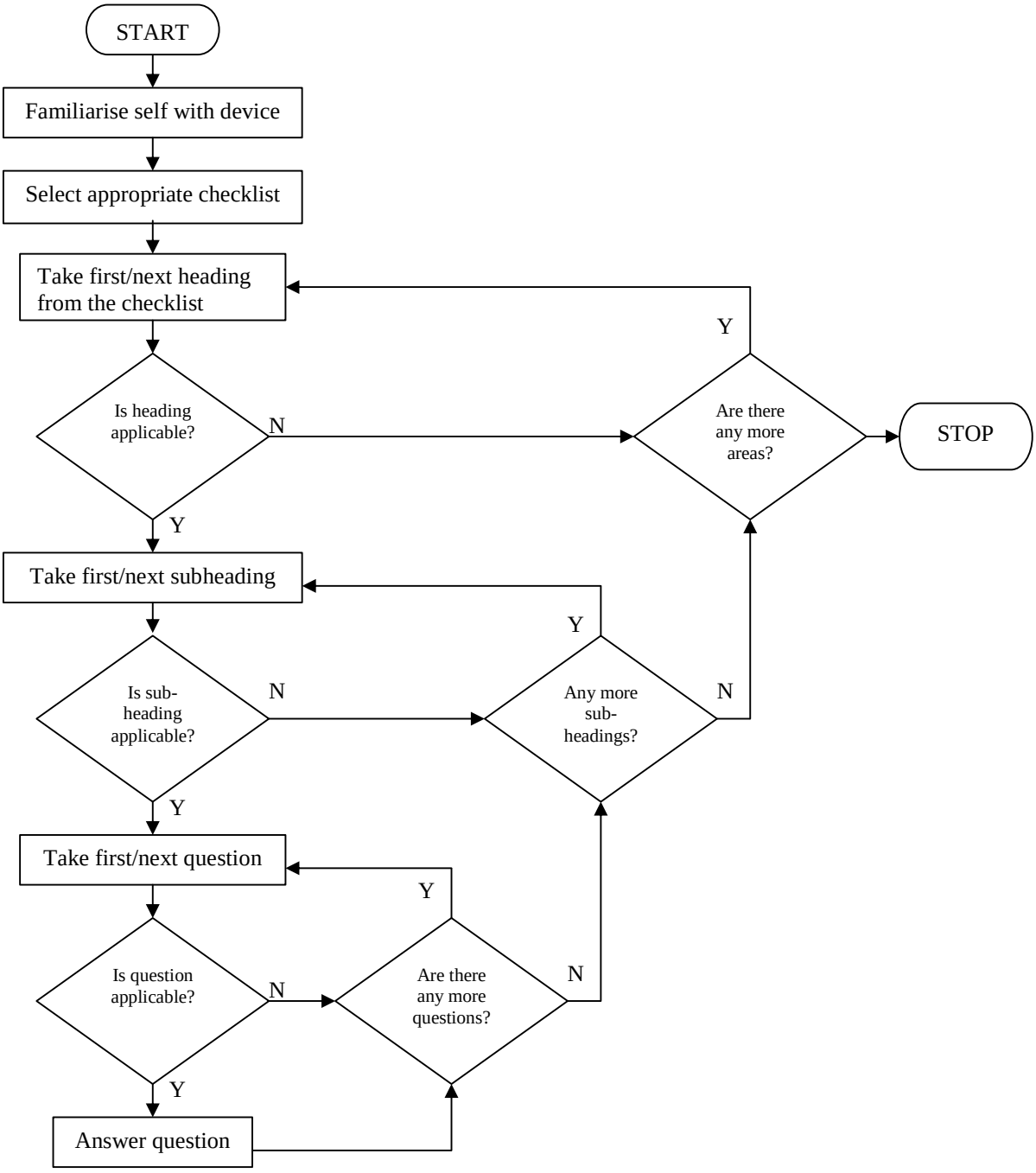
Background and applications
The checklist style approach is a very simple approach whereby the analyst checks the design of a product or system against a pre-defined set of criteria. Checklist style evaluation can occur throughout the life cycle of a product or system, from paper drawings to the finished product. A number of checklists exist in the human factors community, such as Ravden & Johnson's HCI checklist, the Human Engineering Design checklist and various Woodson, Tillman & Tillman (1992) checklists. Checklists can be used to evaluate the usability and design of a device or system in any domain. In the past, checklists have been used in HCI, automotive studies and air traffic control systems. More recently, Ciavarelli (2002) has developed an aviation specific human factors checklist for use in aircraft accident investigations. When using checklists, the analyst using the checklist should have some level of skill or familiarity with the device under evaluation. Performing a checklist analysis is a matter of simply inspecting the device against each point on the chosen checklist. Checklists are also very flexible in that they can be adapted or modified by the analyst according to the demands of the analysis. Stanton & Young (1999) used a section of Ravden & Johnson's HCI checklist in order to evaluate the design of In-Car radios.

Domain of application
Various.

Procedure and advice
Step 1: Select relevant checklist
Firstly, the analyst must decide which form of checklist is appropriate for the product or system under analysis. The checklist used may be simply an existing one or the analyst may choose to adapt an existing checklist to make it more appropriate for the system under analysis. Stanton and Young (1999) used part of Ravden and Johnson's (1989) HCI checklist for In-Car entertainment systems. One of the main features of checklists is that they are very flexible. Checklists can be adapted or modified according to the demands of the analysis. Alternatively, if a suitable checklist is not available, the analyst may choose to create a new checklist specifically for the system/product in question.

Step 2: Check item on checklist against product
The analyst should take the first point on the checklist and check it against the product or system under analysis. For example, the first item in Ravsden & Johnson's checklist asks. 'Is each screen clearly identified with an informative title or description'? The analysts should then proceed to check each screen and its associated title and description. The options given are 'Always', 'Most of the time', 'Some of the time' and 'Never'. Using subjective judgement, the analyst should rate the device under analysis according to the checklist item. Step 2 should be repeated until each item on the checklist has been dealt with.

Flowchart



Advantages

- Checklists are a very simple technique to use.
- Checklists are probably one of the quickest methods available offering an immediately useful output.
- Checklists are based upon established knowledge about human performance (Stanton and Young, 1999)
- The technique requires very little training.
- Resource usage is very low.

- Checklists are very adaptable and can easily be modified in order to use for other devices/systems. Stanton and Young (1999) suggest that the Ravden and Johnson checklist (1989), originally designed for HCI, is easily adapted to cater for the usability of other devices, such as in-car stereos.
- A number of different checklists are available to the human factors practitioner.

Disadvantages

- A checklist type analysis does not account for errors or cognitive problems associated with the device.
- Context is ignored by checklists.
- Checklist type analysis is very subjective. What one analyst classes as bad design may be classed as suitable by another.
- Low consistency.
- Not a very sophisticated approach to system design.

Example

The following example is an extract of an analysis of a Sony Ericcson t68i mobile phone using Ravsden & Johnson's HCI checklist.

Section 1: Visual Clarity

Key:

A = Always, M = Most of the time, S = Some of the time, N = Never

Section 1: Visual Clarity	A	M	S	N	Comments
1. Is each screen clearly identified with an informative title or description		✓			Some screens lack titles
2. Is important information highlighted on the screen? (e.g. cursor position, instructions, errors)	✓				
3. When the user enters information on the screen, is it clear: a) Where the information should be entered? b) In what format it should be entered?		✓			
4. Where the user overtypes information on the screen, does the system clear the previous information, so that it does not get confused with the updated input?					N/A
5. Does information appear to be organised logically on the screen? (e.g. menus organised by probable sequence of selection, or alphabetically)	✓				
6. Are different types of information clearly separated from each other on the screen (e.g. instructions, control options, data displays)				✓	Different information is often grouped into lists
7. Where a large amount of information is displayed on one screen, is it clearly separated into sections on the screen?	✓				
8. Are columns of information clearly aligned on the screen? (e.g. columns of alphanumerics left justified, columns of integers right-justified)					
9. Are bright or light colours displayed on a dark background, and vice versa?	✓				
10. Does the use of colours make the displays clear?	✓				
11. Where colour is used, will aspects of the display be easy to see if used on a monochrome or low resolution screen, or if the user is colour blind?					
12. Is the information on the screen easy to see and read?	✓				
13. Do screens appear uncluttered?		✓			
14. Are schematic and pictorial displays (e.g. figures and diagrams) clearly drawn and annotated)	✓				
15. Is it easy to find the required information on a screen?			✓		Easy to get lost in menu system

Related methods
There are a number of checklists available to the human factors practitioner, such as Woodson, Tillman and Tillman (1992), Ravsden and Johnson (1989) and Ciaverelli (2002).

Approximate training and application times
As checklists are a very simple technique to learn and use, it is estimated that the both training and application times would be low. In an analysis of twelve ergonomics methods, Stanton & Young (1999) report that checklists are one of the quickest techniques to train, practice and apply.

Reliability and Validity
Whilst Stanton and Young (1999) report that checklists performed quite poorly on intra-rater reliability, they also report that inter-rater reliability and predictive validity of checklists was good.

Tools needed
Checklists are a simple pen and paper tool, however, for a checklist analysis, the analyst must have access to some form of the device in question. This could either be the finished article, paper drawings or a prototype version. The relevant checklist is also required.

Bibliography
Ciavarelli, A. (2002) Human Factors Checklist: An Aircraft Accident Investigation Tool, School of Aviation Safety, California. Internet source
Kirwan, B. and Ainsworth, L. K. (eds) (1992) A guide to Task Analysis, London, Taylor and Francis.
Ravden, S. J. And Johnson, G. I. (1989) Evaluating Usability of Human-Computer Interfaces: A practical method, Chirchester, Ellis Horwood.
Stanton, N. A, and Young, M. S. (1999) A guide to methodology in ergonomics: Designing for human use, London, Taylor and Francis.

Heuristic analysis

Background and applications

Heuristic type analysis is probably the simplest technique available to the human factors practitioner. A quick and very easy technique, heuristic analysis involves obtaining analyst(s) subjective opinions on a design concept or product. In conducting a heuristic analysis, an analyst or team of analysts should interact with the design under analysis and make observations regarding the usability, quality, and error potential of the design. Heuristic analysis should be conducted continually throughout the design process in order to evaluate and modify the design concept. The beauty of heuristic analysis lies in its simplicity and the fact that it can be conducted at any stage in the design process. The analysis can be structured further by using pre-determined heuristic criteria or design guidelines. Nielsen & Molich (1990) recommend a simplistic set of heuristics to make the analysis practical.

- Simple and natural dialogue
- Speak the users language
- Minimise user memory load
- Be consistent
- Provide feedback
- Provide clearly marked exits
- Provide shortcuts
- Good error messages
- Prevent errors

It is recommended that a team of analysts should be used when conducting heuristic analysis during the design of command and control systems. Furthermore, SME's or domain experts should also be used to enhance validity.

Domain of application

Generic.

Procedure and advice

Step 1: Define scenario/task under analysis

The first step in a heuristic analysis is to define a representative set of tasks or scenarios for the system or device under analysis.

Step 2: Define heuristic list

In some cases it may be fruitful to determine which aspects are to be evaluated before the analysis begins. Typically, usability (ease of use, effectiveness, efficiency and comfort) and error potential are evaluated. Design guidelines can also be used, such as Wagner et al (1996) or Smith and Mosier's user interface guidelines.

Step 3: Perform task(s)

The analyst(s) should then perform each task from the task list and offer opinions regarding the design and the heuristic categories required. If the analysis concerns a design concept, then a task walkthrough is sufficient. Each opinion offered should be recorded. It may be useful to record the session ensuring that no data is missed.

Step 4: Propose remedies
Once all tasks have been analysed, design remedies for each negative point highlighted should be proposed and recorded.

Advantages
- A very simple technique to apply requiring very little training.
- Heuristic analysis can be very quick.
- Useful output that is immediately useful.
- Very low resource usage.
- Can be used repeatedly throughout the design life cycle.

Disadvantages
- Poor reliability, validity and comprehensiveness.
- Requires a team of SME's in order to be worthwhile.
- Totally subjective.
- Totally unstructured.
- Consistency of such a technique is questionable.

Example
The following example is taken from Stanton & Young (1999). A heuristic analysis of a Sharp in car radio was conducted in order to assess the interface in terms of ease of skill acquisition, effectiveness on task, comfort/satisfaction and flexibility on task.

- On/Off/Volume control is a tad small and awkward, combined with difficult balance control
- Pushbutton operation would be more satisfactory for On/Off, as volume stays at preferred level
- Fader control is particularly small and awkward
- Both of the above points are related to the fact that a single button location has multiple functions – this is too complex
- Treble and Bass controls also difficult and stiff; although these functions are rarely adjusted once set
- Station pre-set buttons are satisfactory; quite large and clear
- Band selector Button and FM Mono-Stereo button should not have 2 functions on each button – could result in confusion if the wrong function occurs. These buttons are the only buttons on the radio which are not self explanatory – the user must consult the manual to discover their function
- Tuning seek and tuning scan buttons are easier to understand and use, although there are still two functions on one button.
- Auto-reverse function is not obvious, although it is an accepted standard
- Illumination – is daytime/nighttime illumination satisfactory? A dimmer control would probably aid matters.

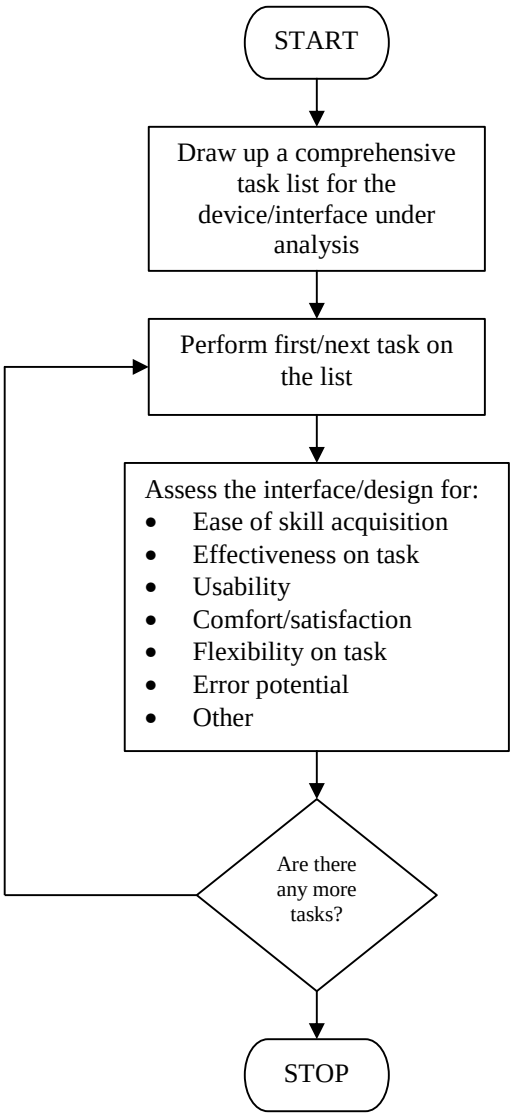
Approximate training and application times
The technique requires no training and the application time is very low.

Reliability and validity
According to Stanton & Young (1999), the unstructured nature of the technique leads to very poor reliability and validity.

Tools needed
Pen and paper. Heuristic analysis can be conducted on paper diagrams of the system under analysis.

Bibliography
Nielsen, J. & Molich, R. (1990) Heuristic evaluation of user interfaces. In J. C Chew & J. Whiteside (eds), *Empowering people: CHI 90 Conference Proceedings* (pp. 249 – 256) Monterey, CA: ACM Press
Stanton, N. A, and Young, M. S. (1999) A guide to methodology in ergonomics: Designing for human use, London, Taylor and Francis.

Flowchart


START
Draw up a comprehensive task list for the device/interface under analysis
Perform first/next task on the list
Assess the interface/design for:
Ease of skill acquisition
Effectiveness on task
Usability
Comfort/satisfaction
Flexibility on task
Error potential
Other
Are there any more tasks?
STOP

Focus Groups
Various

Background and applications
A Focus group is a group interview approach that involves using a group of SME's to discuss a particular design concept or prototype. Originally used in market research, focus groups normally involve a group of SME's and 1 to 2 moderators discussing critical design points of a certain system or product design. The output of a focus group is normally a list of agreed and disagreed statements. Focus groups can be used for almost any purpose, including predicting potential human error in a certain system, usability problems associated with a design concept or to evaluate a prototypical design in terms of usability, workload, error, performance times etc. Hypponen (1999) suggests that focus groups are used to gather raw data regarding user needs in the concept development phase of a design and that they can also be used to clarify issues during the design. Focus groups can also be used as an evaluation tool in order to evaluate existing system design with regard to errors, usability etc. Focus groups were originally used in the IT domain but have been used in numerous different areas to inform the design process.

Domain of application
Generic.

Procedure and advice
There are no set rules for conducting a focus group type analysis. The following procedure is intended to act as a set of guidelines to consider when conducting a focus group type analysis.

Step 1: Define aims and objectives
The first step in conducting a focus group is to clearly define the overall aims and objectives of the focus group. This involves stating explicitly the purpose of the focus group i.e. to discuss the C4i Gold command interface design concept.

Step 2: Determine key discussion topics
Once the overall aim of the focus group has been defined, it should be divided into specific areas that are to be the topic of discussion during the focus group. Using the example above, the 'C4i gold command interface design concept', this could be spilt into the following key discussion areas: Interface layout, Probability of error, task times, Usability, design flaws and design remedies. The key discussion points should be placed in a logical order and this order should be adhered to during the focus group.

Step 3: Assemble focus group
Assembling the correct personnel for a focus group is crucial. For the example outlined above, the focus group would require a number of different personnel. The recommended mix is outlined below.

- Human factors experts
- Military personnel
- Experienced command and control system operators
- Project manager
- HRA/HEI specialist

- Usability specialist
- Designers
- Data recorder
- Controllers from different domains (such as ATC, Police, Ambulance)

Step 4: Introduce design concept
Once the focus group has been assembled, the starting point of the focus group session is to introduce to the group the design concept that is to be the topic of discussion. This would normally take the form of a presentation. Once the presentation is finished, the focus group leader should introduce the first topic of discussion.

Step 5: Introduce first/next topic
The first topic of discussion should be introduced clearly to the group, including what the topic is, why it is important and what is hoped to be achieved by discussing that certain topic. The actual topic should be discussed thoroughly until it is exhausted and a number of points are agreed upon. Step 5 should be repeated until all of the chosen discussion points have been discussed fully.

Step 6: Transcribe data
Once the focus group session has been completed, the data requires transcribing. The analyst should use an audio or video recording of the focus group session in order to do this.

Step 7: Analyse data
Once transcribed, the data then needs to be analysed. This can be done in many ways and is dependent upon the analysis requirements. Typically, the data output from a focus group session is a set of agreed upon statements regarding the design concept.

Advantages
- The make up of the focus group is down to the analyst. A correctly assembled focus group can provide a very powerful input into the design process.
- A focus group can discuss anything from probability of error to interface layout.
- The analyst(s) has complete control of the focus and direction of the analysis and can change this at any time.
- Very powerful data can be elicited from a focus group type analysis.
- Focus group type interviews allow the analyst to quickly survey a great number of opinions.
- People discuss issues more freely in a group context.

Disadvantages
- Assembling the desired focus group is a very difficult thing to do. Getting such a diverse group of experts together at the same location and at the same time is a very difficult.
- The chemistry within the focus group has a huge effect upon the data collected.
- The reliability and validity of focus groups is questionable.
- Large amounts of data are gathered. This is very time consuming to transcribe and analyse.

Related methods
The focus group technique is a group interview technique, thus it is related to interviews.

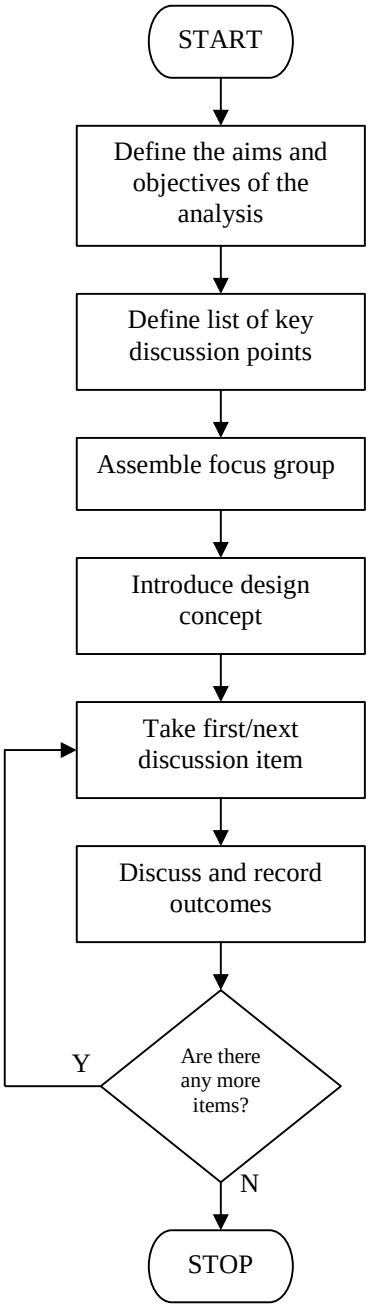
Approximate training and application times
There are no training times associated with a focus group type analysis. Typical focus group session duration is between 90 minutes and two hours. However, this is dependent upon the requirements of the focus groups and it is not unheard for focus group sessions to last days at a time.

Reliability and validity
Whilst no data regarding the reliability and validity of focus groups is available in the literature, it is apparent that it could be questionable.

Tools needed
The tools required conduct a focus group analysis include pen and paper, a video recording device, such as a video recorder and/or an audio recording device, such as a cassette recorder. A PC with a word processing package such as Microsoft Word is required to transcribe the data collected.

Bibliography
Hyponen, H. (1999) Focus Groups. In H. A. Williams, J. Bound & R. Coleman (eds) *The Methods Lab: User Research for Design*. Design for Ageing Network
Stanton, N. A, and Young, M. S. (1999) A guide to methodology in ergonomics: Designing for human use, London, Taylor and Francis.

Flowchart


START
Define the aims and objectives of the analysis
Define list of key discussion points
Assemble focus group
Introduce design concept
Take first/next discussion item
Discuss and record outcomes
Are there any more items?
Y
N
STOP

Walkthrough analysis

Background and applications

Walkthrough analysis is a very simple procedure used by designers whereby experienced system operators perform a walkthrough or demonstration of a task or set of tasks using the system under analysis. Walkthroughs are typically used early in the design process to envisage how a design would work and also to evaluate and modify the design concept. They can also be used on existing systems to demonstrate to system designers how a process is currently performed, highlighting flaws, error potential and usability problems. The appeal in walkthrough type analysis lies in the fact that the scenario or task under analysis does not necessarily have to occur. One of the problems of observational analysis is that the required scenario simply may not occur, or if it does, the observation team may have to spend considerable time waiting for it to occur. Walkthrough analysis allows the scenario to be 'acted out' removing the problems of gaining access to systems and personnel and also waiting for the scenario to occur. A walkthrough involves an operator walking through a scenario, performing (or pretending to perform) the actions that would occur, explaining the function of each control and display used. The walkthrough is also verbalised and the analyst(s) can stop the scenario and ask questions at any point. Walkthrough analysis is particularly useful in the initial stages of task analysis.

Domain of application

Generic.

Procedure and advice

There are no set rules for a walkthrough analysis. The following procedure is intended to act as a set of guidelines for conducting a walkthrough analysis of a proposed system design concept.

Step 1: Define set of representative scenarios
Firstly, a representative set of tasks or scenarios for the system under analysis should be defined. As a general rule, the set of scenarios used should cover every aspect of the system and its interface at least once. The personnel involved in each scenario should also be defined. If the required personnel cannot be gathered for the walkthrough, then members of the design team can be used.

Step 2: Perform walkthrough
The analyst team then simply take each scenario and perform a verbalised walkthrough using the system design under analysis. The scenario can be frozen at any point and questions asked regarding controls, displays, decisions made, situation awareness, error occurrence etc. The walkthrough should be recorded using video recording equipment. Any problems with the design concept encountered during the walkthrough should be recorded and design remedies offered and tested.

Step 3: Analyse data
Once the walkthrough has been performed, the data should be analysed accordingly and used with respect to the goals of the analysis. Walkthrough data is very flexible and can be used for a number of purposes, such as task analysis, constructing timelines and evaluating error potential.

Step 4: Modify design
Once the walkthrough is complete and the data is analysed, the design can be modified based upon problems encountered during the walkthrough. If a new design is proposed, a further walkthrough should be conducted in order to analyse the new design.

Advantages
- When used correctly, a walkthrough can provide a very accurate description of the task under analysis and also how a proposed system design would be used.
- Walkthrough analysis allows the analyst to stop or interrupt the scenario in order to query certain points. This is a provision which is not available when using other techniques such as observational analysis.
- A walkthrough analysis does not necessarily require the system under analysis.
- Walkthrough analysis is a simple, quick and low cost technique.
- Walkthrough analysis would appear to be a very useful tool in the analysis of distributed (team based) tasks.
- Walkthrough analysis can provide a very powerful assessment of a design concept.

Disadvantages
- For the analysis to be fruitful, experienced operators for the system under analysis are required.
- Reliability of the technique is questionable.

Related methods
The walkthrough technique is very similar to verbal protocol analysis and observational analysis.

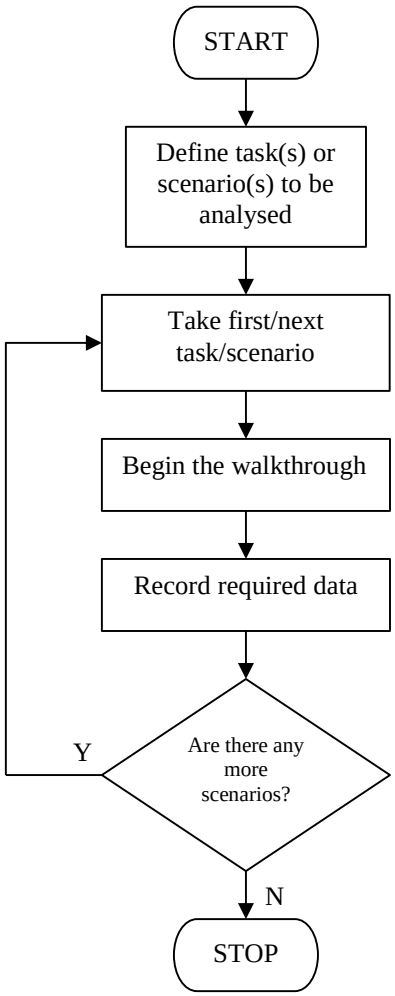
Approximate training and application times
There is no training as such for walkthrough analysis, and the associated application time is dependent upon the size and complexity of the task or scenario under analysis. The application time for walkthrough analysis is typically very low.

Reliability and validity
No data regarding the reliability and validity of the walkthrough technique are available.

Tools needed
A walkthrough analysis can be conducted using pen and paper. It is also recommended that video and audio recording equipment are used to record the walkthrough.

Bibliography
Kirwan, B. and Ainsworth, L. K. (Eds) (1992) A guide to Task Analysis, London, Taylor and Francis.

Flowchart


START
Define task(s) or scenario(s) to be analysed
Take first/next task/scenario
Begin the walkthrough
Record required data
Are there any more scenarios?
Y
N
STOP

Data Collection techniques

Data collection techniques are techniques that are used by the HF practitioner in order to collect specific information regarding tasks or scenarios, error occurrence, system usability, attitudes, opinions etc. The importance of a representation of the existing system or work under analysis cannot be underestimated and is a necessary starting point for any design effort. According to Stanton (2003) the starting point for designing future systems is a description of a current or analogous system, and that any inaccuracies could hinder the design effort. Data collection techniques are typically used as the starting point of HF analysis, and provide the basis for most of the techniques reviewed in this document. Data collection techniques include observation, interviews, questionnaires, analysis of artefacts, usability metrics and the analysis of performance and are probably the most commonly used techniques in human factors and provide extremely useful outputs.

Observational techniques are used to gather data regarding specific tasks or scenarios. A number of different types of observational analysis exist, such as direct observation, and participant observation. Although at first glance simply observing an operator at work seems to be a very simple technique to employ, it is evident that this is not the case, and that careful planning and execution are required (Stanton 2003). Observational techniques also require the provision of technology, such as video and audio recording equipment. The main problems associated with the use of observational techniques are the lengthy data analysis process and the various biases associated with observing operators.

Interviews are extensively used for a number of different purposes. Interviews are typically used to elicit information regarding product usability, error, and attitudes. Semi-structured techniques such as CDM (Klein 2003) and ACTA (Militello & Hutton 2000) are used to elicit data regarding operator decision-making.

Questionnaires offer a very flexible way of quickly collecting large amounts of data from large amounts of subjects. Questionnaires have been used in many forms to collect data regarding numerous issues within human factors and design. Questionnaires can be used to collect information regarding almost anything at all, including usability, user satisfaction, opinions and attitudes. More specifically, questionnaires can be used in the design process to evaluate concept and prototypical designs, to probe perceptions and to evaluate existing system designs. Established questionnaires such as the system usability scale (SUS), the Questionnaire for User Interface Satisfaction (QUIS) and the Software Usability Measurement Inventory (SUMI) are available for practitioners to apply to designs and existing systems. Alternatively, specific questionnaires can be designed and administered during the design process.

The data collection techniques reviewed in this document are outlined below:

1. Interviews
2. Questionnaires
3. Observation analysis

Interviews

Background and applications

Interviews have been used extensively in human factors to gather specific information regarding many different areas, such as system design, system usability, attitudes, job analysis, task analysis, error and many more. Indeed, interviews are probably the most commonly used human factors technique for information gathering. A number of human factors techniques are also interview based, with specifically designed probes or questions, such as the Critical Decision Method (Klein 2003), Applied Cognitive Task Analysis (Militello & Hutton 2000) and cognitive walkthrough analysis (Pocock et al 1992). There are three types of interview available to the human factors practitioner.

1) Structured Interview – The content of the interview i.e. the questions and their order, is pre-determined.
2) Semi-structured Interview – Some of the questions and their order is pre-determined. However, the interviewer also allows flexibility in directing the interview, and new issues or topics can be embarked on.
3) Unstructured Interview – The interview has no structure whatsoever and the interviewer goes into the interview 'blind'.

When conducting an interview, there are three main types of question that the interviewer can use.

1) Open ended question – An open-ended question is one that the interviewee has to answer in more than one word. Open-ended questions are used to elicit more than simple yes/no information. For example, if querying the interviewee about the usability of a certain device, a closed question would be; "Did you think that the system was usable?" This type of question will more often than not elicit merely a yes or no answer. An open-ended question approach to the same topic would be something like, "What do you think about the usability of the system". This type of open-ended question encourages the interviewee to share more that the typical yes/no answer, and gives the interviewer an avenue to gain much deeper, valuable information.
2) Probing question – A probing question is normally used after an open ended or closed question to gather more specific data regarding the interviewee's previous answer. Typical examples of a probing question would be, "Why did you think that the system was not usable?", or "How did it make you feel when you made that error with the system?"
3) Closed questions – A closed question can be used to elicit specific information. Closed questions typically prompt a yes or no reply.

According to Stanton & Young (1999), when conducting an interview, the interviewer should start on a particular topic with an open-ended question, and then once the interviewee has answered, use a probing question to gather further information. A closed question should then be used to gather specific information regarding the topic. Stanton & Young (1999) suggest that the interviewer should open up a topic and probe it until that topic is exhausted. When exhausted, the interviewer should move onto a new topic. This cycle of open, probe and closed question should be maintained throughout the interview.

Domain of application
Generic**.**

Procedure and Advice (Semi-structured interview)
As their are no set rules for the construction and conduction of an interview, the following is procedure should act as flexible guidelines for the human factors practitioner.

Step 1: Define the interview objective
Firstly, before any interview construction takes place, the analyst should clearly define the objective of the interview. For example, when interviewing a civil airline pilot for a study into design induced human error on the flight deck, the objective of the interview would be to discover which error's the pilot had made in the past, with which part of the interface, during which task.

Step 2: Question development
Once the objective of the interview is clear, the development of the questions can begin. The questions should be developed based upon the overall objective of the interview. In the design induced pilot error case, the opening question would be, "What sort of design induced errors have you made in the past on the flight deck?". This would then be followed by a probing question such as, "Why did you make this error?", or "What task were you performing when you made this error?". Once all of the relevant questions are developed, they should be put into some sort of coherent order or sequence. The wording of each question should be very clear and concise, and the use of acronyms or confusing terms should be avoided. Also when developing the interview questions, a data collection sheet should be prepared.

Step 3: Piloting the interview
Once the questions have been developed and ordered, the analyst should then pilot the interview in order to highlight any potential problems or discrepancies. This can be done through submitting the interview to colleagues or even by performing a trial interview with a 'live' subject. This process is very useful in shaping the interview into its most efficient form.

Step 4: Conduct and record the interview.
According to Stanton and Young (1999), the interviewee should use a cycle of open ended, probe and closed questions. The interviewee should persist with one particular topic until it is exhausted, and then move onto a new topic. Below are a set of interview Do's and Don'ts.

Do's	**Don'ts**
Make the relevance of each question clear	Avoid an over-bearing approach
Record the interview	Do not belittle, embarrass or insult interviewee
Be confident	Do not go over 40 minutes in length
Establish a good rapport with interviewee	Do not mislead or bias the interviewee
Communicate clearly	
Be very familiar with the topic of interview	

Step 5: Transcribe the data
Once the interview is completed, the analyst should proceed to transcribe the data. This involves replaying the initial recording of the interview and transcribing fully everything that is said during the interview. This is a lengthy process and requires much patience on behalf of the analyst.

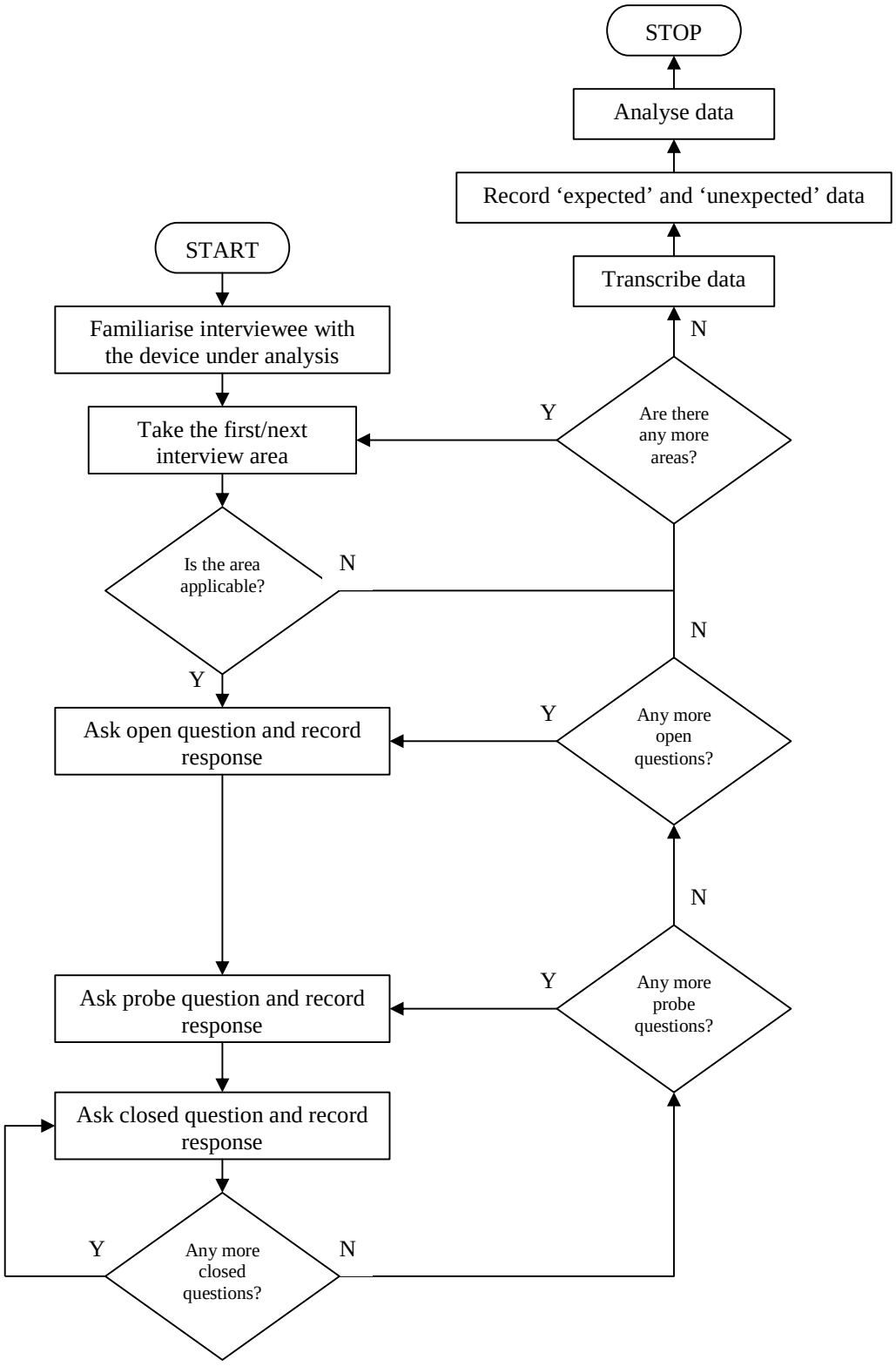
Step 6: Data gathering
Once the transcript of the interview is complete, the analyst should analyse the interview transcript, looking for the specific data that was required by the objective of the interview. This is known as the 'expected' data. Once all of the 'expected data' is gathered, the analyst should re-analyse the interview in order to gather any 'unexpected data', that is any extra data (not initially outlined in the objectives) that is unearthed.

Step 7: Data analysis
Finally, the analysts should then analyse the data using statistical tests, graphs etc. The form of analysis used is based upon the initial objective of the interview.

Advantages
- Interviews can be used to gather data regarding anything e.g. usability of existing systems, potential design flaws, errors etc.
- Interviews can be used at any stage in the design process.
- The use of SME's as interviewee's gives interviews the potential to be very powerful.
- The interviewer has full control over the interview and can direct the interview in way. This allows the collection of specific data.
- Response data can be treated statistically.
- A structured interview offers consistency and thoroughness (Stanton & Young 1999)
- Interviews are a very flexible technique.
- Interviews have been used extensively in the past for a number of different types of analysis.
- Specific, structured human factors interviews already exist, such as ACTA and the Critical Decision Method.

Disadvantages
- The construction and data analysis process ensures that the interview technique is a very time consuming one.
- The reliability and validity of the technique is difficult to address.
- Interviews are susceptible to both interviewer and interviewee bias.
- Transcribing the data is a laborious, time consuming process.
- Conducting an interview correctly is a difficult thing to do.

Flowchart


STOP
Analyse data
Record 'expected' and 'unexpected' data
START
Transcribe data
Familiarise interviewee with the device under analysis
N
Y
Are there any more areas?
Take the first/next interview area
Is the area applicable?
N
N
Y
Y
Any more open questions?
Ask open question and record response
N
Y
Any more probe questions?
Ask probe question and record response
Ask closed question and record response
Y
Any more closed questions?
N

Approximate training and application times
In a study comparing 12 HF techniques, Stanton & Young (1999) reported that interviews took the longest to train of all the methods, due to the fact that the technique is a refined process requiring a clear understanding on the analyst's behalf. In terms of application times, a normal interview could last anything between 10 and 60 minutes. Kirwan & Ainsworth (1992) suggest that an interview should last a minimum of 20 minutes and a maximum of 40 minutes. However, the analysis process associated with interviews is very time consuming, and can last weeks in some cases.

Reliability and validity
A structured interview technique scored poorly in terms of reliability and validity in a study carried out on in-car stereo's (Stanton & Young 1999).

Tools needed
An interview requires a pen and paper and an audio recording device, such as a tape or mini-disc recorder.

Bibliography
Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. Taylor and Francis, UK.
Stanton, N. S & Young, M. S. (1999) A guide to methodology in Ergonomics. Taylor and Francis, UK.

Questionnaires

Background and applications

Questionnaires offer a very flexible way of quickly collecting large amounts of data from large amounts of subjects. Questionnaires have been used in many forms to collect data regarding numerous issues within ergonomics and design. Questionnaires can be used to collect information regarding almost anything at all, including usability, user satisfaction, opinions and attitudes. More specifically, questionnaires can be used in the design process to evaluate concept and prototypical designs, to probe perceptions and to evaluate existing system designs. A multitude of questionnaires are available to the human factors practitioner and the system designer. Established questionnaires such as the system usability scale (SUS), the Questionnaire for User Interface Satisfaction (QUIS) and the Software Usability Measurement Inventory (SUMI) are available for practitioners to apply to designs and existing systems. Alternatively, specific questionnaires can be designed and administered during the design process. This method description will concentrate on the design of questionnaires, as existing questionnaire techniques are described elsewhere in this review.

Domain of application

Generic.

Procedure and Advice

There are no set rules for the design and administration of questionnaires. The following procedure is intended to act as a set of guidelines to consider when constructing a questionnaire.

Step 1: Define study objectives

The first step in the design and administration of a questionnaire should always be the definition of the studies objectives. Before any thought is put into the design of the questions, the objectives of the questionnaire must be clearly defined. The objectives should be defined in depth. The analyst or team of analysts should go further than merely describing the goal of the research i.e. Find out which usability problems exist with current command and control set-ups. Rather, the objectives should contain precise descriptions of different usability problems already encountered and descriptions of the usability problems that are expected. Also, the different tasks involved in command and control systems should be defined and the different personnel should be categorised. What the results are supposed to show and what they could show should also be specified as well as the types of questions (closed, multiple choice, open, rating, ranking etc) to be used. Often this stage of questionnaire construction is haphazardly conducted, and consequently the data obtained normally reflects this. Wilson and Corlett (1999) suggest that enough time is spent on this part of the design only when the questions begin to virtually write themselves.

Step 2: Define the population

Once the objectives of the study are clearly and thoroughly defined, the analyst should define the population i.e. the participants whom the questionnaire will be administered to. Again, the definition of the participant population should go beyond simply describing an area of personnel, such as ‘control room operators’ and should

go as deep as defining age groups, different job categories (control room supervisors, operators, management etc) and different organisations (Transco, Military, Railway Safety, NATS etc). The sample size should also be determined at this stage. Sample size is dependent upon the scope of the study and also the amount of time available for data analysis.

Step 3: Construct the Questionnaire
A questionnaire should be made up of four parts; an introduction, a participant classification section, the information section and an epilogue. The introduction should contain information that informs the participant who you are, what the purpose of the questionnaire is and what the results are going to be used for. One must be careful to avoid putting information in the introduction that may bias the participant in any way. For example, describing the purpose of the questionnaire as 'determining usability problems with existing command and control interfaces' may lead the participant. The classification part of the questionnaire normally contains multiple-choice questions requesting information about the participant, such as age, sex, occupation, experience etc. The information part of the questionnaire is the most crucial part, as it contains the questions regarding the initial objectives of the questionnaire. There are numerous categories of questions that can be used in this part of the questionnaire. Which type of question to be used is dependent upon the analysis and the type of data required. Where possible, the type of question used in the information section of the questionnaire should be consistent i.e. if the first few questions are multiple choice, then all of the questions should be kept as multiple choice. The different types of questions available are displayed in table 34.

Table 35. Types of questions used in questionnaire design

Type of Question	**Example question**	**When to use**
Multiple choice	On how many occasions have you witnessed an error being committed with this system? (0-5, 6-10, 11-15, 16-20, More than 20)	When the participant is required to choose a specific response
Rating scales	I found the system unnecessarily complex (Strongly Agree (5), Strongly Disagree (1)	When subjective data regarding participant opinions is required
Paired Associates (Bipolar alternatives)	Which of the two tasks A + B subjected you to the most mental workload? (A or B)	When two alternatives are available to choose from
Ranking	Rank, on a scale of 1 (Very Poor Usability) to 10 (Excellent Usability) the usability of the device.	When a numerical rating is required
Open ended questions	What did you think of the systems usability?	When data regarding participants own opinions about a certain subject is required i.e. subjects compose their own answers
Closed questions	Which of the following errors have you committed or witnessed whilst using the existing system (Action omitted, action on wrong interface element, action mistimed, action repeated, action too little, action too much)	When the participant is required to choose a specific response
Filter questions	Have you ever committed an error whilst using the current system interface? (Yes or No, if Yes, go to question 10, if No, go to question 15)	To determine whether participant has specific knowledge or experience To guide participant past redundant questions

Each question used in the questionnaire should be short in length, worded clearly and concisely, using relevant language. Also, data analysis should be considered when constructing the questionnaire. For instance, if there is little time available for the data analysis process, then the use of open-ended questions should be avoided, as they are time consuming to collate and analyse. If time is limited, then closed questions should be used, as they offer specific data that is quick to collate and analyse. The size of the questionnaire is also of importance. Too large and participants will not complete the questionnaire, yet a very small questionnaire may seem worthless and could suffer the same fate. Optimum questionnaire length is dependent upon the participant population, but it is generally recommended that questionnaires should be no longer than 2 pages.

Step 4: Piloting the questionnaire
According to Wilson & Corlett (1992), once the questionnaire construction stage is complete, the next stage is to pilot the questionnaire. This is a crucial part of the questionnaire design process, yet it is often neglected by human factors practitioners due to various factors, such as time and financial constraints. During this step, the questionnaire is evaluated by its potential user population and also by other human factors practitioners. This allows any problems with the questionnaire to be removed before the critical administration phase. Often, there is only one shot at the administration of a questionnaire, and so the piloting stage is crucial ensuring the questionnaire is adequate and contains no errors or fallacies. Various problems are encountered during the piloting stage, such as errors within the questionnaire, redundant questions and questions that the participants simply do not understand or find confusing. Wilson and Corlett (1999) suggest that the pilot stage should be carried out in three stages.

- *Individual criticism* – questionnaire should be administered to several colleagues who are experienced in questionnaire construction, administration and analysis. Colleagues should be encouraged to offer criticisms of the questionnaire.
- *Depth Interviewing* – Once the individual criticisms have been attended to and any changes have been made, the questionnaire should be administered to a small sample of the intended population. Once they have completed the questionnaire, the participants should be subjected to an interview regarding the answers that they provided. This allows the analyst to ensure that the questions were fully understood and that the correct (required) data is obtained.
- *Large sample administration* – The redesigned questionnaire should then be administered to a large sample of the intended population. This allows the analyst to ensure that the correct data is being collected and also that sufficient time is available to analyse the data. Worthless questions can also be highlighted during this stage. The likely response rate can also be predicted based upon the returned questionnaires in this stage.

Step 5: Questionnaire administration
Once the questionnaire has been successfully piloted, it is ready to be administered. Exactly how the questionnaire is administered is dependent upon the aims and objectives of the analysis, and also the target population. For example, if the target population can be gathered together at a certain time and place, then the questionnaire should be administered at this time, with the analyst present. This ensures that the questionnaires are completed. However, the grouping of the target population in one

place at in time is a very difficult thing to do and so questionnaires are often administered by post. Although this is quick and cheap, requiring little input from the analyst(s), the response rate is very low, typically 10%. Procedures to circumvent this poor response rate are available, such as offering payment on completion, the use of encouraging letters, donation to charity, contacting non-respondents by telephone and sending shortened versions of the initial questionnaire to non-respondents. All these methods have been shown in the past to improve response rates, but almost all involve extra cost.

Step 6: Data Analysis
Once all (or a sufficient amount) of the questionnaires have been returned or collected, the data analysis process should begin. This is a lengthy process and is dependent upon the analysis needs. Questionnaire data is typically computerised and reported statistically. According to Wilson and Corlett (1999) raw data should first be edited, involving transferring the raw data into a computer programme (e.g. Microsoft Excel) and scanning the data for any erroneous answers (e.g. Male respondent with 25 years experience in control room operation reporting that he is aged between 18 – 25). Open-ended questions can also be coded to reduce the data collected. Once the initial data-editing phase is over, the analyst then has a number of 'treated' data sets, and analysis can begin. Typically, data sets are analysed statistically using programs such as SPSS.

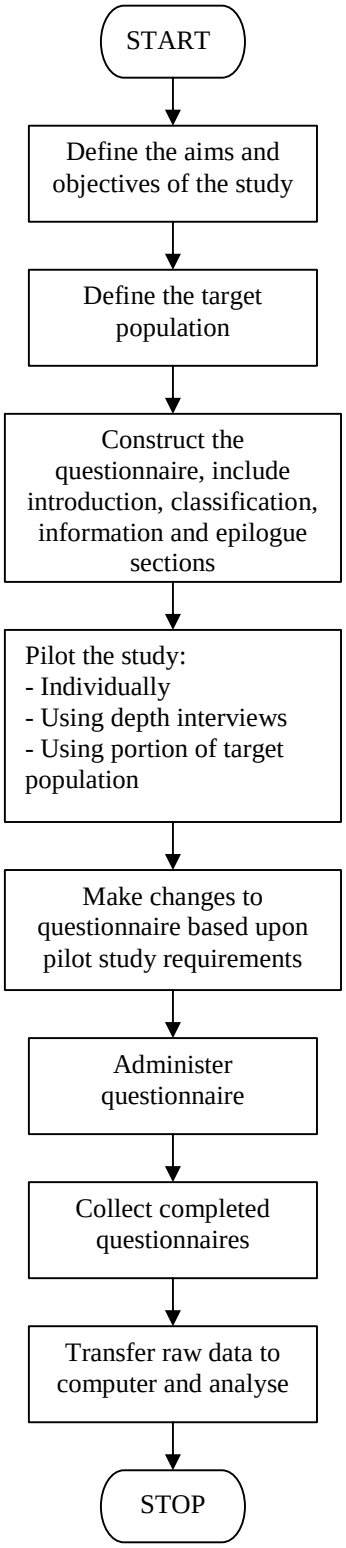
Step 7: Follow up phase
Once the data is analysed sufficiently and conclusions are drawn, the participants who completed the questionnaire should be sent an information pack, informing them of the findings of the questionnaire and also thanking them again for taking part.

Advantages
- When the questionnaire is properly designed, the data analysis phase should be quick and very straightforward.
- Very few resources are required once the questionnaire has been designed.
- Numerous questionnaires already exist in the human factors literature (QUIS, SUMI, SUS etc), allowing the human factors practitioner to choose the most appropriate for the study purposes. This also removes the time associated with the design of the questionnaire. Also, results can also be compared with past results obtained using the same questionnaire.
- Questionnaires offer a very flexible way of collecting high volumes of data from high numbers of subjects. The 'anytime, anyplace' aspect of data collection is very appealing.
- Very easy to administer to large numbers of participants.
- Skilled questionnaire designers can use the questions to direct the data collection.

Disadvantages
- Reliability and validity of questionnaires is questionable.
- The questionnaire design process is a very lengthy one, requiring great skill on the analyst's part.
- Piloting of the questionnaire adds considerable time to the process.
- Typically, response rates are low e.g. around 10% for postal questionnaires.
- The answers provided in questionnaires are often rushed and non-committal.

- Questionnaires are prone to a number of different biases, such as prestige bias.
- Questionnaires offer limited output.

Flowchart



Example
For an example of a human factors orientated questionnaire, the reader is referred to either SUMI (Kirakowski & Corbett 1993) or QUIS (Chin, Diehl & Norman 1988).

Related methods
Questionnaires are a group of techniques that use pre-determined questions on a form to elicit data regarding specific issues. There are numerous questionnaire techniques available to the human factors practitioner. Different types of questionnaires include rating scale questionnaires, paired comparison questionnaires and ranking questionnaires. Questionnaires are also related to the interview technique, in that they utilise open ended and closed questions.

Approximate training and application times
Wilson and Corlett (1999) suggest that questionnaire design is more an art than a science. Although the training time for questionnaire techniques would be minimal, this would not guarantee efficient questionnaire design. Rather, it appears that practice makes perfect, and that practitioners would have to conduct numerous attempts at questionnaire design before becoming proficient at the process. Similarly, although the application time associated with questionnaires is at first glance minimal (completion), when one considers the time expended in the construction and data analysis phases, it is apparent that the total application time is very high.

Reliability and validity
The reliability and validity of questionnaire techniques is highly questionable. Questionnaire techniques are prone to a number of biases and often suffer from the participants merely 'giving the analyst(s) what they want'. Questionnaire answers are also often rushed and non-committal. In a study comparing 12 HF techniques, Stanton and Young (1999) report that questionnaires demonstrated an acceptable level of inter-rater reliability, but also unacceptable levels of intra-rater reliability and validity.

Tools needed
Questionnaires are normally paper based and completed using pen and paper. In the design of the questionnaire a PC is normally used, along with a word processing package such as Microsoft Word. In the analysis of the questionnaire, a spreadsheet package such as Microsoft Excel is required, and a statistical software package such as SPSS is also required to treat the data statistically.

Bibliography
Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. London, Taylor and Francis.
Stanton, N. A & Young, M. S. (1999) A guide to methodology in Ergonomics. London, Taylor and Francis
Wilson, J. R. & Corlett, N. E. (1999) Evaluation of Human Work. Second Edition, UK, Taylor and Francis.

Observational analysis techniques
Various

Background and applications
Observational techniques are a family of techniques that are used to gather data regarding the physical or verbal aspects of a particular task or scenario. Observation has been extensively used in the human factors community for a number of applications, ranging from control room operation to public technology use (Baber and Stanton 1996). The most obvious and widely used form of observational technique is visual observation, whereby an analyst records visually and verbally a particular task or scenario. A number of observational techniques exist, including direct observation, participant observation and remote observation. Baber and Stanton (1996) suggest that there are many observational techniques available, and that these techniques come under three categories; Direct observation, Indirect observation and participant observation. Drury (1999) suggests that there are five different types of information that can be elicited using observational techniques:

1) Sequence of activities
2) Duration of activities
3) Frequency of activities
4) Fraction of time spent in states
5) Spatial movement

As well as visual data, verbal data is also frequently recorded, particularly verbal interactions between team members. Observational techniques can be used at any stage of the design process in order to gather information regarding existing or proposed designs.

Domain of application
Generic.

Procedure and advice
There is no real set procedure for carrying out an observational analysis. The procedure would normally be determined by the type and scope of analysis required. A typical observational analysis procedure can be split into three phases; the observation design stage, the observation application stage and the analysis stage. The following procedure provides the analyst with a general set of guidelines for conducting a 'direct' type observation.

Step 1: Define the objective of the analysis
The first step in observational analysis has to be the definition of the analysis aims and objectives. This should include determining which product or system is under analysis, which environment the observation will take place, which user groups will be observed, what type of scenario's will be observed and what data is required. Each point should be clearly defined and stated before the process continues.

Step 2: Define the scenario(s)
Once the aims and objectives of the analysis are clearly defined, the scenario(s) to be observed should be defined and described further. For example, when conducting an observational analysis of control room operation, which type of scenario is required should be clearly defined. Is normal operation under scrutiny or is the analysis

focussed upon operator interaction and performance under emergency situations. The exact nature of the required scenario(s) should be clearly defined by the observation team.

Step 3: Observation plan
Once the aim of the analysis is defined and also the type of scenario to be observed is determined, the analysis team should proceed to plan the observation. The team should consider what they are hoping to observe, what they are observing, and how they are going to observe it. Any recording tools should be defined and also the length of observations should be determined. Placement of video and audio recording equipment should also be considered. To make things easier, a walkthrough of the system/environment/scenario under analysis is required. This allows the analyst(s) to become familiar with the task in terms of time taken, location and also the system under analysis.

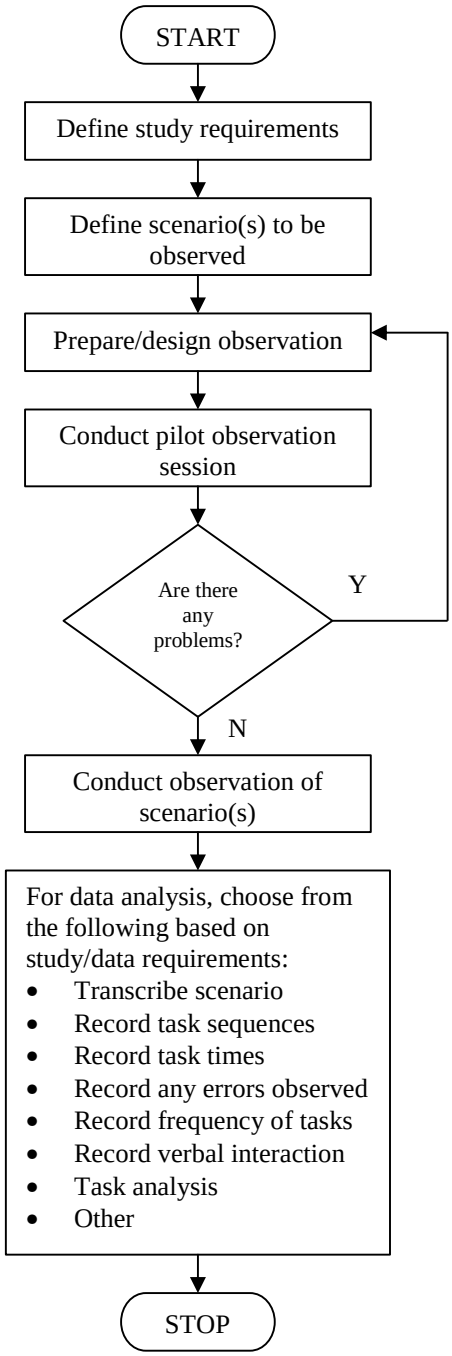
Step 4: Pilot observation
In any observational study a pilot or practice observation is crucial. This allows the analysis team to assess any problems with the data collection, such as noise interference or problems with the recording equipment. The quality of data collected can also be tested and also any effects of the observation upon task performance can be assessed. If major problems are encountered, the observation may have to be re-designed. Steps 1 to 4 should be repeated until the analysis team are happy that the quality of the data collected will be sufficient for their study requirements.

Step 5: Observation
Once the observation has been designed, the team should proceed with the observation. Observation length and timing are dependent upon the scope and requirements of the analysis. Once the required data is collected, the observation should stop and step 6 should be undertaken.

Step 6: Data analysis
Once the observation is complete, the analysis team should begin the data analysis procedure. Firstly, a transcript of the whole observation should be made. This is a very time consuming process but is crucial to the analysis. Depending upon the analysis requirements, the team should then proceed to analyse the data in the format that is required, such as frequency of tasks, verbal interactions, sequence of tasks etc. When analysing visual data, typically user behaviours are coded into specific groups. The software package Observer is used to aid the analyst in this process.

Step 7: Participant feedback
Once the data has been analysed and conclusions have been drawn, the participants involved should be provided with feedback of some sort. This could be in the form of a feedback session or a letter to each participant. The type of feedback used is determined by the analysis team.

Flowchart


START
Define study requirements
Define scenario(s) to be observed
Prepare/design observation
Conduct pilot observation session
Are there any problems?
Y
N
Conduct observation of scenario(s)
For data analysis, choose from the following based on study/data requirements:
Transcribe scenario
Record task sequences
Record task times
Record any errors observed
Record frequency of tasks
Record verbal interaction
Task analysis
Other
STOP

Advantages

- Observation technique data provides a 'real life' insight into man-machine, and team interaction.
- Various data can be elicited from an observational study, including task sequences, task analysis, error data, task times, verbal interaction and task performance.
- Observation has been used extensively in a wide range of domains.
- Observation provides objective information.
- Detailed physical task performance data is recorded, including social interactions and any environmental task influences (Kirwan & Ainsworth 1992).
- Observation is excellent for the initial stages of the task analysis procedure.
- Observation analysis can be used to highlight problems with existing operational systems. It can be used in this way to inform the design of new systems or devices.
- Specific Scenarios are observed in their 'real world' setting.

Disadvantages

- The main criticism of observational techniques centres on their intrusiveness. Knowing that they are being watched tends to elicit new and different behaviours in participants. For example, when observing control room operators, they may exhibit a performance that is exact in terms of training requirements. This may be due to the fact that the operator's do not wish to be caught bending the rules in any way i.e. bypassing a certain procedure.
- Observational techniques are very, very time consuming, both in conducting the actual observation and then in analysing the data collected. When conducting an observation, a certain scenario cannot simply be conjured out of thin air. If an emergency scenario is required, the observation may go on for a number of weeks before the required scenario occurs. Also, the data analysis procedure is even more time consuming. Kirwan & Ainsworth (1992) suggest that when conducting the transcription process, 1 hour of recorded audio data takes on analyst approximately 8 hours to transcribe.
- Cognitive aspects of the task under analysis are not elicited using observational techniques. Verbal protocol analysis is more suited for collecting data on the cognitive aspects of task performance.
- An observational study can be both difficult and expensive to set up and conduct. Many re-iterations may take place before the observation can be carried out. Also, the use of recording equipment ensures that the technique is not a cheap one.
- Causality is a problem. Errors can be observed and recorded during an observation but why the errors occur may not always be clear.
- The analyst has a very low level of experimental control.
- In most cases, a team of analysts is required to perform an observation study.

Related methods

The observational technique described comes from a family of observation techniques, including indirect observation and participant observation. Other related techniques include verbal protocol analysis, critical decision method, applied cognitive task analysis, walkthroughs and cognitive walkthroughs. All of these techniques require some sort of task observation. Observation is also instrumental in task analysis techniques, such as HTA, and in the construction of timeline analyses.

Approximate training and application times
Whilst the training time for an observational analysis are low (Stanton & Young 1999), the application time is normally very high. The data analysis stage can be particularly time consuming. Kirwan & Ainsworth (1992) suggest that in the transcription process, 1 hour of audio recorded data would take approximately 8 hours to transcribe.

Reliability and validity
Observational analysis is beset by a number of problems that can potentially affect the reliability and validity of the technique. According to Baber & Stanton (1996) problems with causality, bias (in a number of forms), construct validity, external validity and internal validity can all arise unless the correct precautions are taken. Whilst observational techniques possess a high level of face validity (Drury 1990) and ecological validity (Baber & Stanton 1996), analyst or participant bias can adversely affect the reliability and validity of the techniques.

Tools needed
For a thorough observational analysis, recording equipment is required. Normally, both visual and audio recording equipment is used. Observational studies can be conducted using pen and paper, however this is not recommended, as crucial parts of data are often not recorded.

Bibliography
Baber, C. & Stanton, N. A. (1996) Observation as a technique for Usability Evaluations. In P. Jordan et al (eds.), *Usability in Industry*, pp 85-94. London, Taylor and Francis.
Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. London, Taylor and Francis.
Stanton, N. A & Young, M. S. (1999) A guide to methodology in Ergonomics. London, Taylor and Francis.

Performance time prediction techniques

Task performance time prediction is used in the design of systems and processes in order to determine whether proposed design concepts offer performance time reductions, and also to offer total performance times associated with a task or set of tasks. Predicted task performance times are compared to existing performance times in order to evaluate the impact of proposed design concepts. Predicted task performance times are also evaluated in order to ensure that task performance with the proposed design meets the associated performance time constraints or requirements. According to Card, Moran & Newell (1983), it is useful for system designers to possess a model enabling the prediction of how much time it takes to accomplish a given task.

The prediction of performance times associated with operator tasks was first attempted in the HCI domain (Card, Moran & Newell 1983). The GOMS family of techniques included the Keystroke Level Model (KLM), which offered a set of standard times for operator actions, such as button press, mental operation and homing. Operator tasks are broken down into unit-tasks and standard times are assigned to each unit-tasks. These unit-task times are then summed to calculate the total performance time. Although initially developed for HCI, the technique has been used elsewhere. Stanton & Young (1998) used KLM to predict the performance time for the operation of two In-Car stereo/radio devices. According to Baber et al (2003), Gray et al (1993) and Lawrence et al (1995) report the potential of critical path analysis (CPA) for predicting the performance time using telephone operator workstations. Timeline analysis techniques have also been used to predict performance time. According to Kirwan & Ainsworth (1992), the American national standards institute defines timeline analysis as,

> "An analytical technique for the derivation of human performance requirements which attends to both the functional and temporal loading for any given combination of tasks"

Typically, observational data is used to construct graphically the performance times associated with operator tasks. Timeline type analysis seems to be potentially suited to analysing team performance times. Kirwan & Ainsworth (1992) also suggest that timelines are useful in assessing task allocation and identifying communications requirements.

Predicting task performance times is a crucial aspect of system design. Design concepts or proposals require evaluation in terms of task performance time, in order to establish improvements in performance time over similar, existing systems. The C4i system design should of course exhibit performance time reductions or gains. Operators of the C4i system will undoubtedly be required to perform under great time constraints, with time a crucial factor in achieving mission success. As a result, any design concept proposed requires testing in terms of task performance times exhibited. Thus, the use of both performance time prediction and measurement techniques is proposed. Performance time measurement techniques are reviewed in work package 1.3.3.

The performance time prediction techniques reviewed in this document are shown below.

1) KLM – Keystroke Level Model

2) Timeline Analysis
3) CPA – Critical Path Analysis

KLM – The Keystroke Level Model

Card, S. K., Moran, T. P. & Newell, A. (1983) The Psychology of Human Computer Interaction. Hillsdale, NJ. Erlbaum.

Background and applications

The Keystroke Level model (KLM) is a very simple technique that is used to predict task execution time in HCI. Part of the GOMS family of methods, KLM uses a number of pre-determined operators to predict expert error free task execution times. KLM uses four physical motor operators, one mental operator and one system response operator. These are:
Keystroking (K) – represents a keystroke or button press (on any button device)
Pointing (P) – represents pointing to a target on a display with a mouse
Homing (H) – represents the hand movement of the user when moving his hands between keys, buttons etc.
Drawing (D) – represents the drawing of straight line segments using a mouse.
Mental operator (M) – represents the users mental preparation to execute a physical operation.
System response operator (R) – represents the system response time.

Each operator has an associated execution time. Total task performance time is equal to the sum of each operator exhibited in the task. Each operator execution time is shown below.

$$T_{execute} = T_k + T_p + T_h + T_d + T_m + T_r$$

Operator/Action	**Execution time**
K – Pressing Key or Button	
Best typist	.08
Good typist	.12
Average skilled typist	.20
Average non-secretary typist	.28
Typing random letters	.50
Typing complex codes	.75
Worst typist (unfamiliar with keyboard)	1.20
P – Pointing with mouse to a target on a display	1.10
H – Homing hands on keyboard, button etc	.40
D – Drawing straight line segments	.9nd + .16/d
M – Mental preparation	1.35
R – System response time	t

The KLM technique also provides a set of heuristic rules for placing the mental operations (M). These are shown below.

Rule 0: Insert Ms in front of all Ks that are not part of argument strings proper (e.g. text or numbers)
Rule 1: If an operator following an M is fully anticipated in an operator previous to M, then delete the M
Rule 2: If a string of MKs belongs toa cognitive unit (e.g. the name of a command) then delete all Ms but the first.
Rule 3: If a K is a redundant terminator (e.g. the terminator of a command immediately following the terminator of its argument) then delete the M in front of it.
Rule 4: If a K terminates a constant string (e.g. a command name), then delete the M infront of it; but if the K terminates a variable string (e.g. an argument string) then keep the M in front of it

Domain of application
HCI.

Procedure and advice
Step 1: Compile task list and set scenario to be analysed
Firstly, the analyst should compile an exhaustive task list for the device or system under analysis. Once the task list is complete, the analyst should select the particular task or set of tasks that are to be analysed.

Step 2: Determine the component operations involved in the task
Once the task under analysis has been defined, the analyst needs to determine the component operations involved in the task. KLM calculates task performance time by summing the component operations involved in the task.

Step 3: Insert physical operations
Any homing or button presses involved in the task should be recorded. The time for each component should be recorded.

Step 4: Insert system response time
Next, the analyst should insert the appropriate system response time. This is normally determined from manufacturer specifications (Stanton & Young 1999). If these are not readily available a domain expert estimate is sufficient.

Step 5: Insert mental operations
Finally, the mental operation times should be inserted. The analyst should use the KLM heuristic rules to place the mental operations.

Step 6: Calculate the total task time
To calculate the total task time, the analyst should add each associated component operation time. The sum of the operation times equals the total task performance time (error free performance). For maximum accuracy, the final sum should be multiplied by 1.755.

Advantages
- KLM is a very easy and quick technique to use.
- KLM requires very little training (Stanton & Young 1999)
- KLM has been used in different domains, such as driving (Stanton & Young 1999)
- KLM can be used to quickly compare the task times for two different devices or systems.
- Gives an immediately useful output of estimated task performance time.
- Encouraging reliability and validity data (Stanton & young 1999).

Disadvantages
- KLM was designed specifically for computer based tasks (HCI). New operators may have to be developed for the technique to be used in other domains.
- KLM only models error free expert performance.
- KLM does not take context into account.

- There is limited validation evidence associated with the use of KLM outside of HCI.
- KLM assumes that all performance is serial and cannot deal with Parallel activity.
- KLM ignores other unit task activity and also variation in performance.
- KLM ignores flexible human activity (Baber and Mellor, 2001)

Related methods
KLM is part of the GOMS family of methods developed for use in the HCI domain. These are NGOMSL, KLM, CMN-GOMS and CPM-GOMS. A HTA for the system or device under analysis is also very useful.

Approximate training and application times
Stanton & Young (1999) suggest that KLM is moderately time consuming to train. Execution time is dependent upon the size of the task under analysis, but is generally low. Stanton & Young also reported that KLM execution times improve considerably on the second application.

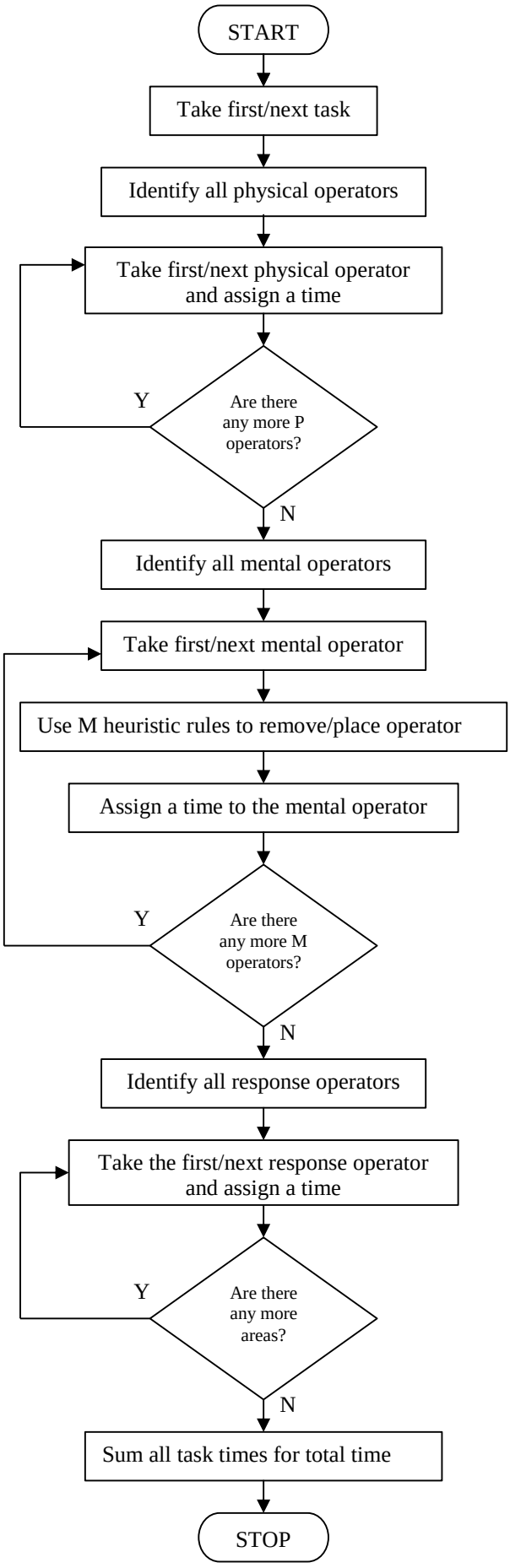
Reliability and validity
Stanton & Young (1999) reported outstanding reliability and validity measures for KLM. Out of twelve HF techniques tested, KLM was the only technique to achieve acceptable levels across the three ratings of inter-rater reliability, intra-rater reliability and validity.

Tools needed
KLM is pen and paper method. The analyst should also have access to the device or system under analysis and also the KLM operator times.

Domain of application
KLM is part of the GOMS family of methods that was developed for use in the HCI domain. There is evidence of its use in the analysis of in-car entertainment systems (Stanton and Young, 1999) and also 'bank deposit reconciliation systems' (Kieras and John, 1994).

Bibliography
Card, S. K, Moran, T. P, Newell, A (1983) The psychology of human computer interaction, Lawrence Erlbaum Associates, New Jersey.
Salvendy, G. (1997) Handbook of human factors and ergonomics, 2nd edition, John Wiley and Sons, Canada.
Stanton, N. A. and Young, M. S (1999) A guide to methodology in Ergonomics: Designing for Human Use, Taylor and Francis, New York.

Flowchart


START
Take first/next task
Identify all physical operators
Take first/next physical operator and assign a time
Are there any more P operators?
Y
N
Identify all mental operators
Take first/next mental operator
Use M heuristic rules to remove/place operator
Assign a time to the mental operator
Are there any more M operators?
Y
N
Identify all response operators
Take the first/next response operator and assign a time
Are there any more areas?
Y
N
Sum all task times for total time
STOP

Example (Source: Stanton and Young, 1999)
The following example is taken from a KLM analysis of a Ford In-Car Radio system.

When using the Ford 7000 RDS EON in-car stereo, to switch the device on the user has to push the on/off button. For the KLM analysis, this would be presented as:

Task	**Execution time(s)**
Switch on	MHKR = 2.65 + 1 = 3.65

i.e. M = the driver thinking about pressing the on/off button, H = the driver positioning his finger over the button, K = the driver actually pressing the button and R = the time it takes for the radio to turn on (system response time).

The above example is a very simple one. A more complicated one, again for the Ford 7000 RDS EON, would be to adjust the treble on the system. To do this, the driver would have to push the bass button twice and then use the volume knob. Using a KLM analysis, this would be presented as:

Task	**Execution time(s)**
Adjust treble	MHKKHKR = 4.15+0.3 = 4.45

i.e. M = the driver thinking about the following actions, H = the driver positioning his finger over the BASS button, KK = the driver pressing the BASS button twice, H = the driver positioning his finger over the volume button, K = the driver turning the volume button and R = the system response time.
The full KLM analysis of the Ford and Sharp in-car radios performed by Stanton and Young (1999) is shown below:

Table 36. KLM output

Task	**Time – FORD**	**Time - SHARP**	**Difference +/-**
Switch unit on	MHKR = 2.65+1 = 3.65	MHKR = 2.65+1 = 3.65	**0**
Adjust Volume	MHKR = 2.65+0.1 = 2.75	MHKR = 2.65+0 = 2.65	**+0.1**
Adjust Bass	MHKHKR = 3.95+0.2 = 4.15	MHKR = 2.65+0 = 2.65	**+1.5**
Adjust Treble	MHKKHKR = 4.15+0.3 = 4.45	MHKR = 2.65+0 = 2.65	**+1.8**
Adjust Balance	MHKKHKR = 4.15+0.3 = 4.45	MHKKR = 2.85+0.1 = 2.95	**+1.5**
Choose new Pre-set	MHKR = 2.65+0.2 = 2.85	MHKR = 2.65+0.2 = 2.85	**0**
Use Seek	MHKR = 2.65+1 = 3.65	MHKR = 2.65+1 = 3.65	**0**
Use Manual Search	MHKHKR = 3.95+1 = 4.95	MHKR = 2.65+1 = 3.65	**1.3**
Store Station	MHKR = 2.65+1 = 3.65	MHKR = 2.65+3 = 5.65	**-2**
Insert Cassette	MHKR = 2.65+1 = 3.65	MHKR = 2.65+1 = 3.65	**0**
Autoreverse and FF	MHKRHKRKR = 4.15+5 = 9.15	MHKRKRK = 3.05+5 = 8.05	**1.1**
Eject Cassette	MHKR = 2.65+0.5 = 3.15	MHKR = 2.65+0.3 = 2.95	**0.2**
Switch Off	MHKR = 2.65+0.5 = 3.15	MHKR = 2.65+0.7 = 3.35	**-0.2**
Total time	**53.65**	**48.35**	**5.3**

As a result of the KLM analysis, it can be concluded that the when performing the set of tasks outlined above, it takes around 5 seconds longer to complete them using the Ford design.

Timeline analysis
Although not a set methodology, timeline analysis is an approach that can be used in order to depict scenarios in terms of tasks and their associated task performance times. Timeline analysis can be used to display the functional and temporal requirements of a task. Timeline analysis can be used both predictively and retrospectively, and the output is typically a graph. Timeline analysis can also combined with workload analysis to represent the workload associated with each task step (Kirwan & Ainsworth 1992). In terms of analysing command and control and team-based tasks, the appeal of timeline analysis lies in the fact that it could potentially depict individual and team task steps over time.

Domain of application
Generic.

Procedure and advice
Step 1: Data collection
The first step in any timeline analysis is to collect specific data from the system under analysis. Task performance times should be recorded for all of the behaviours exhibited in the system. Typically, observational analysis is used during the data collection phase. If the technique is being applied retrospectively, then the analyst(s) should observe the scenario under analysis. If a predictive timeline is required, similar scenarios in similar systems should be observed.

Step 2: HTA
Once sufficient data regarding the task under analysis is collected, a HTA should be conducted. HTA (Annett et al., 1971; Shepherd, 1989; Kirwan & Ainsworth, 1992) is based upon the notion that task performance can be expressed in terms of a hierarchy of goals (what the person is seeking to achieve), operations (the activities executed to achieve the goals) and plans (the sequence in which the operations are executed). The hierarchical structure of the analysis enables the analyst to progressively re-describe the activity in greater degrees of detail. The analysis begins with an overall goal of the task, which is then broken down into subordinate goals. At this point, plans are introduced to indicate in which sequence the sub-activities are performed. When the analyst is satisfied that this level of analysis is sufficiently comprehensive, the next level may be scrutinised. The analysis proceeds downwards until an appropriate stopping point is reached (see Annett et al, 1971; Shepherd, 1989, for a discussion of the stopping rule).

Step 3: Determine performance times
Step 3 allows the analyst(s) to create a performance time database for the analysis. Each task step in the HTA should be assigned a performance time. If the analysis is retrospective, this involves sifting through the data gathered during observations and recording the task performance times for each task. If a predictive timeline is required, then the analyst(s) should record the performance times for similar tasks to that involved in the predicted scenario.

Step 4: Construct the Timeline graph
The timeline graph normally flows from left to right with the time running along the Y-axis and the tasks running up the X-axis.

Advantages

- Timeline graphs can be used to compare the performance times associated with two different systems or designs.
- Timeline analysis could be used to represent team-based tasks and parallel activity.
- Timeline analysis can be used to highlight problematic tasks or task sequences in the design of systems and processes.
- Workload analysis can be mapped directly onto a timeline graph. This makes for a very powerful analysis.
- Timeline analysis is a simple technique requiring little training.
- Requires very few resources once data collection phase is complete.

Disadvantages

- The reliability and validity of the technique is questionable.
- Observation data is often flawed by a number of biases.
- When used predictively, timeline analysis can only model error free performance.
- Initial data collection phase is time consuming and resource intensive.

Approximate training and application times

The training for timeline analysis is very low. The application time is minimal once the initial data collection is complete. The data collection involved is dependent upon the scenario under analysis. For large, complex scenarios, the data collection time associated with timeline analysis is very high.

Reliability and validity

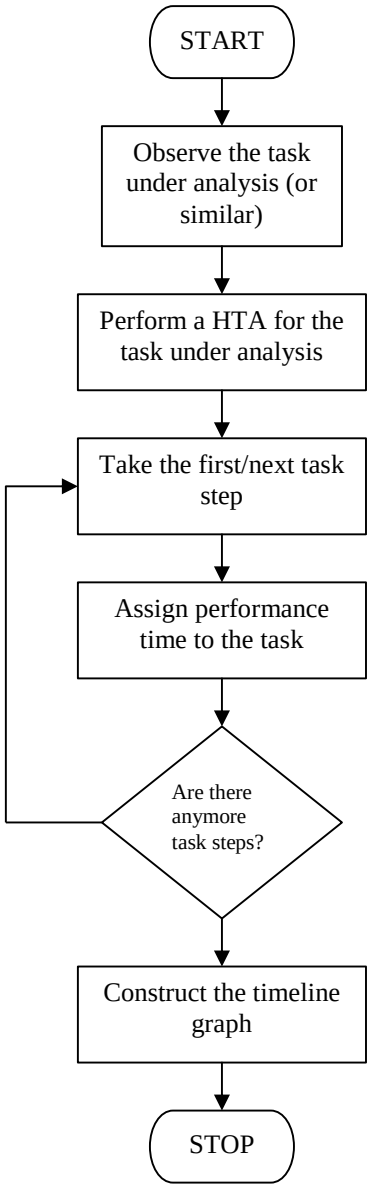
Kirwan and Ainsworth (1992) report that the technique possesses high face validity. No data regarding the reliability and validity of the technique are available in the literature.

Tools needed

Once the data collection phase is complete, timeline analysis can be conducted using pen and paper. The data collection phase (observation) typically requires using video and audio recording devices.

Bibliography

Kirwan, B. & Ainsworth, L. K. (1992) A Guide to Task Analysis. UK, Taylor and Francis.


START
Observe the task under analysis (or similar)
Perform a HTA for the task under analysis
Take the first/next task step
Assign performance time to the task
Are there anymore task steps?
Construct the timeline graph
STOP

Example

The following example is a mock timeline analysis based upon an aviation scenario. For further examples, the reader is referred to Kirwan & Ainsworth (1992).

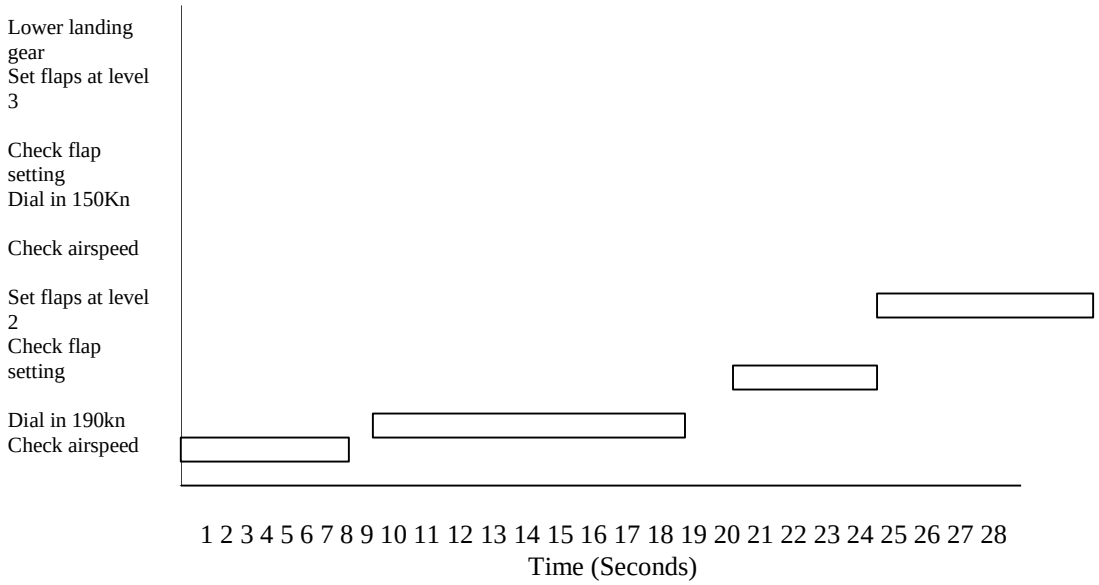


Figure 29. Mock timeline analysis graph for landing task

Critical Path Analysis
See Pages 127-133

Mental Workload Prediction techniques

Mental workload (MWL) is a concept of increasing importance within the human factors domain as operators of complex dynamic systems face ever increasing cognitive demands. Mental workload represents the proportion of resources demanded by a task or set of tasks. Young & Stanton (2001) propose the following definition of MWL:

> "The mental workload of a task represents the level of attentional resources required to meet both objective and subjective performance criteria, which may be mediated by task demands, external support, and past experience." (Young & Stanton, 2001; p. 507)

According to Young (2003) MWL is a core area for research in virtually every field imaginable. Research concerning MWL prediction and assessment is ongoing in a number of fields, including aviation, air traffic control, military operations, driving and control room operation. The assessment or measurement of MWL is of great importance within these domains, informing system and task design and evaluating incidents and error. Typically, human factors practitioners employ a battery of MWL assessment techniques including primary task performance measures, secondary task performance measures (reaction times, embedded tasks), physiological measures (HRV, HR), and subjective ratings (SWAT, NASA TLX) (Young 2003).

A number of MWL assessment techniques exist, which allow the HF practitioner to evaluate the MWL associated with a certain task. However, typical MWL assessment techniques are used to assess workload after or during task performance. These techniques fall into the evaluation category and are evaluated in work package 1.3.3. Such techniques are difficult to employ during the design process, as the system under analysis may not actually exist, and simulation can be extremely costly. Therefore, mental workload prediction techniques can be used. Although in its infancy, a number of techniques exist that can be used to predict mental workload. Also known as analytical techniques, these techniques are available in numerous forms. In the past, models have been used to predict operator workload, such as the timeline model or Wicken's multiple resource model. MWL assessment techniques such as Pro-SWORD have also been tested for there use in predicting operator MWL (Vidulich, Ward & Schueren 1991). Although the use of MWL assessment techniques in a predictive fashion is limited, Salvendy (1997) reports that SME projective ratings tend to correlate well with operator subjective ratings. It is apparent that analytical mental or predictive workload techniques are particularly important in the early stages of system design and development (Hendry et al).

Cognitive task load analysis (CTLA) is a technique used to assess or predict the cognitive load of a task or set of tasks imposed upon an operator. CTLA is typically used early in the design process to aid the provision of an optimal cognitive load for the system design in question. The CTLA is based upon a model of cognitive task load (Neerincx 2003) that describes the effects of task characteristics upon operator mental workload. According to the model, cognitive (or mental) task load is comprised of percentage time occupied, level of information processing and the number of task set switches exhibited during the task.

Pro-SWAT is a variation of the SWAT (Reid & Nygren 1988) subjective workload assessment technique that has been used to predict operator workload. Originally developed to assess pilot workload the SWAT has been used predictively as Pro-SWAT on a number of occasions (Detro 1985, Kuperman 1985, Masline & Biers

1987). SWAT is a multidimensional tool that uses three dimensions of operator workload; *time load, mental effort load and stress load*. After an initial weighting procedure, participants are asked to rate each dimension (time load, mental effort load and stress load, on a scale of 1 to 3. A workload score is then calculated for each dimension and an overall workload score is between 1-100 is also calculated.

The Subjective Workload Dominance Technique (SWORD) is a subjective workload assessment technique that has been used both retrospectively and predictively (Pro-SWORD) (Vidulich, Ward & Schueren 1991). SWORD uses paired comparison of tasks in order to provide a rating of workload for each individual task. Participants are required to rate one tasks dominance over another in terms of workload imposed. When used predictively, tasks are rated for their dominance before the trial begins, and then rated post-test to check for the sensitivity of the predictions. Vidulich, Ward & Schueren (1991) report the use of the SWORD technique for predicting the workload imposed upon F-16 pilots by a new HUD attitude display system.

The mental workload prediction techniques evaluated in this document are shown below.

1. CTLA – Cognitive Task Load Analysis
2. Pro-SWAT
3. Pro-SWORD

Cognitive Task Load Analysis

Background and applications

Cognitive task load analysis (CTLA) is a technique used to assess or predict the cognitive load of a task or set of tasks imposed upon an operator. CTLA is typically used early in the design process to aid the provision of an optimal cognitive load for the system design in question. The technique has been used in its present format in naval domain (Neerincx 2003). The CTLA is based upon a model of cognitive task load (Neerincx 2003) that describes the effects of task characteristics upon operator mental workload. According to the model, cognitive (or mental) task load is comprised of percentage time occupied, level of information processing and the number of task set switches exhibited during the task. According to Neerincx (2003), the operator should not be occupied by one task for more than 70-80% of the total time. The level of information processing is defined using the SRK framework (Rasmussen 1986). Finally, task set switches are defined by changes of applicable task knowledge on the operating and environmental level exhibited by the operators under analysis (Neerincx 2003). The three variables time occupied, level of information processing and task set switches are combined to determine the level of cognitive load imposed by the task. High ratings for the three variables equal a high cognitive load imposed on the operator by the task. The three dimensional model of cognitive task load is shown below (Source: Neerincx 2003)

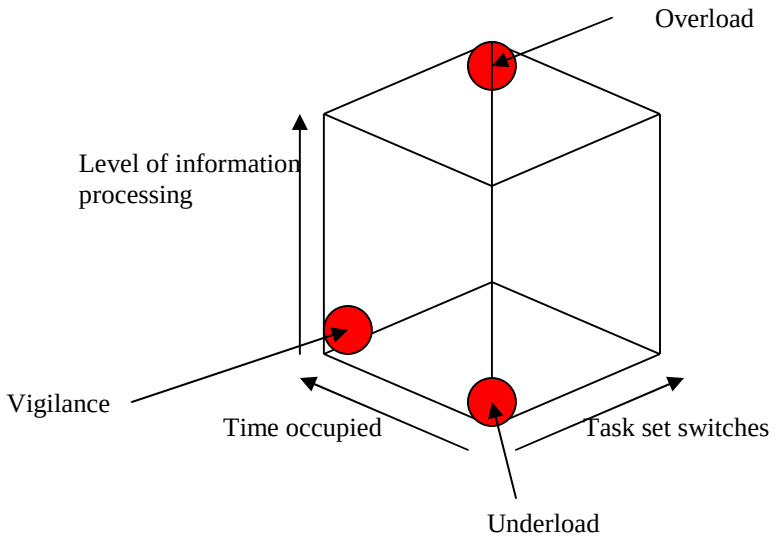


Figure 30. Three dimensional model of cognitive task load (Neerincx 2003)

Domain of application

Maritime

Procedure and advice
The following procedure is adapted from Neerincx (2003).

Step 1: Define task(s) or scenario under analysis
The first step in analysing operator cognitive load is to define the task(s) or scenario(s) under analysis.

Step 2: Data collection
Once the task or scenario under analysis is clearly defined, specific data should be collected regarding the task. Observation, interviews, questionnaires and surveys are typically used.

Step 3: Task decomposition
The next step in the CTLA involves defining the overall operator goals and objectives associated with each task under analysis. Task structure should also be described fully.

Step 4: Create event list
Next, a hierarchical event list for the task under analysis should be created. According to Neerincx (2003), the event list should describe the event classes that trigger task classes, providing an overview of any situation driven elements.

Step 5: Describe scenario(s)
Once the event classes are described fully, the analyst should begin to describe the scenarios involved in the task under analysis. This description should include sequences of events and their consequences. Neerincx (2003) recommends that this information is displayed on a timeline.

Step 6: Describe basic action sequences (BAS)
BAS describe the relationship between event and task classes. These action sequences should be depicted in action sequence diagrams.

Step 7: Describe compound action sequences (CAS)
CAS describe the relationship between event and task instances for situations and the associated interface support. The percentage time occupied, level of information processing and number of task set switches are elicited from the CAS diagram.

Step 8: Determine percentage time occupied, level of information processing and number of task set switches
Once the CAS are described, the analyst(s) should determine the operators percentage time occupied, level of information processing and number of task set switches exhibited during the task or scenario under analysis.

Step 9: Determine cognitive task load
Once percentage time occupied, level of information processing and number of task set switches are defined, the analyst(s) should determine the operator(s) cognitive task load. The three variables should be mapped onto the model of cognitive task load shown in figure 28.

Example

The following example is taken from Neerincx (2003). As the output from a CTLA is large, only extracts are shown below. For a more thorough example of CTLA, the reader is referred to Neerincx (2003).

Table 37. Scenario description table (Source: Neerincx (2003))

Initial state: Ship is en route to Hamburg: there are two operators present on the bridge

Time	Event	Additional Information
21:54	Short circuit	Location: Engine's cooling pump in engine room Details: Short circuit causes a fire in the pump, which is located in the cooling system Consequences: Cooling system will not work, and the engine temperature will increase Source: None (event is not detected by the system
22:03	Fire	Location: Engine room Details: A pump in the engine room is on fire Consequences: Unknown Source: Smoke detector of fire control system
22:06	Max. engine temp	Location: Engine Room Details: The temperature of the engine increased beyond the set point Consequences: The engine shuts down after a period of high temperature Source: Propulsion management system
22:08	Engine shutdown	Location: Engine Room Details: The temperature was too high for the critical period Consequences: The vessel cannot maintain its current speed Source: Propulsion management system

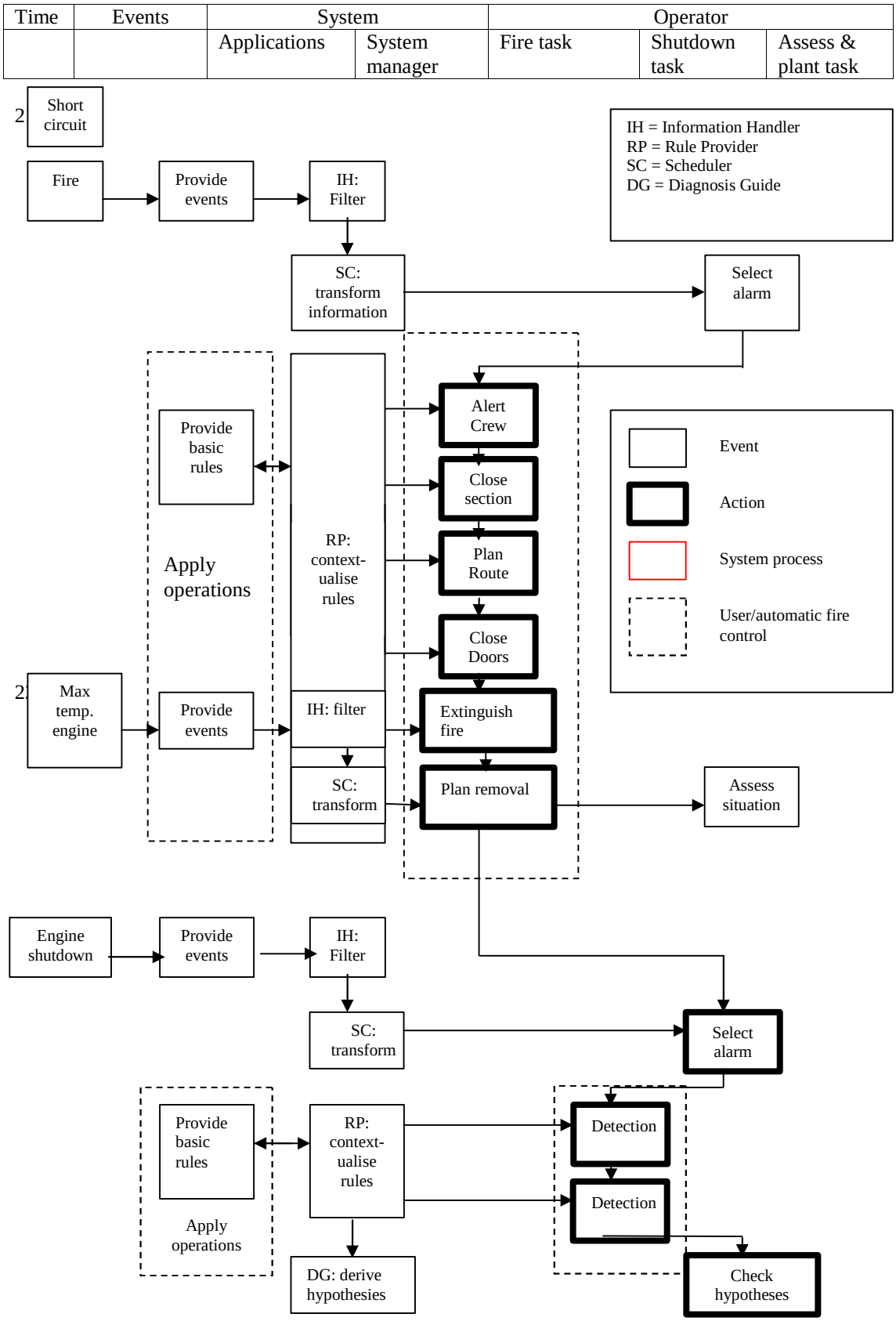


Figure 31. Example Action Sequence Diagram (Source: Neerincx 2003)

Advantages
- The technique is based upon sound theoretical underpinning.
- Can be used during the design of systems and processes to highlight tasks or scenarios that impose especially high cognitive task demands.
- Seems to be suited to analysing control room type tasks or scenarios.

Disadvantages
- The technique appears to be quite complex
- Such a technique would be very time consuming in its application.
- A high level of training would be required.
- There is no guidance on the rating of cognitive task load. It would be difficult to give task load a numerical rating based upon the underlying model.
- Initial data collection would be very time consuming.
- The CTLA technique requires validation.
- Evidence of the use of the technique is limited.

Related methods
The CTLA technique uses action sequence diagrams, which are very similar to operator sequence diagrams. In the data collection phase, techniques such as observation, interviews and questionnaires are used.

Approximate training and application times
It is estimated that the training and application times associated with the CTLA technique would both be very high.

Reliability and validity
No data regarding the reliability and validity of the technique are offered in the literature.

Tools needed
Once the initial data collection phase is complete, CTLA can be conducted using pen and paper. The data collection phase would require video and audio recording equipment and a PC.

Bibliography
Neerincx, M. A. (2003) Cognitive Task Load Analysis: Allocating Tasks and Designing Support. In E. Hollnagel (ed) *Handbook of Cognitive Task Design*. Pp 281-305. Lawrence Erlbaum Associates Inc.

SWAT – Subjective Workload Assessment Technique

G. B.Reid, &.T. E Nygren

Background and applications

The subjective workload assessment technique (SWAT) (Reid & Nygren 1988) is a workload assessment technique that was developed by the US Air force Armstrong Aerospace Medical Research laboratory at the Wright Patterson Air force Base, USA. SWAT was originally developed to assess pilot workload in cockpit environments but more recently has been used predictively (Pro-SWAT)(REFERENCE AND EXAMPLE HERE!). Along with the NASA TLX technique of subjective workload, SWAT is probably one the most commonly used of the subjective techniques to measure operator workload. Like the NASA TLX, SWAT is a multidimensional tool that uses three dimensions of operator workload; *time load, mental effort load and stress load*.

Time load is extent to which a task is performed within a time limit and the extent to which a multiple tasks must be performed concurrently. Mental effort load is the associated attentional demands of a task, such as attending to multiple sources of information and performing calculation. Finally, stress load includes operator variables such as fatigue, level of training and emotional state. After an initial weighting procedure, participants are asked to rate each dimension (time load, mental effort load and stress load, on a scale of 1 to 3. A workload score is then calculated for each dimension and an overall workload score is between 1-100 is also calculated. SWAT uses a three point rating scale for each dimension. This scale is shown in table 37.

Table 38. SWAT three point rating scale

Time Load	**Mental Effort Load**	**Stress Load**
1 – Often have spare time: interruptions or overlap among other activites occur infrequently or not at all	1 – Very little conscious mental effort or concentration required: activity is almost automatic, requiring little or no attention	1 – Little confusion, risk, frustration, or anxiety exists and can be easily accommodated
2 – Occasionally have spare time: interruptions or overlap among activities occur frequently	2 – Moderate conscious mental effort or concentration required: complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity; considerable attention is required	2 – Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload: significant compensation is required to maintain adequate performance
3 – Almost never have spare time: interruptions or overlap among activities are very frequent, or occur all of the time	3 – Extensive mental effort and concentration are necessary: very complex activity requiring total attention	3 – High to very intense stress due to confusion, frustration, or anxiety: high to extreme determination and self-control required

The output of SWAT is a workload score for each of the three SWAT dimensions, time load, mental effort load and stress load. An overall workload score between 1 and 100 is also calculated. Further variations of the SWAT technique have also been developed, including a predictive variation (PRO-SWAT) and a computerised version.

Domain of application
Aviation.

Procedure and advice
Step 1: Scale development
Firstly, participants are required to place in rank order all possible 27 combinations of the three workload dimensions, time load, mental effort load and stress load, according to their effect on workload. This 'conjoint' measurement is used to develop an interval scale of workload rating, from 1 to 100.

Step 2: Performance of Task under analysis
Once the initial SWAT ranking has been completed, the subject should perform the task under analysis. SWAT can be administered during the trial or after the trial. It is recommended that the SWAT is administered after the trial, as on-line administration is intrusive to the primary task. If On-line administration is required, then the SWAT should be administered and completed verbally.

Step 3: SWAT scoring
The participants are required to provide a subjective rating of workload by assigning a value of 1 to 3 to each of the three SWAT workload dimensions.

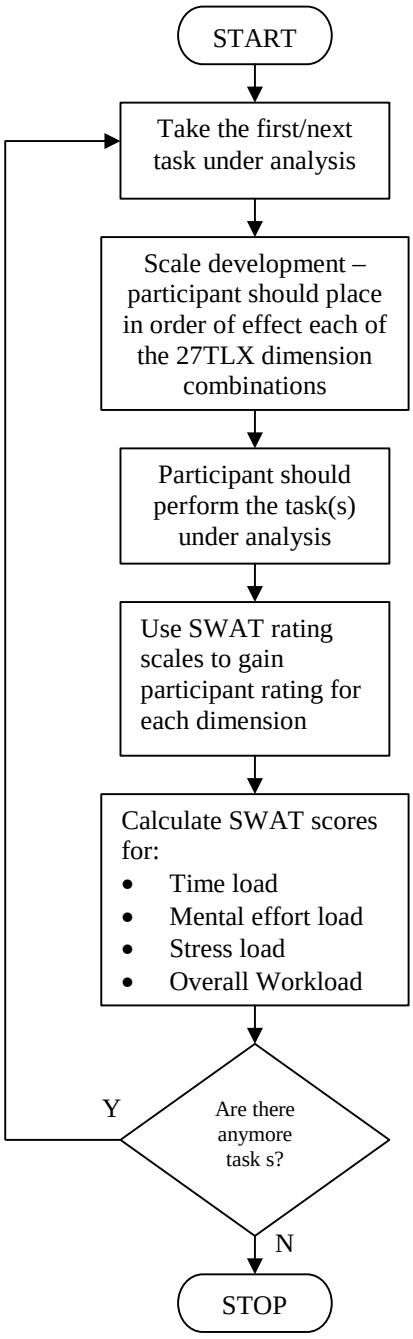
Step 4: SWAT score calculation
For the workload score, the analyst should take the scale value associated with the combination given by the participant. The scores are then translated into individual workload scores for each SWAT dimension. Finally, an overall workload score should be calculated.

Advantages
- The SWAT technique provides a quick and simple technique for estimating operator workload.
- The SWAT workload dimensions are generic, so the technique can be applied to any domain. In the past, the SWAT technique has been used in a number of different domains, such as aviation, air traffic control, command and control, nuclear reprocessing and petro chemical, and automotive domains.
- The SWAT technique is one of the most widely used and well know subjective workload assessment techniques available, and has been subjected to a number of validation studies (Hart & Staveland 1988, Vidulich & Tsang 1985, 1986)
- The Pro-SWAT variation allows the technique to be used predictively.
- SWAT is a multidimensional approach to workload assessment.
- Unobtrusive.

Disadvantages
- SWAT can be intrusive if administered on-line.
- Pro-SWAT has yet to be validated thoroughly.
- In a number of validation studies it has been reported that the NASA TLX is superior to SWAT in terms of sensitivity, particularly for low mental workloads (Hart & Staveland 1988, Hill et al 1992, Nygren 1991).
- SWAT has been constantly criticised for having a low sensitivity for mental workloads (Luximon & Goonetilleke 2001).

- The initial SWAT combination ranking procedure is very time consuming (Luximon & Goonetilleke 2001).
- Workload ratings may be correlated with task performance e.g. subjects who performed poorly on the primary task may rate their workload as very high and vice versa. This is not always the case.
- When administered after the fact, participants may have forgotten high or low workload aspects of the task.
- Unsophisticated measure of workload. NASA TLX appears to be more sensitive.
- The Pro-SWAT technique is still in its infancy.

Flowchart



Related methods
The SWAT technique is similar to a number of subjective workload assessment techniques, such as the NASA TLX, Cooper Harper Scales and Bedford Scales. For predictive use, the Pro-SWORD technique is similar.

Approximate training times and application times
Whilst the scoring phase of the SWAT technique is very simple to use and quick to apply, the initial ranking phase is time consuming and laborious. Thus, the training times and application times are estimated to be quite high.

Reliability and validity
A number of validation studies concerning the SWAT technique have been conducted Hart & Staveland 1988, Vidulich & Tsang 1985, 1986). Vidulich and Tsang (1985, 1986) reported that NASA TLX produced more consistent workload estimates for participants performing the same task than the SWAT (Reid & Nygren 1988) technique did. Luximon & Goonetilleke (2001) also reported that a number of studies have shown that the NASA TLX is superior to SWAT in terms of sensitivity, particularly for low mental workloads (Hart & Staveland 1988, Hill et al 1992, Nygren 1991).

Tools needed
A SWAT analysis can either be conducted using pen and paper. A software version also exists. Both the pen and paper method and the software method can be purchased from various sources.

Bibliography
Cha, D. W (2001) Comparative study of subjective workload assessment techniques for the evaluation of ITS-orientated human-machine interface systems. *Journal of Korean Society of Transportation*. Vol 19 (3), pp 45058
Dean, T. F. (1997) Directory of Design support methods, Defence Technical Information Centre, DTIC-AM. MATRIS Office, ADA 328 375, September.
Hart, S. G., & Staveland, L. E. (1988) Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam. The Netherlands. Elsevier.
Reid, G. B. & Nygren, T. E. (1988) The subjective workload assessment technique: A scaling procedure for measuring mental workload. In P. S. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam. The Netherlands. Elsevier
Vidulich, M. A., & Tsang, P. S. (1985) Assessing subjective workload assessment. A comparison of SWAT and the NASA bipolar methods. *Proceedings of the Human Factors Society 29th Annual Meeting*. Santa Monica, CA: Human Factors Society, pp 71-75.
Vidulich, M. A., & Tsang, P. S. (1986) Collecting NASA Workload Ratings. Moffett Field, CA. NASA Ames Research Center.
Vidulich, M. A., & Tsang, P. S. (1986) Technique of subjective workload assessment: A comparison of SWAT and the NASA bipolar method. *Ergonomics*, 29 (11), 1385-1398.

SWORD – Subjective Workload Dominance Technique

Dr Michael A. Vidulich, Department of Psychology, Wright State University, 3640 Colonel Glen Hwy, Dayton OH 45435-0001.

Background and applications

The Subjective Workload Dominance Technique (SWORD) is a subjective workload assessment technique that has been used both retrospectively and predictively (Pro-SWORD) (Vidulich, Ward & Schueren 1991). Originally designed as a retrospective workload assessment technique, SWORD uses paired comparison of tasks in order to provide a rating of workload for each individual task. Administered post trial, participants are required to rate one tasks dominance over another in terms of workload imposed. When used predictively, tasks are rated for their dominance before the trial begins, and then rated post-test to check for the sensitivity of the predictions.

Domain of application

Generic.

Procedure and advice – Workload assessment

The procedure outlined below is the procedure recommended for an assessment of operator workload. In order to predict operator workload, it is recommended that SME's are employed to predict workload for the task under analysis before step 3 in the procedure below. The task should then be performed and operator workload ratings obtained using the SWORD technique. The predicted workload ratings should then be compared to the subjective ratings in order to calculate the sensitivity of the workload predictions made.

Step 1: Task description

The first step in any SWORD analysis is to create a task or scenario description of the scenario under analysis. Each task should be described individually in order to allow the creation of the SWORD rating sheet. Any task description can be used for this step, such as HTA or tabular task analysis.

Step 2: Create SWORD rating sheet

Once a task description (e.g. HTA) is developed, the SWORD rating sheet can be created. The analyst should list all of the possible combinations of tasks (e.g. AvB, AvC, BvC) and the dominance rating scale. An example of a SWORD dominance rating sheet is shown in table XX.

Step 3: Performance of task

SWORD is normally applied post-task. Therefore, the task under analysis should be performed first. As SWORD is applied after the task performance, intrusiveness is reduced and the task under analysis can be performed in its real world setting.

Step 4: Administration of SWORD questionnaire

Once the task under analysis is complete, the SWORD data collection process begins. This involves the administration of the SWORD rating sheet. The participant should be presented with the SWORD rating sheet immediately after task performance has ended. The SWORD rating sheet lists all possible paired comparisons of the tasks conducted in the scenario under analysis. A 17 point rating scale is used.

Table 39 – Example SWORD rating sheet

Task	Absolute	Very Strong	Strong	Weak	EQUAL	Weak	Strong	Very Strong	Absolute	Task
A										B
A										C
A										D
A										E
B										C
B										D
B										E
C										D
C										E
D										E

The 17 slots represent the possible ratings. The analyst has to rate the two tasks (e.g. task A v's B) in terms of their level of workload imposed, against each other. For example, if the participant feels that the two tasks imposed a similar level of workload, then they should mark the 'EQUAL' point on the rating sheet. However, if the participant feels that task A imposed a slightly higher level of workload than task B did, they would move towards task A on the sheet and mark the 'weak' point on the rating sheet. If the participant felt that task A imposed a much greater level of workload than task B, then they would move towards task A on the sheet and mark the 'Absolute' point on the rating sheet. This allows the participant to provide a subjective rating of one tasks workload dominance over the over. This procedure should continue until all of the possible combinations of tasks in the scenario under analysis are exhausted and given a rating.

Step 5: Constructing the judgement matrix
Once all ratings have been elicited, the SWORD judgement matrix should be conducted. Each cell in the matrix should represent the comparison of the task in the row with the task in the associated column. The analyst should fill each cell with the participant's dominance rating. For example, if a participant rated tasks A and B as equal, a '1' is entered into the appropriate cell. If task A is rated as dominant, then the analyst simply counts from the 'Equal' point to the marked point on the sheet, and enters the number in the appropriate cell. An example SWORD judgment matrix is shown below.

Table 40 – example SWORD matrix

	A	B	C	D	E
A	**1**	**2**	**6**	**1**	**1**
B	**-**	**1**	**3**	**2**	**2**
C	**-**	**-**	**1**	**6**	**6**
D	**-**	**-**	**-**	**1**	**1**
E	**-**	**-**	**-**	**-**	**1**

The rating for each task is calculated by determining the mean for each row of the matrix and then normalising the means (Vidulich, Ward & Schueren 1991). In the example shown in figure XX the

Step 64: Matrix consistency evaluation
Once the SWORD matrix is complete, the consistency of the matrix can be evaluated by ensuring that there are transitive trends amongst the related judgements in the matrix. For example, if task A is rated twice as hard as task B, and task B is rated 3

times as hard as task C, then task A should be rated as 6 times as hard as task C (Vidulich, Ward & Schueren 1991). Therefore the analyst should use the completed SWORD matrix to check the consistency of the participant's ratings.

Advantages

- Easy to learn and use.
- Non intrusive
- High face validity
- SWORD has been demonstrated to have a sensitivity to workload variations (Ried and Nygren 1988)
- Very quick in its application.

Disadvantages

- Data is collected post task.
- SWORD is a dated approach to workload assessment.
- Workload projections are more accurate when domain experts are used.
- Further validation is required.
- The SWORD technique has not been as widely used as other workload assessment techniques, such as SWAT, MCH and the NASA TLX.

Example

Vidulich, Ward & Schueren (1991) tested the SWORD technique for its accuracy in predicting the workload imposed upon F-16 pilots by a new HUD attitude display system. Participants included F-16 pilots and college students and were divided into two groups. The first group (F-16 pilots experienced with the new HUD display) retrospectively rated the tasks using the traditional SWORD technique, whilst the second group (F-16 pilots who had no experience of the new HUD display) used the Pro-SWORD variation to predict the workload associated with the HUD tasks. A third group (college students with no experience of the HUD) also used the Pro-SWORD technique to predict the associated workload. In conclusion, it was reported that the pilot Pro-SWORD ratings correlated highly with the pilot SWORD (retrospective) ratings (Vidulich, Ward & Schueren 1991). Furthermore, the Pro-SWORD ratings correctly anticipated the recommendations made in an evaluation of the HUD system. Vidulich and Tsang (1987) also reported that the SWORD technique was more reliable and sensitive than the NASA TLX technique.

Related methods

SWORD is one of a number of mental workload assessment techniques, including the NASA-TLX, SWAT, MCH and DRAWS. A number of the technique have also been used predictively, such as Pro-SWAT and MCH. Any SWORD analysis requires a task description of some sort, such as HTA or a tabular task analysis.

Approximate training and application times

Although no data is offered regarding the training and application times for the SWORD technique, it is apparent that the training time for such a simple technique would minimal. The application time associated with the SWORD technique would be based upon the scenario under analysis. For large, complex scenario's involving a great number of tasks, the application time would be high as an initial HTA would have to be performed, then the scenario would have to performed, and then the

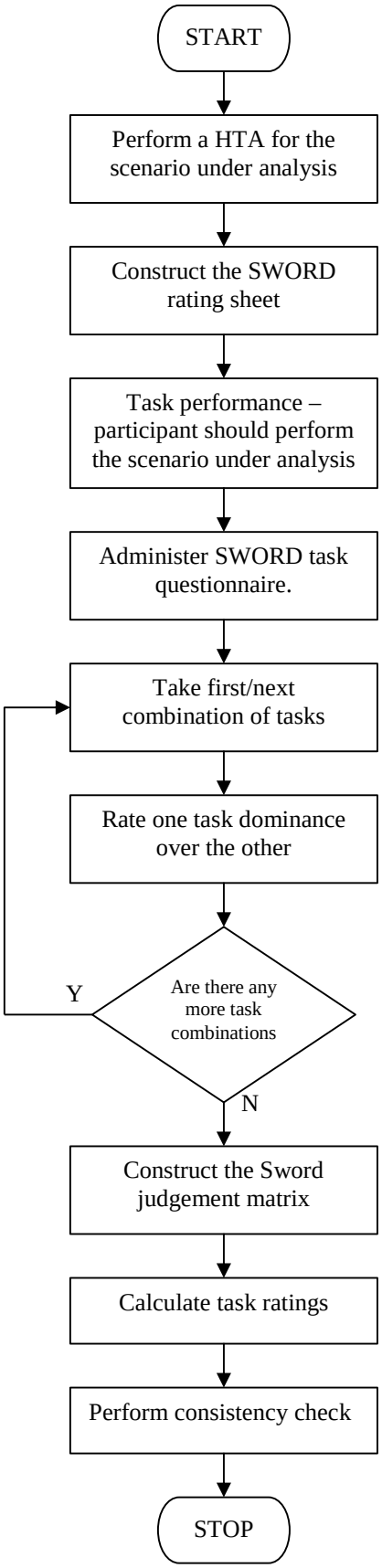
SWORD technique. The actual application time associated purely the administration of the SWORD technique is very low.

Reliability and validity
Vidulich, Ward & Schueren (1991) tested the SWORD technique for its accuracy in predicting the workload imposed upon F-16 pilots by a new HUD attitude display system. In conclusion, it was reported that the pilot Pro-SWORD ratings correlated highly with the pilot SWORD (retrospective) ratings (Vidulich, Ward & Schueren 1991). Furthermore, the Pro-SWORD ratings correctly anticipated the recommendations made in an evaluation of the HUD system. Vidulich and Tsang (1987) also reported that the SWORD technique was more reliable and sensitive than the NASA TLX technique.

Tools needed
The SWORD technique can be applied using pen and paper. Of course, the system or device under analysis is also required.

Bibliography
Vidulich, M. A. (1989) The use of judgement matrices in subjective workload assessment: The subjective WORkload Dominance (SWORD) technique. In proceedings of the Human Factors Society 33rd Annual Meeting (pp. 1406-1410). Santa Monica, CA: Human Factors Society.
Vidulich, M. A., Ward, G. F., & Schueren, J. (1991) Using Subjective Workload Dominance (SWORD) technique for Projective Workload Assessment. Human Factors, 33, Vol 6, pp 677-691.

Flowchart


START
Perform a HTA for the scenario under analysis
Construct the SWORD rating sheet
Task performance – participant should perform the scenario under analysis
Administer SWORD task questionnaire.
Take first/next combination of tasks
Rate one task dominance over the other
Are there any more task combinations
Y
N
Construct the Sword judgement matrix
Calculate task ratings
Perform consistency check
STOP

Bibliography

Ainsworth, L. & Marshall, E. (1998) Issues of quality and practicality in task analysis: preliminary results from two surveys. Ergonomics 41(11), 1604-1617. Reprinted in J. Annett & N..A. Stanton (2000) op.cit. Pp. 79-89.

Annett, J. (In Press) Hierarchical Task Analysis (HTA). In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.

Annett, J., Duncan, K.D., Stammers, R.B., & Gray, M. (1971) Task Analysis. London: HMSO.

Annett, J., Cunningham, D.J., & Mathias-Jones, P. (2000). A method for measuring team skills. Ergonomics, 43(8), 1076-1094.

Annett, J. & Stanton, N.A. (Eds) (2000) Task Analysis. London: Taylor & Francis.

Baber, C. & Stanton, N.A. (1994) Task analysis for error identification. Ergonomics 37, 1923-1941.

Baber, C. & Stanton, N. A. (1996). Human error identification techniques applied to public technology: predictions compared with observed use. *Applied Ergonomics*, 27(2), 119-131.

Baber, C. & Stanton, N. A. (1996) Observation as a technique for Usability Evaluations. In P. Jordan et al (eds.), *Usability in Industry*, pp 85-94. London, Taylor and Francis.

Baber, C. & Stanton, N. A. (1999) Analytical prototyping. In: J. M. Noyes & M. Cook (eds) Interface Technology: the leading edge. Research Studies Press: Baldock.

Baber, C. and Mellor, B.A. (2001) Modelling multimodal human-computer interaction using critical path analysis, International Journal of Human Computer Studies 54 613-636

Baber, C. and Stanton, N. A. (2002) Task Analysis For Error Identification: theory, method and validation. Theoretical Issues in Ergonomics Science 3 (2), 212-227.
Bartlett, F.C. (1932) Remembering: a study in experimental and social psychology, Cambridge: Cambridge University Press

Bass, A., Aspinal, J., Walter, G., & Stanton, N. A. (1995) A software toolkit for hierarchical task analysis. Applied Ergonomics. 26 (2) pp. 147-151.

Bowers, C. A., Morgan, B. B., Salas, E., & Prince, C. (1993) Assessment of co-ordination demand for aircrew co-ordination training. *Military Psychology*, 5(2), 95-112.

Bowers, C. A., Baker, D. P., & Salas, E. (1994). Measuring the importance of teamwork: The reliability and validity of job/task analysis indices for team training design. *Military Psychology*, 6(4), 205-214

Brewer, W.F. (2000) Bartlett's concept of the schema and its impact on theories of knowledge representation in contemporary cognitive psychology, In A. Saito (ed) Bartlett, Culture and Cognition, London: Psychology Press, 69-89

Burford, B. (1993) Designing Adaptive ATMs, Birmingham: University of Birmingham unpublished MSc Thesis

Burke, C. S. (In Press) Team Task Analysis. In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.

Card, S. K., Moran, T. P.; Newell, A. (1983) The Psychology of Human Computer Interaction. Lawrence Erlbaum Associates.

Center for Chemical Process Safety (CCPS) (1994) Guidelines for Preventing Human Error in Process Safety. New York: American Institute of Chemical Engineers. In W. Karwowski & W. S Marras (eds) The Occupational Ergonomics Handbook. (1999) Florida, CRC Press LLC.

Cha, D. W (2001) Comparative study of subjective workload assessment techniques for the evaluation of ITS-orientated human-machine interface systems. *Journal of Korean Society of Transportation*. Vol 19 (3), pp 45058

Ciavarelli, A. (2002) Human Factors Checklist: An Aircraft Accident Investigation Tool, School of Aviation Safety, California. Internet source

Crawford, J. O.; Taylor, C. & Po, N. L. W. (2001) A case study of on-screen prototypes and usability evaluation of electronic timers and food menu systems. International Journal of Human Computer Interaction 13 (2), 187-201.

Dean, T. F. (1997) Directory of Design support methods, Defence Technical Information Centre, DTIC-AM. MATRIS Office, ADA 328 375, September.

Driskell, J. E. & Mullen, B. (In Press) Social Network Analysis. In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.

Drury, C. G. (1990) Methods for direct observation of performance, In Wilson, J. R. And Corlett, E. N (Eds) 'Evaluation of Human Work – A practical ergonomics methodology' Taylor and Francis, London.

Easterby, R. (1984) Tasks, processes and display design. In R. Easterby & H. Zwaga (Eds.), *Information Design* (pp. 19-36). Chichester: Whiley.

Embrey, D. E. (1986). SHERPA: A systematic human error reduction and prediction approach. Paper presented at the International Meeting on Advances in Nuclear Power Systems, Knoxville, Tennessee.

Embrey, D. E. (1993). Quantitative and qualitative prediction of human error in safety assessments. Institute of Chemical Engineers Symposium Series, 130, 329-350.

Flanagan, J. C. (1954). The Critical Incident Technique. *Psychological Bulletin*, 51, 327-358.

Gertmann, D. I., Blackman, H. S., Haney, L. N., Seidler, K. S., and Hahn, H. A. (1992) INTENT: A method for estimating human error probabilities for decision based errors. *Reliability engineering and system safety*, 35, 127-137. In W. Karwowski & W. S Marras (eds) The Occupational Ergonomics Handbook. (1999) Florida, CRC Press LLC.

Gertmann, D. I., Blackman, H. S. (1994) Human Reliability and Safety Analysis Data Handbook. New York: John Wiley & Sons. In W. Karwowski & W. S Marras (eds) The Occupational Ergonomics Handbook. (1999) Florida, CRC Press LLC.

Glendon, A.I. and McKenna, E.F. (1995) Human Safety and Risk Management. London: Chapman and Hall

Go, K., & Carroll, J., M. (2003) Scenario-Based Task Analysis. In D. Diaper & N. Stanton (Eds) *The Handbook of Task Analysis for Human-Computer Interaction, London*, Lawrence Erlbaum Associates.

Gray, W.D., John, B.E. and Atwood, M.E. (1993) Project Ernestine: validating a GOMS analysis for predicting and explaining real-world performance *Human-Computer Interaction 8* 237-309

Greenberg, S. (2003) Working through Task-Centred System Design. In D. Diaper & N. Stanton (Eds) The Handbook of Task Analysis for Human Computer Interaction. Lawrence Erlbaum Associates Inc.

Hahn, A. H. and DeVries, J. A. (1991) Identification of Human Errors of Commission Using SNEAK analysis, Proceedings of the Human Factors Society 35th Annual Meeting, pp 1080-1084. In W. Karwowski & W. S Marras (eds) The Occupational Ergonomics Handbook. (1999) Florida, CRC Press LLC.

Harrison, A. (1997) *A Survival Guide to Critical Path Analysis* London: Butterworth-Heinemann

Hart, S. G., & Staveland, L. E. (1988) Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam. The Netherlands. Elsevier.

Hollnagel, E. (1993). Human Reliability Analysis: context and control. London: Academic Press.

Hollnagel, E. (1998) Cognitive Reliability and Error Analysis Method – CREAM, 1st Edition. Elsevier Science, Oxford, England

Hone, K.S. and Baber, C. (1999) Modelling the effect of constraint on speech-based human computer interaction *International Journal of Human Computer Studies 50* 85-105

Hyponen, H. (1999) Focus Groups. In H. A. Williams, J. Bound & R. Coleman (eds) *The Methods Lab: User Research for Design*. Design for Ageing Network.

Isaac, A., Shorrick, S.T., Kirwan, B., (2002) Human Error in European air traffic management: The HERA project. *Reliability Engineering and System Safety*, Vol. 75 pp 257-272

ISO 9241 (1998) *Ergonomics of office work with VDTs-guidance on usability*, Geneva: International Standards Office

ISO 13407 (1999) *Human-centred design processes for interactive systems*, Geneva: International Standards Office

ISO 9126 (2000) *Software engineering - product quality*, Geneva: International Standards Office

Johnson, P., Diaper, D. & Long, J. 1984. Tasks, skills and knowledge: Task analysis for knowledge-based descriptions. In B. Shackel (Ed.) Interact '84 - First IFIP Conference on Human-Computer Interaction. Amsterdam: Elsevier. Pp. 23-27.

Karwowski, W., & Marras, W. S. (1999) The Occupational Ergonomics Handbook. Florida, CRC Press LLC.

Kennedy, R., & Kirwan, B. (1998) Development of a Hazard and Operability-based method for identifying safety management vulnerabilities in high risk systems, *Safety Science*, Vol. 30, Pages 249-274

Kieras, D. (2003) GOMS Models for Task Analysis. In D. Diaper & N. Stanton (Eds) *The Handbook of Task Analysis for Human-Computer Interaction*. Pp 83-117. Lawrence Erlbaum Associates.

Kim, I. S. (2001) "Human reliability analysis in the man-machine interface design review" *Annals of Nuclear Energy,* 28, 1069 – 1081

Kirwan, B. (1990). Human reliability assessment. In J. R. Wilson & E. N. Corlett (eds.), Evaluation of human work: a practical ergonomics methodology (2nd ed., pp. 921-968). London: Taylor & Francis.

Kirwan, B., Ainsworth, L. K. (1992) A guide to Task Analysis, Taylor and Francis, London, UK.

Kirwan, B. (1992a) Human error identification in human reliability assessment. Part 1: Overview of approaches. *Applied Ergonomics* Vol. 23(5), 299 – 318

Kirwan, B. (1992b). Human error identification in human reliability assessment. Part 2: detailed comparison of techniques. Applied Ergonomics, 23, 371-381.

Kirwan, B. (1994) A guide to Practical Human Reliability Assessment. Taylor and Francis, London.

Kirwan, B. (1996) Human Error Recovery and Assessment (HERA) Guide. Project IMC/GNSR/HF/5011, Industrial Ergonomics Group, School of Manufacturing Engineering, University of Birmingham, March.

Kirwan, B. (1996) "The validation of three Human Reliability Quantification techniques – THERP, HEART and JHEDI: Part 1 – technique descriptions and validation issues" *Applied Ergonomics,* Vol 27, 6, pp 359 – 373

Kirwan, B. (1997) "The validation of three Human Reliability Quantification techniques – THERP, HEART and JHEDI: Part 2 – Results of validation exercise" *Applied Ergonomics,* Vol 28, 1, pp 17 – 25

Kirwan, B. (1997) "The validation of three Human Reliability Quantification techniques – THERP, HEART and JHEDI: Part 3 – Practical aspects of the usage of the techniques" *Applied Ergonomics,* Vol 28, 1, pp 27 – 39

Kirwan, B. (1998a) Human error identification techniques for risk assessment of high-risk systems – Part 1: Review and evaluation of techniques. *Applied Ergonomics,* 29, pp157-177

Kirwan, B. (1998b) Human error identification techniques for risk assessment of high-risk systems – Part 2: Towards a framework approach. *Applied Ergonomics,* 5, pp299 – 319

Klein, G. A., Calderwood, R., & MacGregor, D. (1989) Critical Decision Method for Eliciting Knowledge. *IEEE Transactions on Systems,* Man and Cybernetics, 19(3), 462-472.

Klein, G. & Armstrong, A. A. (In Press) Critical Decision Method. In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.

Lawrence, D., Atwood, M.E., Dews, S. and Turner, T. (1995) Social interaction in the use and design of a workstation: two contexts of interaction In P.J. Thomas (ed) *The Social and Interactional Dimensions of Human-Computer Interaction* Cambridge: Cambridge University Press 240-260.

Lewis, C., & Reiman, J. (1993). Task centred user interface design: A practical introduction. Boulder, CO: University of Colorado. Shareware book available from ftp.cs.colorado.edu/pub/cs/distribs/clewis/HCI-Design-Book

Lewis, C., Polson, P., Wharton, C. & Rieman, J. (1990) Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces. In Proceedings of CHI'90 conference on Human Factors in Computer Systems, pp 235-241, New York: Association for Computer Machinery

Lim, K.Y. & Long, J. (1994) The MUSE Method for Usability Engineering. Cambridge: Cambridge University Press.

Lockyer, K. and Gordon, J. (1991) *Critical Path Analysis and Other Project Network Techniques* London: Pitman

Marshall, A., Stanton, N., Young, M., Salmon, P., Harris, D., Demagalski, J., Waldmann, T., Dekker, S. (2003) Development of the Human Error Template – A new methodology for assessing design induced errors on aircraft flight decks.

Militello, L. G. & Militello, J. B. (2000) Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. In J. Annett & N. S Stanton (Eds) *Task Analysis*, pp 90-113. UK, Taylor & Francis

Neerincx, M. A. (2003) Cognitive Task Load Analysis: Allocating Tasks and Designing Support. In E. Hollnagel (ed) *Handbook of Cognitive Task Design*. Pp 281-305. Lawrence Erlbaum Associates Inc.

Nielsen, J. & Molich, R. (1990) Heuristic evaluation of user interfaces. In J. C Chew & J. Whiteside (eds), *Empowering people: CHI 90 Conference Proceedings* (pp. 249 – 256) Monterey, CA: ACM Press

O'Hare, D., Wiggins, M., Williams, A., & Wong, W. (2000) Cognitive task analyses for decision centred design and training. In J. Annett & N. Stanton (eds) *Task Analysis*, pp 170-190. UK, Taylor and Francis

Olson, J.R. and Olson, G.M. (1990) The growth of cognitive modelling in human-computer interaction since GOMS *Human-Computer Interaction 3* 309-350

Ormerod, T.C., Richardson, J. & Shepherd, A. (1998) Enhancing the usability of a task analysis method: A notation and environment for requirements. Ergonomics 41(11), 1642-1663. Reprinted in Annett & Stanton (2000) op.cit. Pp. 114-135.

Patrick, J., Gregov, A. & Halliday, P. (2000) Analysing and training task analysis. Instructional Science, 28(4), 51-79.

Pennycook, W. A., & Embrey, D. E. (1993) An operating approach to error analysis, in Proceedings of the first Biennial Canadian Conference on Process Safety and Loss Managament. Edmonton, Alberta, Canada. Waterloo, Ontario, Canada: Institute for Risk Research, Univeristy of Waterloo. In W. Karwowski & W. S Marras (eds) The Occupational Ergonomics Handbook. (1999) Florida, CRC Press LLC.

Pocock, S., Harrison, M., Wright, P., Fields, B. (2001) THEA – A Reference Guide. *Unpublished work*

Pocock, S., Harrison, M., Wright, P., Johnoson, P. (2001) THEA: A technique for Human Error Assessment Early in Design, *unpublished work*
Polson, P. G., Lewis, C., Riemant, J. & Wharton, C. (1992) Cognitive walkthroughs: a method for theory based evaluation of user interfaces. International Journal of Man-Machine Studies, *36 pp741-773.*

Ravden, S. J. And Johnson, G. I. (1989) Evaluating Usability of Human-Computer Interfaces: A practical method, Chirchester, Ellis Horwood.

Reason, J. (1990) Human Error. Cambridge, Cambridge University Press.

Reid, G. B. & Nygren, T. E. (1988) The subjective workload assessment technique: A scaling procedure for measuring mental workload. In P. S. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam. The Netherlands. Elsevier

Salmon, P., Stanton, N. A., Young, M. S., Harris, D., Demagalski, J., Marshall, A., Waldman, T., Dekker, S. W. A. (2002). Using existing HEI techniques to predict pilot error: A comparison of SHERPA, HAZOP, and HEIST. Proceedings of HCI-Aero 2002, MIT, Cambridge, MA.

Salmon, P. M, Stanton, N.A, Young, M.S, Harris, D, Demagalski, J, Marshall, A, Waldmann, T, Dekker, S (2003) Predicting Design Induced Pilot Error: A comparison of SHERPA, Human Error HAZOP, HEIST and HET, a newly developed aviation specific HEI method, In D. Harris, V. Duffy, M. Smith, C. Stephanidis (eds) *Human-Centred Computing – Cognitive, Social and Ergonomic Aspects*, Lawrence Erlbaum Associates, London

Sanders, M. S., & McCormick, E. J. (1993) Human Factors in Engineering and Design. UK, McGraw-Hill Publications.

Salvendy, G. (1997) Handbook of human factors and ergonomics, 2nd edition, John Wiley and Sons, Canada.

Seigal, A. I., Bartter, W. D., Wolf, J. J., Knee, H. E., & Haas, P. M. (1984) Maintenance Personnel Performance Simulation (MAPPS) Model: Summary Description. NUREG/CR-3626, U.S Nuclear Regulatory Commission, Washington, D.C.

Shepherd, A. (2002). Hierarchical Task Analysis. London: Taylor & Francis.
Ainsworth, W. (1988) Optimization of string length for spoken digit input with error Correction. International Journal of Man Machine Studies 28 573-581

Shorrock, S.T., Kirwan, B., (1999) The development of TRACEr: a technique for the retrospective analysis of cognitive errors in ATC. In: Harris, D. (Ed), *Engineering Psychology and Cognitive Ergonomics*, Vol. 3, Aldershot, UK, Ashgate Publishing.

Shorrock, S.T., Kirwan, B., (2000) Development and application of a human error identification tool for air traffic control. *Applied Ergonomics*, Vol. 33 pp319-336

Stanton, N. A. (1995). Analysing worker activity: a new approach to risk assessment? Health and Safety Bulletin, 240, (December), 9-11.

Stanton, N. A. (2002) Human error identification in human computer interaction. In: J. Jacko and A. Sears (eds) The Human-Computer Interaction Handbook. New Jersey: Lawrence Erlbaum Associates.

Stanton, N. A. (2002) Human error identification in human computer interaction. In: J. Jacko and A. Sears (eds) The Human Computer Interaction Handbook. (pp. 371-383) Mahwah, NJ: Lawrence Erlbaum Associates.

Stanton, N. A., & Baber, C. (1996). A systems approach to human error identification. Safety Science, 22, 215-228.

Stanton, N. A., & Baber, C. (1996). Task analysis for error identification: applying HEI to product design and evaluation. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland (eds.), Usability Evaluation in Industry (pp. 215-224). London: Taylor & Francis.

Stanton, N. A., & Baber, C. (1998). A systems analysis of consumer products. In N. A. Stanton (ed.), Human factors in consumer products (pp. 75-90). London: Taylor & Francis.

Stanton, N. A. & Stevenage, S. V. (1998). Learning to predict human error: issues of acceptability, reliability and validity. Ergonomics, 41(11), 1737-1756.

Stanton, N. A. & Young, M. (1998) Is utility in the mind of the beholder? A review of ergonomics methods. Applied Ergonomics. 29 (1) 41-54

Stanton, N. A. & Young, M. (1999) What price ergonomics? Nature 399, 197-198 methodology for designing error-tolerant consumer products. Ergonomics, 37, 1923-1941.

Stanton, N. A, and Young, M. S. (1999) A guide to methodology in ergonomics: Designing for human use, London, Taylor and Francis.

Stanton, N. A. & Wilson, J. (2000) Human Factors: step change improvements in effectiveness and safety. Drilling Contractor, Jan/Feb, 46-41.

Stanton, N. A. & Baber, C. (2002) Error by design: methods to predict device usability. Design Studies, 23 (4), 363-384.

Stanton, N. A., & Baber, C. (2002) Error by design. Design Studies, 23 (4), 363-384.

Swann, C. D. and Preston, M. L. (1995) Twenty-five years of HAZOPs. *Journal of Loss Prevention in the Process Industries*, Vol. 8, Issue 6, 1995, Pages 349-353.

Van Welie, M., & Van Der Veer, G. (2003) Groupware Task Analysis. In E. Hollnagel (Ed) *Handbook of Cognitive Task Design*. Pp 447 – 477. Lawrence Erlbaum Associates Inc.

Verplank, B., Fulton, J., Black, A., Moggridge, B. (1993) Observation and Invention – Use of scenarios in interaction design. Tutorial notes for InterCHI 93. Amsterdam.

Vidulich, M. A., & Tsang, P. S. (1985) Assessing subjective workload assessment. A comparison of SWAT and the NASA bipolar methods. *Proceedings of the Human Factors Society 29th Annual Meeting*. Santa Monica, CA: Human Factors Society, pp 71-75.

Vidulich, M. A., & Tsang, P. S. (1986) Collecting NASA Workload Ratings. Moffett Field, CA. NASA Ames Research Center.

Vidulich, M. A., & Tsang, P. S. (1986) Technique of subjective workload assessment: A comparison of SWAT and the NASA bipolar method. *Ergonomics*, 29 (11), 1385-1398.

Vidulich, M. A. (1989) The use of judgement matrices in subjective workload assessment: The subjective WORkload Dominance (SWORD) technique. In proceedings of the Human Factors Society 33rd Annual Meeting (pp. 1406-1410). Santa Monica, CA: Human Factors Society.

Vidulich, M. A., Ward, G. F., & Schueren, J. (1991) Using Subjective Workload Dominance (SWORD) technique for Projective Workload Assessment. Human Factors, 33, Vol 6, pp 677-691.

Walker, G. H (In Press) Verbal Protocol Analysis. In N. A. Stanton, A. Hedge, K, Brookhuis, E. Salas, & H. Hendrick. (In Press) (eds) *Handbook of Human Factors methods*. UK, Taylor and Francis.

Walker, G.H., Stanton, N.A., & Young, M.S. (2001). An on-road investigation of vehicle feedback and its role in driver cognition: Implications for cognitive ergonomics. *International Journal of Cognitive Ergonomics*, 5(4), 421-444

Watts, L. A., & Monk, A. F. (2000) Reasoning about tasks, activities and technology to support collaboration. In J. Annett & N. Stanton (Eds) Task Analysis. UK, Taylor and Francis.

Weber, R.P. (1990). Basic Content Analysis. Sage Publications, London.

Whalley (1988) Minimising the cause of human error. In B. Kirwan & L. K. Ainsworth (eds.) *A Guide to Task Analysis*. Taylor and Francis, London.

Wharton, C., Rieman, J., Lewis, C., Polson, P. The Cognitive Walkthrough Method: A Practitioners Guide

Williams, H., A., Bound, J., & Coleman, R. (1999) the Methods Lab: User Research for Design. Design for Ageing network, London, UK.

Williams, J. C. (1986) HEART – a proposed method for assessing and reducing human error. In 9th Advances in Reliability Technology Symposium, University of Bradford.

Wilson, J.R., & Corlett, N.E. (1995). Evaluation of Human Work: A Practical Ergonomics Methodology. Taylor and Francis, London.

Yamaoka, T. and Baber, C. (2000) 3 point task analysis and human error estimation, Proceedings of the Human Interface Symposium 2000, Tokyo, Japan, 395-398

Appendix 1: HF methods database

HEI/HRA techniques

Table 1. HEI/HRA techniques

Method	Author/Source
AIPA – Accident Investigation and Progression Analysis	Fleming et al (1975)
APJ – Absolute Probability Judgement	Kirwan (1994)
ASEP – Accident Sequence Evaluation Programme	Swain (1987)
ATHEANA – A Technique for Human Error Analysis	Cooper et al (1996) Hollnagel (1998)
CADA – Critical Action Decision Approach	Gall (1990) Kirwan (1992, 1994)
CAMEO/TAT	Fujuta et al (1995) Kirwan (1988)
Confusion Matrice Analysis	Potash et al (1981)
COMET – Commission Event trees	Blackman (1991)
COSIMO – Cognitive Simulation Model	Kirwan (1998a)
CREAM – Cognitive Reliability and Error Analysis Method	Hollnagel (1999)
DYLAM – Dynamic Logical Analysing Methodology	Kirwan (1998a)
EOCA – Error of Commission Analysis	Kirwan (1994, 1998a)
FMEA – Failure Modes and Effects Analysis	Kirwan & Ainsworth (1992)
GASET – Generic Accident Sequence Event Tree	Kirwan (1994)
GEMS – Generic Error Modelling System	Reason (1990)
HET – Human Error template	Marshall et al (2003) Salmon et al (2002, 2003)
HAZOP – Hazard and Operability Study	Swann & Preston (1995)
HCR – Human Cognitive Reliability	Hannaman et al (1984)
HEART – Human Error Assessment Rate Technique	Williams (1986)
HECA – Human Error Criticality Analysis	Karwowski (2000)
HEIST – Human Error Identification in Systems Tool	Kirwan (1994)
HERA – Human Error and Recovery Assessment System	Kirwan (1998)
HMECA – Human Error Mode, Effect and Criticality Analysis	Kirwan (1992a)
HFAM – Human Factors Analysis Methodology	Pennycook & Embrey (1993)
Hit-Line – Human Interaction Timeline	Macwan & Mosleh (1994)
Human Error HAZOP	Whalley (1988) Kirwan and Ainsworth (1992)
HRMS – Human Reliability Management System	Kirwan (1994)
IMAS – Influence Modelling and Assessment System	Embrey (1986)
INTENT	Gertmann et al (1992)
INTEROPS – Integrated Reactor Operator System	Kirwan (1998a)
JHEDI – Justification of Human Error Data Information	Kirwan (1990b)
MAPPS – Maintenance Personnel Performance Simulation	Seigal et al (1984)
MEDA – Maintenance Error Decision Aid	Eurocontrol website
Murphy Diagrams	Pew et al (1981) Kirwan (1994)
Paired Comparisons	Kirwan (1994)
PHEA – Predictive Human Error Analysis	Embrey (1986) Stanton & Young (1999)

PHECA – Potential Human Error Cause Analysis	Whalley (1988)
SAINT – Systems Analysis of Integrated Networks of Tasks	Kirwan (1994)
SCHAZOP	Kirwan & Kennedy (1996a)
SCHEMA – Systematic Critical Human Error Management Approach	Livingston et al (1992)
SHARP – Systematic Human Action Reliability Procedure	Spurgeon et al (1987) Karwowski (2000)
SHERPA – Systematic Human Error Reduction Approach	Embrey (1986) Stanton & Young (1999)
SLIM MAUD	Embrey et al (1984) Kirwan (1992, 1994)
SNEAK	Hahn and de Vries (1991)
SPEAR – System for Predictive Error Analysis and Reduction	CCPS (1993) Karwowski (2000)
SRK Framework	Rasmussen et al (1981) Kirwan (1992a)
STAHR – Socio-Technical Assessment of Human Reliability	Hollnagel (1998) Phillips et al (1983)
TAFEI – Task Analysis for Error Identification	Baber and Stanton (1991, 1996)
TALENT – Task Analysis Linked Evaluation Technique	Kirwan (1998a) Ryan (1988)
THEA – Technique for Human Error Assessment Early in Design	Pocock, Harrison, Wright & Johnson
THERP – Technique for Human Error Rate Prediction	Swain and Guttman (1983)
TOPPE – Team Operations Performance and Procedure Evaluation	Kirwan (1998a)
TRACer - Technique for the Retrospective and Predictive Analysis of Cognitive Errors in ATC	Kirwan and Shorrock (2002)

Task Analysis techniques

Table 2. Task Analysis techniques

Method	**Author/Source**
ACTA – Applied Cognitive Task Analysis	Miltello & Hutton (2000) Annett & Stanton (2000)
CPA – Critical Path Analysis	Baber (1998)
Critical Incident Technique	Flanagan (1954)
Critical Decision Method	Klein (2003)
Cognitive Task Analysis	Klein (2003)
Cognitive Walkthrough technique	Pocock et al (1992)
GOMS – Goals, Operators, Methods and Selection Rules	Card, Newell & Moran (1983)
HTA – Hierarchical Task Analysis	Annett and Duncan (1971) Annett, Duncan and Stammers (1971)
HTA(T)	
Integrated Task Analysis	
TAKD – Task Analysis for Knowledge Descriptios	Diaper
TAG – Task Action Grammers	Karwowski (1999)
Tabular scenario analysis	
Task Decomposition	Kirwan & Ainsworth (1992)
TKS – Task Knowledge Structure	Karwowski (1999)
TTA – Tabular Task Analysis	Human Factors Integration in Future ATM systems – Methods and Tools
TTRAM – Task and training requirements analysis methodology	Swezey et al (2000)
User needs Task Analysis	Wilson and Corlett (1999)

Data collection techniques

Table 3. Data collection techniques

Method	Author/Source
Interviews	Various
Questionnaires	Various
Observational analysis	Various
Talkthrough	Kirwan & Ainsworth (1992)
Verbal Protocol analysis	Walker (In Press)
Walkthrough	Kirwan & Ainsworth (1992)

Situation Awareness measurement techniques

Table 4. Situation Awareness measurement techniques

Method	Author/Source
SAGAT – Situation Awareness Global Assessment Technique	Endsley (1995)
SA-SWORD – Subjective Workload Dominance metric	Vidulich (1989)
SARS – Situation Awareness rating Scales	
SART – Situation Awareness Rating Technique	Taylor (1990)
SALSA	Hauss and Eyferth (2002)
SABARS – Situation Awareness Behavioural Rating Scales	Endsley (2000)
PSAQ – Participant SA questionnaire	Endsley (2000)
SPAM – Situation-Present Assessment Method	Durso et al (1998)
SACRI - Situation Awareness Control Room Inventory	Hogg et al (1995)

Mental Workload assessment techniques

Table 5. Mental Workload assessment techniques

Method	Author/Source
Bedford Scale	Roscoe and Ellis (1990)
CNS – Cognitive Neurometric System	Dean (1997)
Cognitive task load analysis	Neerincx (2002)
DRAWS – Defence Research Agency Workload Scale	Farmer et al (1995) Jordan et al (1995)
ISA – Instantaneous Self Assessment Workload	Jordan (1992)
MACE - Malvern Capacity Estimate	Goillau and Kelly (1996)
MCH – Modified Cooper Harper Scale	Cooper & Harper (1969)
NASA TLX – NASA Task Load Index	Hart and Staveland (1988)
Objective Workload Assessment (WinCrew)	Hadley, Guttman and Stringer (1999) Coolican (1994)
Physiological techniques (HRV, HR etc)	Varius
SWAT – Subjective Workload Assessment Technique	Reid and Nygeren (1998)
SWORD – Subjective WORkload Dominance assessment technique	Vidulich (1989)

Performance time measurement prediction techniques

Table 6. Performance time measurement prediction techniques

Method	Author/Source
KLM – Keystroke Level Model	Card, Moran and Newell (1983) Stanton & Young (1999)
CPA – Critical Path Analysis	Baber (1998)
Timeline Analysis	Kirwan & Ainsworth (1992)

Charting techniques

Table 7. Charting techniques

Method	Author/Source
CPA – Critical Path Analysis	Baber (1996)
DAD – Decision Action Diagrams	Kirwan & Ainsworth (1992) Kirwan (1994)
Event tree Analysis	Kirwan & Ainsworth (1992) Kirwan (1994)
Fault Tree Analysis	Kirwan & Ainsworth (1992) Kirwan (1994)
Information flowcharts	Various
Input-Output diagrams	Various
Murphy Diagrams	Pew et al (1981)
Operational sequence diagrams	Kurke (1961) Sanders & McCorrmick (1992)
Operator action event tree	Various
Petri Nets	Kirwan & Ainsworth (1992)
Process Charts	Kirwan & Ainsworth (1992) Marshall et al (2003)
Signal flow graphs	Kirwan & Ainsworth (1992)

Traditional Design Techniques

Table 8. Traditional Design Techniques

Method	Author/Source
Co-Design	Williams, Bound & Coleman (1999)
Conjoint techniques	Williams, Bound & Coleman (1999)
Ethnography	Williams, Bound & Coleman (1999)
Focus groups	Williams, Bound & Coleman (1999)
Immersion	Williams, Bound & Coleman (1999)
Mentoring	Williams, Bound & Coleman (1999)
Rapid prototyping	Williams, Bound & Coleman (1999)
Role play	Williams, Bound & Coleman (1999)
Scenarios	Diaper & Stanton (2003)
Shadowing	Williams, Bound & Coleman (1999)
Talkthrough analysis	Williams, Bound & Coleman (1999)
Think aloud protocols	Williams, Bound & Coleman (1999)
Time and Motion studies	Various
Walkthrough analysis	Kirwan & Ainsworth (1992)

Interface analysis techniques

Table 9. Interface analysis techniques

Method	Author/Source
Checklists	Various
Heuristics	Stanton & Young (1999)
Interface Surveys	Kirwan & Ainsworth (1992)
Link Analysis	Drury (1990)
Layout Analysis	Stanton & Young (1999)
QUIS – Questionnaire for User Interface Satisfaction	Chin, Diehl & Norman (1988)
Repertory Grids	Kelly (1955)
SUMI – Software Usability Measurement Inventory	Kirakowski
SUS – System Usability Scale	Stanton & Young (1999)
WAMMI – Website Analysis and Measurement Inventory	Internet source

Software based techniques

Table 10. Software based techniques

Technique	Author(s)
Analytica	
ATCS Performance Measurement Database	
ATLAS	
BMD-HMS - Boeing McDonnel Douglas Human Modelling System	
CASHE:PVS – Computer Aided Systems Human Engineering Performance Visualisation System	
CADA – CSERIAC Anthropometric Data Analysis files	
CSSM – Continuous Safety Sampling Methodology	
FAULTrEASE	
FAST – Functional Analysis System Technique	
Hiser Element Toolkit	
IPME - Integrated Performance Modelling Environment	
JACK	
KIT - Key Issues Tool	
MicrSaint	
MIDAS – Man-Machine Integration Design and Analysis Systems	
Observer	
PHA-Pro 5	
PUMA – Performance and Usability Modelling in ATM	
SAM 2000	
WinCrew	

Table 11. Team techniques

h

Method	Author
Team Training methods (Various)	Salas (2003) Various
Distributed Simulation Training for Teams	Andrews (2003)
Synthetic Task Environments for Teams – CERTT's UAV-STE	Cooke and Shope (2003)
Team Building	Salas (2003)
Measuring Team Knowledge	Cooke (2003)
Team Communications Analysis	Jentsch & Bowers (2003)
Questionnaires for Distributed Assessment of Team Mutual Awareness	MacMillan et al (2003)
Team Decision Requirement Exercise	Klinger & Bianka (2003)
TARGETS – Targeted Acceptable Responses to Generated Events or Tasks	Fowlkes & Burke (2003)
BOS – Behavioural Observation Scales	Baker (2003)
Team Situation Assessment Training for Adaptive Co-ordination	Burke (2003)
Team Task Analysis	Burke (2003)
Team Training methods (Various)	Salas (2003) Various
Social Network Analysis	Driskell & Mullen (2003)
CDM	Klien (2000)
CUD – Comms Usage System	Watts and Monk (2000)
Pentanalysis	Diaper, McKearney & Hearne (2000)
TTRAM – Task and Training Requirements Methodology	Swezey et al (2000)
MUSE – Method for Usability Engineering	Lim & Long (1994) Stanton, Hedge, Brookhuis, Salas, Hendrick (2003)
CUD – Comms Usage Diagram	Annett & Stanton (2000)
Team Cognitive Task Analysis	Klien (2000) Stanton, Hedge, Brookhuis, Salas, Hendrick (2003)

Other techniques

Table 12. Other techniques

Technique	Author(s)
EFHA - Early Human Factors Analysis	McLeod & Walters (1999)
MORT – Management Oversight Technique	Johnson (1980)
SACL – The Stress Arousal Checklist	Wilson & Corlett (1999)
SSADM – Structured Systems Analysis and Design Methodology	Weaver (1993)
MUSE – Method for Usability Engineering	Lim & Long (1994)

Appendix 2: Rejected HF techniques

Rejected methods

Table 1 – Rejected Situation Awareness Assessment Techniques

Technique	Author(s)	Justification
SAGAT – Situation Awareness Global Assessment Tool	Endsley (1995)	Evaluation technique - Situation Awareness measurement tool
SART – Situation Awareness Rating Technique	Taylor (1990)	Evaluation technique - Situation Awareness measurement tool
SPAM – Situation Present Assessment Technique	Durso et al (1998)	Evaluation technique - Situation Awareness measurement tool
SABARS – Situation Awareness Behavioural Rating Scales	Endsley (2000)	Evaluation technique - Situation Awareness measurement tool
PSAQ – Participant Situation Awareness Questionnaire	Endsley (2000)	Evaluation technique - Situation Awareness measurement tool
SALSA	Hauss & Eyferth (2003)	Evaluation technique - Situation Awareness measurement tool
SARS – Situation Awareness Rating Scale	Waag & Houck (1994)	Evaluation technique - Situation Awareness measurement tool
SACRI – Situation Awareness Control Room Inventory	Hogg et al (1995)	Evaluation technique - Situation Awareness measurement tool

Table 2 – Rejected Mental Workload assessment techniques

Technique	Author(s)	Justification
SWAT – Subjective Workload Assessment Technique	Reid and Nygeren (1998)	Evaluation technique – Workload assessment
Cognitive task load analysis	Neerincx (2002)	Evaluation technique – Workload assessment
Bedford Scale	Roscoe and Ellis (1990)	Evaluation technique – Workload assessment
CNS – Cognitive Neurometric System	Dean (1997)	Evaluation technique – Workload assessment
DRAWS – Defence Research Agency Workload Scale	Farmer et al (1995) Jordan et al (1995)	Evaluation technique – Workload assessment
ISA – Instantaneous Self Assessment Workload	Jordan (1992)	Evaluation technique – Workload assessment
MACE - Malvern Capacity Estimate	Goillau and Kelly (1996)	Evaluation technique – Workload assessment
Objective Workload Assessment (WinCrew)	Hadley, Guttman and Stringer (1999) Coolican (1994)	Evaluation technique – Workload assessment
MCH – Modified Cooper Harper Scale	Cooper & Harper (1969)	Evaluation technique – Workload assessment

Table 3 – Rejected HEI/HRA techniques

Technique	Author(s)	Justification
THERP – Technique for Human Error Rate Prediction	Swain & Guttman (1983)	Evaluation technique – retrospective HRA technique
SRK framework	Rasmussen (1986)	Error classification taxonomy
GEMS – Generic Error Modelling System	Reason (1990)	Error classification taxonomy
APJ – Absolute Probability Judgement	Seaver & Stillwell (1983)	Evaluation technique – Error quantification technique
Paired Comparisons	Hunns & Daniels (1980)	Evaluation technique – Error quantification
SNEAK	Hahn & De Vries (1980)	Evaluation technique – Error quantification technique
COMET – Commission Event trees	Blackman (1991)	Similar to event tree analysis
SCHEMA – Systematic Critical Human Error Management Approach	Livingston et al (1992)	Evaluation technique – Error quantification technique
EOCA – Error of Commission Analysis	Kirwan (1994)	Similar to HEIST
ATHEANA – A Technique for Human Error Analysis	Cooper et al (1996)	Not available in the public domain
JHEDI – Justification of Human Error Data Information	Kirwan (1990b, 1994)	Not available in the public domain.
SCHAZOP	Kirwan & Kennedy (1996a)	Similar to Human Error HAZOP
FMEA – Failure Modes and Effects Analysis	Henley & Kumamoto	Dated technique.
ASEP – Accident Sequence Evaluation Programme	Swain (1987)	Evaluation technique – retrospective error HRA technique

Table 4 – Rejected Design techniques

Technique	Author(s)	Justification
Repertory Grids	Kelly (1955)	Interview technique – covered in interviews section
Input-output diagrams	Kirwan & Ainsworth (1992)	Networking techniques covered sufficiently
Information flowcharts	Kirwan & Ainsworth (1992)	Networking techniques covered sufficiently
Petri Nets	Murata (1989)	Networking techniques covered sufficiently
Signal Flow Graphs	Beishan (1967)	Networking techniques covered sufficiently
MORT – Management Oversight Technique	Johnson (1980)	N/A
User trials	Various	Evaluation technique
Expert Reviews	Various	Evaluation technique
Group Design reviews	Various	Evaluation technique
SSADM – Structured Systems Analysis and Design Methodology	Weaver (1993)	N/A
MUSE – Method for Usability Engineering	Lim & Long (1994)	N/A
QUIS – Questionnaire for User Interface Satisfaction	Chin et al (1988) Foot et al (1996)	Evaluation technique
SUMI – Software Usability Measurement Inventory	Karwowski et al (1992)	Evaluation technique
WAMMI – Website Analysis and Measurement Inventory		Evaluation technique
SUS – System Usability Scale	Stanton & Young	Evaluation technique

	(1999)	
Ethnography	Various	N/A

Table 5 – Rejected Task analysis techniques

Technique	**Author(s)**	**Justification**
TAKD – Task Analysis for Knowledge Descriptions	Diaper	Evaluation technique
TTRAM – Task and Training Requirements Analysis Methodology	Swezey et al (2000)	Evaluation technique
TTA – Tabular Task Analysis	Various	HTA sufficient
Integrated Task Analysis		HTA sufficient
User needs Task Analysis		HTA sufficient

Table 6 – Rejected Software based techniques

Technique	**Author(s)**	**Justification**
Analytica		Software based tool
ATCS Performance Measurement Database		Software based tool
ATLAS		Software based tool
BMD-HMS - Boeing McDonnel Douglas Human Modelling System		Software based tool
CASHE:PVS – Computer Aided Systems Human Engineering Performance Visualisation System		Software based tool
CADA – CSERIAC Anthropometric Data Analysis files		Software based tool
CSSM – Continuous Safety Sampling Methodology		Software based tool
FAULTrEASE		Software based tool
FAST – Functional Analysis System Technique		Software based tool
Hiser Element Toolkit		Software based tool
IPME - Integrated Performance Modelling Environment		Software based tool
JACK		Software based tool
KIT - Key Issues Tool		Software based tool
MIDAS – Man-Machine Integration Design and Analysis Systems		Software based tool
Micro-Saint		Software based tool
PHA-Pro 5		Software based tool
PUMA – Performance and Usability Modelling in ATM		Software based tool
SAM 2000		Software based tool
WinCrew		Software based tool

Table 7. Team Methods

Method	**Author**	**Justification**
Team Training methods (Various)	Salas (2003) Various	Evaluation technique – work package 1.3.3
Distributed Simulation Training for Teams	Andrews (2003)	Evaluation technique – work package 1.3.3
Synthetic Task Environments for Teams – CERTT's UAV-STE	Cooke and Shope (2003)	Evaluation technique – work package 1.3.3
Team Building	Salas (2003)	Evaluation technique – work package 1.3.3
Measuring Team Knowledge	Cooke (2003)	Evaluation technique – work package 1.3.3
Team Communications Analysis	Jentsch & Bowers (2003)	Evaluation technique – work package 1.3.3
Questionnaires for Distributed Assessment of Team Mutual Awareness	MacMillan et al (2003)	Evaluation technique – work package 1.3.3
Team Decision Requirement Exercise	Klinger & Bianka (2003)	Evaluation technique – work package 1.3.3
TARGETS – Targeted Acceptable Responses to Generated Events or Tasks	Fowlkes & Burke (2003)	Evaluation technique – work package 1.3.3
BOS – Behavioural Observation Scales	Baker (2003)	Evaluation technique – work package 1.3.3
Team Situation Assessment Training for Adaptive Co-ordination	Burke (2003)	Evaluation technique – work package 1.3.3
Pentanalysis	Diaper, McKearney & Hearne (2000)	Evaluation technique – work package 1.3.3
TTRAM – Task and Training Requirements Methodology	Swezey et al (2000)	Evaluation technique – work package 1.3.3
MUSE – Method for Usability Engineering	Lim & Long (1994)	Evaluation technique – work package 1.3.3
SSADM – Structured Systems Analysis and Design Methodology		N/A