# Error Classification for Safety Management: Finding the Right Approach

Steven T. Shorrock

Det Norske Veritas (DNV), Highbank House, Exchange Street, Stockport, Cheshire, SK3 0TE, UK.
steven.shorrock@dnv.com

**Abstract:** Human error identification systems have been criticised for failing to consider the problems of operational incident investigators and system developers. Increasingly esoteric human error modelling and classification approaches have often been met with resistance from the potential user groups that could apply them with the greatest impact. In order to improve the transfer of this technology, error classification techniques must balance a range of criteria - some more practical than have previously been considered. This paper provides an example of one such technique in air traffic management (ATM) - 'TRACEr *lite*' - that has been developed with practical aims in mind, while retaining its conceptual roots.

**Keywords:** incident analysis, human error identification, air traffic management, TRACEr *lite*.

## The Means and Ends of Error Analysis

There is an established need in safety-critical industries to implement systems to manage the human contributions to safety. In response to this need, many well-known techniques for the identification and classification of human errors have emerged. These include SHERPA (Embrey, 1986), GEMS (Reason, 1990), CREAM (Hollnagel, 1993), as well as others integrated into Human Reliability Assessment (HRA) methodologies (e.g. HRMS and JHEDI, Kirwan, 1997; THERP, Swain and Guttman, 1983). Many techniques have been influenced heavily by Rasmussen et al.'s (1981) Skill-, Rule-, and Knowledge-based behaviour framework and Reason's (1990) classification of slips, lapses, mistakes and violations. While these techniques have been primarily associated with the nuclear and process industries, HEI has also been applied in other new sectors, such as manufacturing (Paz Barroso and Wilson, 2000), rail (Vanderhaegen, 1999, 2001), consumer products (Baber and Stanton, 1994), public technology (Baber and Stanton, 1996), and medicine (Nyssen, 2000; Taylor-Adams and Vincent, 2000).

Most methods of analysing human error have focused on particular stages of the system development lifecycle, including prospective methods (e.g. predictive human error identification, or HEI) at the design stages, and retrospective approaches (e.g. incident analysis) during operation. Furthermore, in the development stages, prototyping and real-time simulation may be used to provide evidence of safety. All approaches share a need to analyse human error, and yet accident analysis and performance prediction have been pursued as two largely separate activities, by separate communities (Hollnagel, 2000). This is exemplified in the gulf between the terms used to describe human performance issue before, during and after a design is fielded (Dekker, et al., 1997). Dekker, et al., remark that an "error of omission" may be identified in the design phase. During development, the practitioner may state that "the operator couldn't keep up with what the system was doing". And an analyst may use the phrase "lack of situation awareness" *post hoc*. While incident investigators, designers, trainers and so on are using incompatible language and methods, organisational learning opportunities and efficiencies are being lost.

Some of the predictive techniques have been used as part of the HRA process, which feeds human failure probabilities into Probabilistic Risk Assessment (PRA). Incident and accident investigation and analysis, however, rarely employ the same techniques, or even frameworks as error prediction (even though some of the quantitative data that populate human reliability databases is derived from accident and incident data). Within the same company, those responsible for performance prediction and those responsible for incident investigation and analysis may, in fact, be using entirely different approaches, "despite the obvious fact that they refer to the same reality - namely the occurrence of unexpected events leading to unwanted outcomes" (Hollnagel, 2000, pp. KN1). Safety management requires both active and reactive monitoring approaches to provide feedback on organisational performance both before and after such events (Health and Safety Executive, 1997), and harmonisation of approaches should improve efficiency and quality of research, analysis and communication.

## Whither, Error Analysis?

Despite the emergence of many HEI techniques, and following considerable research and development, HEI methods are still under-utilised in industry. Lucas (2001) has reported that the UK Health and Safety Executive (HSE) is seeing "very little evidence of use of human error prediction methods in

COMAH[1] reports". And yet the HSE "expect that the part that foreseeable human failings play in initiating major accidents and the human reliability of safeguards to be understood and addressed with proportionally the same degree of rigour as for process and engineering issues" (Lucas, 2001). Although formal methods of hazard identification (HAZID) and risk and reliability assessment are being used for technical issues, human factors are relatively neglected. This is despite the fact that proportionally greater improvements in safety may often be achieved by a human focus. Cacciabue (2000) has noted that there has been a dramatic increase in the human contribution to accident development due to the very high reliability and refinement of mechanical and electronic components, and the complexity of the system and the role assigned to the human operator.

Johnson (1999) goes further to assert that human reliability approaches have had little impact upon system development in many industries, largely due to the failure of human factors research seriously to consider the problems of systems development. The problems include poor methodological support, analyst subjectivity, poor support for error prediction, focus on accidents and not incidents, individual operator/system focus, and difficulty in reaching consensus on the contextual sources of latent failures. Experience of HEI in industry suggests that additional causes include complexity, lack of demonstrable added value, and forms of denial (i.e. that there is a problem or solution). According to Johnson, until practical problems are addressed, esoteric models of cognitive and organisational failure will be of little practical benefit. Many studies (most unpublished) have shown practical benefits of such approaches applied by the HF community. But the fact remains that the transfer of this technology to the design and operation of safety-critical, interactive systems has encountered serious problems, and many non-HF specialists involved in safety management may feel 'out in the cold'.

**Requirements and Trade-offs**

This 'impact problem' may be partly because HF specialists and non-HF specialists emphasise different criteria when evaluating such techniques. In this paper, it is proposed that error classification techniques can be evaluated on the following criteria:

1. Comprehensiveness - the ability to discriminate and classify a comprehensive range of errors and influencing factors.
2. Consistency - the degree to which the leads consistent analyses between different users and with the same user over time.
3. Life cycle applicability - the degree to which the technique can be used throughout the formative and summative phases of system design lifecycle.
4. Predictive accuracy - the degree to which the technique accurately predicts potential errors.
5. Theoretical validity - whether the technique is based on a model of human performance, with a theoretically plausible internal structure.
6. Contextual validity - the degree to which the technique adequately captures contextual information.
7. Flexibility - whether the technique enables different levels of analysis according to the project needs, known information or expertise of the user.
8. Usefulness - whether the technique suggests, or can generate, error reduction or mitigation measures.
9. Resource efficiency (training) - the time taken to become proficient in the use of the technique.
10. Resource efficiency (usage) - the amount of time required to collect supporting information and conduct the analysis.
11. Usability - the ease of use of the technique.
12. Auditability - the degree to which the degree lends itself to auditable documentation.

These criteria (with the exception of #4) can be applied to both prospective and retrospective techniques, and are an adaptation of those proposed by Kirwan (1992). Such criteria are generally accepted, though they have been criticised for being partly based on user opinion, face validity and utilisation of the technique (Stanton and Baber, 2002). Stanton and Baber propose a quantitative appraisal approach based on signal detection theory, whereby predicted errors are classified as 'hits', 'misses', 'false alarms' and 'correct rejections'. This approach depends on prior observation of the total pool of possible error types that could occur - a possibility with simple consumer products but not so with large-scale, safety-critical, complex systems. So while this is a useful and interesting approach, it may only be useful for evaluating fairly simple products, is only applicable to predictive HEI, and would seem to focus primarily on predictive accuracy, consistency, and to a slightly lesser extent, comprehensiveness. The remaining criteria are not properly addressed by this approach.

[1] Control Of Major Accident Hazards

Some existing published techniques are biased heavily toward comprehensiveness, consistency, predictive accuracy and theoretical validity. Indeed, Baber and Stanton's study similarly valued the first three of these criteria in the appraisal approach described. When these criteria are valued to the expense of contextual validity, resource usage, usability and flexibility, for example, error classification techniques can be off-putting to those in industry, who feel that their practical needs are not being addressed. Shorrock and Kirwan (2002) note that the main problems of many techniques are:

- low usability (e.g. lack of structure, excessive requirements for supporting analyses, excessive jargon or excessive 'resolution', i.e. distinctions which were not possible to make reliably);
- low contextual validity (particularly important for Performance Shaping Factors - PSFs); and
- limited applicability (e.g. to skill- and rule-based performance only, to small-scale systems or applications only; to retrospective or predictive use only).

Methods that are used by incident investigators and designers generally emphasise practical criteria. These methods provide description and context, and are often quick and easy to use. However, the methods are often limited to a specific domain, highly dependent on experience, and fall short of providing the kind of information that is required to explain (or predict) errors. It stands to reason that if error classification techniques are to be useful, they must satisfy a more diverse range of criteria in a more balanced fashion than has previously been achieved, to meet the needs of the user group.

**Technique for the Retrospective and Predictive Analysis of Cognitive Errors**

One approach that has been developed within National Air Traffic Services (NATS) to gain a better understanding of air traffic controller error is called TRACEr ('Technique for the Retrospective and Predictive Analysis of Cognitive Errors') (Shorrock and Kirwan, 1999, 2002). TRACEr is a model-based approach to HEI, permitting both retrospective and predictive analysis. Shorrock and Kirwan (2002) called this 'the Janus perspective' after the Roman god who gave his name to the month of January. Janus presided over openings, beginnings and doorways, and was often depicted with two faces because he could look into the past and the future at the same time. The original version of TRACEr comprises a modular structure of eight taxonomies describing the context, error and error recovery, employing a series of colour-coded decision-flow diagrams and tables.

The process of developing TRACEr was iterative. The main inputs included:

- a literature review (covering over 70 sources);
- a controlled study of error classification;
- analysis of approximately 30 controller interviews regarding unreported human errors; and
- controller reviews of TRACEr taxonomies.

The principal applications have been:

- Analysis of UK Aircraft Proximity (Airprox) incidents (a mandatory reporting system) occurring within both controlled and unregulated (Shorrock et al., 2000) airspace between 1996 and 1999.
- Analysis of confidential incident/error reports (voluntary reporting system) from the Confidential Human Factors Incident Reporting Programme (CHIRP).
- Prediction and analysis of errors occurring in large-scale real-time simulations as part of the New Scottish Centre (NSC) programme (Shorrock, et al., 2001a).
- Prediction and analysis of errors occurring in small-scale military simulations of reduced separation standards outside controlled airspace (Shorrock et al., 2000).
- Human error prediction for the Final Approach Spacing Tool (FAST) (Evans et al., 1999; Shorrock et al., 2001b).

The original TRACEr is described more fully in Shorrock and Kirwan (2002). In a EUROCONTROL project, TRACEr was also developed for European use in the 'HERA' project - human error in ATM (see Isaac, et al., 2000, 2002a, 2002b), and, in collaboration with the Federal Aviation Administration (FAA), further developed in a joint project resulting in the 'Janus' technique.

**Development of TRACEr *lite***

Initially, TRACEr was designed to be used primarily by HF specialists. However, it became clear that TRACEr could be beneficial to other ATM specialists, such as incident investigators and designers. Operational feedback revealed that TRACEr appeared too complex or time-consuming to apply in an operational environment by non-HF specialists (as with other error classification systems). If such a technique was to be used in practice, a reduced-scope version was needed. This idea was called 'TRACEr *lite*' - an error analysis and classification tool for operational ATC personnel.

Following consultation with potential users of TRACEr *lite*, some practical requirements were determined in conjunction with the intended users. These emphasised resource usage, usability, flexibility, training requirement, and contextual validity. However, it was recognised that the technique should share the core of TRACEr - the theoretical underpinnings and framework, and the ability to be comprehensive at a useful level whist maintaining acceptable consistency.

The framework of TRACEr was reviewed to identify the taxonomies of most benefit to TRACEr *lite,* since the complexity of the original TRACEr was partly a product of the number of taxonomies. This was achieved by reviewing the previous NATS projects that had utilised TRACEr, and surveying other techniques used in industry. In addition, operational ATC personnel were consulted to determine which aspects of TRACEr were of most benefit. Some of the names of the taxonomies were simplified for TRACEr *lite* to make them more accessible to incident investigators and ATC specialists. Following this review, the following TRACEr *lite* taxonomies were proposed:

- *Task error.* The task error describes the error in terms of the task that was not performed satisfactorily. This taxonomy applied to *retrospective* use only, since for predictive use the information would be contained within the task step of a task analysis.
- *External error.* The external error describes the observable manifestation of the actual or potential error, based on logical outcomes of erroneous actions, in terms of timing, sequence, selection and quality. External errors are context-free and independent of cognitive processes (e.g. intention), and are used solely as prompts for *predictive* purposes.
- *Internal error (modes and mechanisms).* Error modes describe what cognitive function failed or could fail, and in what way. Error mechanisms describe the psychological nature of the error modes; the cognitive biases that are known to affect performance. These taxonomies are used for both *retrospective and predictive* use. A small number of error mechanisms are not feasible to predict and so were omitted from the predictive version.
- *Information.* Information keywords describe the subject matter or topic of the error, and relate specifically to the error modes. For instance, what information did the controller misperceive, forget, or misjudge, or miscommunicate (e.g. ‘heading’, ‘flight level’)? This taxonomy applied to *retrospective* use only, since for predictive use the information would be contained within the task step of a task analysis.
- *Performance Shaping Factors (PSFs)* - PSFs are factors that have influenced or could influence the controller’s performance, aggravating the occurrence of errors, or perhaps assisting error recovery. This taxonomy applies to both *retrospective and predictive* use.
- *Recovery* - error detection and correction is considered only for *predictive* purposes since they are more peripheral to incident investigation, and were thought to pose a risk of increasing the complexity of the retrospective method during the implementation phases.

A key process in the conversion from TRACEr to TRACEr *lite* was the simplification of TRACEr’s ‘internal error modes (IEMs)’ and ‘psychological error mechanisms (PEMs)’ to create TRACEr lite’s ‘internal errors (modes and mechanisms)’. A card-sorting task was conducted with eight HF specialists and four air traffic controllers (one of whom was also an ATM tool designer). Each TRACEr IEM and PEM was represented on a single card, with the relevant name and a brief description. The participants of the study were asked to perform one or more of the following tasks:

- sort IEMs and PEMs into user-defined piles (i.e. no criteria imposed).
- sort IEMs and PEMs into piles of 1-3 cards on the basis of similarity, within each cognitive domain.
- sort PEMs according to their strength of affect on internal error modes.

The card sorting exercises led to a reduction of approximately 60% in the number IEMs (to an average of 4 per cognitive domain in TRACEr *lite*), and 40% in the number of PEMs (to an average of 5 per cognitive domain in TRACEr *lite*).

Whilst it was not an original aim of the project, parallel study found that TRACEr *lite* may also be capable of classifying pilot errors (Scaife et al., 2001). This study suggested a number of provisional categories that could be added to TRACEr *lite* to enable pilot errors to be addressed.

The TRACEr *lite* prototype technique was presented to operational incident investigators at Manchester Area Control Centre. This consultation produced a number of observations on the format and method of use of TRACEr *lite*, and on a small number of the categories, which were largely seen as useful. Overall, the feedback revealed that the technique appeared acceptable in terms of the incident investigators’ requirements.

## The TRACEr *lite* Framework

Classifying errors using TRACEr *lite* is represented as four steps for retrospective or predictive use, as shown in Figures 1 and 2. A prototype version of TRACEr *lite* has been represented using Microsoft Excel, integrating both the *'RETRO'* and *'PREDICT'* versions the same package. This contains hyperlinks for navigation and pop-up contextual examples of categories. The final technique may be implemented using a more suitable database platform, and a Web site is being constructed at www.tracer-lite.co.uk. The remainder of this paper focuses on the *retrospective* version: TRACEr *lite RETRO*.

Figure 1 - TRACEr *lite RETRO* prototype front-end

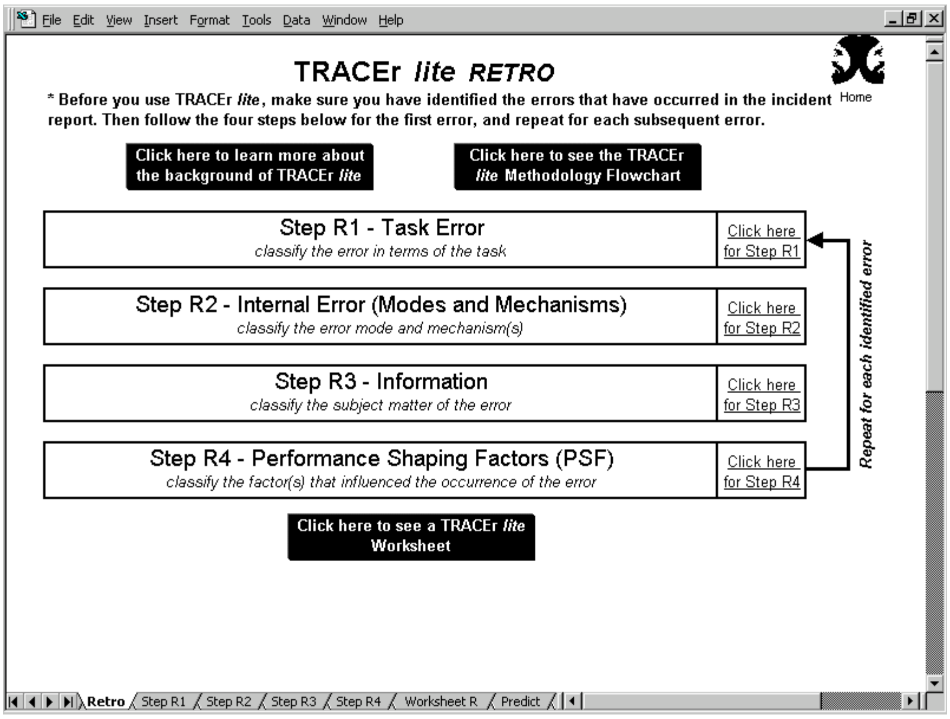
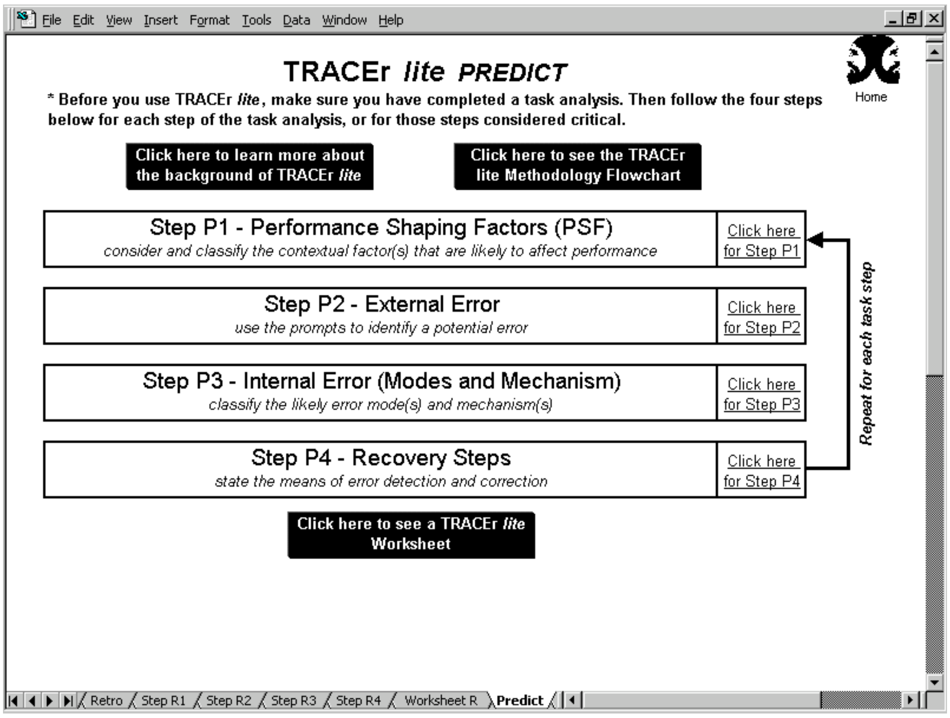



Figure 2 - TRACEr *lite PREDICT* prototype front-end



*Step R1 - Task Error:* The task error taxonomies describe controller and pilot errors in terms of the task that was not performed satisfactorily. Thirteen categories are provided for controller errors, and seven categories are provided for pilot errors. The categories are shown in Table 1 below.

Table 1 - TRACEr *lite RETRO* Task Error list

| Controller Task Errors | Pilot Task Errors |
|---|---|
| Separation error | Pilot-controller communications error |
| Controller-pilot communications error | Aircraft handling error |
| Radar monitoring error | Visual observation |
| Aircraft observation / recognition error | Flightdeck co-ordination/ communications error |
| Co-ordination error | Operational materials checking error |
| Control room communications error | Training, supervision, or examining error |
| Aircraft transfer error | HMI input & functions use error |
| Hand-over / Take-over error | Other pilot task error |
| Flight progress strip use error | |
| Operational materials checking error | |
| Training, supervision, or examining error | |
| HMI input & functions use error | |
| Other controller task error | |

These categories provide a high-level and clearly comprehensible view of error, and a generic structure that will be resilient to changes in the ATC task (e.g. due to increased electronic assistance). Task errors also provide an organising structure that may be required for periodic reports of error trends.

*Step R2 - Internal Error (Error Modes and Mechanisms):* Error modes and mechanisms describe in more detail how the error occurred. The modes and mechanisms are structured around four error domains:

- Perception - did the controller/pilot mis-see or mishear, or fail to see or hear something?
- Memory - did the controller/pilot forget or misrecall information, or forget to do something?
- Decision making - did the controller/pilot make an error projecting required separation, or make an error of planning or decision making?
- Action - did the controller/pilot perform an action in a way not intended, or inadvertently say something that was incorrect or unclear?

Error modes provide the next level of detail on an error, describing how the controller's/pilot's performance failed to achieve the desired result. One error mode is used for each error identified in the report. For instance, error modes within the 'Perception' error domain include 'mishear', 'mis-see', 'no detection (visual)' and 'no detection (auditory)'. TRACEr *lite* contains 14 error modes - three or four for each domain.

Error mechanisms describe in greater depth the psychological underpinnings of error. Once an error mode has been selected, an associated error mechanism may be selected. TRACEr *lite (RETRO)* contains 21 error mechanisms - five or six for each domain. Example error mechanisms within the 'Perception' domain include 'expectation', 'confusion', 'discrimination failure', 'tunnel vision', 'overload' and 'distraction/preoccupation'. The TRACEr *lite RETRO* error modes and mechanisms are shown in Table 2.

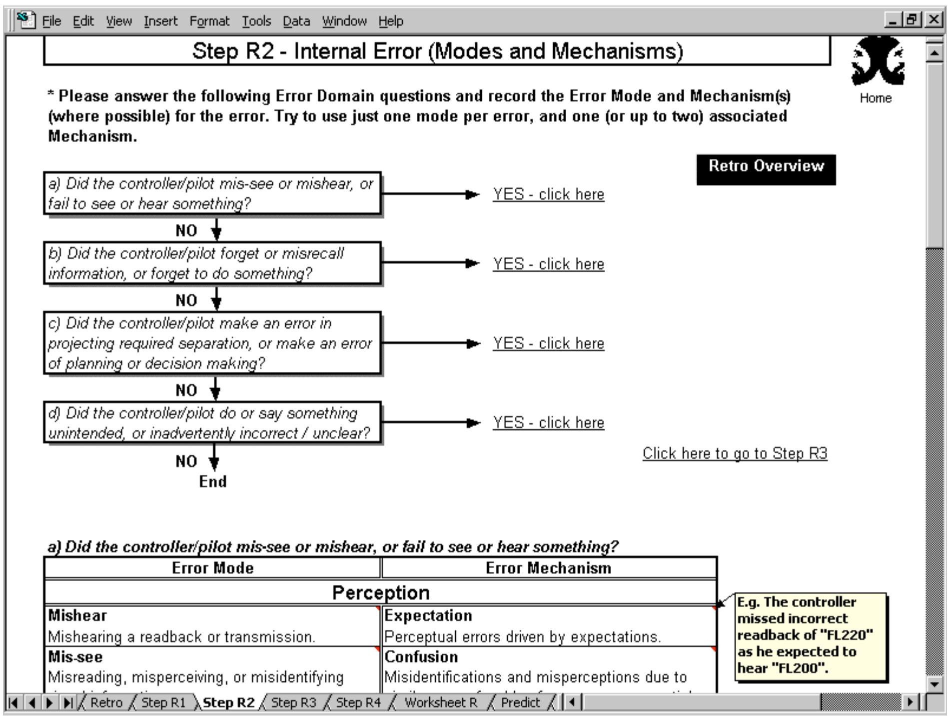
Error mechanisms can better enable the consideration of measures to reduce or mitigate errors, because the internal cause of the error can be analysed. For example, if a controller misidentified an aircraft on radar, this may have been due to 'confusion' (i.e. visually similar callsigns). Such errors could lead to attempts to increase the distinctiveness of lettering, and in the meantime raise awareness of the issue with controllers. However, the error may have been due to 'overload' (i.e. a lot of traffic on radar, and insufficient time to check properly). These types of errors may lead to attempts to filter the amount of information displayed, split the sector, etc. It will often be possible to determine error mechanisms during incident investigation, but when analysing historical reports, it will not always be possible.

A screen shot of Step R2 of TRACEr *lite RETRO* is shown in Figure 3.

Table 2 - TRACEr *lite RETRO* Internal Error (Modes and Mechanisms) taxonomy

| Error Mode | Error Mechanism |
|---|---|
| ***Perception*** | |
| Mishear | Expectation |
| Mis-see | Confusion |
| No detection (auditory) | Discrimination failure |
| No detection (visual) | Tunnelling |
| | Overload |
| | Distraction / Preoccupation |
| ***Memory*** | |
| Omitted or late action | Confusion |
| Forget information | Overload |
| Misrecall information | Insufficient learning |
| | Mental Block |
| | Distraction / Preoccupation |
| ***Decision Making*** | |
| Misprojection | Misinterpretation |
| Poor decision or poor plan | Failure to consider side- or long-term effects |
| Late decision or late plan | Mind set |
| No decision or no plan | Knowledge problem |
| | Decision freeze |
| ***Action*** | |
| Selection error | Variability |
| Unclear information | Confusion |
| Incorrect information | Intrusion |
| | Distraction / Preoccupation |
| | Other slip |

Figure 3 - TRACEr *lite RETRO* prototype screenshot for Step R2 Internal Error (Modes and Mechanisms)



*Step 3 - Information:* The information taxonomy describes the subject matter or topic of the error. For instance, what information did the ATCO misperceive, forget, or misjudge, or miscommunicate? The information classification relates specifically to the error mode, e.g.: for "The controller misprojected the required heading to maintain separation", error mode = 'misprojection'; information = 'heading'.

This is an important taxonomy because it highlights specific areas for error reduction. For instance, it is little use in knowing that a large number of memory failures occur if one cannot pinpoint what information is being forgotten. The information taxonomy contains over 50 keywords. The structure of the information taxonomy, with some example keywords, is shown in Table 3.

Table 3 - Information taxonomy structure

| **Information (structure)** |
|---|
| **ATC/Pilot Activities and Aircraft Information** |
| 1. Controller Materials (e.g. briefing material, flight progress strip) |
| 2. Pilot Materials (e.g. flight plan, charts) |
| 3. Controller Activities (e.g. transfer, co-ordination) |
| 4. Variable Aircraft Information and Pilot Activities (e.g. route, speed) |
| 5. Other |
| **Airspace and Other Keywords** |
| 6. Time and Location (e.g. sector, destination) |
| 7. Airport (e.g. runway, ground vehicles) |
| 8. Other |

*Step 4 - Performance Shaping Factors (PSF):* Performance shaping factors (PSFs) are those factors, either internal to the controller or pilot, or relating to the task and operational environment, that affect performance (i.e. 'aggravated' an error), directly or indirectly. PSFs can often be determined following the classification of the error mode and mechanism. Approximately 60 PSFs are included in TRACEr *lite*. The structure of the information taxonomy, with some example keywords, is shown in Table 4.

Table 4 - Performance Shaping Factors taxonomy structure

| **Performance Shaping Factors (structure)** |
|---|
| 1. Traffic and airspace (e.g. traffic load, sector design) |
| 2. Pilot-controller communications (e.g. controller RT standards, pilot language or accent) |
| 3. Procedures (e.g. complexity, accuracy) |
| 4. Training and experience (e.g. mentoring, time on sector) |
| 5. Workplace design, HMI and equipment factors (e.g. Mode C/SSR, flight progress strip display) |
| 6. Ambient environment (e.g. noise, lighting) |
| 7. Personal factors (e.g. alertness/fatigue, confidence) |
| 8. Social and Team factors (e.g. team relations, sector manning) |

TRACEr *lite*'s modular structure allows the user to describe the error at a level for which there is supporting evidence. For example, if the error mechanism is not known, the user can describe the task error and error mode. When strung together, the classifications from each step form a picture of the situation, and a multi-layered view of controller error. *E.g. The controller was distracted by a phone call and forgot to transfer an aircraft to the next frequency.*

| *Step* | *Taxonomy* | *Classification* |
|---|---|---|
| R1 | Task error: | Aircraft transfer error |
| R2 | Error Domain: | Memory |
| | Error Mode: | Omitted or late action |
| | Error Mechanism: | Distraction |

R3 Information: Transfer
R4 PSF: Traffic load; Alertness/fatigue

Furthermore, TRACEr *lite* is compatible with TRACEr, such that analyses using the two techniques can be mapped and cross-referenced, and more complex cognitive errors can, if required, be initially classified using TRACEr *lite*, then revisited using TRACEr by an HF specialist and incident investigator to derive additional detail

**Future Developments**

Retrospectively, TRACEr *lite* is being used by incident investigators to help investigate a number of real incidents in Manchester Area Control Centre. If successful, TRACEr *lite* may be implemented operationally, using a suitable platform. NATS have developed a post-incident checklist for use by operational ATC personnel at the London Terminal Control Centre (LTCC) and London Area Control Centre (LACC) to record information based on the structure and simplified content of TRACEr *lite* to facilitate information transfer. This structure has also been adapted for use during post-incident interviews at both centres.

TRACEr *lite* has been linked with, and used in conjunction with, a safety model for ATM (Scaife, et al., 2001), to analyse future ATM technology impacts. Finally within ATM, TRACEr and TRACEr *lite* are being tested in an evaluation study in mid-2002 to test coding reliability and user opinion.

TRACEr *lite* will be adapted for potential use in the rail sector in the UK, and is open to potential adaptation for other sectors. Meanwhile, adaptations of the original TRACEr are flourishing in European ATM and potentially U.S. ATM (Isaac, et al., 2000, 2002a, 2002b).

A website is currently under construction at www.tracer-lite.co.uk, and expected to be on-line during autumn-winter 2002. In the meantime, interested readers should contact the author for a copy of TRACEr *lite*.

**Conclusions**

Error classification techniques must redress the balance of criteria if they are to be used by operational personnel and designers without formal training in HF. More emphasis now needs to be paid to contextual validity, flexibility, resource efficiency, and usability. Without this balance and focus on users and requirements, techniques will not have the kind of impact that one might expect, or worse, might not be used at all. Techniques also need to adopt a 'Janus perspective' (Shorrock and Kirwan, 2002), using a common framework and shared taxonomies for prospective and retrospective use, if maximum use is to be made of the feedforward and feedback loops that are available. This paper forms one attempt find the right approach in the domain of ATM. Preliminary feedback from operational users indicates that TRACEr *lite* is reaching the required balance.

**Acknowledgements**

Some of the work described in this paper was done whilst the author was with National Air Traffic Services (NATS). The author would like to thank Dr. Barry Kirwan and Professor John Wilson in particular for their support and guidance. The author is also indebted to the NATS Human Factors Unit, and the controllers, designers and incident investigators involved in the development and testing of TRACEr and TRACEr *lite*. The views expressed in this paper are those of the author, and not necessarily those of any affiliated organisations.

**References**

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

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

Cacciabue, P.C. (2000). Human factors impact on risk analysis of complex systems. Journal of Hazardous Materials, 71, 101-116.

Dekker, S., Fields, B. and Wright, P. (1997). Human error recontextualised. Paper presented at Workshop on Human Error and Systems Development. Glasgow University, Scotland, 19th-22nd March 1997.

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

Evans. A, Slamen A.M. and Shorrock S.T. (1999). Use of human factors guidelines and human error identification in the design lifecycle of NATS future systems. Proceedings of the Eurocontrol/FAA Interchange Meeting, France, 27th-29th April 1999.

Health and Safety Executive. (1997). Successful Health and Safety Management. Norwich: HSE Books.

Hollnagel, E. (1993). Human Reliability Analysis: Context and Control. London: Academic Press.

Hollnagel, E. (2000). The human factors in systems reliability - is human performance predictable? Paper presented at the Human Factors and Medicine Panel (HFM) Workshop held in Sienna, Italy, 1-2 December 1999. NATO Research and Technology Organisation, Meeting Proceedings 32.

Isaac, A., Shorrock, S.T., Kirwan, B., Kennedy, R., Andersen, H. and Bove, T. (2000). Learning from the past to protect the future - the HERA approach. The 24th European Association for Aviation Psychology Conference, Crieff, UK, 4th-8th September 2000.

Isaac, A., Engelen, P. and Polman, M. (2002a). Human error in European air traffic management: from theory to practice. Workshop on the Investigation and Reporting of Incidents and Accidents, Glasgow University, Scotland, 17th-20th July 2002.

Isaac, A., Shorrock, S.T. and Kirwan, B. (2002b). Human error in European air traffic management: the HERA project. Reliability Engineering and System Safety, 75 (2), 257-272.

Johnson, C. (1999). Why human error modeling has failed to help systems development. Interacting with Computers, 11, 517-524.

Kirwan, B. (1992). Human error identification in human reliability assessment. Part 2: detailed comparison of techniques. Applied Ergonomics, 23 (6), 371-381.

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

Lucas, D. (2001). Human error prediction and controls: demonstrations made in COMAH safety cases. Proceedings of an IBC Conference on Human Error, London, 27th-28th February 2001.

Nyssen, A.S. (2000). Analysis of human errors in Anaesthesia. Our methodological approach: From general observations to targeted studies in simulator. In: Vincent, C., de Mol, B. (Eds.) Safety in Medicine. Elsevier, Oxford.

Paz Barroso, M., Wilson, J.R. (2000). Human error and disturbance occurrence in manufacturing and a toolkit for practical analysis. Cognition, Technology and Work, 2, 51-61.

Rasmussen, J., Pedersen, O.M., Carnino, A., Griffon, M., Mancini, C. and Gagnolet, P. (1981). Classification System for Reporting Events Involving Human Malfunctions, RISO-M-2240, DK-4000, Riso National Laboratories, Roskilde, Denmark.

Reason, J.T. (1990). Human Error. Cambridge, UK: Cambridge University Press.

Scaife, R., Smith, E. and Shorrock, S.T. (2001). A practical framework for identifying human safety issues in ATM. IBC Conference on Human Error. London, UK, February 2001.

Shorrock, S.T. and Kirwan, B. (1999). TRACEr: a technique for the retrospective of cognitive errors in ATM. In D. Harris (Ed.) Engineering Psychology and Cognitive Ergonomics: Volume Three - Transportation Systems, Medical Ergonomics and Training. Aldershot, UK: Ashgate.

Shorrock, S.T. and Kirwan, B. (2002). Development and application of a human error identification tool for air traffic control. Applied Ergonomics, 33, 319-336.

Shorrock, S.T., Kirwan, B., MacKendrick, H. and Kennedy, R. (2001b). Assessing human error in air traffic management systems design: methodological issues. Le Travail Humain, 64 (3), 269-289.

Shorrock, S.T., Kirwan, B., MacKendrick, H., Scaife, R. and Foley, S (2001a). The practical application of error analysis in UK air traffic management. Paper presented at People in Control: An International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres, UMIST, Manchester, UK, 18th-21st June 2001.

Shorrock, S.T., Kirwan, B., Scaife, R. and Fearnside, P. (2000). Reduced vertical separation outside controlled airspace. Third Annual Conference on Aviation Safety Management. May 2000.

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

Swain, A.D. and Guttmann, H.E. (1983). Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications. NUREG/CR-1278, USNRC, Washington, DC 20555.

Taylor-Adams, S., Vincent, C. (2000). Clinical accident analysis: understanding the interactions between the task, individual, team and organisation. In: Vincent, C., de Mol, B. (Eds.) Safety in Medicine. Elsevier, Oxford.

Vanderhaegen, F. (1999). APRECIH: a human reliability analysis method - application to railway system. Control Engineering Practice, 7, 1395-1403.

Vanderhaegen, F. (2001). A non-probabilistic prospective and retrospective human reliability analysis method - application to railway system, Reliability Engineering & System Safety, 71 (1), 1-13.