Health and Safety Executive

Review of human reliability assessment methods

Prepared by the **Health and Safety Laboratory** for the Health and Safety Executive 2009

RR679
Research Report

Health and Safety Executive

Review of human reliability assessment methods

Julie Bell & Justin Holroyd
Health and Safety Laboratory
Harpur Hill
Buxton
Derbyshire
SK17 9JN

Human reliability assessment (HRA) involves the use of qualitative and quantitative methods to assess the human contribution to risk. There are many and varied methods available for HRA, with some high hazard industries developing 'bespoke', industry focused methods.

It was considered that it would be useful for HSE to be up to date with developments in the field of quantitative HRA methods and to have knowledge of the capability of the tools and an understanding of their strengths and weaknesses. Therefore, this project was commissioned to further HSE knowledge and expertise, and to form a view on the 'acceptability' of the various tools for use in risk assessments.

This report and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

HSE Books

First published 2009

ACKNOWLEDGEMENTS

The authors would like to thank their HSE colleagues Jane Bowie, Robert Miles and Martin Anderson for their comments on the use of HRA methods in HSE inspection and enforcement activity. They would also like to thank the authors who provided additional information about the HRA tools they have developed, specifically Oliver Straeter and Eric Hollnagel.

FOREWORD

This document is an information resource on human reliability assessments (HRA) based on published research material and the opinion of the authors. It provides a summary of those tools and methods considered to be of potential use to analysts undertaking a HRA assessment in the major hazard sector. It is not intended to be a comprehensive guide to HRA but a useful starting point on which to build knowledge. There are other tools available that were not covered by this review.

CONTENTS

1 INTRODUCTION ... 1
1.1 Background ... 1
1.2 Aims ... 1
1.3 Objectives ... 1

2 LITERATURE SEARCH ... 2
2.1 Methodology ... 2
2.2 Search Results ... 2
2.3 Tool summaries ... 2

3 SUMMARY OF THE HRA METHODS REVIEWED ... 4
3.1 Summary of Tools considered useful to HSE major hazard directorates. 8
3.2 List of Tools considered not useful to HSE major hazard directorates .. 65
3.3 Conclusion ... 68

4 APPENDIX A ... 69

5 APPENDIX B ... 71

6 REFERENCES ... 72

EXECUTIVE SUMMARY

Human reliability assessment (HRA) involves the use of qualitative and quantitative methods to assess the human contribution to risk. There are many and varied methods available for HRA, with some high hazard industries developing 'bespoke', industry focused methods.

It was considered that it would be useful for HSE to be up to date with developments in the field of quantitative HRA methods and to have knowledge of the capability of the tools and an understanding of their strengths and weaknesses. Furthermore, there is potential that methods could be used out of context or inappropriately, and hence it is considered that HSE should form a view on the 'acceptability' of such tools for use in risk assessments.

Objectives

- Undertake a review of the literature to identify the range of qualitative and quantitative HRA techniques available and to carry out an assessment of their strengths and weaknesses. This will include simulation methods.

- Prepare a summary of the methods.

Main Findings

A total of 72 potential human reliability related tools and acronyms were identified within the search timeframe. Of these, 37 were excluded from any further investigation and 35 were identified as potentially relevant to HSE major hazard directorates and were investigated fully.

Of the 35 potentially relevant HRA tools, 17 are considered to be of potential use to major hazards directorates. For each of these 17 methods, a brief summary was prepared that includes:

- What they claim to offer and how they work (their scope, approach and information on the underlying models of the methods);

- The advantages and disadvantages of the method based on objective information available in the research literature;

- A comment on their potential application and major hazard sectors for which they would be suitable (if appropriate);

- A comment on their validity; and

- A note of the resources required for their use.

Each summary is based on the published literature, opinion and in some cases, personal communication with the authors of the methods. It should be noted that the published information, for many tools, provides an incomplete picture and therefore, there may be other information that has not been included in the reviews or the decision making process on the suitability of tools for HSE major hazard directorates.

1 INTRODUCTION

1.1 BACKGROUND

Human reliability assessment (HRA) involves the use of qualitative and quantitative methods to assess the human contribution to risk. There are many and varied methods available for HRA, with some high hazard industries developing 'bespoke', industry focused methods.

It was considered that it would be useful for HSE to be up to date with developments in the field of quantitative HRA methods and to have knowledge of the capability of the tools and an understanding of their strengths and weaknesses. In addition, the number of HRA specialists in HSE and HSL is extremely limited, and their knowledge is often restricted to one domain or tool. Therefore, this project was commissioned to further HSE knowledge and expertise, and improve consistency across major hazard directorates where appropriate. Furthermore, there is potential that methods could be used out of context or inappropriately, and hence it is considered that HSE should form a view on the 'acceptability' of such tools for use in risk assessments.

1.2 AIMS

The aims of this project are:

- To provide HSE with a review of the HRA literature;
- To form a view on the suitability of such tools for application in the major hazard sector, and to provide information for duty holders via the HSE human factors (HF) WebPages.

1.3 OBJECTIVES

- Undertake a review of the literature to identify the range of qualitative and quantitative HRA techniques available and to carry out an assessment of their strengths and weaknesses. This will include simulation methods.
- Prepare a summary of the methods.

This report provides an overview of the literature review that was undertaken by HSL in response to HSE's knowledge requirement. It provides a summary of each of the most relevant HRA tools identified.

2 LITERATURE SEARCH

2.1 METHODOLOGY

A literature search was undertaken to identify published sources of information regarding qualitative and quantitative HRA methods, including simulation studies and, particularly, any review papers. HRA specialists in HSL/ HSE provided many relevant papers and other information resources such as the Internet were searched by the HSL project team. In addition, a search was commissioned from the HSE Information Centre where a team of Information Specialists searched for appropriate papers from a range of databases, using search terms specified by the project team. The search was designed to collate source papers that detailed HRA methods, and subsequent validation and review papers. The aim of the literature review was to draw upon knowledge from existing published reviews, and not necessarily to assess each tool from source references.

A time limit of one month was placed on the literature search period because it was recognised that the main tools would be identified within that timeframe, and prolonged searching would provide diminishing returns.

2.2 SEARCH RESULTS

A total of 72 potential human reliability related tools and acronyms were identified within the search timeframe. After reading the details of these 72 results, 37 were excluded from any further investigation and 35 were identified as potentially relevant to HSE major hazard directorates and were investigated fully. The decision to exclude 37 was based on a high level assessment of the available information to determine the likelihood of applicability to HSE major hazard directorates. A more detailed analysis of the tools might have resulted in more being selected as potentially relevant but this was not possible within the scope of the work. The reasons for exclusion included; the acronym related to a tool for managing crew numbers, aviation specific tools, human computer interaction methods, human reliability databases, a programme of risk assessment models rather than a specific tool (e.g. ASP, accident sequence precursor) or human error identification tools, all of which are outside the scope of the current work (refer to Appendix A for the list of 37 tools and the reasons for their exclusion).

2.3 TOOL SUMMARIES

The remaining 35 HRA tools were reviewed in detail and 17 were considered to be of potential <u>use</u> to major hazards directorates. For each of these methods, a brief summary was prepared that includes:

- What they claim to offer and how they work (their scope, approach and information on the underlying models of the methods);
- The advantages and disadvantages of the method based on objective information available in the literature;
- A comment on their potential application and major hazard sectors for which they would be suitable (if appropriate);

- A comment on their validity; and
- A note of the resources required for their use.

Each tool summary is based on the published literature and in some cases, personal communication with the authors of the methods. In most cases, the information provided is a direct quote from published sources but the opinion of the HSL research team is also included.

It should be noted that the published information about many of the tools provides an incomplete picture and therefore, there may be other information that has not been included in the reviews or the decision making process on the suitability of tools for the HSE major hazard directorate.

A definition of each of the review elements is provided in Appendix B.

3 SUMMARY OF THE HRA METHODS REVIEWED

Table 1: Acronym and full title of the 35 tools identified for review

Tool	In full
ASEP	Accident Sequence Evaluation Programme
AIPA	Accident Initiation and Progression Analysis
APJ	Absolute Probability Judgement
ATHEANA	A Technique for Human Error Analysis
CAHR	Connectionism Assessment of Human Reliability
CARA	Controller Action Reliability Assessment
CES	Cognitive Environmental Simulation
CESA	Commission Errors Search and Assessment
CM	Confusion Matrix
CODA	Conclusions from occurrences by descriptions of actions
COGENT	COGnitive EveNt Tree
COSIMO	Cognitive Simulation Model
CREAM	Cognitive Reliability and Error Analysis Method
DNE	Direct Numerical Estimation
DREAMS	Dynamic Reliability Technique for Error Assessment in Man-machine Systems
FACE	Framework for Analysing Commission Errors
HCR	Human Cognitive Reliability
HEART	Human Error Assessment and Reduction Technique
HORAAM	Human and Organisational Reliability Analysis in Accident Management
HRMS	Human Reliability Management System
INTENT	Not an acronym
JHEDI	Justified Human Error Data Information
MAPPS	Maintenance Personnel Performance Simulation
MERMOS	Method d'Evaluation de la Realisation des Missions Operateur pour la Surete (Assessment method for the performance of safety operation.)

Tool	In full
NARA	Nuclear Action Reliability Assessment
OATS	Operator Action Tree System
OHPRA	Operational Human Performance Reliability Analysis
PC	Paired comparisons
PHRA	Probabilistic Human Reliability Assessment
SHARP	Systematic Human Action Reliability Procedure
SLIM-MAUD	Success likelihood index methodology, multi-attribute utility decomposition
SPAR-H	Simplified Plant Analysis Risk Human Reliability Assessment
STAHR	Socio-Technical Assessment of Human Reliability
TESEO	Tecnica empirica stima errori operatori (Empirical technique to estimate operator errors)
THERP	Technique for Human Error Rate Prediction

- Of these 35 tools, 17 were considered to be of potential use to HSE (refer to Table 2), and 18 were not considered to be of use (refer to Table 3)
- Table 2 is organised into those tools that are potentially useful and publicly available, and those that are potentially useful but are proprietary and/ or not publicly available.
- Section 3.1 includes a summary of each of these tools (presented in the same order as in Table 2).
- Section 3.2 provides a list of the 18 tools that were not considered to be of use to HSE along with a brief reason for their exclusion.

Table 2: A list of the 17 tools considered to be of potential use to HSE major hazard directorates

Availability	Type	Tool	Comment	Domain
Publicly available	1st generation[1]	THERP	A comprehensive HRA approach developed for the USNRC	Nuclear with wider application
		ASEP	A shortened version of THERP developed for the USNRC	Nuclear
		HEART	Relatively quick to apply and understood by engineers and human factors specialists. The method is available via published research papers. (A manual is available via British Energy).	Generic
		SPAR-H	Useful approach for situations where a detailed assessment is not necessary. Developed for the USNRC. Based on HEART.	Nuclear with wider application
	2nd generation[1]	ATHEANA	Resource intensive and would benefit from further development. Developed by the USNRC	Nuclear with wider application
		CREAM	Requires further development. Available in a number of published references.	Nuclear with wider application
	Expert judgement	APJ	Requires tight controls to minimise bias, otherwise validity may be questionable. Viewed by some as more valid that PC and SLIM.	Generic
		PC	Requires tight controls to minimise bias, otherwise validity may be questionable	Generic
		SLIM-MAUD	Requires tight controls to minimise bias of the SLIM element, otherwise validity can be questionable. The SLIM element is publicly available.	Nuclear with wider application
Not publicly available	1st generation	HRMS	Comprehensive computerised tool. A proprietary method.	Nuclear
		JHEDI	Faster screening technique than HRMS, its parent tool. A proprietary method	Nuclear
		INTENT	Narrow focus on errors of intention. Little evidence of use but potentially useful. Available by contacting the authors.	Nuclear

[1] *Refer to Section 3.1*

		Tool	Comment	Domain
Not publicly available	2nd generation	CAHR	A database method that is potentially useful. Available by contacting the authors (CAHR website).	Generic
		CESA	Potentially useful. Available by contacting the authors.	Nuclear
		CODA	Requires further development and CAHR or CESA may be more useful. Available by contacting the authors.	Nuclear
		MERMOS	Developed and used by EdF, its development is ongoing. A proprietary tool.	Nuclear
	3rd generation	NARA	A nuclear specific version of HEART (different author to the original). A proprietary tool.	Nuclear

3.1 SUMMARY OF TOOLS CONSIDERED USEFUL TO HSE MAJOR HAZARD DIRECTORATES

The follow sections provide a summary of those tools identified as being of potential use to HSE major hazard directorates. The tools are classified as first, second and third generation and expert judgment methods.

First Generation Methods

These tools were the first to be developed to help risk assessors predict and quantify the likelihood of human error. They include pre-processed tools like THERP and also expert judgement approaches such as APJ. First generation approaches tend to be atomistic in nature; they encourage the assessor to break a task into component parts and then consider the potential impact of modifying factors such as time pressure, equipment design and stress. By combining these elements the assessor can determine a nominal human error potential (HEP). First generation methods focus on the skill and rule base level of human action and are often criticised for failing to consider such things as the impact of context, organisational factors and errors of commission. Despite these criticisms they are useful and many are in regular use for quantitative risk assessments.

Second Generation Methods

The development of 'second generation' tools began in the 1990s and is on-going. They are an attempt to consider context and errors of commission in human error prediction, however due to the lack of uptake in the UK the benefits of the second generation over first generation approaches is yet to be established. They have also yet to be empirically validated.

Kirwan (2007) reports that the most notable of the second generation tools are ATHEANA, CREAM, MERMOS and CAHR but that MERMOS is the only one that is in regular use. However, this approach appears to be unused outside of EdF (Electricite de France) where the method was developed. The literature shows that second generation methods are generally considered to be still under development but that in their current form they can provide useful insight to human reliability issues.

New tools are now emerging based on earlier first generation tools such as HEART, and are being referred to as **third generation** methods.

Expert judgement based methods

Expert judgement methods became popular in the mid 1980s and remain so, particularly in less safety critical environments than major hazard industries. These tools provide a structured means for experts to consider how likely an error is in a particular scenario. The validity of some approaches (e.g. SLIM and PC) has been questioned in some papers, but they continue to be used and to inform the development of new tools.

3.1.1 Technique for Human Error Rate Prediction (THERP)

Factual information	
Origins of the tool	THERP was developed by Swain (when working for Sandia National Laboratories). Swain & Guttmann (1983) then prepared the THERP handbook for the US Nuclear Regulatory Commission.
Description of the tool	The THERP handbook presents methods, models and estimated human error probabilities (HEPs) to enable qualified analysts to make quantitative or qualitative assessments of occurrences of human errors in nuclear power plants (NPPs) THERP is a total methodology for assessing human reliability that deals with task analyses (e.g. documentation reviews and walk/ talk through), error identification and representation, as well as the quantification of HEPs (Kirwan, 1994). THERP is often referred to as a 'decomposition' approach in that its descriptions of task, have a higher degree of resolution than many other techniques. It is also a logical approach and one that puts a larger degree of emphasis on error recovery than most other techniques. Essentially, the THERP handbook presents tabled entries of HEPs that can be modified by the effects of plant specific Performance Shaping Factors (PSFs), using other tables (Swain, 1987). The key elements to complete the quantification process are described by Kirwan et al (1997) as: • Decomposition of tasks into elements • Assignment of nominal HEPs to each element • Determination of effects of PSF on each element • Calculation of effects of dependence between tasks • Modelling in an HRA event tree • Quantification of total task HEP To arrive at the overall failure probability, the exact failure equation involves summing probabilities of all failure paths in the event tree. When all the HEPs are .01 or smaller, the exact failure equation can be approximated by summing only the primary failure paths, ignoring all the success limbs. The accuracy of the approximation decreases as the number of terms or the values of the HEPs increase (Swain & Guttmann, 1983).
Validation	Kirwan et al (1997) carried out an independent validation of THERP along with two other methods (HEART and JHEDI). They found that no one technique out performed the others, and all three achieved a reasonable level of accuracy. The method has been extensively used in the nuclear industry, particularly in the USA.
Domain usage/ applicability to other domains	THERP was developed for probabilistic risk assessments of nuclear power plants but has been applied to other sectors such as offshore and medical.

Resources required to complete the assessment	Training on THERP is required before application. Specifically, Swain & Guttmann (1983) state that the method is intended to assist trained risk analysts in quantifying human reliability. Assessors can require one to two weeks of training to become accredited. THERP can be resource intensive (Kirwan, 1994) but the available information does not give a clear indication about the average length of an assessment or how many people are required. Swain (1987), when explaining the reasons for the abbreviated version of THERP, known as ASEP, the accident sequence evaluation program human reliability analysis procedure said that, "… the THERP handbook is thorough, for its fullest application it requires considerable manpower and time on the part of a team of experts, including a human reliability specialist, systems analysts, plant personnel and others".
Availability of the tool & support	The THERP Handbook was produced for the US Nuclear Regulatory Commission (NUREG/CR-1278) and is publicly available via their website. There are a number of consultancies that provide training in human reliability assessments and may be able to provide training in THERP.
References used for the summary	Kirwan, B. (1994). A guide to practical human reliability assessment. Taylor & 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*, 27, (6), 359-373 Kirwan, B., Kennedy, R., Taylor-Adams, S. and Lambert, B. (1997). The validation of three human reliability quantification techniques, THERP, HEART and JHEDI: Part II – results of validation exercise. *Applied Ergonomics,* 28 (1), 17-25. Swain AD and Guttmann HE (1983). Handbook of human reliability analysis with emphasis on nuclear power plant applications. US Nuclear Regulatory Commission), Washington, DC. NUREG/CR-1278 Swain, A. D. (1987). Accident Sequence Evaluation Program Human Reliability Analysis Procedure, NUREG/CR-4772 (US Nuclear Regulatory Commission, Washington, DC .1987)
Opinion	
Pros and cons	The following is taken from Kirwan (1994). Pros • THERP is well used in practice • It has a powerful methodology that can be audited • It is founded on a database of information that is included in the THERP handbook. Cons • THERP can be resource intensive and time consuming.

	• It does not offer enough guidance on modelling scenarios and the impact of PSFs on performance. • The level of detail that is included in THERP may be excessive for many assessments.
Suitability for MH sectors	THERP was designed for nuclear industry application but is a generic tool that can be applied in other sectors.
Related methods	ASEP is a shortened version of THERP, which is often used as a screening tool for identifying tasks that require a full THERP analysis.

3.1.2 Accident Sequence Evaluation Program (ASEP) Human Reliability Analysis Procedure

Factual information	
Origins of the tool	Swain (1987) developed ASEP for the US Nuclear Regulatory Commission. ASEP was developed because, "The Nuclear Regulatory Commission (NRC) expressed a need for an HRA method that would provide estimates of human error probabilities (HEPs) and response times for tasks performed during normal operating conditions and post-accident operating conditions and that would be sufficiently accurate for probabilistic reliability analysis and yet require only minimal expenditure of time and other resources" (Swain, 1987).
Description of the tool	ASEP is an "Abbreviated and slightly modified version of THERP. ASEP comprises pre-accident screening with nominal human reliability analysis, and post-accident screening and nominal human reliability analysis facilities. ASEP provides a shorter route to human reliability analysis than THERP by requiring less training to use the tool, less expertise for screening estimates, and less time to complete the analysis." (Everdij and Blom, 2008). The four procedures that comprise the ASEP HRA procedure are described as follows: • Pre-accident tasks: those tasks which, if performed incorrectly, could result in the unavailability of necessary systems or components in a complex plant such as an nuclear power plant (NPP) to respond appropriately to an accident. • Post-accident tasks: those tasks, which are intended to assist the plant to cope successfully with an abnormal event, that is to return the plant's systems to a safe condition. • Screening HRAs: Screening probabilities and response times are assigned to each human task as an initial type of sensitivity analysis. If a screening value does not have a material effect in the systems analysis, it may be dropped from further consideration. Screening reduces the amount of detailed analyses to be performed. HRAs at this stage deliberately use conservative estimates of HEPs, response times, dependence levels, and other human performance characteristics. • Nominal HRAs: The regular probabilistic risk assessment carried out on tasks identified during the screening process. These use what the HRA team judges to be more realistic values, but still somewhat conservative (i.e. pessimistic) to allow for the team's inability to consider all possible sources of error and all possible behavioural interactions.

	The Swain (1987) report provides the details for each of the steps in these four procedures.
Validation	Swain (1987) carried out ASEP user trials and reported the HEPs identified by using ASEP agree with the estimate HEPs produced using THERP. No further research to validate ASEP was identified during the literature search.
Domain usage/ applicability to other domains	This is a nuclear specific tool that has been successfully applied in the nuclear industry.
Resources required to complete the assessment	ASEP requires less training than THERP (which can take assessors one to two weeks of training to become accredited in).
Availability of the tool & support	The ASEP procedure was produced for the US Nuclear Regulatory Commission (NUREG/CR- 4772) and is publicly available.
References used for the summary	Everdij M.H.C. and Blom H.A.P. (2008) Safety Methods Database. http://www.nlr.nl/documents/flyers/SATdb.pdf Gore et al (1997) Conservatism of the Accident Sequence Evaluation Program HRA Procedure, http://www.blackwell-synergy.com/doi/abs/10.1111/j.1539-6924.1997.tb01283.x Kirwan, B. (1994). A guide to practical human reliability assessment. Taylor & Francis, London. Swain, A. D. (1987). Accident Sequence Evaluation Program Human Reliability Analysis Procedure, NUREG/CR-4772. US Nuclear Regulatory Commission, Washington, DC.
Opinion	
Pros and cons	Very little information was identified about the relative pros and cons of ASEP. Kirwan (1994) noted that ASEP is quicker to carry out than THERP and can be computerised. It tends to be used as a screening approach to identify those tasks that require a more detailed analysis using THERP.
Suitability for MH sectors	This is a nuclear specific tool and, therefore, not suitable for other major hazard sectors.
Related methods	THERP.

3.1.3 Human Error Assessment and Reduction Technique (HEART)

<table>
<tr><th colspan="2">Factual information</th></tr>
<tr><td>Origins of the tool</td><td>HEART was first outlined in a conference paper by Williams (1985), while he was working for the Central Electricity Generating Board. The method was described in further detail in subsequent papers (e.g. Williams, 1986 and 1988)</td></tr>
<tr><td>Description of the tool</td><td>
HEART is designed to be a quick and simple method for quantifying the risk of human error. It is a general method that is applicable to any situation or industry where human reliability is important.
The method is based on a number of premises.

Basic human reliability is dependent upon the generic nature of the task to be performed.
In ‘perfect’ conditions, this level of reliability will tend to be achieved consistently with a given nominal likelihood within probabilistic limits.
Given that these perfect conditions do not exist in all circumstances, the human reliability predicted may degrade as a function of the extent to which identified Error Producing Conditions (EPCs) might apply.

There are 9 Generic Task Types (GTTs) described in HEART, each with an associated nominal human error potential (HEP), and 38 Error Producing Conditions (EPCs) that may affect task reliability, each with a maximum amount by which the nominal HEP can be multiplied.
The key elements of the HEART method are: Classify the task for analysis into one of the 9 Generic Task Types and assign the nominal HEP to the task. Decide which EPCs may affect task reliability and then consider the assessed proportion of affect (APOA) for each EPC. Then calculate the task HEP.
An example HEART calculation is as follows:
GTT classified as Task F (restore or shift a system to original or new state following procedures, with some checking).
Nominal HEP = 0.003 (5^{th} & 95^{th} percentile bounds 00008 – 0.009)
<table>
<tr><th>EPCs</th><th>Total HEART Affect</th><th>APOA</th><th>Assessed Affect</th></tr>
<tr><td>Inexperience</td><td>x 3</td><td>0.4</td><td>(3-1) x 0.4 + 1 = 1.8</td></tr>
<tr><td>Opposite Technique</td><td>x 6</td><td>1</td><td>(6-1) x 1.0 + 1 = 6.0</td></tr>
<tr><td>Risk Misperception</td><td>x 4</td><td>0.8</td><td>(4-1) x 0.8 + 1 = 3.4</td></tr>
</table>
Assessed nominal likelihood of failure
0.003 x 1.8 x 6 x 3.4 = 0.11
Similar calculations may be performed to calculate the predicted 5^{th} and 95^{th} percentile bounds, which in this case would be 0.07 – 0.58.
As a total probability of failure can never exceed 1.00, if the multiplication of factors takes the value above 1.00, the probability of failure has to be assumed to be 1.00 and no more (Williams, 1992).
</td></tr>
</table>

	The HEART assessment can shed light on the key points of task vulnerability and suggest areas for improvement. The method includes error reduction strategies or remedial measures linked to the error producing conditions, so that as an assessment is completed the predicted contribution to error reduction can be anticipated, both in terms of EPCs and GEPs (Williams, 1992).
Validation	HEART is one of the few HRA methods that have been empirically validated. Kirwan et al (1997) carried out an independent empirical validation of HEART along with two other methods (THERP and JHEDI). The large-scale comparative validation study involved thirty HRA practitioners who each assessed thirty tasks (i.e. HEPs for which the true values were known, though not by the assessors). Ten assessors applied THERP, ten applied HEART and ten applied JHEDI. The results showed a significant correlation in each case between the assessed values and the true values. Kirwan et al (1997) found that no one technique outperformed the others, and all three achieved a reasonable level of accuracy. The study did highlight, however, that the consistency of use could be improved for all three techniques (Kirwan, 1997). Kirwan (1988) and Kennedy et al (2000) have also completed two other validation studies with similar results. HEART has been extensively used in the UK nuclear industry, and also in most other industries (chemical, aviation, rail, medical etc.). The underlying HEART model has subsequently been used to inform the development of some other tools in the area of HRA.
Domain usage/ applicability to other domains	HEART is a cross sector tool that is applicable to any domain where human reliability is important. As already noted, it has been successfully applied in many industries including nuclear, chemical, aviation, rail and medical.
Resources required to complete the assessment	HEART is a pre-processed form of HRA that is designed to be a relatively quick method to apply, and is generally easily understood by engineers and human factors specialists. However, as elements of the technique are highly subjective, like many other methods, it benefits from having a human factors analyst involved in the process. Any additional information that can be found about a task (e.g. from incident data and subject matter experts etc.) should also be considered in the analysis. Williams (1992) states "in order to be able to apply HEART technology to best effect, … assessors are likely to need a good standard of education…. What they need in particular, is an ability to see operations from the human perspective, and appreciation of statistics and an understanding of the nature of human variability." An individual can carry out the assessment but it is expected that they will discuss the task with operators and/ or other suitably experienced individuals to ensure understanding of the task both in theory and in reality. A number of consultants offer training courses about HEART and other HRA methods.
Availability of the tool & support	The methodology is in the public domain having been outlined in a number of conference papers by Williams and has been reviewed in several HRA publications. A detailed HEART user manual was written for Nuclear Electric (now British Energy) in 1992 but

	this is not publicly available. The manual is available on request from BE who will provide a copy, subject to specific conditions (i.e. not used for commercial gain and not a competitor of BE etc.).
References used for the summary	Kennedy, R., Kirwan, B., and Summersgill, R. (2000) making HRA a more consistent science. In Foresight & Precaution, Eds. Cottam, M., Pape, R.P., Harvey, D.W., and Tait, J. Balkema, Rotterdam. Kirwan, B. (1996). The validation of three human reliability quantification techniques, THERP, HEART and JHEDI: Part 1 technique descriptions and validation issues. *Applied ergonomics*, 27, (6), 359-373 Kirwan, B., Kennedy, R., Taylor-Adams, S. and Lambert, B. (1997). The validation of three human reliability quantification techniques, THERP, HEART and JHEDI: Part II – results of validation exercise. *Applied ergonomics*, 28 (1), 17-25. Kirwan B. (1997) The development of a nuclear chemical plant human reliability management approach: HRMS and JHEDI. Reliability Engineering & System Safety, Volume 56, Issue 2, Pages 107-133. Kirwan, B. (1988) A comparative evaluation of five human reliability assessment techniques. In Human Factors and Decision Making. Sayers, B.A. (Ed.) London: Elsevier, pp. 87-109. Williams, J.C. (1985). HEART – A Proposed Method for Achieving High Reliability in Process Operation by means of Human Factors Engineering Technology. In Proceedings of a Symposium on the Achievement of Reliability in Operating Plant, Safety and Reliability Society, 16 September 1985, Southport. Williams, J.C (1986). A proposed Method for Assessing and Reducing Human error. In Proceedings of the 9th Advance in Reliability Technology Symposium, University of Bradford, 1986, pp. B3/R/1 – B3/R/13. Williams, J. C. (1988). A Data-based method for assessing and reducing Human Error to improve operational experience. In Proceedings of IEEE 4th. Conference on Human Factors in power Plants, Monterey, California, -9 June 1988, pp. 436-450. Williams, J.C. (1992). Toward an Improved Evaluation Analysis Tool for Users of HEART. In Proceedings of the International Conference on Hazard identification and Risk Analysis, Human Factors and Human Reliability in Process Safety. 15-17 January 1992. Orlando, Florida. Williams, J.C (1992). A User Manual for the HEART Human Reliability Assessment Method. Prepared for Nuclear Electric plc. (C2547-1.001). Not in the public domain.
Opinion	
Pros and cons	Pros • A versatile, quick and simple human-reliability-calculation method, which also gives the user (whether engineer or ergonomist) suggestions on error reduction. • Requires relatively limited resources to complete an assessment.

	Cons • Error dependency modelling is not included. • Requires greater clarity of description to assist users when discriminating between generic tasks and their associated EPCs; there is potential for two assessors to calculate very different HEPS for the same task. • Lack of information about the extent to which tasks should be decomposed for analysis. • Potential for double counting (some elements of EPCs are implicit in the task description) • Subjective nature of determining the assessed proportion of affect.
Suitability for MH sectors	HEART is suitable for use in all MH sectors.
Related methods	NARA & CARA are tools that Kirwan has developed using a modified version of the HEART model, for sector specific purposes.

3.1.4 Simplified Plant Analysis Risk Human Reliability Assessment (SPAR-H)

Factual information	
Origins of the tool	SPAR-H was developed for the US Nuclear Research Commission, Office of Regulatory Research. In 1994, in support of the Accident Sequence Precursor Program (ASP), the USNRC (in conjunction with the Idaho National Laboratory, INL), developed the Accident Sequence Precursor Standardized Plant Analysis Risk Model (ASP/SPAR). The method was used in the development of nuclear power plant (NPP) models and, based on experience gained in field-testing, was updated in 1999 and re-named SPAR-H (Standardized Plant Analysis Risk-Human Reliability Analysis method), Gertman et al (2004a).
Description of the tool	Gertman et al (2004a) report that SPAR-H does the following: • Decomposes probability into contributions from diagnosis failures and action failures; • Accounts for the context associated with human failure events (HFEs) by using performance shaping factors (PSFs), and dependency assignment to adjust a base-case HEP; • Uses pre-defined base-case HEPs and PSFs, together with guidance on how to assign the appropriate value of the PSF; • Employs a beta distribution for uncertainty analysis, which can mimic normal and log normal distributions, but it has the advantage that probabilities calculated with this approach range from 0 to 1; and • Uses designated worksheets to ensure analyst consistency. The SPAR-H method assigns human activity to one of two general task categories: action or diagnosis. • Action tasks – carrying out one or more activities indicated by diagnosis, operating rules or written procedures. For example, operating equipment, performing line-ups, starting pumps, conducting calibration or testing, carrying out actions in response to alarms, and other activities performed during the course of following plant procedures or work orders. (Generic error rate of 0.001) • Diagnosis tasks – reliance on knowledge and experience to understand existing conditions, planning and prioritising activities, and determining appropriate courses of action. (Generic error rate 0.01) The base error rates for the two task types associated with the SPAR-H method were calibrated against other HRA methods. They are said to represent the top-level distinction between tasks that are often used in HRA. Eight PSFs were identified as being capable of influencing human performance and are accounted for in the SPAR-H quantification process.

The potential beneficial influence, as well as the detrimental influence, of these factors is included in the method.
PSFs are:

- Available time
- Stress and stressors
- Experience and training
- Complexity
- Ergonomics (& Human Machine Interface)
- Procedures
- Fitness for duty
- Work processes

When developing the basic SPAR-H model, only three of the eight PSFs are evaluated: time available, stress and stressors, and complexity. The remaining five PSFs are generally considered to be event, plant or personnel specific and would be evaluated when a plant-specific model is being developed.

SPAR-H is also reported to address dependency (described as the negative influence of a human error on subsequent errors as influenced by crew numbers, time, location and cues). The ratings of the various combinations of dependency contributory factors were examined and given a rating based on their combined effect on dependency among tasks, these correspond to zero, low, moderate, high or complete dependency among tasks.

A major component of the SPAR-H method is the SPAR-H worksheet, which simplifies the estimation procedure. The process for using the worksheet differs slightly, depending on whether the analyst is using the method to build SPAR models, perform event analysis, or perform a more detailed HRA analysis. HEPs are determined by a multiplicative calculation (i.e. Probability task failure x PSF1 X PSF2 x PSF3) and in previous versions of the SPAR-H method, it was possible to assign PSF levels that can result in a calculation of a mean that would be numerically larger than one. The worksheet, which accompanies the latest version, includes an adjustment factor to avoid probability estimates greater than one.

$$HEP = \frac{NHEP \cdot PSF_{composite}}{NHEP \cdot (PSF_{composite} - 1) + 1}$$

Forester et al (2006) consider that SPAR-H is not a full-scope HRA method in the sense that it does not provide guidance for identifying or modelling HFEs within the context of the PRA.

SPAR-H segregates HFEs into diagnosis failures and action failures, and quantifies the two failure types separately. Nominal HEPs are assigned to both and adjusted to reflect the impact of each of eight PSFs. In doing so, and as is done in many other HRA methods, each PSF is examined against specific guidance provided to assess the influence of each PSF (e.g.,

	complexity is high, moderate, or nominal) and then an associated multiplier is used to adjust the nominal HEP based on the PSF evaluation. The PSFs and associated multiplicative values used in SPAR-H were arrived at through an extensive review of HRA methods available at the time of its development, and are based on incorporating much of what is found in those other methods. SPAR-H also allows modelling of dependencies between HFEs, using the dependence model from THERP.
Validation	Various NRC groups have indirectly evaluated the reliability of the SPAR-H method over the years. These reviewers have indicated different areas for improvement and clarification that have been incorporated in the current version of the method (Gertman et al, 2004). The task types and PSF elements were adapted from other HRA methods available at the time (i.e. HEART, CREAM, THERP and ASEP). While this may give some perceived validation to the SPAR-H method, it is not always clear how decisions about the HEPs are made (Forester et al, 2006).
Domain usage/ applicability to other domains	SPAR-H was developed for the nuclear sector and has been successfully applied to risk informed regulatory activities. No evidence was found of the method being used in other sectors.
Resources required to complete the assessment	Gertman (2004) reports that the time to apply the method is relatively brief, when compared to some of the second generation HRA methods such as MERMOS, ATHEANA, or CAHR. It is designed to be a relatively quick method to apply.
Availability of the tool & support	The full manual/ reference document, summarising the method, is publicly available via the US NRC website (NUREG/CR-6883)
References used for the summary	Gertman, D., Blackman, H., Marble, J., Byers, and Smith, C. (2004a). The SPAR-H human reliability analysis method. NUREG/CR-6883. Idaho National Laboratory, prepared for U. S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research Washington, DC 205555-0001 Gertman, D.I., Boring, R.L., Marble, J.L., and Blackman, H.S. (2004b). Mixed model usability evaluation of the SPAR-H human reliability analysis method. Fourth American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Controls and Human-Machine Interface Technologies (NPIC&HMIT 2004), Columbus, Ohio. September, 2004. Gertman, D.I., Blackman, H.S., Marble, J.L., Smith, C and Boring, R.L. (2004c). The SPAR H Human Reliability Analysis Method. Fourth American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Controls and Human-Machine Interface Technologies (NPIC&HMIT 2004), Columbus, Ohio. September, 2004. Forester J., Kolaczkowski A., Lois E. and Kelly D. (2006) Evaluation of Analysis Methods Against Good Practices. Final Report. NUREG-1842. U.S. Nuclear Regulatory Commission Office of Nuclear Regulatory

	Research Washington, DC 20555-0001
Opinion	
Pros and cons	The following are taken from Forester et al (2006): Pros • A simple underlying model makes SPAR-H relatively easy to use and results are traceable. • The eight PSFs included cover many situations where more detailed analysis is not required. • The THERP-like dependence model can be used to address both subtask and event sequence dependence. Cons • The degree of resolution of the PSFs may be inadequate for detailed analysis. • No explicit guidance is provided for addressing a wider range of PSFs when needed, but analysts are encouraged to use more recent context developing methods if more detail is needed for their application, particularly as related to diagnosis errors. • Although the authors checked the SPAH-H underlying data for consistency with other methods, the basis for selection of final values was not always clear. • The method may not be appropriate where more realistic, detailed analysis of diagnosis errors is needed.
Suitability for MH sectors	SPAR-H was developed for the nuclear industry but the underlying principles and HEP data are applicable to other domains.
Related methods	SPAR-H is based on the HEART approach and uses data from CREAM, THERP and ASEP.

3.1.5 A Technique for Human Event Analysis (ATHEANA)

Factual information	
Origins of the tool	A consortium of HRA specialists (see the ATHEANA team, below) developed ATHEANA for the US NRC. There are two key documents to consult in reference to ATHEANA, the first is, A Technique for Human Event Analysis (ATHEANA) - Technical Basis and Methodological Description. NUREG/CR-6350. U.S. Nuclear Regulatory Commission, Brookhaven National Laboratory, Upton, NY, April. 1996. Prepared by Cooper, S.E., Ramey-Smith, A.M., Wreathall, J., Parry, G.W., Bley, D.C. Luckas, W.J., Taylor, J.H., Barriere, M.T. The second is a full manual; Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA). NUREG-1624, May 2000. Division of Risk Analysis and Applications. Office of Nuclear Regulatory Research, US NRC, Washington DC. The ATHEANA Team: Barriere, M.T. – Brookhaven National Laboratory (BNL) Bley, D.C. – Buttonwood Consulting, Inc. Cooper, S.E. – Science Applications International Corporation (SAIC) Forester, J. – Sandia National Laboratories Kolaczkowski, A. - Science Applications International Corporation (SAIC) Luckas, W.J., - Brookhaven National Laboratory (BNL) Parry, G.W. – Halliburton NUS Corporation (NUS) Ramey-Smith, A.M. – NRC Project Manager Thompson, C. - NRC Whitehead, D. - Sandia National Laboratories Wreathall, J. – John Wreathall & Company, Inc.
Description of the tool	The following extracts are taken from NUREG-1624. ATHEANA is a second-generation tool, which is described as a method for obtaining qualitative and quantitative HRA results. The premise of the method is that significant human errors occur as a result of "error-forcing contexts" (EFCs), defined as combinations of plant conditions and other influences that make an operator error more likely. It provides structured search schemes for finding such EFCs, by using and integrating knowledge and experience in engineering, probabilistic risk assessment (PRA), human factors, and psychology with plant specific information and insights from the analysis of serious accidents. The tool can be used for both retrospective and prospective analyses. Main reasons for developing ATHEANA were: • Human events modelled in previous HRA/ PRA models were not considered to be consistent with the significant roles that operators have played in actual operational events;

- The accident record and advances in behavioural sciences both supported a stronger focus on the contextual factors, especially plant conditions, in understanding human error;
- Advances in psychology were integrated with the disciplines of engineering, human factors and PRA in modelling human failure events.

There are 10 steps in the ATHEANA HRA process but the essential elements are:

- Integration of the issues of concern into the ATHEANA HRA/PRA perspective
- Identification of human failure events and unsafe actions that are relevant to the issue of concern.
- For each human failure event or unsafe action, identification of (through a structured and controlled approach) the reasons why such events occur (i.e. elements of an EFC – plant conditions and performance shaping factors).
- Quantification of the EFCs and the probability of each unsafe action, given its context.
- Evaluation of the results of the analysis in terms of the issue for which the analysis was performed.

EFCs are identified using four related search schemes:

- The first three searches identify plant conditions and rules that involve deviations from the same base case.
- A search for physical deviations from the expected plant response.
- A search of formal procedures that apply normally or that might apply under the deviation scenario identified in the first search.
- A search for support system dependencies and dependent effects of pre-initiating event human actions.
- A "reverse" search for operator tendencies and error types. In this search, a catalogue of error types and operator tendencies is examined to identify those that could cause human failure events or unsafe actions of interest. Then plant conditions and rules associated with such inappropriate response are identified.

ATHEANA uses a quantification model for the probability of human failure events (HFEs) based upon estimates of how likely or frequently the plant conditions and PSFs comprising the EFCs occur.

The three basic elements considered in the quantification process are:

1. The probability of the EFC – a combination of plant conditions and performance shaping factors judged likely to give rise to the unsafe human action (UA). Plant-specific information is available for quantification purposes from the following sources,
    - Statistical analysis of operating experience
    - Engineering calculations
    - Quantitative judgements from experts
    - Qualitative judgements from experts

    The probability of some PSFs can be estimated from historical records, other PSFs may be linked to a variety of factors – the judgements of plant experts, coupled with that of the analysts applying ATHEANA, forms the basis for assessing the likelihood of these PSFs occurring.

2. The probability of the UA
    The preferred situation is one in which operator trainers can provide expert judgement as an input to the quantification of unsafe actions. However, if this input is not available then modelling methods are the next best choice (suggested modelling methods are HEART and SLIM).

	3. The probability of not recovering from the initial UA. This third stage focuses on several recovery issues that may prevent the unsafe action from continuing to the point of core damage. These issues are: • The occurrence of alarms and other indications following the unsafe action that may raise questions as to the correctness of actions taken or not taken. • Opportunities for new operators (i.e. those not involved in the unsafe action) to question the on-going response. • The potential for consequential changes in the plant state to lead to new alarms and indications. The process relies heavily on judgement based on the knowledge used in the previous steps in the quantification. Subsequently, an expert elicitation approach for performing ATHEANA quantification was developed (Forester, et al 2004) because "… significant judgement must be exercised by the analysts performing the quantification. In fact, a significant amount of creativity and insight on the part of the analyst would be necessary to use existing HRA quantification methods to address the error-forcing conditions identified using ATHEANA. As a result, the originators of ATHEANA have recently adopted a facilitator-led group consensus expert elicitation approach for quantifying human actions and treating uncertainty". It was also proposed in this paper that a library of results could be built, and that if a database were available, "…it might be possible to synthesize a more direct quantification approach." Boring et al (2005) in a review of atomistic versus holistic HRA methods, summarised ATHEANA as something of a hybrid method of expert elicitation, but mostly resembles a holistic approach. The ATHEANA analysis is open-ended in terms of identifying UAs and EFCs. It does not use an atomistic, pre-defined list of PSFs for characterising the HFE. Instead, it encourages analysts to explore the event from multiple angles to arrive at a thorough set of UAs and EFCs. No formal guidance is given regarding how the event contributors are used to shape the overall event probability. While guidance is given for defining the scale to be used in the uncertainty distribution, the actual combination of contributors into a single distribution is left to the analysts' discretion.
Validation	No empirical validation of ATHEANA has been undertaken. A peer review of ATHEANA, its documentation, and the results of an initial test of the method were held over a two-day period in June 1998, Seattle, Washington. The four reviewers were: Hollnagel, E., Cacciabue, P.C., Straeter, O. and Lewis, S.R. In addition, approximately 20 other individuals with an interest in HRA and ATHEANA also attended the peer review meeting and were invited to provide comments. The reviewers' general opinion of ATHEANA was that the method represents a significant improvement in HRA methodology; it is a useful and usable method; and it is a "good alternative to first-generation HRA approaches." However, the method did not go far enough and therefore needed to be improved and extended. (Published in NUREG-1624, Appendix F). Some of the comments made during the review process are presented as pros and cons on the following page. No research was identified to show that ATHEANA had been used in any research other than that carried out by the authors.

Domain usage/ applicability to other domains	ATHEANA was developed for and applied in the nuclear industry. There is no evidence of it being applied in any other domain.
Resources required to complete the assessment	"The ATHEANA method is very cumbersome and presumably very costly." (Peer review comment from NUREG-1624) Training in the use of the method and an understanding of the operating context is required.
Availability of the tool & support	The ATHEANA methodology is publicly available via the US NRC website.
References used for the summary	Boring, R., Boring, L. and Gertman, D.I. (2005). Atomistic and holistic approaches to human reliability analysis in the US nuclear power industry. Safety and Reliability, Vol. 25, No. 2, pp. 21 – 37. Forester, J., Bley, D., Cooper, S., Lois, E., Siu, N., Kolaczkowski, A. and Wreathall, J. (2004) Expert elicitation approach for performing ATHEANA quantification. Reliability Engineering & System Safety 83, (2004) pp.207 – 220 Powers, D.A. (Chairman of the 468th Meeting of the Advisory Committee on Reactor Safeguards, December 2-4) (1999). Letter to the Executive Director for Operations US NRC about NUREG-1624, Revision 1, "Technical basis and implementation guidelines technique for human event analysis (ATHEANA). www.nrc.gov/reading-rm/doc-collections/acrs/letters/1999/4681870.httml US Nuclear Regulatory Commission (USNRC). A Technique for Human Event Analysis (ATHEANA) - Technical Basis and Methodological Description. NUREG/CR-6350 Brookhaven National Laboratory, Upton, NY, April. 1996. Prepared by Cooper, S.E., Ramey-Smith, A.M., Wreathall, J., Parry, G.W., Bley, D.C. Luckas, W.J., Taylor, J.H., Barriere, M.T. US Nuclear Regulatory Commission (USNRC). Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA). NUREG- 1624. Division of Risk Analysis and Applications. Office of Nuclear Regulatory Research, Washington DC. May 2000
Opinion	
Pros and cons	The ATHEANA methodology has made a good attempt at dealing with subjects that first generation tools did not address such as errors of commission. It has a good qualitative element but the quantitative element is lacking and relies on expert judgement. The following comments are taken from the peer review provided in NUREG-1624. Pros • [ATHEANA is an] approach, which attempts to solve the problem of including EOC [errors of commission] in PSA in an extensive way. • It provides a systematic way of exploring how action failures can occur. This is something that conventional HRA methods do not do well, if at all. • ATHEANA can be used to develop detailed qualitative insights into conditions that may cause problems. It may generate a solid basis for redesign of working procedures, training, and interface, and it may be used as a tool for scenario generation.

	• It focuses on the important issues of context and cognition. • If properly applied, the methods that comprise ATHEANA should be able to yield significantly more insight into the nature of human actions. Cons • The ATHEANA method is very cumbersome and presumable very costly. The guidance is too complex and depends too much on subject matter experts. • The quantification method is weak, and the quantitative results are unsubstantiated. The quantification is excessively dependent on expert judgement, hence possibly has low reliability as a method. • The qualitative results are good, but these might have been obtained in other ways, perhaps more efficiently. • The effectiveness of the ATHEANA methodology results from forming a diverse, experienced project team to perform a comprehensive, broad-ranging analysis. Few organisations, however, are in a position to undertake such an extensive analysis without clearly defined, commensurate benefits.
Suitability for MH sectors	ATHEANA was developed for the nuclear industry, however the approach is suitable for application in other industries.
Related methods	None.

3.1.6 Cognitive Reliability and Error Analysis Method (CREAM)

Factual information	
Origins of the tool	Erik Hollnagel (1993) developed CREAM and the method is still under development.
Description of the tool	Hollnagel describes CREAM as fully bidirectional i.e. the same principles can be applied for retrospective analysis as well as performance prediction. The model is based on a fundamental distinction between competence and control. A classification scheme clearly separates genotypes (causes) and phenotypes (manifestations), and furthermore proposes a non-hierarchical organisation of categories linked by means of the sub-categories called antecedents and consequents. The distinction between competence and control is based upon Hollnagel's COCOM (contextual control) model: • Competence includes a person's skills and knowledge, • Control is viewed as running along a continuum from a position where the individual has little/no control to where they have complete control. Several aspects of the context are identified; these are called Common Performance Conditions (CPCs). Genotypes are separated out into three categories. • The first category contains genotypes that have a direct or indirect link to behaviour (e.g. emotional state and personality). • The second category contains factors that relate to man-machine interaction and man-machine interface. • The third category includes genotypes that are typified by the organisation such as the local environment (e.g. noise and temperature). Phenotypes are the consequences of the operator's actions or omissions of actions, and will in many cases be the starting point for any analysis. Basic phenotypes (error modes) are divided into four sub-groups: • Action at the wrong time • Action of the wrong type • Action at the wrong object • Action in the wrong place. Phenotypes and genotypes are classified into general consequents and for each general consequent there are numerous general and specific antecedents. For example, for the genotype 'communication' the general consequent is 'communication failure', the general antecedents are 'distraction', 'functional impairment' and 'inattention'. The specific antecedents are 'noise', 'presentation failure' and 'temporary incapacitation'. For the purpose of HRA (the CREAM basic method) the first step is a task analysis. Based on this a list of operator activities is produced, from which a CPC analysis is carried out. There are nine CPCs:

	1. Adequacy of organisation. 2. Working conditions. 3. Adequacy of the man-machine interface and operational support. 4. Availability of procedures/plans. 5. Number of simultaneous goals. 6. Available time. 7. Time of day. 8. Adequacy of training and experience. 9. Quality of crew collaboration. For each activity a CPC level is determined, for example adequacy of training and experience is described as high experience, low experience or inadequate. The expected effect of these levels of experience on performance is respectively- improved, not significant and reduced. The method goes on to describe a way of quantifying these descriptors. The sum of the performance reliability (i.e. improved, not significant and reduced) for each CPC gives a combined CPC score (e.g. for the nine CPCs the result may be [9,0,0], which would be the least desirable situation as all CPCs indicate reduced performance reliability, whereas [0,2,7] describes a much more desirable situation). This basic CREAM method can be used as a screening process to decide whether or not to continue with a HRA. The next stage of extended analysis requires a cognitive demands profile to be built. This involves describing each cognitive activity in terms of observation, interpretation, planning and execution (i.e. COCOM functions) and plotting this in graphical form. Based on the phenotype-genotype classification, it is possible to create a complete list of cognitive function failures, however for practical purposes a subset of the list would be produced. For a defined subset each of the cognitive functions (observation errors, interpretation errors, planning errors and execution errors) would have identifiable potential cognitive function failures, the distribution of which would once again be graphically represented and a Cognitive Failure Probability (CFP) would be calculated for each. Finally, a weighting factor is applied to the CFP scores depending on whether contextual influences (CPCs) are determined to be weak, medium or strong.
Validation	The process of studying validity and reliability of CREAM is ongoing (Everdij and Blom, 2008). Collier (2003) found several problems with “both the CREAM technique and the data needed to complete the analysis”. It was felt that further development was needed before this kind of analysis can be reliable and valid, either in a research setting or as a practitioner’s tool in a safety assessment”. More recently, Marseguerra et al (2007) have applied traditional/basic CREAM and fuzzy CREAM (based on fuzzy logic i.e. a form of algebra employing a range of values from ‘true’ to ‘false’ that is used in making decisions with imprecise data) to a contextual scenario of an actual train crash. They found distinct advantages to applying fuzzy CREAM in that it allows for a more systematic and transparent definition of the underlying

	model and a more explicit treatment of the ambiguity involved in its evaluation. Hollnagel confirmed that the development of this tool is relatively limited, particularly with regards to HRA (personal communication dated 25/5/08).
Domain usage/ applicability to other domains	CREAM has been applied in the nuclear industry (Everdij and Blom, 2008) and to a rail crash scenario (Marseguerra et al (2007) but there is no evidence of extensive use.
Resources required to complete the assessment	There is no clear information available on the skills or knowledge required by assessors to use this method, however as with most methods, human factors knowledge would be advantageous.
Availability of the tool & support	Hollnagel (1998) provides comprehensive details on the principles of CREAM, classification and the methods of assessment, both retrospective and prospective. He has also provided (via email) a link with access to a CREAM Navigator developed at the University of Illinois at Urbana-Champaign (UIUC) - http://www.ews.uiuc.edu/~serwy/cream/v0.6beta/.
References used for the summary	Collier, S (2003) A Simulator Study of CREAM to Predict Cognitive Errors. In Proceedings of the International Workshop. Building the new HRA. Errors of commission form research to application. Nuclear Energy Agency. Pages 56-75. Everdij M.H.C. and Blom H.A.P. (2008) Safety Methods Database http://www.nlr.nl/documents/flyers/SATdb.pdf Hollnagel, E. (1993) Human reliability analysis: Context and control. Academic Press Hollnagel, E. (1998) Cognitive Reliability and Error Analysis Method. Elsevier Hollnagel, E (2008) Personal email communication (25/5/08)– link to University of Illinois at Urbana-Champaign (UIUC) - Kirwan, B. (1998) Human error identification techniques for risk assessment of high risk systems – Part 1: Review and evaluation of techniques, Applied Ergonomics, Vol. 29, No 3, pages 157-177 Marseguerra, M., Zio, E. and Librizzi, M. (2007) Human Reliability Analysis by Fuzzy "CREAM" Risk Analysis Vol. 27 No 1 pages 137–154
Opinion	
Pros and cons	Very few references were available that provided any level of critical review. The only comments that were identified are 10 years old and are as follows: In discussing Cognitive psychological approaches such as CREAM, Kirwan (1998) notes that "these approaches are potentially of most interest to psychologists and others who want to predict the more sophisticated error forms associated with misconceptions, misdiagnosis, etc. They attempt to explore the error forms arising from 'higher-level' cognitive behaviours".

	Kirwan (1998) also states that, "more development is clearly needed in this category, and could be linked to cognitive task analysis approaches". He also reports that the development of such approaches "…is limited, and new approaches are required, whether building on systems such as GEMS, or more novel hybrids such as the prototype CREAM technique which is still under development".
Suitability for MH sectors	CREAM was developed for use in the nuclear industry, however the underlying method is generic and, therefore, it is suitable for use in other major hazard sectors.
Related methods	CREAM is related to SHERPA and COCOM (Everdij and Blom, 2008)

3.1.7 Absolute Probability Judgements (APJ)

Factual information	
Origins of the tool	The mathematical material for APJ is based upon the work of Seaver and Stillwell (1983). Kirwan's Reliability Assessor's Guide (1988) outlined the examples and formats that could be utilised when formulating an APJ approach.
Description of the tool	The APJ approach is conceptually the most straightforward human reliability quantification approach. It simply assumes that people can remember, or better still, estimate directly the likelihood of an event, in this case, a human error (Kirwan 1994). There are different APJ approaches that can be applied to determine human reliability. A 'single expert APJ' would require one expert to make their own judgements on the chances of a human error. Kirwan's favoured approach is a 'group APJ' of which he has identified four distinct types. • Aggregated individual method. This is where individuals make their estimates individually and then a geometric mean of these estimates is calculated. • Delphi method. For this method individuals make their estimates independently of each other, but the assessments are then shared, allowing the experts to reassess their own estimates based on the new information. The human error probability scores are statistically aggregated (i.e. the geometric mean is calculated). • Nominal group technique. This method is similar to the Delphi method, the difference is that the experts are given the opportunity to discuss their estimates and confidentially re-evaluate their assessment. These scores are then statistically aggregated. • Consensus-group method. With this method the experts meet and discuss their estimates, following which a consensus on an agreed estimate must be reached. If this is not possible then a statistical aggregation of the individual estimates is calculated. Kirwan (1994) has identified eight distinct steps to the APJ procedure. 1) Select the subject-matter experts (SMEs). The SMEs should be familiar with the tasks being assessed and as many assessors as practically possible should be included. When discussions are required to reach a consensus the number of assessors would preferably be around 4-6 people. 2) Prepare the task statements. The clearer the task definitions the less they are open to individual interpretation. 3) Prepare the response booklets. The response booklets should include scale values that reflect the estimated range of the true probabilities of the tasks.

	4) Develop instructions. The response booklet should include clear instructions for the SMEs. 5) Obtain judgements. SMEs are encouraged to work through the response booklets, starting with the tasks that they feel confident assessing. 6) Calculate inter-judge consistency. To determine the levels of consistency between SMEs an analysis of variance would be performed. 7) Aggregate the individual estimates. This is achieved by calculating the geometric mean. 8) Uncertainty-bound estimation. Uncertainty bounds may be calculated using a form of Seaver and Stillwell's (1983) formulae or alternatively SMEs can be asked to estimate the confidence intervals, the estimates being aggregated statistically.
Validation	Embrey and Kirwan (1983) carried out a comparative validation of the expert judgement approaches APJ, PC and SLIM. The results showed that APJ had some degree of accuracy but that PC and SLIM needed further development and an improved calibration process. Peer review comments include the following; Humphreys (1988) reviewed APJ (amongst other tools) against a set of criteria (accuracy, validity, usefulness, effective use of resources, acceptability, and maturity). APJ was as good as any of the other methods reviewed (SLIM, PC, TESEO, THERP, HEART, IDA and HCR) with an average rating of 'moderate' against the criteria. Kosmowski et al (1994) report "some studies give support for the validity of this method". However, they do not give specific details about the studies that provide this support. Everdij and Blom (2008) express concern over the scarcity of statistical data and the reliance of APJ on expert judgement.
Domain usage/ applicability to other domains	APJ has been used in the nuclear and offshore industries (Everdij and Blom, 2008).
Resources required to complete the assessment	APJ requires experts, who must have detailed knowledge of the area they are being asked to assess, with at least ten years of practicing in their particular field or job. A second requirement is that the experts must also have some normative expertise – i.e. they must be familiar with probability calculus – or otherwise they will not be able to express their expertise in a coherent quantitative form. It is preferable, if experts are meeting and sharing their expertise, and discussing their arguments in a group, to make use of a facilitator (Kirwan 1994).
Availability of the tool & support	As part of Kirwan's (1994) summary of APJ, he outlines a step-by-step procedure to be followed for an APJ analysis. No other information was found on the availability of the tool.

<table>
<tr><td>References used for the summary</td><td>Embrey,D.E. and Kirwan, B. (1983) A comparative evaluation of three subjective human reliability quantification techniques. The Annual Ergonomics Society Conference Proceedings, Coombes, K. (ed) Taylor and Francis, London, pp 137-142.

Everdij M.H.C, Blom H.A.P., (2008) Safety Methods Database http://www.nlr.nl/documents/flyers/SATdb.pdf

Humphreys P. (1988) Human reliability assessors guide, Safety and Reliability Directorate UKAEA (SRD) Report No TRS 88/95Q

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

Kosmowski, G. Degen, J. Mertens, B. Reer (1994) Development of Advanced Methods and Related Software for Human Reliability Evaluation within Probabilistic Safety Analyses.

Seaver D.A. and Stillwell W.G. (1983) Procedures for using expert judgement to estimate human error probabilities in nuclear power plant operations. NUREG/CR-2743, Washington, DC 20555.

Williams, J.C. (1989) Validation of human reliability assessment techniques. Reliability Engineering, 11, 149-162.</td></tr>
<tr><td colspan="2">Opinion</td></tr>
<tr><td>Pros and cons</td><td>Pros -

The principal advantages of the APJ approach identified by Kirwan (1994) are as follows:

• The technique has been shown to provide accurate estimates in a wide variety of fields (e.g. weather forecasting).
• The method is relatively quick to use, and yet it also allows as much detailed discussion as the experts think fit; this kind of discussion, if documented, can often itself be qualitatively useful.
• Discussion can also be turned towards a consideration of how to achieve error reductions. In such a situation, the group becomes like a HAZOP group, and can develop some highly credible and informed suggestions for improvements. This development is also beneficial where the group members are themselves operational staff, since this fact would improve the chances of such recommendations being accepted and then properly implemented.

Kosmowski et al (1994) also comments that the method is relatively quick to use, if the data acquiring process is well organized and the experts possess a high level of expertise.

Cons -

The principal disadvantages of the APJ approach, as outlined by Kirwan (1994), are as follows:</td></tr>
</table>

	• The APJ technique is prone to certain biases, as well as to personality/group problems and conflicts, which, if not effectively countered (e.g. by a 'facilitator'), can significantly undermine the validity of the technique. • Since the technique is often likened to 'guessing', it enjoys a somewhat low degree of apparent, or 'face' validity. • The technique is critically dependent on the selection of appropriate experts, but there are few really practically useful criteria for the selection of 'good' experts.
Suitability for MH sectors	This method is suitable for a wide range of industries including those in the major hazards sectors.
Related methods	APJ is said to be another name for Direct Numerical Estimation. It can be used with Paired Comparisons and is similar to SLIM (Everdij and Blom, 2008; Kosmowski et al, 1994).

3.1.8 Paired Comparisons (PC)

Factual information	
Origins of the tool	The PC method is credited to a number of sources but it may have originated from Rook (1964), who developed an approximate model for THERP that used rank ordering of tasks in terms of error-likeliness and rank-ordering of system consequences of errors. Later, Swain (1967) proposed the use of psychological scaling methods, especially paired comparisons, to evaluate the worth of different design concepts and developmental models.
Description of the tool	The paired comparisons method (Hunns, 1982) is borrowed from the domain of psychophysics (a branch of psychology). It is a means of defining preferences between items (human errors) and asks experts to make judgements, albeit of a relatively simple variety. The PC method differs from Absolute Probability Judgements (APJs) in that subject matter experts (SMEs) make simple comparative judgements rather than absolute judgements. Each expert compares all possible pairs of error descriptions and decides, in each case, which of the two errors is more probable. For n tasks, each expert makes n(n-1)/2 comparisons. When comparisons made by different experts are combined, a relative scaling of error likelihood can then be constructed. This is then calibrated using a logarithmic calibration equation, which requires that the human error probabilities (HEPs) be known for at least two of the errors within the task set. Paired comparisons are relatively easy for the experts to carry out, and the method usefully determines whether each expert has been consistent in their judgements; inconsistency of judgement would suggest a lack of substantive expertise (Kirwan 1994). Kirwan (1994) details the 16 step procedure to be followed when carrying out this technique: 1) Define the tasks involved. The tasks should be defined simply, unambiguously and comprehensively. 2) Incorporate the calibration tasks. At least two of the task descriptions, for which the Human Error Probabilities are known, should be included in the task set. 3) Select the expert judges. The SMEs should have experience of the tasks being assessed. 4) Prepare the exercise. Each pair of tasks should be presented on its own, so that the expert only considers one pair at any one time. 5) Brief the experts. The SMEs should be briefed on the purpose of the study and the nature of the task to be assessed. 6) Carry out paired comparisons. When the SMEs are carrying out the paired comparisons it is useful for them to have the support of the analyst so that they are able to seek and receive clarification of the tasks.

	7) Derive the raw frequency matrix. Each cell in the matrix indicates the number of SMEs that considered one event as being more likely than an alternative event. 8) Derive the proportion matrix. The next step is to normalise the scores by determining the proportion of SMEs that have concluded that one event was more likely than the other. 9) Derive transformation X-matrix. The next step is to convert these scores into their equivalent unit normal deviate, using normal distribution tables. 10) Derive the column-difference Z-matrix. This is a simple calculation of the differences between the adjacent column values. 11) Calculate the scale values. The average column differences are converted into a linear scale by setting the most preferred task to zero (which represents the highest probability of error). 12) Estimate the calibration points. Ideally error probabilities for the calibration of tasks should be estimated from the frequencies obtained via actual observations, otherwise the APJ approach can be adopted. 13) Transform the scale values into probabilities. The scale values are transformed into human error probabilities (HEPs) via a method of simultaneous equations and using the logarithmic relationship: $HEP = ax + b$. 14) Determine the within-judge level of consistency - Experts can exhibit internal consistencies that need to be accounted for. The method proposed by Seaver and Stillwell (1983) results in a correlation coefficient which if large enough to be statistically significant means that the results of the PC should be rejected. 15) Determine the inter-judge level of consistency. As with APJ, to determine the levels of consistency between SMEs an analysis of variance could be performed. 16) Estimate the uncertainty bounds. If the estimates of statistical uncertainty bounds are required then Seaver and Stillwell (1983) should be consulted.
Validation	There is no good evidence of predictive validity, so indirect forms of validity (in the other checklist items) must be depended upon. In the Comer et al (1984) study, a detailed task analysis was performed, using the following SMEs: systems analysts, trainers, and operations personnel. In this study, face validity was high; a convergent validity ranging from 0.6 to 0.9, as measured by the inter-correlations between HEP estimates was found (Swain 1989). Embrey and Kirwan (1983) carried out a comparative validation of the expert judgement approaches APJ, PC and SLIM. The results showed that APJ had some degree of accuracy but that PC and SLIM needed further development and an improved calibration process.

	Comer et al (1984) reported that both the APJ and PC techniques appeared sufficiently accurate for HRA purposes. Humphreys (1988) reviewed PC (amongst other tools) against a set of criteria (accuracy, validity, usefulness, effective use of resources, acceptability, and maturity). PC was as good as any of the other methods reviewed (SLIM, APJ, TESEO, THERP, HEART, IDA and HCR) with an average rating of 'moderate' against the criteria. Everdij and Blom (2008) state that PC is not restricted to human error only. It can be used together with APJ. They point out that expert judgement is very often used, especially where statistical data is scarce, but needs to be treated with special care. They add that there are well-proven protocols for maximising and testing the validity of PC. In summary, the PC approach has been in use for several decades and has received positive or negative feedback depending on the reviewer.
Domain usage/ applicability to other domains	The PC method has been applied to the transport and nuclear industries (Everdij and Blom, 2008).
Resources required to complete the assessment	Kirwan (1994) states that the number of assessors should be preferably more than ten. They require relevant experience of the tasks to be assessed, but they do not require a considerable amount of experience in terms of probability theory or mathematical/statistical concepts. It is of critical importance that the experts fully understand the nature of the tasks being assessed, as well as the PC procedure itself. It will therefore be useful to discuss all the tasks involved with the experts. Kosmowski et al (1994) question whether assessors might require a background in mathematical/statistical concepts given that the PC method applies a human reliability evaluation where multidimensional events involve complex human activities.
Availability of the tool & support	From Kirwan's (1994) description, the PC approach is more a framework upon which bespoke tasks are defined, compared and measured than a particular method. As such, a 'tool' is not available.
References used for the summary	Comer, M.K., Seaver, D.A., Stillwell, W.G. & Gaddy, C.D. (1984) Generating human reliability estimates using expert judgement, Vol.2; The Maxima Corp. Maryland – U.S Nuclear Regulatory Commission, Contract Report, NUREG/CR – 3688-2. Embrey,D.E. and Kirwan, B. (1983) A comparative evaluation of three subjective human reliability quantification techniques. The Annual Ergonomics Society Conference Proceedings, Coombes, K. (ed) Taylor and Francis, London, pp 137-142. Everdij M.H.C. and Blom H.A.P. (2008) Safety Methods Database. http://www.nlr.nl/documents/flyers/SATdb.pdf Hunns, D.M. (1982) The method of paired comparisons. In: A.E. Green, Editor, High risk safety technology, Wiley, Chichester.

	Hunns, D. M., and Daniels, B. K., (1980) The method of paired comparison, Proceedings of 6th Symposium on Advances in Reliability Technology, Report NCSR R23. Kirwan, B. (1994) A Guide to Practical Human Reliability Assessment, Taylor and Francis, London. Kosmowski, G. Degen, J. Mertens, B. Reer, B. (1994) Development of Advanced Methods and Related Software for Human Reliability Evaluation within Probabilistic Safety Analyses. Lyons, M., Adams, S., Woloshynowych, M. and Vincent C (2004) Human reliability analysis in healthcare: A review of techniques International Journal of Risk & Safety in Medicine 16, pages 223–237 223. IOS Press. Swain, A.D. (1967) Field Calibrated Simulation in Proceedings of the Symposium on Human Performance Quantification in Systems Effectiveness. Naval Material Command and the National Academy of Engineering, Washington, DC. Swain A.D. (1989) Comparative Evaluation of Methods for Human Reliability Analysis. Gesellschaft für Reaktorsicherheit, Garching
Opinion	
Pros and cons	Kirwan (1994) provides a succinct summary of the pros and cons of PC. Pros Kirwan summarises the advantages of the PC approach as follows: • Subjective knowledge can be profitably extracted from comparative judgements, provided that the assumptions of the methods are upheld and the value of the human judgement proves greater than the value of any confusion arrived at via direct numerical assessments. • The technique makes it possible to determine whether or not individual judges are qualified to make judgements about a particular datum or data set. • Since the techniques can work with a minimum of two empirically estimated HEP values, this enables the most effective use to be made even of scarce amounts of empirical data. • Even without calibration, the technique provides a useful means of deriving a measure of the relative importance of different human errors or human events. This technique was used, for example, to prioritise different types of ship-collision threats to offshore platforms in the UK sector (Technica, 1985). • With a small number of tasks and a set of rapidly available experts, the technique can be applied fairly quickly, especially when carried out on a computer (see Hunns and Daniels, 1980). • Experts do not have to carry out the comparisons as a group. Not having to do so will eliminate the logistical problems of bringing experts together for the comparison.

	Cons Kirwan (1994) summarises the disadvantages as follows: • The tasks being considered may be too complex to allow an easy comparison. • The tasks may not be homogeneous. • The comparisons made may not be independent of each other. • If the number of comparisons is large, the judges may become tired and therefore, carry out later comparisons differently from earlier ones. Swain (1989) pointed out that the usefulness of this technique relies heavily on the availability of valid calibrators.
Suitability for MH sectors	The PC approach is a generic one that can be applied to any sector.
Related methods	Everdij and Blom (2008) report that PC can be used together with Absolute Probability Judgements.

3.1.9 SLIM-MAUD

Factual information	
Acronym and full name	**SLIM-MAUD** Success likelihood index method using multi-attribute utility decomposition
Origins of the tool	SLIM was first developed by Embrey (1983) for the US Nuclear Regulatory Commission. The key source reference is Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B. and Rea, K. (1984) SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgment. NUREG/CR-3518.
Description of the tool	SLIM-MAUD is an expert judgement methodology. • SLIM is a set of procedures for making expert judgements when developing Human Error Probability (HEP) estimates. • MAUD is a computer based procedure and is essentially a multi-attribute utility decomposition (MAUD) version of SLIM. Kirwan (1994) describes it as a sophisticated approach, which helps to ensure that the expert group prevent biases that could affect their judgements. A summary description of the method taken from Embry et al (1984): The basic rationale underlying SLIM is that the likelihood of an error occurring in a particular situation depends on the combined effects of a relatively small set of performance shaping factors (PSFs). It is assumed that an expert judge (or judges) is able to assess the relative importance (or weight) of each PSF with regard to its effect on reliability in the task being evaluated. It is also assumed that, independent of the assessment of relative importance, the judge(s) can make a numerical rating of how good or how bad the PSFs are in the task under consideration, where “good” or “bad” mean that the PSFs will either enhance or degrade reliability. Having obtained the relative importance of weights and ratings, these are multiplied together for each PSF and the resulting products are then summed to give the Success Likelihood Index (SLI). The SLI is a quantity, which represents the overall belief of the judge(s), regarding the positive or negative effects of the PSFs on the likelihood of success for the task under consideration. It is assumed that, as a result of their knowledge and experience, the judge(s) have a correct idea of the effects of the PSFs on the likelihood of success, the SLI will then be related to the probability of success that would be observed in the long run in the situation of interest (i.e. the actuarially determined probability). SLIs are transformed into HEPs using a suggested logarithmic relationship of the form, Log p(success) = a (SLI) + b Where a and b are empirically derived constants. In order to produce an empirical calibration relationship between the SLI and the log of the success probability, at least two tasks must be available for which the probability of success is known, in the task set being evaluated. If this is the case, the constants a and b in the above equation can then be used to transform any SLI value produced by the judge(s) into a log probability of success for the task. The log probability of success is readily convertible into the probability of success. An estimate of the HEP or likelihood of task failure, the ultimate goal of SLIM, is found by simply subtracting the success probability from one.

	The bounding HEPs are taken from THERP (Boring & German, 2005) Kirwan (1994) provides practical guidance on applying SLIM, but notes that in many cases strict guidance cannot be given (e.g. with regards to how many PSFs should be used, and what is the minimum/ maximum number of PSFs?) The SLIM procedure goes through the following stages: 1. The selection of the expert panel 2. The definition of situations and subsets 3. The elicitation of PSFs 4. The rating of the tasks on the PSF scale 5. The ideal-point elicitation, and scaling calculations 6. Independence checks 7. The weighting procedures 8. The calculation of the SLIs 9. The conversion of the SLIs into probabilities 10. The uncertainty-bound analysis 11. The sensitivity analysis for error reduction analysis purposes 12. The documentation process The process is more complicated than represented in this list; hence the computerised version using MAUD was developed. SLIM has evolved since its initial development, for example Chien et al (1988) developed a failure likelihood index method (FLIM). Park and Lee (2008) developed AHP-SLIM, a type of "HEP estimation using an analytic hierarchy process (AHP), which quantifies the subjective judgement and confirms the consistency of collected data". SLIM has also been used as the basis for the development of other tools such as HEPI a human error probability index developed for the offshore muster process (Khan et al, 2006).
Validation	Kirwan (1994) notes that the level of accuracy associated with SLIM is, due to the lack of data, indeterminate. However, its theoretical validity is at a reasonably high level, and if the HEPs are calibrated with other, 'known' HEPs, these are likely to fall within the right 'ballpark'. Embrey and Kirwan (1983) carried out a comparative validation of the expert judgement approaches APJ, PC and SLIM. The results showed that APJ had some degree of accuracy but that PC and SLIM needed further development and an improved calibration process. Humphreys (1988) reviewed SLIM (amongst other tools) against a set of criteria (accuracy, validity, usefulness, effective use of resources, acceptability, and maturity). SLIM was as good as any of the other methods reviewed (APJ, PC, TESEO, THERP, HEART, IDA and HCR) with an average rating of 'moderate' against the criteria. The method, like other expert judgement tools, does have supporters and has been developed to address the early problems that were identified with the method.

	In summary, SLIM was developed by recognised experts in the field of human reliability but has received both positive and negative feedback depending on the reviewer.
Domain usage/ applicability to other domains	SLIM has been used in the nuclear and chemical industries (Everdij and Blom, 2008).
Resources required to complete the assessment	SLIM requires an expert panel to carry out an assessment, this should comprise • 2 operators • 1 human factors specialist • 1 reliability analyst Kirwan (1994) notes, "SLIM requires a small panel of experts, but as with any expert group there will be the problem of defining what an expert is.. and the not inconsiderable task of assembling several experts for a few days to make the judgements. These considerations, plus the cost of SLIM and the training required, therefore means that SLIM puts a fairly heavy burden on resources."
Availability of the tool & support	The SLIM element of the tool is publicly available, however Kirwan (1994) notes that using the SLIM method will produce slightly different HEPs to those calculated using SLIM-MAUD because of the mathematics in the MAUD software. The MAUD element of the tool is proprietary to the Decision Analysis Unit of the London School of Economics, but can be purchased for use with SLIM.
References used for the summary	Boring, R., Boring, L. and Gertman, D.I. (2005). Atomistic and holistic approaches to human reliability analysis in the US nuclear power industry. Safety and Reliability, Vol. 25, No. 2, pp. 21 – 37. Chien, S.H., Dykes, A.A., Stetkar, J.W., Bley, D.C. (1988) Quantification of Human Error Rates Using a SLIM-Based Approach. In: Proceedings of the Institute of Electrical and Electronics Engineers 4th Conference on Human Factors and Power Plants, pp. 297-302. Embrey (1983) The Use of Performance Shaping Factors and Quantified Expert Judgment in the Evaluation of Human Reliability: An Initial Appraisal. NUREG/ CR-2986, Brookhaven National Laboratory. Embrey,D.E. and Kirwan, B. (1983) A comparative evaluation of three subjective human reliability quantification techniques. The Annual Ergonomics Society Conference Proceedings, Coombes, K. (ed) Taylor and Francis, London, pp 137-142. Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B. and Rea, K. (1984) SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgment. Volume I: Overview of SLIM-MAUD. NUREG/CR-3518. Prepared for the US NRC Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B. and Rea, K. (1984) SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgment. Volume II: Detailed analysis of the technical issues. NUREG/CR-3518. Prepared for the US NRC Everdij M.H.C. and Blom H.A.P. (2008) Safety Methods Database http://www.nlr.nl/documents/flyers/SATdb.pdf

	Humphreys P. (1988) Human reliability assessors guide, Safety and Reliability Directorate UKAEA (SRD) Report No TRS 88/95Q Kirwan, B. (1994). A guide to practical human reliability assessment. Taylor & Francis, London. Khan, F.I., Amyotte, P.R., and DiMattia, D.G. (2006). HEPI: A new tool for human error probability calculation for offshore operation. Safety Science 44, 313-334. Park, K.S. and Lee, J. (2008) A new method for estimating human error probabilities: AHP-SLIM. Reliability Engineering and System Safety, 93, 578 – 587. Williams, J.C (1985). Validation of Human Reliability Assessment Techniques. Reliability Engineering 11, 149 – 62.
Opinion	
Pros and cons	The following pros and cons are taken from Kirwan (1994): Pros • SLIM is generally a plausible approach. • It allows gross cost-benefit evaluations to take place. • It is a flexible technique Cons • At the time of writing (1994), the choosing of PSFs was somewhat arbitrary. • There is a chronic lack of data with which to calibrate the SLIs. It is reported that Embrey used the Absolute Probability Judgement (APJ) technique to calibrate SLIM. • It requires a panel of experts and it is resource intensive. • Like other expert judgement techniques it is prone to biases, which can significantly undermine the validity of the technique.
Suitability for MH sectors	SLIM is a flexible tool that is essentially a set of procedures for eliciting expert opinion; therefore it is suitable for application in major hazard sectors. The mixed reviews of SLIM suggest that if the method is rigorously applied and the experts are sufficiently monitored to reduce bias, then it can be useful.
Related methods	Some of the SLIM related methods are AHP-SLIM, HEPI, and FLIM.

3.1.10 Human Reliability Management System (HRMS)

Factual information	
Origins of the tool	Developed primarily by Kirwan in the late eighties (e.g. Kirwan & James 1989, Kirwan 1990) to inform the design process for BNFL THORP (British Nuclear Fuels Ltd., Thermal Oxide Reprocessing Plant). JHEDI (Justification of Human Error data Information) was developed alongside HRMS as a quicker screening technique but still based on the HRMS methodology.
Description of the tool	"[HRMS is] a fully-computerised HRA system that contains a human error identification (HEI) module, which is used by the assessor on a previously prepared and computerised task analysis". Kirwan (1994) HRMS is described as a system for managing human reliability and was developed to inform the design process for BNFL THORP It is one of two techniques that were developed: - The Human Reliability Management System (HRMS), and - The Justification of Human Error Data Information (JHEDI) technique, The latter is essentially a quicker screening version of the former. Both techniques can be used to carry out task analysis, error analysis, and performance shaping factor-based quantification, but JHEDI involves less detailed assessment than HRMS. Additionally, HRMS can be utilised to determine error reduction mechanisms, based on the way the performance shaping factors (PSFs) are contributing to the assessed error probabilities. As HRMS is fully computerised the assessments are, therefore, documented and auditable. The method is based on industry error data, which is context specific, and supplemented with expert judgement by the tool's authors. The resulting database of HEPs (human error potential) is combined with a means of extrapolating from the real data to the desired task failure probabilities. This is based on a comparison between PSF profiles for the real, and the to-be-assessed, task or scenario. The methodology is based on other techniques that were available at the time of development. Specifically, the quantification system (PHOENIX, the Prediction of Human Operator Error using Numerical Index eXtrapolation) is based on the Success Likelihood Index Method (SLIM), Influence Diagrams Approach (IDA) and Human Error Assessment and Reduction Technique (HEART). PHOENIX comprises - 7 Task types categories – the assessor selects the task type that most closely resembles the task being assessed, and within each task category there are a number of error types that the assessor could choose from.

	- 6 PSFs - about which the assessor is asked up to 50 questions to determine the strength of impact (this removes the need for subjective judgements by assessors) The HEP associated with the selected task type is multiplied with the PSF value, however the model allows for interactions between PSFs and with the task types, something that previous tools were unable to offer. HRMS also allows for error reduction analysis by providing a sensitivity analysis to determine the factors that, if improved, will maximise the reduction in risk. This is done automatically in HRMS by comparing the original assessed profile with a possible profile including an improved PSF (the most sensitive PSF). Kirwan (1997)
Validation	In 1994, the HRMS method had not been validated and was not commercially available (Kirwan, 1994). It has only been applied to one (nuclear) PSA assessment and no further information was identified to counter this view. Everdij and Blom (2008) state that HRMS is apparently not in current use or else used rarely.
Domain usage/ applicability to other domains	The data underpinning the method is specific to the nuclear industry – "*it can only be used to quantify HEPs for tasks related to its contextual origins, namely nuclear chemical plant operation*." Kirwan (1997).
Resources required to complete the assessment	The HRMS is resource-intensive (takes days to complete), and requires a significant amount of training (Kirwan, 1994).
Availability of the tool & support	HRMS is a BNFL proprietary tool and is not commercially available (Kirwan, 1994).
References used for the summary	Everdij M.H.C and Blom H.A.P, (2008) Safety Methods Database. http://www.nlr.nl/documents/flyers/SATdb.pdf Kirwan, B. and James, N.J. (1989). A Human Reliability Management System. In: Reliability Volume 89 (Brighton Metropole) Kirwan, B. (1990). A resources flexible approach to human reliability assessment for PRA. In: Safety and Reliability Symposium, Elsevier Applied Sciences, pp. 114-135. Kirwan, B. (1994). A guide to practical human reliability assessment. Taylor & Francis, London Kirwan, B. (1997) The development of a nuclear chemical plant human reliability management approach: HRMS and JHEDI. Reliability Engineering & System Safety, Volume 56, Issue 2, Pages 107-133
Opinion	
Pros and cons	Pros • HRMS is useful for scenarios that need to be assessed in depth. • It deals with the whole HRA process from task analysis to error reduction and documentation.

	• Founded on industrial data • PSF rating questions are factual rather than judgemental. • HRMS has the potential to derive or learn extrapolation rules. Cons • It is not empirically validated but has been used successfully in the nuclear industry. • It is resource-intensive. • It is viewed to be a HRA expert's tool, rather than as a general tool for reliability experts.
Suitability for MH sectors	Suitable only for nuclear chemical plant operation and used primarily in design stage assessments.
Related methods	JHEDI (which is a derivative of HRMS).

3.1.11 Justified Human Error Data Information (JHEDI)

Factual information	
Origins of the tool	JHEDI was primarily developed by Kirwan in the late eighties (e.g. Kirwan & James 1989, Kirwan 1990) to inform the design process for BNFL THORP (British Nuclear Fuels Ltd., Thermal Oxide Reprocessing Plant). JHEDI was developed to provide a faster screening technique than that of its 'parent', the HRMS approach, of which it is a derivative. (Kirwan, 1990)
Description of the tool	Both HRMS and JHEDI can be used to carry out task analysis, error analysis, and performance shaping factor-based quantification, but JHEDI involves a less detailed assessment than HRMS. For example, it utilises the same form of task analysis as HRMS, but is shorter, typically involving 10-15 steps instead of the 10-30 sometimes seen with HRMS. The quantification system has also been simplified; the task types are reduced to the contextual ones from HRMS but still retain approximately fifty HEPs and associated descriptions. The HEPs have been made more conservative (i.e. pessimistic) to allow for simplicity within JHEDI (i.e. account for a measure of dependence) and also requires less PSF questions to be answered. JHEDI uses multipliers in a similar way to HRMS but there are less of them and some extrapolation rules are not used. JHEDI is therefore conceptually similar to THERP and HEART, in that both of these techniques have a nominal HEP which is increased according to levels of PSF identified to be present in the situation, via multiplication. As with HRMS, however, the PSF rating process is arguably more straightforward, as it asks factual questions which can be substantiated, rather than asking for subjective judgement by the assessor. Like HRMS, JHEDI is based on actual industry data, which is context specific, and supplemented with expert judgement by the tool's authors.
Validation	Kirwan (1997) carried out a large-scale comparative validation study. Thirty HRA practitioners each assessed thirty tasks (i.e. HEPs for which the true values were known, though not by the assessors). Ten assessors applied THERP, ten applied HEART and ten applied JHEDI. The results showed a significant correlation in each case between the assessed values and the true values. JHEDI is reported to have achieved higher accuracy and precision scores than HEART and THERP, but there was not a statistically significant difference between the three. The study did highlight, however, that the consistency of use could be improved for all three techniques. (Kirwan 1997)
Domain usage/ applicability to other domains	Due to the nature of the underpinning data, JHEDI is only applicable to the UK nuclear chemical industry and, specifically, reprocessing tasks.

Resources required to complete the assessment	The JHEDI system is relatively rapid, and may only take up to half a day for a scenario with several HEPs, assuming that the assessor knows the task requirements reasonably well. JHEDI is a derivative of HRMS, and while HRMS is resource intensive (days to complete), JHEDI is relatively quick and requires little training. (Kirwan, 1994)
Availability of the tool & support	Kirwan (1994) suggests that assessors can access JHEDI but not HRMS, but, like HRMS, JHEDI is a proprietary tool.
References used for the summary	Kirwan, B. (1990). A resources flexible approach to human reliability assessment for PRA. In: Safety and Reliability Symposium, Elsevier Applied Sciences, pp. 114-135. Kirwan, B. (1994). A guide to practical human reliability assessment. Taylor & 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,* 27, (6), 359-373. Kirwan B. (1997) The development of a nuclear chemical plant human reliability management approach: HRMS and JHEDI. Reliability Engineering & System Safety, Volume 56, Issue 2, Pages 107-133. Kirwan, B The validation of three human reliability quantification techniques – THERP, HEART and JHEDI: Part III – Practical aspects of the usage of the techniques, Applied Ergonomics, Vol 28, No 1, pp. 27-39, 1997. Kirwan, B., Kennedy, R., Taylor-Adams, S. and Lambert, B. (1997). The validation of three human reliability quantification techniques, THERP, HEART and JHEDI: Part II – results of validation exercise. *Applied ergonomics*, 28 (1), 17-25. Kirwan, B. (1998). Human error identification techniques for risk assessment of high risk systems - Part 1: review and evaluation of techniques. Applied ergonomics; 1998, v., vv29, no. 3, pp. 157-177
Opinion	
Pros and cons	Pros • JHEDI is a computerised tool and the data can therefore be easily recorded and audited. • Models such as HEART (which has validity) are the basis for the approach. • JHEDI uses real data rather than simulated data. Cons • This is a BNFL proprietary tool and therefore is not publicly available.

	There is insufficient information to comment further on the pros and cons of JHEDI.
Related methods	HRMS (JHEDI is a derivative of HRMS).

3.1.12 INTENT (not an acronym)

Factual information	
Origins of the tool	Gertman et al (1990) presented INTENT as a method for estimating human error probabilities for decision based errors.
Description of the tool	Gertman et al (1990) view errors of intention as an important subset of errors of commission because they are related to cognitive functions (e.g. problem solving). The fact that errors of intention can result from a wide range of factors (e.g. poor training) makes it difficult to model and quantify them. Their aim was to incorporate errors of intent into probabilistic safety assessment. In order to achieve this, Gertman et al compiled a number of potential errors of intention pertinent to nuclear power plants, by using a variety of analytical techniques. Four categories of error of intention were identified: (1) Action consequence - these are errors of intention that take account of the relationship between consequences and decision making. (2) Crew response set – these errors reflect the influence that inhibition, experience and training have on performance. (3) Attitudes leading to circumvention – this category includes errors that are rooted in the manner in which individuals view the world. (4) Resource dependencies – this category is made up of internal resources (e.g. memory capacity) and external resources (e.g. operating procedures). It is these four categories that form the basis of an INTENT assessment. The INTENT user is offered a choice of 20 nominal errors from these four categories. For each error, INTENT gives lower bound and upper bound estimates of the occurrence probability, which are based upon expert opinion. INTENT also includes a set of eleven (very brief and general, e.g. workload) performance shaping factors (PSFs) whose weighting factors were also determined by expert estimates. Once the user of the INTENT method has selected the errors (for which probabilities are to be estimated) from among the 20 nominal errors, estimates for the 11 PSFs are then given on a 5-point scale. Taking the weighting factors into account it is then possible to calculate a reliability index from these estimates. In summary, INTENT involves six stages: 'Compiling errors of intention, quantifying errors of intention, determining human error probabilities (HEP) upper and lower bounds, determining performance shaping factors (PSF) and associated weights, determining composite PSF, and determining site specific HEP's for intention'. (Everdji and Blom, 2008).

Validation	INTENT has not been empirically validated, however Kirwan (1998) states that '…the approach is essentially experience based: incident experience tempered with assessor experience. This means it has a certain degree of context validity'.
Domain usage/ applicability to other domains	INTENT was developed for the nuclear industry; there is no evidence of it being applied to other sectors.
Resources required to complete the assessment	Gertman et al (1990) do not explicitly identify the training, numbers of assessors, level of expertise, data and information requirements.
Availability of the tool & support	The references do not provide any information on the availability of INTENT or provide sufficient information to apply the method without further information.
References used for the summary	Everdij M.H.C., Blom H.A.P. (2008) Safety Methods Database. Gertman, D. I., Blackmann, 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 & System Safety 35 pages.127-136 Kosmowski, G. Degen, J. Mertens, B. Reer (1994) Development of Advanced Methods and Related Software for Human Reliability Evaluation within Probabilistic Safety Analyses. Published by Berichte des Forschungszentrums Julich; 2928, Institut fur Sicherheitsforschung und Reaktortechnik. ISSN 0944-2952
Opinion	
Pros and cons	Pros • Kirwan (1998) states that INTENT and similar methods are relatively easy to use whether by a novice or more experienced practitioner. • Kosmowski (1994) states that INTENT may provide data that can account for rare, high consequence failures due to errors of intention. Cons • Kirwan (1998) notes that while appearing simple and straightforward, methods like INTENT rely on the skill of the assessor, and the degree to which the assessor understands the task being assessed. • Kosmowski (1994) states that the generated list of errors in INTENT may not be exhaustive, but it does provide a foundation on which to build a more complete database as field data becomes more available.
Suitability for MH sectors	Unclear, as there is nothing contained within any of the references to suggest that INTENT is suitable for application in any other domain apart from nuclear.

3.1.13 NARA (Nuclear Action Reliability Assessment)

<table>
<tr><th colspan="2">Factual information</th></tr>
<tr><td>Origins of the tool</td><td>NARA (2005) was developed for the nuclear power company, British Energy by a consortium of Human Reliability Analysis (HRA) experts:
Kirwan, B. (Eurocontrol)
Gibson, H., & Kennedy, R. (University of Birmingham)
Edmunds, J., Cooksley, G. (Corporate Risk Associates)
Umbers, I. (British Energy)</td></tr>
<tr><td>Description of the tool</td><td>NARA has been developed using the HEART methodology as its basis, using more recent data, and tailored to the needs of UK Nuclear Power Plant (NPP) Probabilistic Safety Assessments (PSAs) and HRAs (Kirwan et al, 2005).

Like HEART, NARA consists of Generic Task Types (GTTs) and Error Producing Conditions (EPCs). The key elements of the approach are: Classify the task for analysis into one of the Generic Task Types and assign the nominal Human Error Probability (HEP) to the task. Decide which EPCs may affect task reliability and then consider the assessed proportion of affect (APOA) for each EPC. Then calculate the task HEP.

The same formula, used in HEART, is used in NARA for deriving the HEP:

$HEP = GTT \times \{[EPC1 - 1) \times APOA1 + 1] \times [(EPC2 - 1) \times APOA2 + 1] \times \ldots [(EPCn-1) \times APOAn+1]\}$

The list of GTTs contained in NARA is partly a sub-set of the original HEART GTTs and partly a further refinement of GTT definitions to more accurately encompass the actions being considered in nuclear PSAs. These GTT data were then contrasted with data within CORE-DATA (a database of HEPs), additional data available to the project team but not yet included in CORE-DATA and also the original HEART data sources were reviewed.

A review of the HEART EPCs was undertaken and resulted in several EPCs remaining unchanged from the original HEART method, while others have been modified (e.g. increasing or decreasing their maximum effect).

There are 14 GTTs and 18 EPCs quantified in NARA.

New features of NARA (as compared with HEART) are;
An approach to quantifying operator reliability in relation to long time-scale events.
A prototype approach to error of commission quantification.
Work is ongoing in the area of determining dependence approaches for NARA applications.
More guidance has been developed for use of the APOA process.
Benchmarks for usage of the technique are being developed.</td></tr>
</table>

	All of the above have been taken directly from Kirwan (2005).
Validation	The HEART approach, which underpins NARA, has been empirically validated several times. NARA itself has not yet been validated but Kirwan (2005) reports that NARA itself has been presented in two workshops. The first was to British Energy HRA assessors and PSA specialists, and the second to the Nuclear Installations Inspectorate, which included other interested parties (such as representatives from other parts of the nuclear industry). The impact of using NARA in a PSA compared to HEART has been evaluated. Kirwan & Gibson (2007) report that NARA has been "successfully peer reviewed by the nuclear regulator and industry, and independent HRA experts in a formal peer review process". However, no published details of this peer review have been found.
Domain usage/ applicability to other domains	NARA is nuclear specific. It is a British Energy proprietary tool and is a nuclear specific modified version of HEART. Therefore it is not applicable to other domains.
Resources required to complete the assessment	The resources required to complete a NARA assessment are not specified in Kirwan (2005), however, it is anticipated that they will be comparable to those required for HEART. HEART is a quick method to apply and is easily understood.
Availability of the tool & support	NARA is a British Energy proprietary tool and is not publicly available.
References used for the summary	Kirwan, B., Gibson, H., Kennedy, R., Edmunds, J., Cooksley, G. and Umbers, I. (2005), Nuclear action reliability assessment (NARA): a data-based HRA tool. Safety & Reliability, vol. 25. no. 2 pp 38 – 45. (2005) Kirwan, B and Gibson, H (2007) CARA: A Human Reliability Assessment Tool for Air Traffic Safety Management — Technical Basis and Preliminary Architecture. Pp 197 –214 in The Safety of Systems Proceedings of the Fifteenth Safety-critical Systems Symposium, Bristol, UK, 13–15 February 2007.
Opinion	
Pros and cons	NARA is based on HEART, which is an established, validated method. The method is nuclear specific. There is insufficient information available to comment further.
Suitability for MH sectors	NARA is nuclear specific and therefore not suitable for other sectors.
Related methods	HEART and CARA.

3.1.14 Connectionism Assessment of Human Reliability (CAHR)

Factual information	
Origins of the tool	CAHR was developed at the Technical University of Munich and the Gesellschaft fur Anlagen-und Reaktorsicherheit (GRS) between 1992 and 1997 (Sträter, 1997). Sträter has since moved to EuroControl and is developing CAHR for use in Air Traffic Management (ATM).
Description of the tool	CAHR is a second-generation tool that combines event analysis and assessment in order to use past experience as the basis for human reliability assessment. It is a database system used for analysing 'operational disturbances', which are caused by inadequate human actions or organisational factors. CAHR has a generic underlying model that is applicable to all observable events and to allow the collection of all information on human error from events. The information is stored in a database that contains a generic knowledge base for the event analysis that is extendable by the description of further events. The knowledge base contains information about the system state and the tasks as well as for error opportunities and performance influencing factors (PIFs). There are three key elements to the tool (Sträter, 2000): - A framework for structured data collection (both retrospective and prospective information) - A method for qualitative analysis - A method for HRA (quantitative analysis). The philosophy underlying this tool is: - The focus of analysis is the working system and not the human. - Human error results from the interrelation of several situational and causal factors of the working system. - The method uses a fixed structure but no fixed taxonomy. - Strict differentiation between observable information (phenotypes) and causes (genotypes) in the event analysis and description. The method is underpinned by a connectionism algorithm which is a term coined by modelling human cognition on the basis of artificial intelligence models. It refers to the idea that human performance is affected by the interrelation of multiple conditions and factors rather than singular ones that may be treated in isolation (Everdij and Blom, 2008).
Validation	Sträter reports that the data obtained by event evaluation (i.e. now contained in the CAHR database) has been compared to human reliability data provided in the THERP data tables, in French PSA studies (Le Bot, 2004) and also the INTENT method (Gertman et al, 1992). The comparison with the THERP data revealed "good numerical agreement of the data from practical operational experience and the data from the THERP catalogue" (Sträter, 2000). As the method relies on detailed descriptions of events, Linsenmaier (2006) is reported to have investigated the validity of the event analysis process

	with favourable results. Similarly, Theis (2002) is reported to have shown that cross domain assessments are possible i.e. that data collected in one domain can validly be used for the assessment of human behaviour in a different domain. There is no published evidence of the CAHR being used by anyone other than Sträter and colleagues.
Domain usage/ applicability to other domains	CAHR was developed to be applied in the nuclear industry but has been applied to various areas including occupational health and safety, shipping, car industry, aviation safety and software ergonomics.
Resources required to complete the assessment	One to two days of training/ familiarisation is required for each of the two elements in CAHR. The effort required to use the two elements of the tool is as follows: 1. Analysing events mode – this takes about 1-2 hours per event (on average) to put into the CAHR database. 2. For the prospective (HRA mode) – little effort is required once the task has been specified, the error mode identified and a database query made (5~10 min per task).
Availability of the tool & support	The CAHR website includes a basic demonstration of the tool. CAHR is available by contacting the author directly to negotiate access to CAHR.
References used for the summary	CAHR website http://www.cahr.de/tools/CAHR.htm. Dang, V.N., Reer, B., Sträter, O., Hirschberg, S. (2000) A Comparison of Emerging methods for Errors of Commission Based on Applications to the Davis-Besse (1985) Event. In: Kondo, S & Furua, K. (eds) PSAM 5 – Probabilistic Safety and Management. Universal Academy Press. Tokyo, Japan. Everdij M.H.C. and Blom H.A.P. (2008) Safety Methods Database http://www.nlr.nl/documents/flyers/SATdb.pdf Gertman, D. I., Blackmann, 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 & System Safety 35 pages.127-136 Le Bot, P. (2004). Human reliability data, human error and accident models - illustration through the Three Mile Island accident analysis. *Reliability Engineering & System Safety*, 83, 153-167. Linsenmaier, B & Sträter, O. (2000). Recording and Evaluation of Human Factor Events with a View to System Awareness and Ergonomic Weak Points within the system, and the Example of Commercial Aeronautics. In: HFES (Ed), Human Factors and Ergonomics Society, Santa Monica CA, Mira Digital Publishing: South Jefferson, St. Lois, MO. Sträter, O. (2000). *Evaluation of Human Reliability on the Basis of Operational Experience.* (Koln/ Germany: GRS 2000) Included on the

	CAHR website. (A translation of the German report ref: Report GRS-138, published 1997). Sträter, O. (2006) The use of incidents for human reliability management. Safety & Reliability, Vol. 26, No. 2 pp 26 – 47. US NRC NEA/CSNI/R(2002)3. Proceedings of the International Workshop, Building the new HRA: Errors of Commission from research to application (2002) US NRC, Rockville, Maryland, USA, 7-9 May 2001. NEA/CSNI/R(2002)3. Specific paper: Overview about the CAHR Method and its Application in Assessing Errors of Commission presented by Sträter, pp. 117 – 129.
Opinion	
Pros and cons	Pros – CAHR is one of a number of tools that attempt to collate information from previous events and build an extensive database. It is an attempt to move towards analysing errors of commission and to capture the complexity of human behaviour. Cons – Dang et al (2000) carried out a comparison of emerging methods for errors of commission-based applications (including CAHR); he suggested that there was still work to be done to establish a proven, practical methodology for analysing potential errors of commission. There is no published evidence that this situation has changed, with regard to CAHR, since that time.
Suitability for MH sectors	Though developed to inform nuclear PSAs, the underlying method is reported to be generic and is, therefore, suitable for application in other major hazard sectors.
Related methods	CAHR is a stand-alone database method.

3.1.15 Commission Errors Search and Assessment (CESA)

Factual information	
Origins of the tool	Sträter, Dang and Hirschberg (1999) undertook a review of emerging Errors of Commission (EOC) methods, as a joint project performed with Gesellschaft fur Anlagen-und Reaktorsicherheit (GRS) and the Paul Scherrer Institut (PSI). One outcome of this work was that the PSI proposed an EOC identification method. The result was CESA, which integrates aspects from the search schemes of some HRA methods (e.g. ATHEANA) with concepts from CODA.
Description of the tool	The CESA method is strongly based on importance screening. The identification process prioritises plant systems with a high achievement worth and scenarios are prioritised based on the size of the contribution to the core damage frequency. In this sense a trade-off is made between the scenarios with a high safety impact against the completeness of the search. The intention is to bias the search towards EOC situations that are risk-significant and credible. (Reer et al 2004). The first step of the CESA method is to catalogue key action responses to the plant events to be reviewed. This catalogue is then used in a systematic search of context-action combinations, to obtain a set of situations with EOC opportunities; these situations are then analysed in detail (Reer and Dang, 2006).
Validation	To date, no validation work has been published. However, Reer and Dang (2006) state that future work will address the verification of the method data, the elaboration of a user manual, and a large-scale pilot application.
Domain usage/ applicability to other domains	CESA has been applied in the nuclear industry (Everdij and Blom, 2008)
Resources required to complete the assessment	From the available references, it is not possible to identify what resources are required to carry out a CESA assessment.
Availability of the tool & support	A description of the CESA methodology is publicly available but the necessary details for carrying out an assessment are not.
References used for the summary	Dang, V.N., Reer, B., Sträter, O., Hirschberg, S. (2000) A Comparison of Emerging methods for Errors of Commission Based on Applications to the Davis-Besse (1985) Event. In: Kondo, S & Furua, K. (eds) PSAM 5 – Probabilistic Safety and Management. Universal Academy Press. Tokyo, Japan. Reer, B. and Dang V.N. (2006) A Method for Quantifying Errors of Commission on the Basis of Operational Event Data. http://safe.web.psi.ch/documents/pdfs/EOC_HRA_Meth_AnnexPaper-2006-01-25.pdf Reer B., Dang V.N. and Hirschberg S. (2004) The CESA method and its application in a plant-specific pilot study on errors of commission. Reliability Engineering & System Safety. Volume 83, Issue 2, Pages 187-205.

	Everdij M.H.C. and Blom H.A.P. (2008) Safety Methods Database http://www.nlr.nl/documents/flyers/SATdb.pdf
Opinion	
Pros and cons	As CESA is a relatively new method there were very few references available and no critical review. However, the following comments were identified. Pros - Reer and Dang (2006) found that the CESA identification process is feasible and effective: it is able to identify plausible situations in which EOCs may occur. Cons – Reer et al (2004) report that the quantification of the risk contribution using CESA is uncertain and this uncertainty is larger than would be typical of other HRA methods.
Suitability for MH sectors	CESA has been used in the nuclear industry (Everdij and Blom, 2008), but given the integration of search schemes from more generic tools (e.g. ATHEANA) it may be suitable for other major hazard sectors.
Related methods	CODA was identified as a related HRA method (Reer, Dang and Hirschberg, 2004).

3.1.16 Conclusions from Occurrences by Descriptions of Actions (CODA)

Factual information	
Origins of the tool	CODA was first outlined in a conference paper by Reer (1997).
Description of the tool	The CODA method uses an open list of guidelines based on insights from previous retrospective analyses. The general approach is to compile a short story that includes all unusual occurrences and their essential context without excessive technical details. The analysis should then focus on the potential major occurrences first (Everdij and Blom, 2008). The method presents a list of criteria (i.e. items for data acquisition) that are easy to obtain (i.e., objectively, as free from judgment as possible) and which have been proved to be useful for causal analysis. For example, for each incorrect human response that has occurred the analyst will look for: the critical (committed or omitted) action; the underlying goal or plan if it is self-evident; the anticipated correct response and its consequence; the underlying task and sub-task; the underlying sequence of events. Various guidelines are provided for the causal analysis of each situation; these are mainly holistic, comparative and generalizing in nature. For instance, it is recommended to consider an occurrence not separately, but within the context of a 'wide-enough' defined sequence of events; similar situational patterns from other events or findings; or the common presence of several items. A number of event cases have been used to demonstrate that CODA is able to identify cognitive tendencies (CTs) as typical attitudes or habits in human decision-making (Reer, 1997).
Validation	No empirical validation or peer review papers were identified in relation to CODA.
Domain usage/ applicability to other domains	Everdij and Blom (2008) report that CODA has been used in the nuclear industry.
Resources required to complete the assessment	Reer (1999) provides guidance for assessors in terms of a list of key term definitions, a step-by-step procedure for structuring the analysis together with rules for information gathering and an extensible taxonomy of CTs that drive human behaviour. However, there is no information on other aspects of the resources required (e.g. the number of assessors).
Availability of the tool & support	A description of the CODA methodology is publicly available but the necessary details for carrying out an assessment are not. CODA may be available from Dr Bernhard Reer, Paul Scherrer Institut, CH-5232 Villigen PSI, Switzerland.
References used for the summary	Everdij M.H.C. and Blom H.A.P. (2008) Safety Methods Database http://www.nlr.nl/documents/flyers/SATdb.pdf Reer B. (1997), "Conclusions from Occurrences by Descriptions of Actions (CODA)", in: B.M. Drottz Sjöberg (Ed.), New Risk Frontiers, Proceedings of the 1997 Annual Meeting of the Society for Risk Analysis-Europe, Stockholm.

	Reer B. (1999), "Retrospective Event Analysis Using CODA and Perspectives on Human Error Prediction", in: Modarres M. (Ed.), Risk-Informed Performance-Based Regulation. Proceedings of the International Topical Meeting on Probabilistic Safety Assessment (PSA '99), 22-26 Aug. 1999, Washington DC, USA. American Nuclear Society. LaGrange Park, Illinois, USA. Vol. 2, 1337-1345.
Opinion	
Pros and cons	Very few references were identified that related to CODA, however, the following comments were identified in Sträter et al (1999) Pros - • As CODA is partly an action-centred approach it has the advantage of treating human interventions in a neutral and flexible manner. • CODA incorporates three search processes: actions, system-failures and scenarios. By integrating these approaches the search procedure may be optimised towards low analytical effort and high completeness. Cons – It was not possible to identify any cons as part of this literature search.
Suitability for MH sectors	It has not been possible to determine whether CODA is suitable for major hazard sectors or not.
Related methods	CESA is related to CODA (Reer, Dang and Hirschberg, 2004).

3.1.17 Méthode d'Evaluation de la Réalisation des Missions Opérateur pour la Sûreté (MERMOS).

Factual information	
Acronym and full name	Méthode d'Evaluation de la Réalisation des Missions Opérateur pour la Sûreté (**MERMOS**). In English: Assessment method for the performance of safety operation.
Origins of the tool	MERMOS was developed for Electricité de France (EdF). Le Bot et al (1997) first outlined the method and subsequent papers provide more detail for example Bieder et al (1998) and Le Bot et al (1999).
Description of the tool	Le Bot et al (1999) – MERMOS is a Human Reliability Assessment Method. It is reported to be an improvement of EdF's previous methods, and is designed to guide EdF's analysts in taking human factors aspects into account in the 'level 1' Probabilistic Reliability Assessment (PRA) for units of the N4 series (the most recent type of French reactors). The method only considers emergency operation during the four hours after the incident initiator, because it is assumed that four hours after the initiator the crisis support team will prevent or recover any human failure. The main underlying concept of MERMOS is the 'human factor mission'. For each initiator [of an emergency], a functional analysis will determine the 'missions' that have to be performed to recover or mitigate the accident. The human factors mission refers to safety critical actions that the operating system has to initiate and carry out to handle the situation. MERMOS considers that the performance of the human factors mission is the responsibility of what is termed the "emergency operations system" (EOS) – this comprises the operating crew, operating procedures, the man-machine interface, the formal organization and the workplace. Instead of assuming that human error is a decisive element of failure, in MERMOS it is embedded within the EOS as one of the determinants of inadequate performance. Other underlying concepts include: • CICAs – the actual functioning of the system is modelled with the help of a concept, named CICA (Important Characteristics of Emergency Operation). CICAs refer to particular ways of operating the plant adopted by the EOS in the course of the emergency situation. It takes into account organizational aspects such as task prioritisation or distribution among people and systems. • Strategy, Action, Diagnostic - the three functions involved in the performance of HF missions assumed by the EOS. The method is divided into two modules: **Module 1** – Identification and definition of the HF mission through functional analysis.

	From a functional analysis of the state of the plant after the initiator, the analyst describes the characteristics of each HF mission and their context in a standard form. It is necessary to verify it and enrich it with elements from other information sources such as emergency operating procedures and simulator feedback. **Module 2** – Qualitative and quantitative analysis of the HF missions. The aim of the MERMOS qualitative analysis is to identify as many scenarios as possible leading to the HF mission failure. The mission failure event occurs if one of the scenarios described in the quantitative analysis occurs and leads to failure. The failure scenarios of a mission are a set of events and these sets are exclusive. The **probability of a mission failure is therefore the sum of the probabilities of occurrence** of the failure of each of the scenarios described for a mission. The probability of occurrence of the mission failure, according to a given scenario, can be worked out from the probability of occurrence of each event in the scenario. The failure of an HF mission is analysed as the result of identifiable and non-identifiable scenarios. P (failure of HF mission) = $\Sigma_{\text{identifiable sc.}}$ P(identifiable scenario failure) + Pr Where Pr = Pr.action + Pr.diagnostic + Pr.strategy Occurrence of a failure scenario is governed by the simultaneous existence of associated CICAs and of the situational features determining the failure of the mission, once the CICAs have been initiated. However, once the relevant CICAs have been brought together, the failure scenario is not certain to occur: PSC/ CICAs /Cica designates its probability of occurrence. **The probably of occurrence of failure scenarios for a given HF mission can be calculated as follows**: P(failure scenario) = $P_{\text{SC/CICAs}} \times P_{\text{CICAs / SITU}} \times P_{\text{SITU}}$ $P_{\text{SC/CICAs}}$ designates the probability of appearance of the scenario, given that the associated CICAs are all present $P_{\text{CICAs / SITU}}$ designates the probability of simultaneous existence of the CICAs, given that the associated characteristics of the situation are all present. This can be assessed by statistics from simulator tests. P_{SITU} designates the probability of presence of the properties of the situation participating in the appearance of CICAs To make the method more user friendly to the analysts, the developers are building a database of HF missions and failure scenarios.

Validation	MERMOS has been through an EdF validation process but no reviews or evidence of extensive use were identified within the current literature search. External consultants were employed to comment on MERMOS during its development and made recommendations for improving the method, for example, by simplifying the quantifications that were incorporated in the method. Bieder, et al (2000), reported on work to implement MERMOS as part of a PSA and involved 200 HF missions. They reported that, "The probabilities calculated with the MERMOS process proved consistent with previous HRA results." "The main lesson learnt… MERMOS is not intrinsically more difficult to implement than other methods, but its implementation requires a certain expertise, particularly in the field of real or at least simulated emergency operations (beyond the knowledge of procedures)." Le Bot, et al (2002) Provides an overview of the methodological validation of MERMOS and concludes that it has been through both a theoretical validation (i.e. scientific validation) and industrial validation but that, "No objective and universal criterion is available to assess a PHRA method, and validation must still be performed case by case based on criteria which remain to be defined with respect to the objective of the implementation of the method". Le Bot, P. (2004). Performed a retrospective analysis of the Three Mile Island (TMI) accident using MERMOS to illustrate the modelling of operation and reliability that is proposed by the method. He concluded that, "The retrospective analysis of the TMI accident shows that the operators' mode of operation resulted from an overall logic, reflecting emergency operation logic upon which the design of the operating system was based". This emphasises the shift that MERMOS methodology makes from focusing on individual human error, as many 1st generation tools did, onto the operating system.
Domain usage/ applicability to other domains	MERMOS has only been applied to the nuclear industry.
Resources required to complete the assessment	Le Bot et al (2002) report that expertise in prescribed and actual (or simulated) emergency operation, or in normal operation, is not only necessary to perform the analyses, but ensures their validity as well. The average resources required for MERMOS are: • Standard mission (first analysis) - 4 days • Mission reusing a standard mission - 1 day • Collection of data and description of the mission - 1 day • Quantification -1 day.
Availability of the tool & support	MERMOS is a proprietary tool, however, representatives from the HSE might be able to access to the tool on application to EdF.

References used for the summary	Bieder, C., Vidal, S. and Le Bot, P. (2000). Feedback from the actual implementation of the MERMOS method. PSAM 5 – Probabilistic Safety Assessment and Management pp. 1529 – 1534 (Universal Academy Press, Inc.) Le Bot, P., Cara, F. and Bieder, C. (1999). A second generation HRA method: What it does and doesn't do. In: (Washington, DC: American Nuclear Society), Volume II, pp. 852-860. Le Bot, P., Pesme, H., and Ruiz, F. Methodological Validation of MERMOS by 160 Analyses. PSA '02. 2002 International Topical meeting on Probabilistic Safety Assessment. October 2002, Detroit, Michigan. Le Bot, P. (2004). Human reliability data, human error and accident models - illustration through the Three Mile Island accident analysis. *Reliability Engineering & System Safety*, **83**, 153-167. Le Bot, P., Desmares, E. Bieder, C. Cara, F., Bonnet, J-L. (1997). MERMOS: An EdF project to update the PHRA (Probabilistic Human Reliability Assessment) methodology, OECD Nuclear Energy Agency Specialists Meeting on Human Performance in Operational Events, Chattanooga, USA (October 13-17 1997)
Opinion	
Pros and cons	Pros • A second-generation tool that attempts to deal with important underlying concepts of HRA. It moves away from focusing on individual error, instead considering the operating system as a whole. • EdF report good results from using the tool. Cons • The tool is an EdF proprietary tool and therefore not freely available for review and use. • It is currently only for emergency operations and does not propose a normal operation model which is needed to assess the performance of operators in non-emergency situations. • As a relatively new, second generation, proprietary tool, its validity and reliability have yet to be established.
Suitability for MH sectors	MERMOS is a nuclear specific tool and, therefore, is not applicable to other major hazard sectors.
Related methods	MERMOS is an update of EdF's PHRA (Probabilistic Human Reliability Assessment) approach.

3.2 LIST OF TOOLS CONSIDERED NOT USEFUL TO HSE MAJOR HAZARD DIRECTORATES

The following table is provided as a brief explanation of why 18 methods were not considered to be relevant to HSE major hazard directorates.

Table 3: A list of those tools not of use to the HSE major hazard directorate

Tool	Date	Reason for exclusion	Availability	Domain
AIPA	1974	Accident Initiation and Progression Analysis. The references relate to work in the 70's and there is no evidence of the approach being used or developed since that time. It is only of interest in a historical sense. Swain (1989) thought it was not acceptable for PSA use.	The tool is no longer used	Nuclear
CARA	2007	Controller Action Reliability Assessment. The air traffic management specific version of HEART (different author to the original); this is not applicable to the work of HSE.	Not publicly available	Air traffic mgmt
CES	1987	A computer simulation method that could potentially be useful to HSE but there is no evidence of use or development since the late 80's.	Not publicly available	Nuclear
CM	1981	Confusion Matrix. An expert judgement tool that has value as a qualitative analytical tool rather than a quantitative one. Other expert judgement tools are better than CM.	Publicly available	Nuclear
COGENT	1992	COGnitive EveNt Tree. This is an extension of the THERP event tree model and is of little use in itself. It was developed to bring cognitively based approaches into the human error identification process as an attempt to enhance first generation HRA approaches. Various versions have been developed over the past 16 years.	Available via the COGENT website	Generic
COSIMO	1992	A computer simulation method that could be potentially useful to HSE but like CES, there is no evidence of development since the early 90's	Not publicly available	Nuclear
DREAMS	1993	Dynamic Reliability Technique for Error Assessment in Man-machine Systems. Only one published paper was identified in the literature search and there is no evidence of the approach being used or developed since. There is insufficient information to make an informed decision about the usefulness of this method.	Not publicly available	Nuclear
DNE	1983	An expert opinion tool that seems to be another name for APJ and it would be duplication to also include DNE. In addition, there is insufficient information available,	A publicly available method but	Generic

Tool	Date	Reason for exclusion	Availability	Domain
		compared with APJ, to write a meaningful summary. APJ has been included in Section 3	very little information	
FACE	1999	Framework for Analysing Commission Errors. Only one published paper was identified in the literature search and there is no evidence of the approach being used or developed since. There is insufficient information to make an informed decision about the usefulness of this method.	Not publicly available	Nuclear
HCR	1984	Human Cognitive Reliability. Its developers, EdF, no longer support this approach. PHRA is an update of HCR, which in turn has been superseded by MERMOS.	No longer available	Nuclear
HORAAM	2000	Human and Organisational Reliability Analysis in Accident Management Only one published paper was identified in the literature search and there is no evidence of the approach being used or developed since. There is insufficient information to make an informed decision about the usefulness of this method.	Proprietary tool	Nuclear
MAPPS	1985	Maintenance Personnel Performance Simulation. There is no evidence of development since the 80's and little evidence of application. The review papers suggest that the validity of the method is questionable.	Unclear. It may be available via USNRC	Nuclear
OATS	1981	Operator Action Tree System Swain (1989) reported that there was no evident that OATS has predictive or convergent validity, but that face validity was good. Elements of the method are based on time reliability curves, which have been invalidated in two studies.	Not publicly available	Nuclear
OHPRA	1993	Operational Human Performance Reliability Analysis. Only one published paper was identified in the literature search and there is no evidence of the approach being used or developed since. There is insufficient information to make an informed decision about the usefulness of this method.	Not publicly available	
PHRA	1990	The EdF PHRA approach has been superseded by MERMOS and was similarly, an EdF proprietary tool.	Proprietary tool	Nuclear
SHARP	1984	Systematic Human Action Reliability Procedure A model for the HRA process, not a tool itself. Potentially useful but there is no evidence of use or development since the 80's.	Unclear. It may be available via NUS Corporation	Generic

Tool	Date	Reason for exclusion	Availability	Domain
STAHR	1982	Socio-Technical Assessment of Human Reliability An expert judgment method. Only one published research paper was identified in the literature search. Review papers indicated that STAHR is not sufficiently consistent or valid to use for HRA purposes.	Unclear. It may be available via USNRC	Nuclear and offshore
TESEO	1980	Tecnica empirica stima errori operatori Authors of review papers question the theoretical background of this method, and it is not considered to be accurate.	Not publicly available	Chemical/ nuclear

3.3 CONCLUSION

The aim of this review was to provide HSE with a summary of the HRA literature in order to bring HSE up to date with developments in the field of HRA methods, to have knowledge of the capability of the tools and an understanding of their strengths and weaknesses. In addition, the project was designed to improve consistency across major hazard directorates where appropriate and form a view on the 'acceptability' of such tools for use in risk assessments.

The review identified 17 tools that are considered to be of use to HSE major hazard directorates. Most of these tools are established methods and only a minority of 'new' tools were identified. The 'new' tools that are emerging, and referred to as third generation tools, are developments of first generation methods with the addition of specific industry data. While all the tools included in this review have recognised limitations, there are no significant objections to any of the tools, all of which can provide a useful insight to risk assessment.

Some of the selected tools are nuclear specific, however most of them are generic tools and can be applied to any sector. There is no need, therefore, to distinguish between tools for the different sectors. However, different tools may be appropriate depending on the 'maturity' of the site with regard to quantified human risk assessment. For example, first generation tools may be most appropriate for those sites just beginning to consider the quantification of human risks because, while they may not give information on issues such as dependency or errors of commission, they will give a basic insight to the issue. Second generation tools may be more useful for those sites that have historically used first generation methods and now need more insight to the risk (e.g. those in the nuclear sector). Currently, only one third generation tool was identified as being relevant to HSE major hazard directorates and it is nuclear specific.

In summary, this document is an information resource on human reliability assessments (HRA) based on published research material and the opinion of the authors. It provides a summary of those tools and methods considered to be of potential use to analysts undertaking a HRA assessment in the major hazard sector. It is not intended to be a comprehensive guide to HRA but a useful starting point on which to build knowledge.

4 APPENDIX A

Table 4: A summary of those tools considered to be outside the project scope

Tool	Reason for exclusion	In Full (where available)
3D-SART	Not a HRA method	3D Situation Awareness Rating Technique
ACWA	Not a HRA method	Applied Cognitive Work Analysis
Adaptive user model	Not a HRA method	Not an acronym
AEA	Insufficient information in the timescale.	Action Error Analysis
AEMA	Insufficient information in the timescale.	Action Error Mode Analysis
AHP-SLIM	Included in the review as part of SLIM-MAUD	Analytic Hierarchy Process - Success Likelihood Index Methodology
Air-MIDAS	Not a HRA method	Air Man Machine Integrated Design and Analysis System
ASP	Not a HRA method	Accident Sequence Precursor
CBHRA	Insufficient information in the timescale.	Condition-based human reliability assessment
CESA	Insufficient information in the timescale.	Commission Errors Search and Assessment
CORE-DATA	Not a HRA method	Computerised Human Error Data Base
COSIMO	Insufficient information in the timescale.	
CREWSIM	Not a HRA method	
DETAM	Insufficient information found in the timescale.	
DIAS	Insufficient information in the timescale.	Dynamic Interaction Analysis Support
EOCA	Insufficient information in the timescale.	Errors of commission analysis
HECA	Insufficient information in the timescale.	Human error criticality analysis
HEPI	Insufficient information in the timescale.	Human error probability index
HERA	Not a HRA method	Human Event Repository and Analysis
HERA-PREDICT	Insufficient information in the timescale.	
HEROS	Insufficient information in the timescale.	The human error rate assessment and optimizing system
HERTES	Insufficient information in the timescale.	Human Error Reduction Technique for Engineering Systems
HPM	Insufficient information in the timescale.	Human performance modeling
HRAET	Insufficient information in the timescale.	Human Reliability Analysis Event Tree
IDA	Insufficient information in the timescale.	
IDACrew	Insufficient information in the timescale.	
MONACO	Information to suggest this is part of the MERMOS work, but no details were found.	
MSFM	Swain classifies this as a part-HRA method and insufficient information was found to suggest it should be included in the review.	Multiple Sequential Failure Mode

Tool	Reason for exclusion	In Full (where available)
PROCRU	Not a HRA method	Procedure-Orientated Crew Model
QMAS	Not a HRA method	A qualitative quality management assessment system
SAINT	Swain states that this is not a HRA method but is a structure for a computer modal of human performance.	Systems Analysis of Integrated Networks of Tasks
SCHEMA	Not a HRA method	System for critical human error management and assessment OR Systematic critical human error management approach
SLIM	Included in the review as SLIM-MAUD	Success Likelihood Index Methodology
STARS	Insufficient information in the timescale.	
SYRAS	Insufficient information in the timescale.	Quantitative System Risk Analysis System
TALENT	Not a HRA method	Task Analysis-Linked Evaluation Technique
TRACEr	A human error identification method, not a HRA method	

5 APPENDIX B

A brief summary of each of the 17 methods considered to be of use to HSE major hazard directorates was prepared. The following table provides a definition of each of the review elements.

Table 5: Definition of each review element

Factual information (information collected from published literature)	
Acronym and full name	
Origins of the tool	Author/ institution/ company/ date
Description of the tool	Author declared basis for the tool • What the author claims to offer with the tool • Mechanical basis for the tool • The scope • Approach • Information on the underlying model of the method
Validation	Scientific validation will not be available for all tools. Therefore, evidence of a wider, less rigorous 'validation' process will be included (e.g. peer review, comparison with data from other tools and extensive use of the tool) as evidence that the tool provides meaningful/ useful/ relevant information. The 'maturity' of the tool will also be taken into consideration. This will be evidenced by available information to show the acceptability/ utility of the method. If we are unable to find evidence that the tool has been used beyond the developmental process, then it is not possible to verify the utility of the tool within the current scope of work.
Domain usage/ applicability to other domains	Information about whether the tool is applicable for HSE cross sector use, has the potential for cross sector use or is applicable only to one major hazard sector will be collected. Any tools that explicitly state that they are only applicable to one sector and that sector is not a major hazard sector will be excluded.
Resources required to complete the assessment	Details of the numbers of assessors, level of expertise, data and information requirements and training required will be collected where possible.
Availability of the tool & support	Information about the availability of the tool (i.e. is it a proprietary tool or freely available etc.) will be collected.
References used for the summary	A list of all references used to provide the factual information will be included.
Opinion (a mixture of information from published literature/ web pages etc. and the opinion of the HSL project team members. References are provided where appropriate.)	
Pros and cons	These will be based on published research, but may include comment from the HSL project team. It is recognised that this section will be biased by perspective of user/ authors.
Related methods	Information about whether the tool is linked to others (e.g. other tools used to inform the development of the new tool, whether it is best used in conjunction with another tool etc.)

6 REFERENCES

Bello G.C. and Colombari V. (1980) The Human Factors in Risk Analyses of Process Plants: The Control Room Operator Model (TESEO), Reliability Engineering, Volume 1, pages 3-14.

Bieder, C., Vidal, S. and Le Bot, P. (2000). Feedback from the actual implementation of the MERMOS method. PSAM 5 – Probabilistic Safety Assessment and Management pp. 1529 – 1534 (Universal Academy Press, Inc.)

Boring, R., Boring, L. and Gertman, D.I. (2005). Atomistic and holistic approaches to human reliability analysis in the US nuclear power industry. Safety and Reliability, Vol. 25, No. 2, pp. 21 – 37.

Cacciabue, P.C., Carpignano, A. & Vivalda, C (1993) Dynamic Reliability Technique for Error Assessment in Man-Machine Systems. International Journal of Man-Machine Studies, 38 (3) pages 403-428.

Cacciabue, P.C., Decortis, F., Drozdowicz, B., Masson, M. and Nordvik, J.P (1992) COSIMO: A cognitive simulation model of human decision making and behavior in accident management of complex plants. Systems, Man and Cybernetics, IEEE Transactions on Volume 22, Issue 5, Pages 1058 – 1074.

CAHR website http://www.cahr.de/tools/CAHR.htm.

Chien, S.H., Dykes, A.A., Stetkar, J.W., Bley, D.C. (1988) Quantification of Human Error Rates Using a SLIM-Based Approach. In: Proceedings of the Institute of Electrical and Electronics Engineers 4th Conference on Human Factors and Power Plants, pp. 297-302.

Collier, S (2003) A Simulator Study of CREAM to Predict Cognitive Errors. In Proceedings of the International Workshop. Building the new HRA. Errors of commission form research to application. Nuclear Energy Agency. Pages 56-75.

COGENT home page: http://cogent.psyc.bbk.ac.uk

Comer, M.K., Seaver, D.A., Stillwell,W.G. & Gaddy, C.D. (1984) Generating human reliability estimates using expert judgement, Vol.2; The Maxima Corp. Maryland – USNuclear Regulatory Commission, Contract Report, NUREG/CR – 3688-2.

Concord Associates Inc website - http://www.concordassoc.com/main.aspx?PID=68

Dang, V.N., Reer, B., Stater, O., Hirschberg, S. (2000) A Comparison of Emerging methods for Errors of Commission Based on Applications to the Davis-Besse (1985) Event. In: Kondo, S & Furua, K. (eds) PSAM 5 – Probabilistic Safety and Management. Universal Academy Press. Tokyo, Japan.

Embrey (1983) The Use of Performance Shaping Factors and Quantified Expert Judgment in the Evaluation of Human Reliability: An Initial Appraisal. NUREG/ CR-2986, Brookhaven National Laboratory.

Embrey,D.E. and Kirwan, B. (1983) A comparative evaluation of three subjective human reliability quantification techniques. The Annual Ergonomics Society Conference Proceedings, Coombes, K. (ed) Taylor and Francis, London, pp 137-142.

Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B. and Rea, K. (1984) SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgment. Volume I: Overview of SLIM-MAUD. NUREG/CR-3518. Prepared for the US NRC

Embrey, D.E., Humphreys, P., Rosa, E.A., Kirwan, B. and Rea, K. (1984) SLIM-MAUD: An approach to assessing human error probabilities using structured expert judgment. Volume II: Detailed analysis of the technical issues. NUREG/CR-3518. Prepared for the US NRC

Everdij M.H.C., Klompstra M.B., Blom H.A.P. and Fota O.N. (1996) MUFTIS work package report 3.2, final report on safety model, Part I: Evaluation of hazard analysis techniques for application to en-route ATM, NLR TR 96196L, (MUFTIS3.2-1)

Everdij M.H.C. and Blom H.A.P. (2008) Safety Methods Database. http://www.nlr.nl/documents/flyers/SATdb.pdf

Fleming, K.N., Raabe, P.H., Hannaman, G.W., Houghton, W.J., Pfremmer, R.D and Dombek, F.S. (1975) HTGR Accident Initiation and Progression analysis Status Report, Volume II. AIPA Risk Assessment Methodology, GA/A13617 Vol. II UC-77, General Atomic Co., San Diego, CA.

Forester, J., Bley, D., Cooper, S., Lois, E., Siu, N., Kolaczkowski, A. and Wreathall, J. (2004) Expert elicitation approach for performing ATHEANA quantification. Reliability Engineering & System Safety 83, (2004) pp.207 – 220.

Forester J., Kolaczkowski A., Lois E. and Kelly D. (2006) Evaluation of Analysis Methods Against Good Practices. Final Report. NUREG-1842. U.S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research Washington, DC 20555-0001

Gertman D.I. (1992) Representing cognitive activities and errors in HRA trees. EGG-HFRU-10026.

Gertman, D. I., Blackmann, 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 & System Safety 35 pages.127-136

Gertman, D., Blackman, H., Marble, J., Byers, and Smith, C. (2004a). The SPAR-H human reliability analysis method. NUREG/CR-6883. Idaho National Laboratory, prepared for U. S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research Washington, DC 205555-0001.

Gertman, D.I., Boring, R.L., Marble, J.L., and Blackman, H.S. (2004b). Mixed model usability evaluation of the SPAR-H human reliability analysis method. Fourth American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Controls and Human-Machine Interface Technologies (NPIC&HMIT 2004), Columbus, Ohio. September, 2004.

Gertman, D.I., Blackman, H.S., Marble, J.L., Smith, C and Boring, R.L. (2004c). The SPAR H Human Reliability Analysis Method. Fourth American Nuclear Society International Topical Meeting on Nuclear Plant Instrumentation, Controls and Human-Machine Interface Technologies (NPIC&HMIT 2004), Columbus, Ohio. September, 2004.

Gibson, W.H. and Kirwan, B. (2007) CARA: A Human Reliability Assessment Technique for Air Traffic Management Safety Assurance. www.eurocontrol.int/eec/public/standard_page/DOC_Conf_2007_007.html

Gibson, W.H. and Kirwan, B. (2008) Application of the CARA HRA Tool to Air Traffic Management Safety Cases.
http://www.eurocontrol.int/eec/public/standard_page/DOC_Conf_2008_002.html

Gore et al (1997) Conservatism of the Accident Sequence Evaluation Program HRA Procedure, http://www.blackwell-synergy.com/doi/abs/10.1111/j.1539-6924.1997.tb01283.x

Hall R. E., Fragola J., Wreathall J. (1982) Post Event Human Decision Errors: Operator Action Tree / Time Reliability Correlation. NUREG/CR-3010 (United States Nuclear Regulatory Commission), Washington, DC (USA)

Hannaman, B.W. and Spurgin, A.J. (1984). Systematic Human Action Reliability Procedure (SHARP). (Palo Alto, CA: Electric Power Research Institute), EPRi-NP-3583 (quoted by Swain, 1989).

Hannaman, G.W., Spurgin, A.J., and Lukic, Y.D. (1984) Human Cognitive Reliability Model for PRA Analysis. NUS-4531. Palo Alto, California: Electric Power Research Institute.

Hannaman, G.W., Spurgin, A.J., Lukic, Y. (1985) A Model for Assessing Human Cognitive Reliability in PRA studies. IEEE Third Conference on Human Factors in Nuclear Power Plants, Monterey, California, June 23-27, 1985. Institute of Electronic and Electrical Engineers, New York (USA).

Health and Safety Commission (1991) ACSNI Study Group on Human Factors. Second report: Human Reliability Assessment – a critical overview. London: HMSO.

Hollnagel, E. (1993) Human reliability analysis: Context and control. Academic Press

Hollnagel, E. (1998) Cognitive Reliability and Error Analysis Method - CREAM. Elsevier

Humphreys P. (1988) Human reliability assessors guide, Safety and Reliability Directorate UKAEA (SRD) Report No TRS 88/95Q

Hunns, D.M. (1982) The method of paired comparisons. In: A.E. Green, Editor, High risk safety technology, Wiley, Chichester.

Joksimovich, V. and Spurgin A.J. (1987) Benchmark of Systematic Human Action Reliability Procedure (SHARP): Final report. NUS Corp. Report number EPRI-NP-5546

Kennedy, R., Kirwan, B., and Summersgill, R. (2000) making HRA a more consistent science. In Foresight & Precaution, Eds. Cottam, M., Pape, R.P., Harvey, D.W., and Tait, J. Balkema, Rotterdam.

Khan, F.I., Amyotte, P.R., and DiMattia, D.G. (2006). HEPI: A new tool for human error probability calculation for offshore operation. Safety Science 44, 313-334.

Kirwan, B. (1988) A comparative evaluation of five human reliability assessment techniques. In Human Factors and Decision Making. Sayers, B.A. (Ed.) London: Elsevier, pp. 87-109.

Kirwan, B. and James, N.J. (1989). A Human Reliability Management System. In: Reliability Volume 89 (Brighton Metropole)

Kirwan, B. (1990). A resources flexible approach to human reliability assessment for PRA. In: Safety and Reliability Symposium 1990 Elsevier Applied Sciences, pp. 114-135.

Kirwan, B. (1994). A guide to practical human reliability assessment. Taylor & 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, 27, (6), 359-373

Kirwan B. (1997) The development of a nuclear chemical plant human reliability management approach: HRMS and JHEDI. Reliability Engineering & System Safety, Volume 56, Issue 2, Pages 107-133.

Kirwan, B., Kennedy, R., Taylor-Adams, S. and Lambert, B. (1997). The validation of three human reliability quantification techniques, THERP, HEART and JHEDI: Part II – results of validation exercise. Applied ergonomics, 28 (1), 17-25.

Kirwan, B (1997) The validation of three human reliability quantification techniques – THERP, HEART and JHEDI: Part III – Practical aspects of the usage of the techniques, Applied Ergonomics, Vol. 28, No 1, pp. 27-39.

Kirwan, B. (1998) Human error identification techniques for risk assessment of high risk systems – Part 1: Review and evaluation of techniques, Applied Ergonomics, Vol 29, No 3, pages 157-177

Kirwan, B., Gibson, H., Kennedy, R., Edmunds, J., Cooksley, G. and Umbers, I. (2005), Nuclear action reliability assessment (NARA): a data-based HRA tool. Safety & Reliability, vol. 25. no. 2 pp 38 – 45. (2005)

Kirwan, B and Gibson, H (2007) CARA: A Human Reliability Assessment Tool for Air Traffic Safety Management — Technical Basis and Preliminary Architecture. Pp 197 –214 in The Safety of Systems Proceedings of the Fifteenth Safety-critical Systems Symposium, Bristol, UK, 13–15 February 2007.

Kosmowski, G. Degen, J. Mertens, B. Reer, B. (1994) Development of Advanced Methods and Related Software for Human Reliability Evaluation within Probabilistic Safety Analyses. Published by Berichte des Forschungszentrums Julich; 2928, Institut fur Sicherheitsforschung und Reaktortechnik. ISSN 0944-2952

Le Bot, P., Desmares, E. Bieder, C. Cara, F., Bonnet, J-L. (1997). MERMOS: An EdF project to update the PHRA (Probabilistic Human Reliability Assessment) methodology, OECD Nuclear Energy Agency Specialists Meeting on Human Performance in Operational Events, Chattanooga, USA (October 13-17 1997)

Le Bot, P., Cara, F. and Bieder, C. (1999). A second generation HRA method: What it does and doesn't do. In: (Washington, DC: American Nuclear Society), Volume II, pp. 852-860.

Le Bot, P., Pesme, H., and Ruiz, F.Methodological Validation of MERMOS by 160 Analyses. PSA ’02. 2002 International Topical meeting on Probabilistic Safety Assessment. October 2002, Detroit, Michigan.

Le Bot, P. (2004). Human reliability data, human error and accident models - illustration through the Three Mile Island accident analysis. Reliability Engineering & System Safety, 83, 153-167.

Linsenmaier, B & Sträter, O. (2000). Recording and Evaluation of Human Factor Events with a View to System Awareness and Ergonomic Weak Points within the system, at the Example of Commercial Aeronautics. In: HFES (Ed), Human Factors and Ergonomics Society, Santa Monica CA, Mira Digital Publishing: South Jefferson, St. Lois, MO.

Lyons, M., Adams, S., Woloshynowych, M. and Vincent C (2004) Human reliability analysis in healthcare: A review of techniques International Journal of Risk & Safety in Medicine 16, pages 223–237 223. IOS Press

Marseguerra, M., Zio, E. and Librizzi, M. (2007) Human Reliability Analysis by Fuzzy "CREAM" Risk Analysis Vol 27 No 1 pages 137–154

Moieni P., Spurgin A.J. and Singh A. (1994) Advances in human reliability analysis methodology. Part I: Frameworks, models and data. Reliability Engineering and System Safety, 44.

Nuclear Energy Agency, Committee on the Safety of Nuclear Installations. Critical Operator Actions: Human Reliability Modelling and Data Issues. Principal Working Group No. 5 - Task 94-1 NEA/CSNI/R(98)1

Park, K.S. and Lee, J. (2008). A new method for estimating human error probabilities: AHP-SLIM. Reliability Engineering and System Safety, 93, 578 – 587.

Phillips .L. D. and Humphreys . P. Embrey .D. E. and Selby D. L., (1985) A Socio-Technical Approach to Assessing Human Reliability (STAHR). In: Pressurized Thermal Shock Evaluation of the Calvert Cliffs Unit 1 Nuclear Power Plant, Appendix D. NUREG/CR-4022, Washington, DC. Originates from Phillips .L. D. and Humphreys . P. Embrey .D. E. and Selby D. L., (1983) A Socio-Technical Approach to Assessing Human Reliability (83-4). Oak Ridge, TN: Oak Ridge National Laboratory.

Potash, L. M. Stewart, M. Dietz, P. E. Lewis, C. M. Dougherty Jr. E. M., (1981) Experience in Integrating the Operator Contributions in the PRA in Actual Operating Plants. In: Proceedings of the ANS/ENS Topical Meeting on Probabilistic Risk Assessment, Port Chester, NY, Vol. II, pages 1054-1063. American Nuclear Society, LaGrange Park (USA).

Powers, D.A. (Chairman of the 468th Meeting of the Advisory Committee on Reactor Safeguards, December 2-4) (1999). Letter to the Executive Director for Operations US NRC about NUREG-1624, Revision 1, "Technical basis and implementation guidelines technique for human event analysis (ATHEANA). www.nrc.gov/reading-rm/doc-collections/acrs/letters/1999/4681870.htmml

Reason J. (1990). Human Error. Cambridge University Press, Cambridge (GB)

Reer B. (1997), "Conclusions from Occurrences by Descriptions of Actions (CODA)", in: B.M. Drottz Sjöberg (Ed.), New Risk Frontiers, Proceedings of the 1997 Annual Meeting of the Society for Risk Analysis-Europe, Stockholm.

Reer B. (1999), "Retrospective Event Analysis Using CODA and Perspectives on Human Error Prediction", in: Modarres M. (Ed.), Risk-Informed Performance-Based Regulation. Proceedings of the International Topical Meeting on Probabilistic Safety Assessment (PSA '99), 22-26 Aug. 1999, Washington DC, USA. American Nuclear Society. LaGrange Park, Illinois, USA. Vol. 2, 1337-1345.

Reer, B. and Dang V.N. (2006) A Method for Quantifying Errors of Commission on the Basis of Operational Event Data.
http://safe.web.psi.ch/documents/pdfs/EOC_HRA_Meth_AnnexPaper-2006-01-25.pdf

Reer B., Dang V.N. and Hirschberg S. (2004) The CESA method and its application in a plant-specific pilot study on errors of commission. Reliability Engineering & System Safety. Volume 83, Issue 2, Pages 187-205.

Scarborough A., Bailey L., Pounds J. (2005) Examining ATC Operational Errors Using the Human Factors Analysis and Classification System. Civil Aerospace Medical Institute Federal Aviation Administration.

Seaver D.A. and Stillwell W.G. (1983) Procedures for using expert judgement to estimate human error probabilities in nuclear power plant operations. NUREG/CR-2743, Washington, DC 20555.

Senders J.W., Moray N., and Smiley A. (1985) Modelling Operator Cognitive Interactions in Nuclear Power Plant Safety Evaluation. Report prepared for the Atomic Energy Control Board, Ottawa, Canada.

Sträter, O. (2000). Evaluation of Human Reliability on the Basis of Operational Experience. (Koln/ Germany: GRS 2000) Included on the CAHR website. (A translation of the German report ref: Report GRS-138, published 1997).Sträter, O. (2006) The use of incidents for human reliability management. Safety & Reliability, Vol. 26, No. 2 pp 26 – 47.

Swain, A.D. (1967) Field Calibrated Simulation in Proceedings of the Symposium on Human Performance Quantification in Systems Effectiveness. Naval Material Command and the National Academy of Engineering, Washington, DC.

Swain, A. D. (1987). Accident Sequence Evaluation Program Human Reliability Analysis Procedure, NUREG/CR-4772. US Nuclear Regulatory Commission, Washington, DC.

Swain A.D. (1989) Comparative Evaluation of Methods for Human Reliability Analysis. Gesellschaft für Reaktorsicherheit, Garching.

Swain AD and Guttmann HE (1983). Handbook of human reliability analysis with emphasis on nuclear power plant applications. US Nuclear Regulatory Commission), Washington, DC. NUREG/CR-1278.

Swanson P.J and Hass P.M. (1993) Operational Human Performance Reliability Analysis (OHPRA), A Tool for the Assessment of Human Factors Contribution to Safe and Reliable Process Plant Operation, CA/TR-93003, Concord Associates Inc.

US Nuclear Regulatory Commission (USNRC). A Technique for Human Event Analysis (ATHEANA) - Technical Basis and Methodological Description. NUREG/CR-6350 Brookhaven National Laboratory, Upton, NY, April. 1996. Prepared by Cooper, S.E., Ramey-Smith, A.M., Wreathall, J., Parry, G.W., Bley, D.C. Luckas, W.J., Taylor, J.H., Barriere, M.T.

US Nuclear Regulatory Commission (USNRC). Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA). NUREG- 1624. Division of Risk Analysis and Applications. Office of Nuclear Regulatory Research, Washington DC. May 2000

US NRC NEA/CSNI/R(2002)3. Proceedings of the International Workshop, Building the new HRA: Errors of Commission from research to application (2002) US NRC, Rockville, Maryland, USA, 7-9 May 2001. NEA/CSNI/R(2002)3. Specific paper: Overview about the CAHR Method and its Application in Assessing Errors of Commission presented by O Sträter, GRS, pp. 117 – 129

Wei Wu; Yoshikawa, H. A pilot study on human cognitive reliability (HCR) by human model simulation. Intelligent Information Systems. IIS apos;97. Proceedings Volume 8, Issue 10, pages 95 – 99

Williams, J.C. (1985). HEART – A Proposed Method for Achieving High Reliability in Elsevier, pp. 87-109.Process Operation by means of Human Factors Engineering Technology. In Proceedings of a Symposium on the Achievement of Reliability in Operating Plant, Safety and Reliability Society, 16 September 1985, Southport.

Williams, J.C (1986). A proposed Method for Assessing and Reducing Human error. In Proceedings of the 9th Advance in Reliability Technology Symposium, University of Bradford, 1986, pp. B3/R/1 – B3/R/13.

Williams, J. C. (1988). A Data-based method for assessing and reducing Human Error to improve operational experience. In Proceedings of IEEE 4th. Conference on Human Factors in power Plants, Monterey, California, -9 June 1988, pp. 436-450.

Williams, J.C. (1989) Validation of human reliability assessment techniques. Reliability Engineering, 11, 149-162.

Williams, J.C. (1992). Toward an Improved Evaluation Analysis Tool for Users of HEART. In Proceedings of the International Conference on Hazard identification and Risk Analysis, Human Factors and Human Reliability in Process Safety. 15-17 January 1992. Orlando, Florida.

Williams, J.C (1992). A User Manual for the HEART Human Reliability Assessment Method. Prepared for Nuclear Electric plc. (C2547-1.001). Not in the public domain.

Wreathall, J. (1982) Operator Action Trees. An Approach to Quantifying Operator Error Probability During Accident Sequences, NUS-4159, NUS Corp., Gaithersburg, D.

Woods, D.D. and Roth E.M. (1987) Cognitive environment simulation: An artificial intelligence system for human performance assessment: Cognitive reliability analysis technique. NUREG/CR-4862-Vol.3.

Woods, D.D., Pople H.E. and Roth E.M (1990) The Cognitive Environment Simulation as a tool for modelling human performance and reliability. NUREG/CR-5213-Vol.1.

HSE

Health and Safety Executive

Review of human reliability assessment methods

Human reliability assessment (HRA) involves the use of qualitative and quantitative methods to assess the human contribution to risk. There are many and varied methods available for HRA, with some high hazard industries developing 'bespoke', industry focused methods.

It was considered that it would be useful for HSE to be up to date with developments in the field of quantitative HRA methods and to have knowledge of the capability of the tools and an understanding of their strengths and weaknesses. Therefore, this project was commissioned to further HSE knowledge and expertise, and to form a view on the 'acceptability' of the various tools for use in risk assessments.

RR679