Available online at www.sciencedirect.com
SCIENCE DIRECT®

Safety Science 43 (2005) 147–171

SAFETY SCIENCE

www.elsevier.com/locate/ssci

# Designing and evaluating a human factors investigation tool (HFIT) for accident analysis

R. Gordon [*,1], R. Flin, K. Mearns

*Industrial Psychology Research Centre, University of Aberdeen, Scotland*

---

## Abstract

In an attempt to improve the investigation of the human factors causes of accidents in the UK offshore oil and gas industry, a Human Factors Investigation Tool (HFIT) was developed with the sponsorship of the UK Regulator, the Health and Safety Executive, and four exploration-related companies. The tool was developed on a theoretical basis with reference to existing tools and models and it collects four types of human factors information including (a) the action errors occurring immediately prior to the incident, (b) error recovery mechanisms, in the case of near misses, (c) the thought processes which lead to the action error and (d) the underlying causes. The investigation tool was evaluated on the basis of (i) an inter-rater reliability assessment, (ii) usability assessment, (iii) case studies and (iv) an evaluation system developed by Benner [Benner, L. 1985. Rating accident models and investigation methodologies. Journal of Safety Research 16, 105–126] Evaluation system. Although there is a need for further validation and analysis of HFIT using more realistic accident scenario exercises, some validation of the tool has been possible. In addition, it has been shown, in a small sample of accident investigations, that HFIT was found to be useful for the development of remedial actions, one of the main objectives of the tool.

---

[*] Corresponding author. Address: Centre de Bois des Bordes, BP 15, F-91222 Bretigny-sur-Orge Cedex, France. Tel.: +33 01 69 88 78 67; fax: +33 01 69 88 78 90.

*E-mail address:* rachael.gordon@eurocontrol.int (R. Gordon).

[1] Now at Eurocontrol Experimental Center, Brétigny-sur-Orge, France.

0925-7535/$ - see front matter
doi:10.1016/j.ssci.2005.02.002

## 1. Introduction

The collection and analysis of accurate accident data is essential for improving workplace safety, although is only one of several possible diagnostic sources (see Dekker, 2004 for a recent critique of over-reliance on accident and error data). Despite the importance of accident analysis, many industries still have accident reporting systems that are vulnerable to under reporting, have incomplete recordings and do not necessarily provide a complete picture of the conditions under which accidents take place (Stoop, 1997). For example in the offshore oil industry, there are currently no standard accident reporting systems in existence, instead companies tend to develop or purchase their own specific systems. Most of the oil companies operating on the UK Continental Shelf (UKCS) base their accident reporting systems on the International Safety Rating System (ISRS) developed by the International Loss Control Institute (ILCI; Bird and Germain, 1985), which (along with other systems in use) lacks a firm theoretical framework for psychological factors. Although information produced from these accident reporting forms can be extensive, the quality and quantity of data concerning human factors causes of accidents is generally poor; such as the sparse inclusion of human factors codes and the lack of understanding of these codes.

Accident investigation methods which are based on more robust human factors accident causation models allow safety managers to make a broader interpretation of their accident statistics in order to reduce the likelihood of future accidents. This paper describes the development and evaluation of a human factors incident investigation tool (HFIT), based on the dominant psychological theories of accident causation, which has the potential to improve the quality of human factors incident data.

### *1.1. Background research*

Prior to the development of HFIT, two prototype human factors reporting forms were developed, tested and evaluated in the offshore oil industry, and provided part of the basis for HFIT (see Mearns et al., 1997; Gordon et al., 2000). One reporting form contained 11 open questions regarding the causes of an incident and the other reporting form contained 'yes'/'no'-choice questions. The forms were completed by the witnesses to the incident and the relevant line management. Both were found to extract additional and more specific information regarding the human factors causes of accidents than the company's original report. However, it was felt that in order to gather more comprehensive and accurate data, the human factors investigation of offshore incidents could be further improved.

A review of the theories of accident causation and an analysis of 18 incident reporting systems provided the basis for HFIT (Gordon, 2002) and are listed in Table 1.

The theoretical basis of HFIT includes the Model of Human Malfunction by Rasmussen et al. (1981), the Human Information Processing Model by Wickens (1992) and Kontogiannis (1999) system for measuring error recovery. Three of the incident

Table 1
List of incident reporting systems reviewed for the development of HFIT

| | Reference |
|---|---|
| 1. *Reactive Incident Reporting Systems* | |
| Management oversight risk tree (MORT) | Johnson (1980) |
| Nuclear regulatory commission (NRC) | West et al. (1991) |
| Maintenance error decision aid (MEDA) | Boeing (1995) |
| Maintenance error investigation (MEI) | Baachi et al. (1997) |
| TapRoot | Paradies et al. (1996) |
| Human performance investigation process (HPIP) | Paradies et al. (1993) |
| Incident reporting system (IRS) | IAEA (1998) |
| Human performance enhancement system (HPES) | Bishop and La Rette (1988) |
| Safety through organisational learning (SOL) | Fahlbruch and Wilpert (1997) |
| Human factors analysis and classification system (HFACS) | Wiegmann and Shappell (1999) |
| Technique for retrospective analysis of cognitive errors (TRACEr) | Kirwan et al. (1999) |
| IFE incident investigation system | Green et al. (2000) |
| 2. *Combined pro-active and reactive investigation systems* | |
| Tripod (BETA and DELTA) | Hudson et al. (1994) |
| Aircraft dispatch and maintenance safety (ADAMS) | McDonald (1998) |
| 3. *Confidential incident reporting systems* | |
| Aviation safety reporting system (ASRS) | Reynard et al. (1986) |
| British airways human factors reporting (HFR) programme | O'Leary (1999) |
| Confidential human factors incident reporting program (CHIRP) | CHIRP (2000) |
| Confidential incident reporting and analysis system (CIRAS) | Wright and Davies (2002) |

reporting systems examined were found to be of particular relevance: the system for analysing aircraft dispatch and maintenance incidents (ADAMS, 1998), a taxonomy developed for Air Traffic Management incident analysis, called Technique for Retrospective Analysis of Cognitive Errors (TRACEr, Kirwan et al., 1999) and an incident investigation system developed for Phillips oil company (IFE, Green et al., 2000). Causal codes identified from a review of the human factors common to safety climate surveys and accident analysis studies (Gordon et al., under review) were used to check that relevant codes were included within the investigation tool. This stage of background research also identified the increasing use of psychological concepts relating to threat management (Helmreich et al., 1999) and situation awareness (Endsley and Garland, 2000) in human factors analysis systems for the aviation industry. Both of these constructs were relevant to the accidents occurring in the offshore oil industry and consequently these were also incorporated into the HFIT system.

## 2. Underlying model and structure of HFIT

On the basis of the above review and analysis, the structure of HFIT is developed on a sequential model of the incident trajectory where incidents (accidents and near

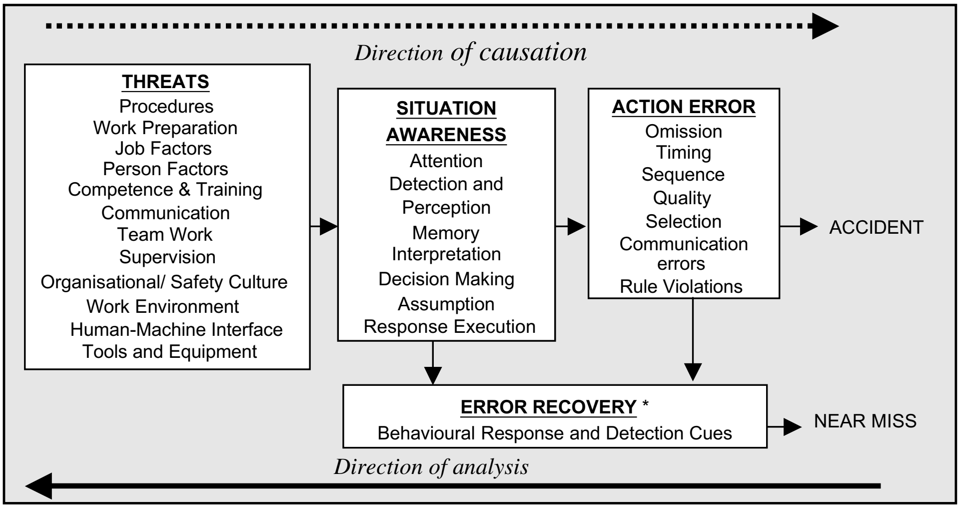

Fig. 1. HFIT model of incident causation and direction of analysis. (*) Can be analysed after the 'action error' or 'situation awareness' categories.

misses) are seen as the product of a number of different causes organised into four categories. As Fig. 1 illustrates, the behaviours immediately prior to the incident are described as the first category called 'Action Errors', which personnel at the sharp-end enact. These action errors are generally preceded and caused in part by a reduction in awareness of their situation, so Situation Awareness is the second category. The reduction in situation awareness is often related to 'Threats' to safety from the work environment or are conditions that may have been in the system for some time, but have not been identified nor rectified (third category). If the error, or reduced situation awareness is detected and recovered from before an accident occurs (error recovery), a near miss results. So a fourth category called Error Recovery is included that could occur during the action error or situation awareness stages.

The four categories, contain a total of 28 elements, listed in Fig. 1. Each of these elements are further described in Fig. 2, although only some examples are given at the 'sub-element' and 'item' levels. Action error elements are divided into 22 further 'items', situation awareness elements are described by 21 'items' and the error recovery elements contain 7 items. The 12 threat elements are divided into 'sub-elements' ($n = 43$) and 'items' ($n = 271$) and these are described in more detail in Gordon et al. (2002).

The following sections describe each of the four categories in more detail.

### 2.1. Action errors

This category is based on task-based taxonomies (such as Swain and Guttman, 1983) that describe the observable errors occurring immediately prior to the incident, but do not provide any causal information as to why or how the incident

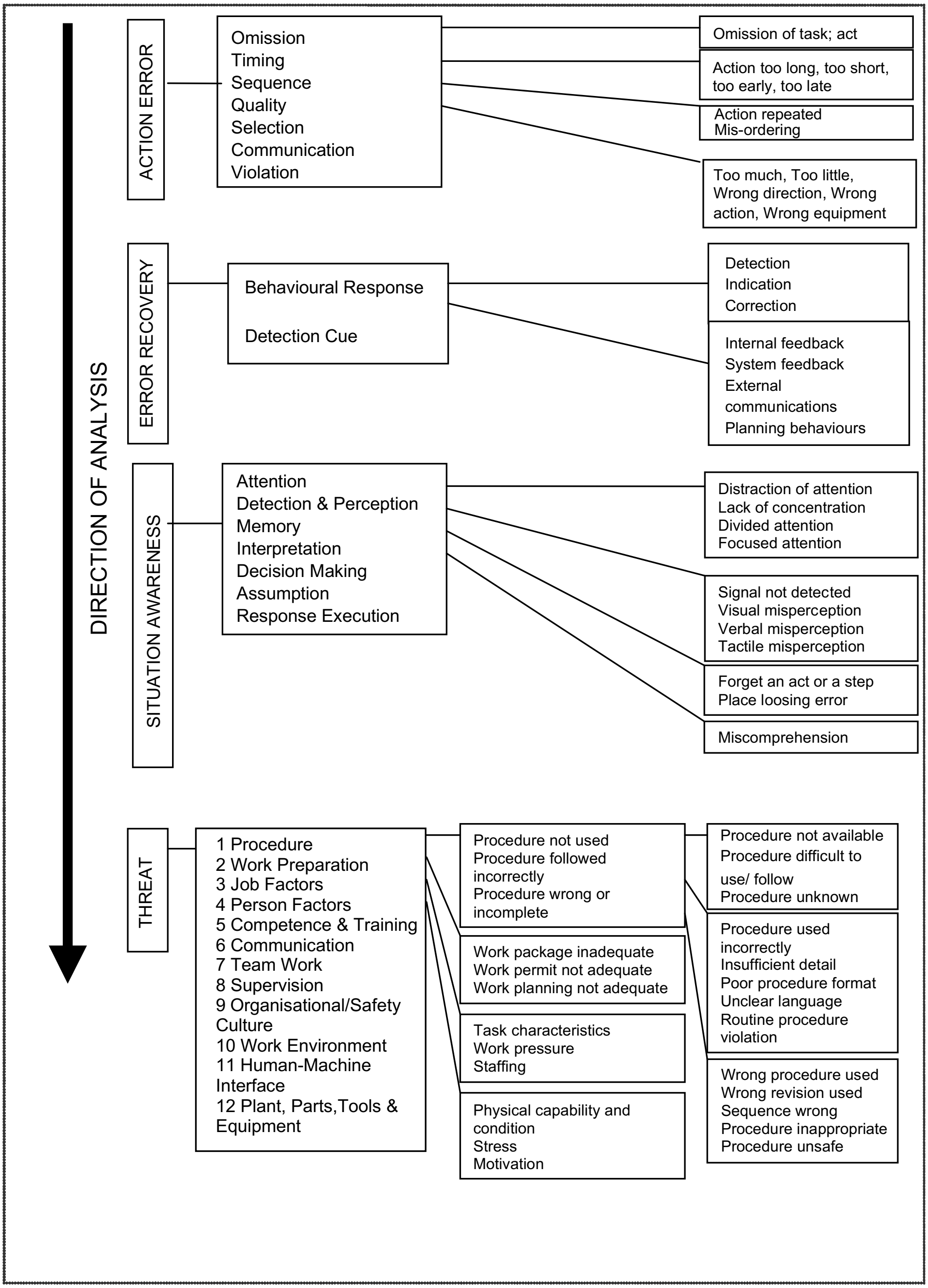

Fig. 2. Structure of HFIT.

happened. Such errors are referred to as External Error Modes in Rasmussen's (1981) taxonomy and as phenotypes by Hollnagel (1993). A taxonomy developed for Air Traffic Management incident analysis (TRACEr: Kirwan et al., 1999), contains a revised Swain and Guttman (1983) error mode taxonomy (consisting of: omissions; timing errors; sequence; quality; selection and communication errors). This has been revised slightly (at the item level) and used in HFIT. Action errors have been included in HFIT in order to understand the precise nature of the error before the causes for the error are investigated. It has been relabelled as 'action errors', as the original label (External Error Mode) uses jargon language (which is not user-friendly for non-human factors experts). It was thought that because the described errors are about errors of action, the term 'action errors' provides a clearer label. Despite this, some basic human factors training is required for potential users of the tool (see Section 3.2). This category contains six elements:

- Omissions—task or part of task not performed,
- Timing errors—action too short; too long; too early; too late,
- Sequence errors—action repeated; mis-ordering,
- Quality errors—action too much; too little; in wrong direction; wrong action right equipment,
- Selection errors—correct on wrong equipment/parts,
- Communication errors—information not transmitted/recorded; unclear information; incomplete information; incorrect,
- Violations—unintended; exceptional; routine; general.

### 2.2. Error recovery

Error recovery is thought to be an important supplementary safety goal since the 'zero accident policy' postulated by many oil companies (although remaining the ultimate safety goal) may be difficult to achieve in complex socio-technical systems (Kontogiannis, 1999). In some industries, systems are being developed which focus on preventing the consequences of human error by providing opportunities for error recovery (Helmreich et al., 1999).

A simplified version of the error recovery framework developed by Kontogiannis (1999) was used in HFIT. The first element, 'behavioural response', contains three questions regarding the possible recovery process of the error: (i) detect (i.e. realise or suspect that an error is about to occur), (ii) indicate (i.e. notify others in the team) and (iii) correct (i.e. modify an existing plan or develop a new one). The second element, 'detection cues', contained four questions regarding how the error was detected. This included 'internal feedback', 'system feedback' 'external communication' and 'planning behaviours'. This stage of the incident analysis would normally be undertaken after the action errors have been identified, although it could also be undertaken after the 'situation awareness' section.

### 2.3. Situation awareness

Information processing theory is one of the most widely used models in human error research and is perhaps the most useful cognitive error model for industrial applications. It states that people perceive information via their senses, interpret this information and make decisions concerning its meaning and relevance based on their previous understanding and current interpretation. (Wickens and Hollands, 2000). Both ADAMS (1998) and Kirwan et al. (1999) used Wickens (1992) Human Information Processing Model to collect data on cognitive failures. This approach has been included in HFIT, although it has been relabelled as 'situation awareness' (Banbury and Tremblay, 2004) since many of the items under this heading refer to the cognitive awareness of the individual. Situation Awareness has been defined as "the perception of the elements in the environment within a volume of space and time, the comprehension of their meaning, and the projection of their status in the near future" (Endsley and Garland, 2000, p. 5). Situation awareness issues (e.g. loss of concentration, distraction) are frequently mentioned in accident reports from the offshore drilling industry (Sneddon et al., 2005). The category within HFIT is divided into seven elements that are based on the systems used by ADAMS (1998), Kirwan et al. (1999) and Wickens (1992):

- attention—distraction; lack of concentration; divided attention; focussed attention,
- detection/perception—signal not detected; visual, verbal, tactile misperception,
- memory—forget or miss a step; failure to consider all factors; place losing error,
- interpretation—miscomprehension,
- decision making—apply incorrect/inappropriate/partial solution,
- assumption—relating to task, equipment, parts, systems, procedures,
- response execution—stereotype take-over, motor variability.

### 2.4. Threats

Threats are defined as situations that can encourage the occurrence of errors. This label has been taken from the work of Helmreich et al. (1999) from their research into threat and error management in the aviation industry. The 12 elements of threat are based on the content of the two human factors reporting forms briefly described previously (Mearns et al., 1997) and are described below.

*Policies, standards and procedures*—refers to the formal instructions or guidance that personnel need to carry out a task or job, such as work-cards, checklists, maintenance manuals, operating procedures, emergency operating procedures. This also includes the content and use of the company's management documents (e.g. general guidelines for planning and carrying out of training, maintenance, production, construction and development of plant/systems, planning systems and the company's goals).

*Work preparation*—This category includes problems associated with granting work permits, preparation of the plant before starting work (e.g. isolations, pressure testing), as well as planning of time and resources that could have contributed to the incident.

*Job factors*—This section deals with problems in actually carrying out the task, such as the nature of the task itself, the pressure to carry out the job and the level of staffing.

*Person factors*—This section deals with problems related to the individuals carrying out the task, such as the individual's physical capability or condition, stress or their motivation to carry it out.

*Competence and training*—Competence is the combination of skills and knowledge of a job or task. Lack of training can be a contributory factor to an undesired incident because a task that had consequences for the incident was not being correctly carried out.

*Communication*—Problems with communication can occur between individuals, work teams and managers. This category covers both technical methods (radio, telephone etc.) as well as building up communication to secure clear and distinct information.

*Team work*—This category includes shared situational awareness (do they have the same common goals/expectations for the job?), team decision-making, and the issue of roles and responsibilities.

*Supervision*—This category includes supervision during completion of the task, such as the level of work supervision, the roles and responsibilities, the supervisor's instruction and their leadership.

*Organisational and safety culture*—This includes the level of management commitment, whether or not there is a learning organisation, the reporting culture of the organisation, as well as the use of incentives (see Reason, 1997).

*Work environment*—This category examines the external & internal environments (e.g. extremes in temperatures) that can lead to incidents such as, problems with the manual handling of the task.

*System-equipment interface*—This section includes the legibility, labelling, user-friendliness and accessibility of equipment as well as increasing levels of automation.

*Tools and equipment*—This section involves the design and use of tools and equipment, the plant and parts design, the systems in place for design, maintenance and testing, and protective systems.

## 3. Procedure

### 3.1. Process of using HFIT

The Human Factors Investigation Tool, HFIT, was developed in a flowchart paper-based format, and after initial testing by potential users, it was developed as a computer-based tool (see Gordon et al., 2002). The paper version of HFIT is 54 pages long. It was designed for use by investigators of incidents. Fig. 3 illustrates

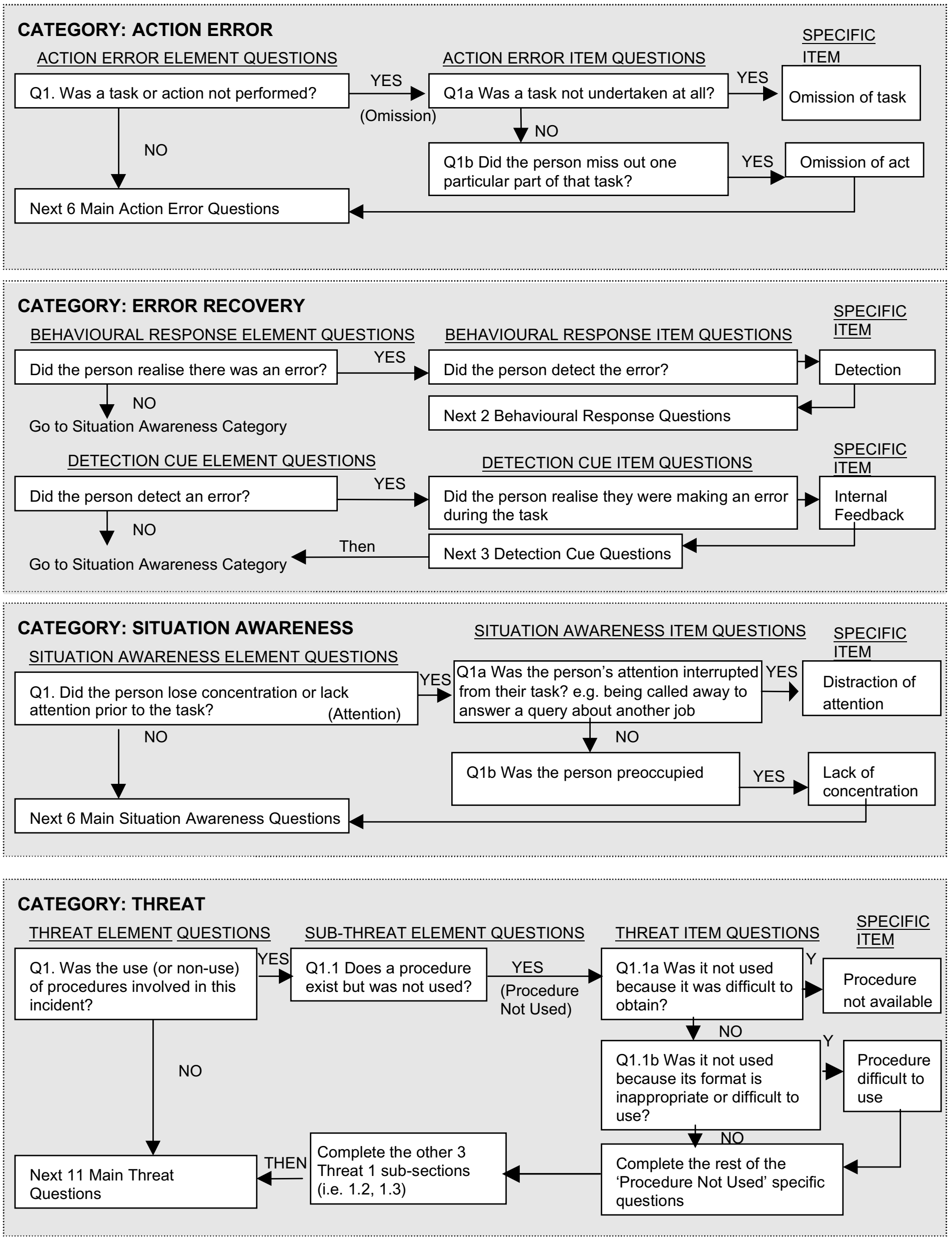

Fig. 3. HFIT process.

the process of investigating each category. The tool can be used in a number of different ways, first as an interview tool, where the investigator goes through the questions with each witness in turn. Secondly, the tool can be used after the witness interviews have taken place and the investigator/s use the tool themselves, keeping in mind what they found from the interviews. Finally, it can be used retrospectively on incidents that have been previously investigated using other investigation tools. The tool has not yet been tested to see which of the first two systems would be most effective for investigating incidents.

Before the investigator uses HFIT to investigate the causes of the incident, as with other investigation methods, information regarding the incident needs to be gathered, such as the people, objects and equipment involved in the incident and their actions. These actions can be plotted on a time line, which can help to establish whether or not there are gaps in the understanding of the accident sequence. The critical events (i.e. those which could have prevented the incident from occurring had they taken place) are identified and these are targeted using the investigation questionnaire. The causes of the critical events are analysed and appropriate remedial actions are implemented to prevent reoccurrence.

The process begins with the action error category, where the investigator asks a series of yes/no questions. The process is illustrated in Fig. 3, where the investigator begins at the element level and if they answer in the affirmative, they go to the item level. If they answer in the negative, they go to the next element question. Once they finish answering all the action error questions, they go onto the situation awareness element questions and follow the same procedure. After the situation awareness section has been completed, the investigator completes both the threat and error recovery sections, where the threat section contains an additional step (sub-elements).

### 3.2. Training the users

A one-day training course was developed to provide accident investigators ($n = 35$) with information about general human factors principles, instruction on how to use HFIT and scenario exercises to practise using HFIT. The participants generally had engineering backgrounds and previous training and experience in accident investigation. Some of the participants had some previous human factors training (such as Crew Resource Management). They were recruited from the four participating companies and the Health and Safety Executive (UK regulator) (HSE).

A total of five training courses were held, consisting of an introductory section and 10 modules: action errors; error recovery; situation awareness; job threats; person threats; competence and training; communication; team work; supervision; organisational and safety culture. Some of the Threat elements were not included, due in part to time constraints, hence the topics which engineers generally find more difficult to understand were the focus. After each section, respondents were asked to use HFIT to investigate the potential causes of an accident scenario. This allowed participants to practise using HFIT and become more familiar with it. In addition, the responses given by participants to the accident scenario were used to evaluate rater consistency.

The course could be deemed a success if the participants came away from the course with a better awareness and knowledge of the human factors causes of incidents and a good understanding of how to use HFIT to investigate incidents. An evaluation questionnaire was distributed to the participants, which is a standard measure for training evaluation (Goldstein and Ford, 2002) and results from this survey are described in more detail in Gordon et al. (2002). In total, 27 evaluation forms were completed, and overall, training was rated as either satisfactory or good on a 5-point scale (1 = very poor, 2 = poor, 3 = satisfactory; 4 = good; 5 = excellent) regarding their satisfaction with the following five indicators: their level of interest in the topic (mean = 3.7); the presentation of the materials ($m = 3.6$), the structure of the teaching ($m = 3.6$), the standard of the course materials ($m = 3.7$) and the relevance of the topic to their job ($m = 3.9$). Over the five training courses, small modifications to the course were made, where more time was spent using HFIT and less time was spent lecturing about the human factors principles and theories, which helped to improve the course ratings.

### *3.3. Implementing the human factors investigation tool*

Data from accidents and incidents were collected from one of the four participating companies over a 5 month period between July and December 2001 in order to evaluate the effectiveness of the HFIT reporting system for collecting human factors information. Two of the companies did not use HFIT to investigate incidents because they had not taken part in the HFIT training and one company reported that they had no incidents since the training. For the company that collected incident data using HFIT, their investigators were asked to use the paper version of HFIT whenever they felt it could support any incident investigations they were involved in. Initially, the participating company used HFIT after an investigation had been completed using traditional techniques. This was intended to test the HFIT method and demonstrate the integrity of the process and outcomes from HFIT to the users.

### *3.4. Computer interface and database development*

A computer programme was developed in Microsoft Access (1998 and 2000 versions) for HFIT. This tool can guide accident investigators through the relevant questions ultimately leading them to the causes of the incident. The user is given the option of answering either ‘yes’ (that it is a possible cause) or ‘no’ (that is not a cause) to each question that appears on the screen. Each cause is recorded on the screen as the user proceeds through the investigation, which allows them to follow their line of investigation. The computer-based version of HFIT can be used by investigators during the interview process with the witnesses, or after they had interviewed the witnesses (see implementation for further details). After the HFIT questionnaire has been completed, investigators are invited to write comments or ‘evidence’ to support each of the causes they found by describing why they thought that cause contributed to the incident. In addition, the investigation team can include possible ‘remedial actions’ beside each of the causes. Finally, the data from the investigation can be exported to either Word (in the form of individual reports) or Excel (for analysis with other incidents).

## 4. Evaluation

In order to evaluate the effectiveness of HFIT for collecting human factors information, the following four evaluation methods were used to assess HFIT and are discussed in turn in this section.

### *4.1. Accident scenario exercise*

Accident investigators ($n = 25$) from the four participating companies (described above in Section 3.2) coded the causes of a specified incident (an actual incident) using HFIT during the HFIT training course. This was undertaken in order to determine the level of agreement between the investigators with regard to the causes of incidents, to determine the inter-rater reliability of the tool.

The incident scenario exercise comprised a one-page offshore accident scenario that the investigators were asked to read. After each section of HFIT was described in the training course, investigators used HFIT to determine which elements, sub-elements and items contributed to the incident from the Action Error, Situation Awareness, Error Recovery and Threat categories. These responses were recorded on a Response Sheet. Investigators were able to choose as many of the causes they thought may have contributed to the incident. The investigators' responses from the Scenario exercise were recoded as 'yes, a cause' = 1 and 'no, not a cause' = 0, and entered into Excel (97) and SPSS (Statistical Package for Social Sciences). These responses were compared to the 'investigation findings' which were a combination of the original investigation findings and a re-analysis of the incident by the HFIT developer and an original member of the investigation team using HFIT. Inter-rater reliability scores for each item, and the 'investigation findings' are described below.

#### *4.1.1. Inter-rater reliability*

Inter-rater reliability is the extent to which different raters give the same response for the same observed performance (Howell, 2002). In this case, the test was to find out the extent to which 25 investigators attribute the same causes (by responding 'yes' or 'no') to an accident scenario. Inter-rater reliability scores were calculated for the HFIT Action Error Items ($n = 22$), the Situation Awareness Items ($n = 21$), the Error Recovery Items ($n = 7$) and the Threat Sub-Elements ($n = 42$) in order to determine the consistency of the 25 investigators' responses in terms of which categories, elements and items they selected. Threat Items ($n = 271$) were not included in the analysis due to insufficient time in the training course. It is hypothesised that if the investigators' responses are consistent with each other, this may indicate a shared understanding of the questions, suggesting that the questions may be comprehensible to the investigators. An index developed by James et al. (1984) called the within group inter-rater reliability measure ($r_{wg}$) was used to test this hypothesis. The scores

Table 2
Inter-rater reliabilities for the main HFIT sections

| Action errors | Number of items in scale[a] | % of investigators who found this item to be a cause of the incident | Inter-rater reliability ($r_{wg}$) of individual element |
|---|---|---|---|
| Omission | 3 | 100 | 1 |
| Timing | 5 | 44 | 0 |
| Sequence | 3 | 24 | 0.24 |
| Quality | 5 | 68 | 0.09 |
| Selection | 2 | 0 | 1 |
| Communication | 6 | 100 | 1 |
| Violation | 5 | 68 | 0.09 |
| *Situation awareness* | | | |
| Attention | 5 | 76 | 0.24 |
| Detection and perception | 5 | 40 | 0 |
| Memory | 3 | 24 | 0.24 |
| Interpretation | 2 | 28 | 0.16 |
| Decision making | 5 | 52 | 0 |
| Assumption | 5 | 92 | 0.69 |
| Response execution | 3 | 20 | 0.33 |
| Error recovery | | | |
| Behavioural response | 4 | 76 | 0.24 |
| Recovery cue | 5 | 60 | 0 |
| *Threats* | *Number of sub-elements in scale*[a] | | |
| Procedures | 6 | 72 | 0.16 |
| Work preparation | 4 | 56 | 0 |
| Job factors | 4 | 44 | 0 |
| Person factors | 4 | 20 | 0.33 |
| Competence and training | 4 | 12 | 0.56 |
| Communication | 4 | 80 | 0.33 |
| Team work | 6 | 72 | 0.16 |
| Supervision | 4 | 76 | 0.24 |
| Organisational/safety culture | 5 | 16 | 0.44 |
| Work environment | 4 | 8 | 0.69 |
| Human–machine interface | 4 | 16 | 0.44 |
| Plant, parts, tools and equipment | 6 | 0 | 1 |

[a] Including element.

for the elements are displayed in Table 2 (Column 4). This index is defined as the proportional reduction in error variance of a distribution of obtained responses compared to a distribution representing a random response pattern in which the frequency of the responses is equal for each possible point on the scale ($n = 2$). In this case, there were 2 possible responses: 'yes' and 'no'. The equation for $r_{wg}$ is: $r_{wg} = 1 - (S_x^2/\sigma EU^2)$ where $S_x^2$ equals the variance of the observed and $\sigma EU^2$ equals the population variance of a discrete rectangular distribution of the responses. The equation for this is: $\sigma EU^2 = (A^2 - 1)/12$, where $A$ is the number of possible alternatives in the rating scale. Values of $r_{wg}$ can vary from 0 to 1, where a score of 1 denotes perfect reliability between investigators. When the variance of the obtained ratings is random, then $r_{wg} = 0$, reflecting no agreement between investigators.

Overall, the results indicate the overall level of agreement between investigators was low. The causal codes that were selected by over three-quarters of the investigators were omissions (action error); communication (action error); attention (situation awareness); assumption (situation awareness); behavioural response (error recovery); communication (threat) and supervision (threat) indicating the highest consistency between investigators. Additionally, a high number of investigators (68%–72%) agreed on the following causes: quality (action error); violation (action error); procedures (threat) and teamwork (threat).

#### *4.1.2. Agreement between investigators responses and 'investigation findings'*

In order to measure how "accurate" the investigators were in coding the causes of the incident, their responses were compared to the 'investigation findings' (See Table 2, Column 1). They selected 33 codes, 19 of which were also selected by more than 50% of investigators. For 10 out of the 33 of codes, more than 75% of investigators chose the same codes (see Table 2, Column 2).

The most common elements that were chosen by investigators were omissions (where 100% of investigators chose this category), communication errors (100%); behavioural response (76%), attention (76%), assumptions (92%), as well as procedural (72%), communication (80%), team work (72%) and supervision threats (76%). At the item level, the most common responses were omission: task not performed (76%), communication: information not transmitted (76%); error recovery: detection (76%); lack of concentration (56%), divided attention (56%) and assumption relating to previous task (56%). The percentages of investigators who agreed with the causes in the threat section were smaller, indicating less agreement between investigators (the best agreement between the investigators' responses and the 'investigation findings' at the sub-element level of the 'threat' category, was 'location of communication threat', 48%).

The relationship between the inter-rater-reliabilities and the percentage of investigators who found the items to be causal indicates that there was high correlation between the investigators when the majority of them either agreed that the item was a cause or when the majority disagreed that that the item was a cause. Furthermore, inter-rater-reliabilities were very small (about 0) when around only about 33%–66% of investigators agreed (or disagreed) that the item was a cause.

Out of the 33 elements, sub-elements and items described in the 'investigation findings' to be the cause of the scenario, 10 were chosen by less than 33% of investigators: only 24% of investigators chose memory: forget an act or a step (16%); work planning not adequate (32%), task characteristics (16%); staffing (28%); communication misunderstood (20%); shared situation awareness-specific event (16%); shared situation awareness-in general (16%); co-operation (28%) and instruction (32%). The results from each category are described in Table 3.

Overall, the majority of investigators chose at least 50% of the codes described in the 'investigation findings'. The average number of codes attributed to the accident scenario by the 25 investigators was 30.6 (range 8–56), where the 'investigation findings' attributed 33 codes to the accident scenario. Some of these codes attributed by the investigators were not identified in the 'investigation findings' ($n = 12.2$ codes;

Table 3
'Human factors investigation findings' and common responses

| Human factors investigation findings | % of investigators who found this item to be a cause of the incident |
|---|---|
| **Action errors** | |
| *Omission* | 100 |
| Task not performed | 76 |
| *Communication* | 100 |
| Information not transmitted | 76 |
| **Error recovery** | |
| *Behavioural Response* | 76 |
| Detection | 76 |
| Indication | 52 |
| *Recovery cue* | 60 |
| System feedback | 48 |
| **Situation awareness** | |
| *Attention* | 76 |
| Distraction of attention | 52 |
| Lack of concentration | 56 |
| Divided attention | 56 |
| *Memory* | 24 |
| Forget an act or step | 16 |
| *Assumption* | 92 |
| Assumption relating to previous task | 56 |
| **Threats** | |
| *Procedures* | 72 |
| Procedure followed incorrectly | 40 |
| *Work preparation* | 56 |
| Work planning not adequate | 32 |
| *Job factors* | 44 |
| Task characteristics | 16 |
| Staffing | 28 |
| *Communication* | 80 |
| Location of communication threat | 48 |
| Communication misunderstood | 20 |
| *Team work* | 72 |
| Shared situation awareness (event specific) | 16 |
| Shared situation awareness (in general) | 20 |
| Co-operation | 28 |
| *Supervision* | 76 |
| Level of supervision | 44 |
| Instruction | 32 |

Labels in bold refer to the categories; in *italics* refer to the elements; indented in italics refer to the items/sub-elements.

range 2–33), indicating that on average, 40% of the codes attributed by investigators were not in the 'investigation findings'.

In conclusion, it would seem the level of agreement between investigators responses and the 'investigation findings' is generally fairly low when using HFIT to code an accident scenario. However, this is not unexpected, since the investigators had only minimal training and practice using the tool. In addition, the accident scenarios were very simple with regard to the amount of detail given and the inability of the investigators to ask further questions of the people involved in the incident.

### *4.2. Evaluations by the users*

Participants were asked for their opinions on the operation and value of the system at the end of the trial period using three methods of data collection: (i) user evaluation form; (ii) written feedback after investigators completed an investigation and (iii) information from informal discussions with HFIT users. The findings below are in reference to the paper version for HFIT. Out of the 35 investigators on the HFIT training course, 15 provided evaluations at the end of the HFIT training course. The user evaluation form was divided into four main sections with a total of 37 questions. The four sections included: (i) ease of use (13 questions); (ii) validity of results (4 questions); (iii) identification of causes of the incident (17 questions), and (iv) comparison with traditional accident analysis techniques (3 questions). This was developed as an Excel spreadsheet.

#### *4.2.1. Ease of use*

Overall, the comments indicate that users found HFIT useful for investigating incidents. Some investigators commented that they found it difficult to use at first, but after some practice with the tool they found it much easier to navigate through the flow charts. Investigators felt that they received sufficient training to be able to use HFIT. Some comments from the investigators include: "*Easy to use in paper form but I got the impression it was leading round and round at one point .... until I came to the root cause*"; "*I found the investigation part quite easy as you are just following the flowcharts going from step to step*" and "*HFIT in this scenario proved to be very successful and lent itself to the investigation process. Only took 1 h and 20 min*" Although another investigator felt that "*If there are a lot of people involved in the investigation it would be very time consuming.*"

There were some comments for improving the comprehension of the questions within HFIT: "*Some of the terminology is above some of the general users, and needs to be understood by all users*". However, other investigators felt there were hardly any questions which were difficult to answer. Some investigators felt that it was difficult to monitor their progress using HFIT. In order to aid the monitoring during use of the paper version, a progress sheet was developed to help investigators track their progress.

#### *4.2.2. Identification of causes and validity of results*

The majority of investigators reported that HFIT addresses the key causes of incidents, although this will require further testing to verify. "*Use of the tool provided greater and more detailed questions along any threads identified that could contribute to the corrective action*".

"*Although the investigation can be very time consuming, if you go through all the steps there is nothing that would be missed*". "*The HFIT did lead us to some aspects of the operation that would not normally have been considered*". Another investigator felt that the investigation technique was "*very non-confrontational*".

### *4.3. Individual case studies*

The causal analysis of individual incidents were evaluated in terms of the causes attributed to incidents using the company's original reporting system compared to the causes attributed using HFIT. Furthermore, in order to investigate whether HFIT aided in the development of remedial actions, the remedial actions and the incident causes have been compared.

In total six case studies were provided, although only three of the incidents could be used as individual case studies as the others had incomplete recordings of the causes identified using HFIT. These are described in Gordon et al. (2002) and only summary results are described in this paper. Each case study was analysed with respect to the following information:

(i) Brief description of the event.
(ii) Immediate and underlying causes from the original report.
(iii) Findings from HFIT (action errors, situation awareness and threats).
(iv) Remedial actions for the original report.
(v) Links between the original report and HFIT causes.
(vi) Link between the original report and HFIT causes and the remedial actions.

Three case studies were collected in order to assess HFIT in terms of its ability to generate further human factors data and remedial actions for incident investigation.

The results indicate that HFIT may have helped to improve the analysis of the incidents. Additional codes were identified from the HFIT analysis that could not be coded using the company's original coding system. A total of eight, nine and four additional causes were identified from the three case studies over and above the company's own reporting system. The HFIT analysis (in addition to using their own analysis) was used by the company to help develop the remedial actions in each of the case studies. This was noticeable from the comparison of the results in the final investigation report with the HFIT results, where the causes reported in the investigation report did not always directly link to the remedial actions, whereas they did link to the HFIT results. However, it was not possible to identify precisely which remedial actions had been developed based on the HFIT analysis. It was clear, however, that not all of the causes identified by HFIT were developed into remedial actions.

The incidents that were analysed using HFIT ($n = 6$) provide useful information regarding how effectively the tool was used. In the main, the tool appears to be used effectively, although the method of recording the results was not always complete. For example in one case study, communication was found to be a threat (Communication not effective). Using HFIT correctly, further analysis should have taken the investigator to the point where they understood why the communication was not effective. In addition, the sub-element: team situation awareness was not recorded accurately by the investigator, as they missed out the term: 'situation awareness', although this was deemed to be the cause of the incident. A reason for this maybe that the term 'situation awareness' is not familiar to the other users of the system (who have either not had the human factors training), or the term is not apparent or 'user-friendly' to the investigators. This information may not have been available to the investigator, or they neglected to record the data at the item stage. It is important that the data are recorded, so that if the incidents are ever reviewed, the findings and the evidence for the findings are documented. It is also important to record the data if they are to be used to analyse trends of the causes of accidents. The computer version of HFIT automatically records the causes, and hence the progression of the analysis is recorded as well as evidence for the causes (i.e. the reasons why they came to the conclusions they did).

### *4.4. Benner's evaluation system*

Benner's (1985) model and method evaluation system was used to evaluate HFIT using 10 criteria (e.g. comprehensiveness and ability to define remedial actions, the criteria are provided in Tables 3 and 4). The two human factors reporting forms (described briefly in the background section) were also evaluated using these criteria, and have been included in the tables. Benner's (1985) evaluation scheme used a three-point rating scale (where a rating of 2 = would satisfy; 1 = might satisfy; 0 = cannot satisfy) and the maximum score for any model or method was 20. The two human factors reporting forms and HFIT were assessed on each of the 10 criteria by deciding whether the models could satisfy the criteria. The ratings of the methods were derived by deciding whether the methods could satisfy the criteria, both conceptually and in their application within the company. It must be noted that the developer of the three reporting tools also undertook the Benner evaluation, which could have introduced a bias (Table 5).

The overall evaluation score for HFIT was 33 out of a possible 40 points (which compared favourably to the two forms which scored 21 and 25 respectively).

Although the number of the questions about the HFIT method (Table 4) have not scored full points (14/20), Forms 1 and 2 scored even lower (11 and 12 respectively). This may be in part because the measure is very stringent, and very few investigation systems would be able to score highly on their ability to support personal initiatives (q.3) or 'truth-test' the data (q.9). The other three questions that scored poorly were about providing information about duties under a standard with regard to the enforcement programme (Table 4, q.7); about the compatibility of HFIT with 'pre-investigations' (or safety analyses) of potential accidents (Table 4, q.10); and about the theoretical consistency of HFIT with the company's safety programme concepts (Table 3, q.6). These three aspects could be improved by further refinement by closely liasing with individual companies.

Table 4
Evaluation of the HFIT model (and Form1 and Form2 models) according to Benner's evaluation system

| | Score (HFIT) | Score (Form1) | Score (Form2) |
|---|---|---|---|
| 1. How realistically is the accident described? *The causes of the accident include the proximal (action errors, situation awareness) and distal factors (threats). In the database version, investigators are asked to write a brief summary of the events leading up to the incident. In addition, there is section for witnesses to write statements* | 2 | 1 | 1 |
| 2. How well does the model define the aspects of an incident? *HFIT includes 271 item-level codes, providing the investigator with a very detailed and specific set of causes* | 2 | 1 | 1 |
| 3. How well does the model demonstrate the company's safety mission? *The company's safety mission is to gather more data regarding the human factors causes* | 2 | 2 | 2 |
| 4. How comprehensive is the model at encompassing the development and consequences of an accident? *HFIT is designed to take the investigator through the causes of the incident, beginning with what happened immediately prior to the incident through to the threats that exist at the work site and in the system* | 2 | 1 | 1 |
| 5. Is the model a technically sound framework that can test the quality, validity and relationships of data developed during an investigation? *With sufficient data, it would be possible to test the quality and validity of the data found using HFIT. The database version could allow for relationships to be tested within Excel* | 2 | 1 | 1 |
| 6. Is the model theoretically consistent with or provide consistency for the company's safety programme concepts? *The company includes human factors into their safety programme—this form helps to enhance it* | 1 | 1 | 1 |
| 7. Does the model provide for direct identification of safety problems so that prompt correction can be made? *HFIT provides a systematic method for investigators to source the causes of the incident, and with the specific 'item' questions, identification of safety problems are readily identified* | 2 | 1 | 2 |
| 8. Does the model make it possible to link accident descriptions to the work process in which the accident occurred? *In the database version, after each cause has been identified, investigators are encouraged to provide explanations for the causes chosen. Here, the investigators would identify the work process and accident link. This may also be described in the narrative description* | 2 | 1 | 1 |
| 9. Does the model show interactions among all parties and things, rather than oversimplification? *HFIT captures a large set of human factors issues, which could be used to show interactions* | 2 | 0 | 1 |
| 10. Does the model enable investigators and others to see the relevance of the model to any accident under investigation easily and credibly? *Yes, the human factors issues covered in this model could be applied to many accidents* | 2 | 1 | 2 |
| Total score | 19/20 | 10/20 | 13/20 |

Table 5
Evaluation of the HFIT method according to Benner's evaluation system

| | Score (HFIT) | Score (Form1) | Score (Form2) |
|---|---|---|---|
| 1. Does the method encourage employees to participate in investigations and to have their views heard? *HFIT can be used to interview witnesses directly, thereby encouraging people to have their views heard* | 2 | 2 | 2 |
| 2. Does the method produce blameless outputs and identify the full scope of the accident, including the role of management and supervisors? *HFIT starts with the errors that occurred immediately prior to the incidents, although the main part of the investigation is taken up with the threats further back in the system* | 2 | 1 | 1 |
| 3. Does the method support personal initiatives? | 0 | 0 | 0 |
| 4. Does the method support timely discovery process? *The process is very thorough and has been found to really get to the underlying causes of the incident between 1.5 and 2 h—the systematic and thoroughness of the process supports timely discovery* | 2 | 2 | 2 |
| 5. Does the method increase the competence and safety effectiveness of personnel, such as used in training? *Yes, there is a section that identifies training issues* | 2 | 2 | 2 |
| 6. Does the method show definitive corrections so that remedial actions can be defined, evaluated and selected? *Yes, the specific questions provide the investigator with more detailed accounts of possible causal factors. The database version encourages investigators to provide preliminary remedial actions* | 2 | 1 | 2 |
| 7. Does the method provide information about duties under a standard with regard to the enforcement programme? *No* | 0 | 0 | 0 |
| 8. Does the method provide a practical way to produce consistent, reliable accident reports, hence encouraging the company to take responsibility, to fulfill their occupational safety and health mandates? *The systematic process encourages consistent, reliable reports* | 2 | 1 | 1 |
| 9. Does the method allow for accidents to be technically "truth-tested" to assure the quality of the information? *Yes, other witnesses who are asked for their version using HFIT will provide more information as well as the possibility for 'truth-testing' the data* | 1 | 1 | 1 |
| 10. Is the method compatible with 'pre-investigations' (or safety analyses) of potential accidents? *There is a set of questions in the HFIT that ask about the planning of the job that includes risk assessments. Although this may not be compatible with the system used in the companies* | 1 | 1 | 1 |
| Total score | 14/20 | 11/20 | 12/20 |

## 5. Conclusions

The Human Factors Investigation Tool has been subjected to a preliminarily evaluation in order to determine its effectiveness for analysing the causes of incidents in the offshore oil industry using an accident scenario exercise, three case studies, user evaluations and Benner's (1995) evaluation system.

The accident scenario exercise provided information regarding only a very small proportion of the causal codes in HFIT. Overall, the results indicated that the overall level of agreement between investigators was low, perhaps due to them having received minimal training and not being very familiar with the tool. However, the following six elements were thought to be causes of the incident scenario by over three-quarters of investigators possibly indicating that these elements are better understood and usable by the participating investigators: two action error elements: 'omission' and 'communication'; two situation awareness elements: 'attention' and 'assumption', one error recovery element (response behaviour) and two threat elements: 'communication' and 'supervision'.

Using the accident scenario data, the investigators' responses were compared to those of the human factors investigation findings. The scores were found to differ greatly across investigators, with some investigators being very close to 'the investigation' findings while others were very different. The level of agreement between investigators and 'the investigation' findings using HFIT to code an incident is generally low (overall mean = 0.38). This result could be due to the very large number of possible causal codes, investigators minimal training and practice with HFIT. In addition, the amount of information contained within a paper-based incident scenario is very limited, therefore making the exercise quite different from an actual investigation.

The inter-rater reliability tests indicated that investigators were more likely to agree with each other at the more general level (i.e. element level) rather than with regard to the specifics (i.e. the item or sub-element levels). At the sub-element, element and category levels, investigators were in least agreement with the 'investigation findings' regarding threats, situation awareness and error recovery respectively. This may indicate that threats and situation awareness problems are better understood at the more general level and error recovery is better understood at the more specific level.

The results from the three case studies have provided some initial evidence that HFIT improved the analysis of the incidents, where additional codes were identified from the HFIT analysis that could not be coded by the company's original coding system and the company used the findings from HFIT to develop the remedial actions. It must be noted that this is a prototype with some encouraging preliminary findings.

In order to improve the reliability of the tool (i.e. the agreement between investigators) it would be necessary for the investigators to share the same understanding of the categories, elements, sub-elements and items. In order for this, investigators may require more on-the-job training with the tool, and perhaps some sort of calibration between investigators during training. The question format of HFIT should in fact enhance the reliability as detailed questions explain the meaning of each label. Due to the nature of accident investigations (many possible contributing factors) it is very difficult to obtain "clean" reliable data from incident investigations. Reliability could be enhanced by 'team' investigations, rather than individual investigations (individual investigations were undertaken in this experiment); and investigators having more familiarity with the tool. In order to assess the reliability and validity of HFIT more fully, more incident data needs to be collected using HFIT. In addition, interventions could be designed and implemented on the basis of recommendations that come from using the tool and these interventions could then be evaluated to see if they reduce accident and incident rates (or certain types of incidents). Although these tests were out of the scope of this study, they are planned for future work.

Some problems with the method of recording the results were experienced, making it difficult to retrace where the results could have originated. It is important that the data are recorded, so that if the incidents are ever reviewed, the findings and the evidence for the findings are documented. It is also important to record the data if they are to be used to analyse trends of the causes of accidents. This problem could be eliminated using the computer version, which automatically collects and records each level of the data.

The case studies provided examples of where the investigator stopped at the subelement level (at least when recording the data), and when further analysis should have taken the investigator to the 'endpoint' where they understood more about the threat, which could have aided them in the development of the remedial actions. Furthermore, the case studies revealed that a possible reason why investigators were sometimes not recording the complete findings from the HFIT analysis was that they felt others reading the findings may not understand some of the labels used to describe the causes. In order to evaluate the reliability and validity of HFIT more fully, more incidents need to be investigated using HFIT. The tool is currently being used by one of the participating companies in Alaska and Brazil and it is hoped that after a longer period of data gathering, statistical comparisons between the HFIT and pre-HFIT incident data could be undertaken.

The Human Factors Investigation Tool was developed with the intention of it being used by engineers not necessarily expert in human factors. However, in order for investigators to feel confident using the tool, investigators were first given basic human factors training and training to understand the structure of the tool and to practice using it. Overall, the investigators indicated that the training was either 'satisfactory' or 'good' and as the trainers became more practised, the participants' satisfaction ratings improved. Although participants indicated that the training was good or satisfactory, the human factors training should be further evaluated to ensure that it is providing participants with information that will help them to investigate incidents with regard to the human and organisational causes.

The implementation of HFIT into the incident investigation procedures of the participating companies indicated a very poor result, where only one out of the four participating companies collected data using HFIT. Lack of time and resources and no incidents to report were the reasons given for this poor response. In order for companies to implement HFIT, management support for the tool needs to be expressed to the potential users, encouraging them to make use of the tool and presenting potential users with examples of how this tool can aid their investigations. One of the main issues seems to be the cost and resources implications for implementing new tools especially for large, international organisations. In addition, keeping track of the use of the tools has proved difficult, as personnel in these companies often move positions and many of the companies have merged.

The evaluation of HFIT using Benner's (1985) system indicated an improvement over the previous two reporting forms, although further refinement of the tool would be necessary in order for it to be compatible with individual company's safety management systems.

Although this tool was developed specifically for the oil industry, many of the elements could be used in other industries. The specific questions relating to the subelements and items may need to be customised for the particular industry, as examples from the oil industry are provided. Each of the categories: 'action errors', 'situation awareness', 'error recovery' and 'threats' could be applied to other industries. HFIT has been used in the shipping industry where it was recently trialed on 3 accidents. The four main HFIT categories were found to be transferable within this domain as were the majority of items within them.

Although this paper has highlighted the need for further validation and analysis of HFIT either using more realistic accident scenario exercises or gathering more data from incident analysis, it has been possible to gain a better understanding of the clarity of some elements. In particular, it has shown in a small sample of accident investigations, that HFIT has helped to identify additional human factors causes from the traditional incident investigations.

## Acknowledgement

This research was sponsored by four oil companies and the Offshore Safety Division of the UK Health and Safety Executive. The views presented are those of the authors and should not be taken to represent the position or policy of the funding bodies.

## References

ADAMS Consortium, 1998. ADAMS Reporting Form and End-User Manual. BRPR-CT95-0038, European Commission, Ispra, Italy.

Baachi, M.B., Cacciabue, C., O'Connor, S., 1997. Reactive and pro-active methods for human factors studies in aviation maintenance. Paper Presented at the 9th Symposium on Aviation Psychology, Columbus, OH.

Banbury, S., Tremblay, S., (Eds.), 2004. A Cognitive Approach to Situation Awareness: Theory and Application. Aldershot: Ashgate.

Benner, L., 1985. Rating accident models and investigation methodologies. Journal of Safety Research 16, 105–126.

Bird, F.E., Germain, L., 1985. Practical Loss Control Leadership. Institute Publishing (Division of International Loss Control Institute), Loganville, GA.

Boeing, 1995. Available from <http://www.boeing.com/news/releases/1995>.

CHIRP, 2000. Available from <http://www.chirp.co.uk/air_transport/Chirp_Summary.htm>.

Dekker, S., 2004. Ten Questions about Human Error. LEA, Mahwah, NJ.

Endsley, M.R., Garland, D.J. (Eds.), 2000. Situation Awareness: Analysis and Measurement. Erlbaum, Mahwah.

Fahlbruch, B., Wilpert, B., 1997. Event analysis as problem solving process. In: Hale, A., Wilpert, B., Freitag, M. (Eds.), After the Event: from Accident to Organisational Learning. Elsevier Science, Oxford, pp. 113–129.

Goldstein, I.L., Ford, J.K., 2002. Training in Organisations: Needs Assessment, Development and Evaluation, fourth ed. Wadsworth Group, Belmont, CA.

Gordon, R., 2002. The contribution of human factors to incidents in the UK offshore oil industry: development of a human factors investigation tool. Unpublished PhD thesis. Department of Psychology, University of Aberdeen, Scotland.

Gordon, R., Flin, R., Mearns, K., under review. A comparison of the causes of accidents from safety climate surveys and accident analysis studies.

Gordon, R., Mearns, K., Flin, R., 2000. The development and evaluation of a human factors accident and near miss reporting form for the offshore oil industry. In: Mearns, K. (Ed.), Factoring the Human into Safety: Translating Research into Practice, vol. II. HSE Books, London.

Gordon, R., Mearns, K., Flin, R., 2002. The development and evaluation of a human factors investigation tool (HFIT) for the offshore oil industry. In: A Joint HSE/Oil Industry Sponsored Project (HSE Report Reference D3933). HSE Books, London.

Green, M., Morisseau, D., Seim, L.A., Skriver, J., 2000. Development of an incident investigation process. In: Proceedings of the Society of Petroleum Engineers Conference on Health, Safety and Environment in Exploration and Production, Stavanger, Norway. Society of Petroleum Engineers, Richardson, TX.

Helmreich, R., Klinect, J., Wilhelm, J., 1999. Models of threat, error and CRM in flight operations. In: Proceedings of the 10th International Symposium on Aviation Psychology. The Ohio State University, Columbus, OH, pp. 677–682.

Hollnagel, E., 1993. The phenotype of erroneous actions. International Journal of Man–Machine Studies 39, 1–32.

Howell, D.C., 2002. Statistical Methods for Psychology, fifth ed. Wadsworth Group, London.

Hudson, P., Primrose, M.J., Edwards, C., 1994. Implementing tripod-DELTA in a major contractor. SPE Paper 27302. In: Proceedings of the the SPE International Conference on Health, Safety and Environment, Jakarta, Indonesia. Society of Petroleum Engineers, Richardson, TX.

James, L.R., Demaree, R.G., Wolf, G., 1984. Estimating within-group interrater reliability with and without response bias. Journal of Applied Psychology 69 (1), 85–98.

Johnson, W., 1980. MORT Safety Assurance Systems. Occupational Safety and Health 4.

Kirwan, B., Shorrock, S., Isaac, A., 1999. Human error in European air traffic management: the HERA project. In: Conference on Human Error, Safety and System Development (HESSD). Liege, Belgium.

Kontogiannis, T., 1999. User strategies in recovering from errors in man–machine systems. Safety Science 32, 49–68.

McDonald, N., 1998. Human factors and aircraft dispatch and maintenance safety. Paper presented at the Nouvelle Revue D'aeronautique et d'astronautique. 3 Aero Days Post-Conference Proceedings.

Mearns, K., Flin, R., Fleming, M., Gordon, R., 1997. Human and Organisational Factors in Offshore Safety (Offshore Technology Report OTH 543). UK Health and Safety Executive, London.

O'Leary, M., 1999. The British Airways human factors reporting programme. Paper presented at the Human Error, Safety and System Development Conference Liege, Belgium.

Paradies, M., Unger, L., Busch, D., 1996. TapRooT. Root Cause Tree Users Manual. System Improvements Inc, Knoxville, Tennessee.

Paradies, M., Unger, L., Haas, P., Terranova, M., 1993. Development of the NRC's Human Performance Investigation Process (HPIP) (NUREG/CR-5455 SI-92-101 vol. 2). System Improvements, Inc. and Concord Associates, Inc. Washington, DC.

Rasmussen, J., Pedersen, O.M., et al., 1981. Classification System for Reporting Events involving Human Malfunctions. Ispra, Italy, Joint Research Centre.

Reason, J., 1997. Managing the Risks of Organizational Accidents. Ashgate, Aldershot.

Reynard, W.D., Billings, C.E., Cheaney, C.E., Hardy, R., 1986. The Development of the NASA Aviation Safety Reporting System (Reference Publication 1114): NASA.

Sneddon, A., Mearns, K., Flin, R., under review. Situation awareness and safety in the drilling industry.

Stoop, J., 1997. Accident scenarios as a tool for safety enhancement strategies in transportation systems. In: Hale, A., Wilpert, B., Freitag, M. (Eds.), After the Event: from Accident to Organisational Learning. Elsevier Science, Oxford, pp. 77–93.

Swain, A.D., Guttman, H.E., 1983. A Handbook of Human Reliability Analysis with emphasis on Nuclear Power Plant Applications. Washington, DC 20555, USNRC.

West, G.J., Eckenrode, R.J., Goodman, P.C., 1991. 2–6 September Investigation of events involving human performance. Paper presented at the Proceedings of the Human Factors Society 35th Annual Meeting, San Fransisco, CA.

Wickens, C.D., 1992. Engineering Psychology and Human Performance. HarperCollins, New York.

Wickens, C., Hollands, J., 2000. Engineering Psychology and Human Performance, third ed. Prentice Hall, Upper Saddle River, NJ.