

## **“How do I save it?” Usability evaluation of a systems theory-based incident reporting software prototype by novice end users**

Eryn Grant<sup>1</sup>, Natassia Goode<sup>1</sup>, Paul M. Salmon<sup>1</sup>, Michael G. Lenné<sup>2</sup>, Bridie Scott-Parker<sup>1</sup> and Caroline F. Finch<sup>3</sup>

<sup>1</sup> University of the Sunshine Coast Accident Research, University of the Sunshine Coast, Locked Bag 4, Maroochydore DC, QLD, 4558, Australia

<sup>2</sup> Monash University Accident Research Centre, Monash University, Building 70, Melbourne, VIC, 3800, Australia

<sup>3</sup> Australian Centre for Research into Injury in Sport and its Prevention (ACRSIP), Federation University Australia, PO Box 663, Ballarat, VIC, 3353, Australia

{[egrant1@usc.edu.au](mailto:egrant1@usc.edu.au), [ngoode@usc.edu.au](mailto:ngoode@usc.edu.au), [psalmon@usc.edu.au](mailto:psalmon@usc.edu.au), [michael.lenne@monash.edu](mailto:michael.lenne@monash.edu), [bscottpa@usc.edu.au](mailto:bscottpa@usc.edu.au), [c.finch@federation.edu.au](mailto:c.finch@federation.edu.au)}

**Abstract:** The level of usability achieved by software tools is a key factor that determines their success and indeed uptake by end users. This paper describes a study that was undertaken to evaluate the usability of a prototype incident reporting software tool. The study involved novice end users completing a series of tasks using the software tool and then completing Ravden and Johnson’s Human Computer Interaction (HCI) checklist. The findings identify aspects of the system that pose particular challenges for participants. Participants appeared to lack a clear understanding of the relationship between the information required from them, and the underpinning accident analysis method of the software tool. This is perhaps unsurprising, considering that most incident reporting systems do not include these functions. The findings indicate that the tool requires better levels of intuitiveness to assist users in complex tasks so the focus is on awareness of accident causation methods rather than task instructions. The implications for the design of incident reporting software tools are discussed.

## **1 Introduction**

Incident reporting systems are fundamental to understanding the causes of accidents and injuries in most domains [1, 2]. In recent times the importance of designing incident reporting systems in line with contemporary models of accident causation has been emphasized[3]. To collect data about causal factors from across the overall work system, incident reporting systems need to be underpinned by a system-theory model of accident causation and corresponding analysis method [4]. However the difficulty

with such incident reporting systems often lies in their system usability, difficult terminology and perceived effort and utility, which create frustration and deter people from using software systems to report incidents [5, 6].

The authors recently developed a prototype incident reporting system for the led outdoor recreation sector in Australia [3]. The system, known as UPLOADS (Understanding and Preventing Led Outdoor Accidents Data System), is underpinned by Jens Rasmussen's seminal risk management framework and accompanying Accimap methodology (Rasmussen, 1997). The tool was developed as part of a major project attempting to address the need to collect reliable and accurate data on near misses and injury causing incidents [7, 8]. In addressing this need, the UPLOADS prototype was designed with an underpinning human factors methodology, which provided a whole systems based structure for the collection of information on near misses and injury causing incidents

A key part of UPLOADS development involves testing its usability with end-users. The aim of the study described in this paper was to perform a usability analysis of the UPLOADS incident reporting system. The study involved end-users performing a series of tasks with the UPLOADS system and then completing a HCI checklist [9] designed to assess the usability of software tools.

## **2 Method**

### **2.1 Design**

The study was a usability evaluation. It involved participants using the UPLOADS software prototype to input simulated information and other relevant records. During this process participants recorded information about their experience of the software. The study was approved by the University of the Sunshine Coast Human Ethics Committee.

### **2.2 Recruitment**

Managers and staff from the outdoor education and recreation sector who play a key role in risk management within their organisation were invited to participate. Through these invitations 26 people volunteered to participate.

### **2.3 Materials**

Materials for the study included an instruction booklet, a response booklet, a stopwatch and a copy of the software for participants to install on their own computer system.

**Software Tool.** The software was developed in FileMaker Pro 12 and Java. The software consisted of: four linked databases for collecting data (incidents; staff; people; and course); a tool for classifying the contributing factors, and relationships, identified in incident reports; a tool for summarizing the contributing factor and relationships data collected; and a tool for exporting deidentified data (e.g. names removed) to send to the research team.

**Instruction and response booklets.** The instruction booklet provided participants with information specific to operating the software as well as providing task information to facilitate using the functions of the incident reporting tool including installing the software and other general information about the software including log in information and general navigating instructions.

The booklet contained descriptions of a series of tasks for participants to complete. Each task provided “mock” information, which allowed them to complete each step in the task description. Participants were directed at the beginning and end of each task to start/stop their stopwatch. The following tasks were contained in the instruction booklet:

- Task 1 asked for demographic information to gain an understanding of the people participating in the evaluation and the organisations they work within;
- Task 2 add a template for a course;
- Task 3 add a record for a staff member;
- Task 4 add a record for a participant;
- Task 5 importing a “people record” into the database;
- Task 6 enter a course record;
- Task 7 entering an incident report;
- Task 8 coding causal factors and relationships;
- Task 9 exporting data for analysis; and
- Task 10 generate a graphic from the data recorded in the database.
- Task 11 completing the HCI Checklist

The final section Tasks 12-14 involved recording any technical issues while using software and additional comments. The response booklets recorded the answers for each Task.

**HCI Checklist.** In order to appreciate the incident reporting tool’s usability from a novice end users perspective a usability assessment was completed using Ravden and Johnsons [9] Human Computer Interaction (HCI) checklist. Checklists are seen as both convenient and reliable this is because they enable the end user to compare and contrast a device, in this case a software prototype, against a set of conditions predefined in a checklist. They are intended to elicit the user’s subjective experience of the device through a simple rating scheme (i.e. is it good or bad) and space to provide comments about the interface and usability. The usability study assessed the 10

themes however 1-9 are represented in this study. The 10 themes are designed to assess the overall usability of a device or system according to the following dimensions:

1. *Visual Clarity*: information displayed on screen should be clear, well-organised, unambiguous and easy to read.
2. *Consistency*: The way the system looks and works should be consistent at all times
3. *Compatibility*: The way the system looks and works should be compatible at all times
4. *Informative feedback*: Users should be given clear, informative feedback on where they are in the system, what actions they have taken, whether these actions have been successful and what actions should be taken next.
5. *Explicitness*: The way the system works and is structured should be clear to the user.
6. *Appropriate Functionality*: The system should meet the needs and requirements of the users when carrying out tasks
7. *Flexibility and Control*: The interface should be sufficiently flexible in structure, in the way information is presented and in terms of what the user can do, to suit their needs and requirements of all users, and to allow them to feel in control of the system.
8. *Error Prevention and Correction*: The system should be designed to minimise the possibility of the user error, with in built facilities for detecting and handling those that do occur; users should be able to check their inputs and to correct errors or potential error situations before the input is processed.
9. *User Guidance and Support*: Informative, easy-to-use and relevant guidance and support should be provided, both on the computer and (via an online help facility) and in hard copy document form, to help the user understand and use the system.
10. *System Usability*: When using the system did you experience problems with any of the following

## 2.4 Procedure

Instruction and response booklets were placed in blank envelopes along with a postage paid return envelope for the completed booklets and the prototype software tool on a USB stick. These were delivered by mail to the people who volunteered via email. Completed booklets were returned directly to the University of the Sunshine Coast.

**Data analysis.** The qualitative feedback from each Task was coded according to the themes in the HCI checklist. One researcher independently analysed task related feedback responses and coded to them to the themes in the HCI checklist that best

described the issue using a qualitative software tool. A second researcher checked the coding of the feedback responses to HCI checklist themes. The few disagreements identified in coding feedback responses to HCI checklist themes were resolved through discussion.

### 3 Results

#### 3.1 Sample

The participants returned 22 of 26 complete booklets representing an 84.6% response rate. The participants were predominately male (12 male, 9 female, 1 unclassified). The median age was 42.5 years (range = 36) and median years' of experience working in the outdoor education and recreation sector was 12 (range = 30).

#### 3.2 Evaluation of the Software tool

As per Stanton et al [10] radar plots were used to display the participant ratings for each task as it related to the usability of the software tool. The centre point of the axis demonstrates a low usability score, the farther away from the centre of the axis the higher the usability score.

In the following sections, the ratings from the HCI Checklist and task-specific feedback are presented according to the following themes: visual clarity; compatibility; consistency; informative feedback, explicitness; appropriate Functionality; flexibility and control, error prevention and correction and user guidance and support.

**Visual Clarity.** Median ratings for visual clarity in the HCI Checklist are presented in figure 1. Participants' ratings indicate that there were no major issues with visual clarity. All of the sub-dimensions were rated as present at least "most of the time" or "always". In total 30 feedback responses were coded to visual clarity. In particular, the organisation of information during tasks underpinned by accident analysis methods was frequently identified as problematic (n=7). Many participants thought that the presentation of information was repetitive and too dense, "Way too much information to sort through to select correct category" (Participant 13).

Participants suggested that the coding section of the database required a better layout that separated information into smaller portions to make them more manageable (n = 5). In addition, participants commented that the format for entering information, such as dates and times was unclear (n = 3).

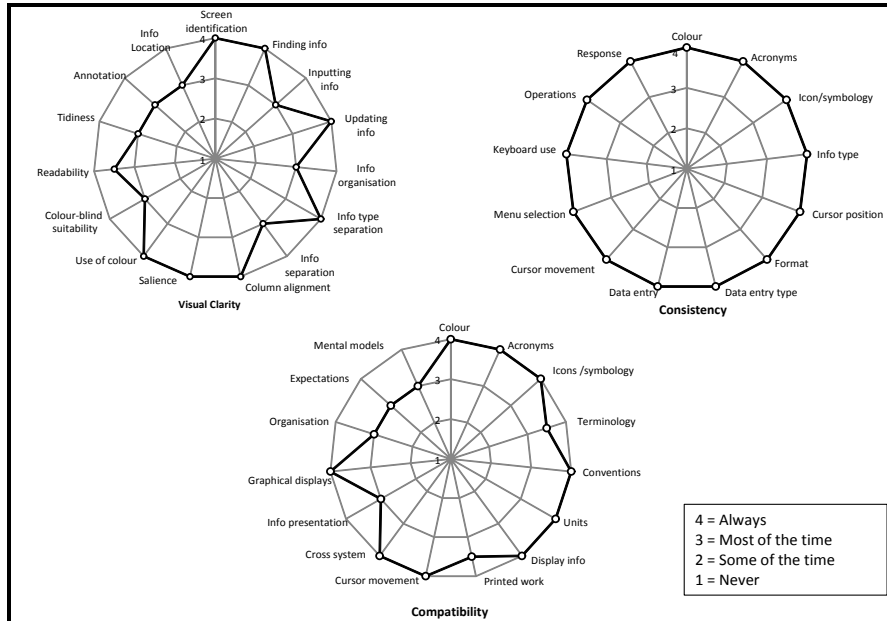


Fig. 1. Median results of Visual Clarity, Consistency and Compatibility

**Consistency.** The median consistency in the HCI Checklist are presented in figure 1. Overall the system scored highly for consistency; all sub dimensions were rated as “always”.

In total, 11 feedback responses were coded to consistency. The task-specific feedback was mostly compatible with HCI Checklist ratings. A small number of comments identified issues with moving through the database field using the tab key (n=2) and inconsistencies between calendars and drop down lists (n = 2). One participant commented that the layout of some screens were too long (n=1).

**Compatibility.** The median compatibility in the HCI Checklist are presented in figure 1. Overall the system scored highly for compatibility.

In total, 66 feedback responses were coded to Compatibility. Three major issues were identified. First, a number of feedback responses indicated terminology in tasks was not understood, (n=19). Second, a number of participants commented that alphabetical ordering, time display and date inputs did not follow normal conventions (n=13). For example the date input required more information than participants felt was necessary for some tasks, “I’d like to be able to just put the year, but it wouldn’t let me” (participant 14). Third, a number of participants were unable to use the Acci-map program, as they did not have a latest version of Java (n=7).

**Informative feedback.** The median informative feedback results are presented in figure 2. Overall the system scored highly for informative feedback. Mostly sub di-

mensions were rated as “always” present, the remaining rated as present “most of the time”.

In total 43 feedback responses were coded to informative feedback. Participant comments revealed significant issues with system responses. Specifically, the software did not have a clear response to indicate successful or unsuccessful completion of tasks (n=14), “Not sure if completed record has been saved” (Participant 10). Several comments referred to this issue. Similarly feedback indicated that the system did not always provide adequate instructions or prompts about what to do (n=6), for example, “ending ‘add records’ seems to end with ‘what now?’” (Participant 19).

**Explicitness.** The median explicitness results are presented in figure 2. Overall the system scored soundly for explicitness. Mostly sub dimensions were rated as present “most of the time” with one sub dimension rates as “always’ present.

In total 23 feedback responses were coded to explicitness. Minor issues were identified with the organisation of information such as the location of buttons or pages required to complete tasks (n = 7) “A couple of times I had to really look”. Some participants also commented that options in menus or task instructions were not meaningful, participants felt as if they were making the wrong choice or needed clarification (n =6). In relation to exporting data from the software, some participants stated that the system was not well organised (n = 4).

**Appropriate functionality.** The median results for appropriate functionality are presented in figure 2. Overall the system scored soundly for appropriate functionality. Mostly sub dimensions were rated as present “most of the time” with only two sub dimension rated as “always” present.

In total 33 feedback responses were coded to appropriate functionality. Feedback indicated task specific terminology needs to be defined at an early stage in the to prevent confusion. This was a strong issue associated with participant comments; “Terminology log would be useful” (Participant 8) (n=10). Participants also indicated they were not provided with all the options they felt were necessary mostly for inputting small details associated with task (e.g. dates/times) (n=5). Other comments suggested lengthy tasks could be broken down into manageable stages within the software particularly in tasks that require more effort than others, “I thought it is very timely as I would tick a box and provide a description and scroll down more and would see similar question which also relates to the incident. It might just be me, but I find it a little confusing sometimes.” (Participant 12)(n=4).

**Flexibility and control.** The median results for flexibility and control are presented in figure 3. Overall the system scored soundly for flexibility and control. Mostly sub dimensions were rated as present “most of the time” with one sub dimension rated slightly below “always” and one sub dimension rated as “always” present.

In total 40 feedback responses were coded to Flexibility and Control. Several comments (n=10) indicated that they system needs improved auto fill options, this included ‘default’ settings, calendar inputs and saving templates for later use. Similarly the need for more shortcuts was noted (n=8).

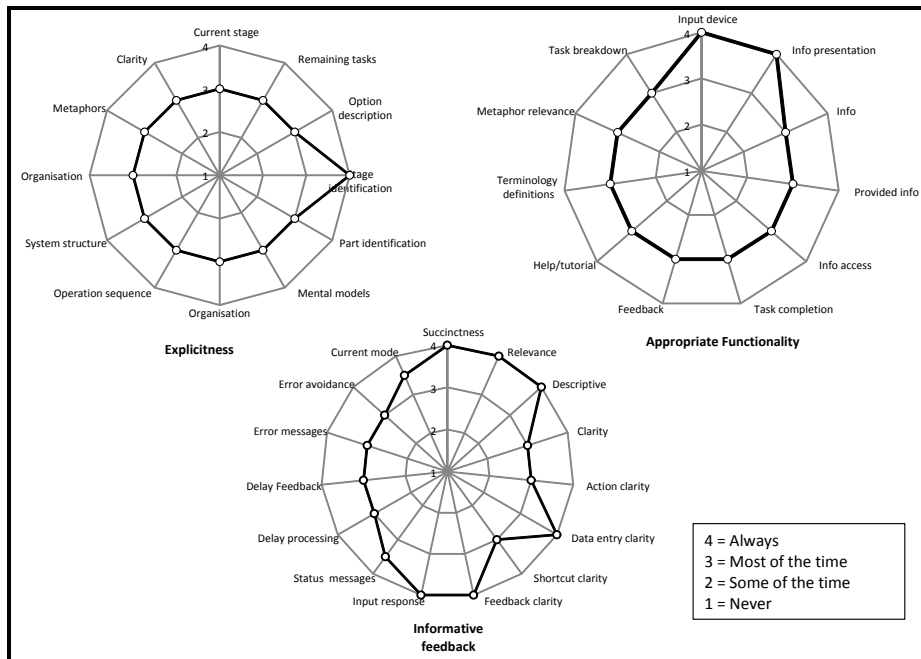


Fig. 2. Median results for Informative feedback, Explicitness and Appropriate Functionality

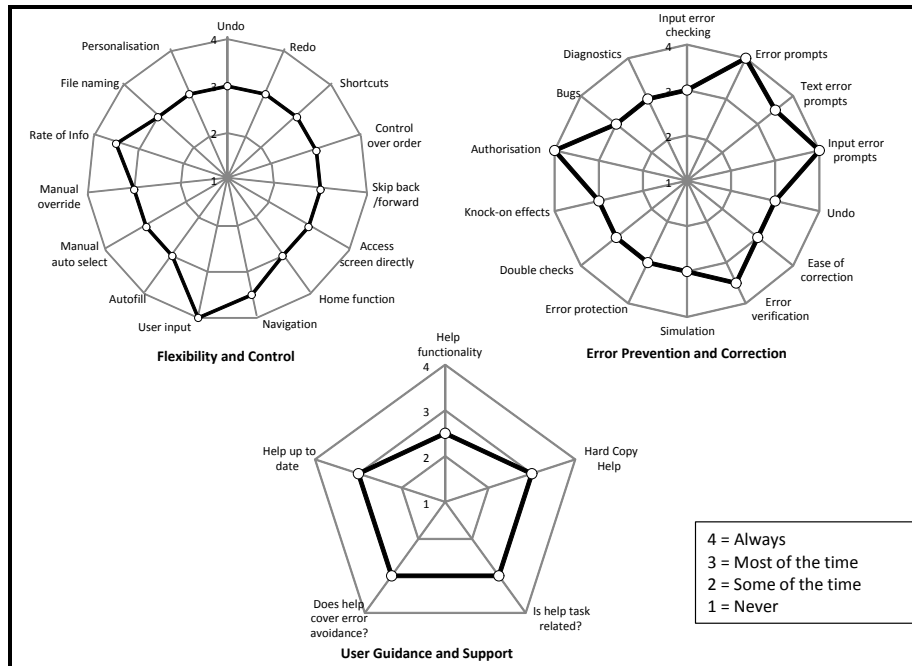
This feedback suggested simpler options for inputting large amounts of text and options to auto populate known fields, more dropdown menus, for example, “Forget the skip button...would be better if have info drop down if yes is selected”. Feedback also suggested that options to redo information needed to be more explicit in the software as it was not obvious to participants, “Frustrating that you can't go back and amend course name if you make mistake”.

**Error prevention and correction.** The median results for error prevention and correction are presented in figure 3. Overall error prevention and correction scored well. Mostly sub dimensions were rated as present “most of the time” with a smaller number rated as “always” present.

In total 27 feedback responses were coded to Error Protection and Correction. Feedback indicated prompts did not consistently report errors when known errors were present, for example participants in the graphic task did not get advice from the system that it was not operational, “The graph never came up. I wasn't going to wait any longer. I expect that java was not working properly” (OE Participant 14) (n=5). Other feedback indicated that the system does not adequately protect against common errors, “On one occasion I entered the system and without thinking changed a course by typing over the current screen rather than selecting new record. This was painful as it amended all previously entered incidents for original course.” (Participant 22) (n=5)



Additionally, feedback indicated that the system only validates user inputs on specific data (e.g. phone numbers and dates) and did not include a spell check function (n=4).



**Fig. 3.** Median results for Flexibility and Control, Error Prevention and Correction and User Guidance and Support

**User guidance and support.** The median results for user guidance and support are presented in figure 3. Overall user guidance and support scored soundly. Mostly sub dimensions were rated as present “most of the time” with one sub dimension scoring in between “most of the time” and “some of the time”.

In total 9 feedback responses were coded to User Guidance and Support. Participant feedback perceived that help functionality to be non-existent “I’m unaware of any help facility within the system” (participant 21) and a manual with images to demonstrate how to perform difficult tasks in the software would be beneficial (n=6). Feedback also indicated that user guidance may require an initial training stage to overcome the complexities of tasks in the software, “It requires training for people to use correctly” (participant 18) (n=1).

## 4 Discussion

The aim of this study was to assess the usability of a prototype incident reporting system underpinned by a systems accident analysis method. Taken together, participants’ usability ratings on the HCI checklist indicated moderate to high usability ratings regarding the 9 key themes assessed by the checklist. Notably, however, the

feedback provided by participants for each of the study tasks did not always follow these ratings. The majority of the discrepancies seem to reflect a knowledge gap associated with the tasks related to the systems theory accident analysis method. That is, participants had trouble executing tasks when they required deeper knowledge of the accident analysis method underpinning the system, rather than just general knowledge about using databases. This is an interesting aspect of developing theoretically appropriate incident reporting tools and is representative of the widely reported accident analysis and research-practice gap [14].

Predominately the software worked well in the areas of, consistency, explicitness, error prevention and correction and user guidance and support. Several 'simple' functions demonstrated issues with operational aspects of the tool. These issues included validating user inputs regarding saving data, shortcut clarity and an improved help facility. Mostly participant ratings and feedback within the checklist showed a need for improved communication for user actions (e.g. pop ups and save info) or available actions (e.g. shortcuts). Participants completed all tasks, except for the graphic task due to a JAVA issue on participants' own computers.

Problems were identified in relation to visual clarity, compatibility, informative feedback, appropriate functionality and flexibility and control. The major issues were executing tasks with an underpinning accident analysis method. The HCI Checklist revealed that the software's ability to assist with learning complex tasks was somewhat limited. Feedback responses indicated coding the contributing factors proved frustrating because of a perceived repetitiveness and information density (see Fig. 1 visual clarity). Equally participants had difficulty with terminology of complex task descriptions and category selections and appreciating the purpose of the tasks (see Fig.2 appropriate functionality). While these issues were not anticipated, it is not altogether unexpected as it is consistent with usability problems associated with complex systems [11].

While at face value differences between ratings and feedback may seem inconsistent, it demonstrated how the software tool usability is unable to bridge the knowledge gap by simply being "easy to use". Tasks designed with the underpinning accident analysis method were often seen as burdensome for the information they required. The extent that the checklist ratings and feedback contrasted at points illustrates a design issue for the software tool, as it was unable to fully support the transition of higher-level knowledge regarding accident causation methods and instead ill focused on task instruction.

To better position the incident reporting tools in terms of usability it would be more appropriate to see them as complex systems. This means that the tasks do not necessarily have to be "easy to use" however they ought to be "easy to learn". The nature of complex systems means they rely on information external to the system to make sense, the user is essential (and central) to the system for the knowledge they bring with them. For Reddish (2007) such systems must be designed with the user context in mind in order to provide them with the ability to see the results of the methods employed and reasoning for their particular area and their particular problem. This suggests a higher level of intuition is required from the software to bridge

this knowledge gap and facilitate higher-level learning for accident causation methods and limit the amount of instruction orientated to tasks [12].

Other incident reporting tools have relied heavily on tutorials and face-to-face instruction to bridge this gap [13]. The reliance on face-to-face instruction is time consuming and costly for both the instructor and the user community with no assurance of success, and therefore not the best option. The software must facilitate a better relationship between research and practice to successfully bridge the gap that currently exists between the two [14, 15]. Future research into how the software may help to bridge the research practice gap and understand more about the constraints and interdependencies of the system is needed [16].

The study had some notable limitations. Firstly, assumptions about participant hardware meant the graphic task was not uniformly completed by all participants. Secondly, the 'help' functionality in the software was limited. Had a better "help" system been present some features may have been more easily navigated and general usability may have been less troublesome.

In conclusion, the intention of the prototype software was to support the outdoor education and recreation community in applying a systems approach to accident causation and help organisations collect reliable and accurate data on incidents [8, 17]. The study found the general usability of the software tool is sound with only minor fixable issues. Tasks underpinned by the systems accident analysis methods were more challenging for participants, providing evidence that the research practice gap in this area has implications for software tool design. The HCI checklist provided a method to distinguish between the functionality of the software and the ability of novice end users to adapt to using the underpinning human factors method. While the incident reporting tool is operationally fine tasks underpinned by a systems approach need an improved interface design that aids user communities in gaining higher-level knowledge of accident causation methods.

**Acknowledgements.** This project was supported by funding from the Australia Research Council (ARC) in partnership with the Australian Camps Association, Outdoor Educators' Association of South Australia, United Church Camping, Outdoors Victoria, Outdoor Council of Australia, Recreation South Australia, Outdoor Recreation Industry Council, Outdoors WA, YMCA Victoria, The Outdoor Education Group, Girl Guides Australia, Queensland Outdoor Recreation Federation, Christian Venues Association, Parks Victoria, Victoria Department of Planning and Community Development, Outdoor Education Australia and the Department of National Parks, Recreation, Sport and Racing Australia. Caroline Finch was supported by a NHMRC Principal Research Fellowship (ID: 1058737). The Australian Centre for Research into Injury in Sport and its Prevention (ACRISP) is one of the International Research Centres for Prevention of Injury and Protection of Athlete Health supported by the International Olympic Committee (IOC). Paul Salmon's contribution to this article was funded through his current Australian Research Council Future Fellowship (FT140100681).

## 5 References

1. Dekker, S., *Drift into failure: from hunting broken components to understanding complex systems*. 2011: Ashgate Publishing, Ltd.
2. Reason, J., *Managing the risks of organizational accidents*. Vol. 6. 1997: Ashgate Aldershot.
3. Salmon, P.M., et al., *Translating Ergonomics legacies in practice: A new Rasmussen inspired incident reporting and learning system*. . *Applied Ergonomics*, Under review.
4. Salmon, P.M., M. Cornelissen, and M.J. Trotter, *Systems-based accident analysis methods: A comparison of Accimap, HFACS, and STAMP*. *Safety science*, 2012. **50**(4): p. 1158-1170.
5. Shojania, K.G., *The frustrating case of incident-reporting systems*. *Quality and Safety in Health Care*, 2008. **17**(6): p. 400-402.
6. Furuta, K., et al., *An incident reporting support system for airline cabin crew*. *Medicine Sciences and Bioengineering*, 2015: p. 3.
7. Salmon, P.M., et al., *Injury causation in the great outdoors: A systems analysis of led outdoor activity injury incidents*. *Accident Analysis & Prevention*, 2014. **63**: p. 111-120.
8. Goode, N., et al. *Trial of a new incident reporting system for the outdoor sector (UPLOADS): initial data and lessons learnt*. in *Proceedings of the 6th International Outdoor Education Research Conference*. 2014. Outdoor Education Australia.
9. Ravden, S. and G. Johnson, *Evaluating usability of human-computer interfaces: a practical method*. 1989, West Sussex, UK: Ellis Horwood.
10. Stanton, N.A., et al., *Human factors in the design and evaluation of central control room operations*. 2009: CRC Press.
11. Redish, J., *Expanding usability testing to evaluate complex systems*. *Journal of Usability Studies*, 2007. **2**(3): p. 102-111.
12. Norman, D.A., *The design of everyday things*. 2002: Basic books.
13. Gordon, R., R. Flin, and K. Mearns, *Designing and evaluating a human factors investigation tool (HFIT) for accident analysis*. *Safety Science*, 2005. **43**(3): p. 147-171.
14. Underwood, P. and P. Waterson, *Systemic accident analysis: Examining the gap between research and practice*. *Accident Analysis & Prevention*, 2013. **55**(0): p. 154-164.
15. Chung, A.Z.Q. and S.T. Shorrock, *The research-practice relationship in ergonomics and human factors – surveying and bridging the gap*. *Ergonomics*, 2011. **54**(5): p. 413-429.
16. Clegg, C.W., *Sociotechnical principles for system design*. *Applied Ergonomics*, 2000. **31**(5): p. 463-477.
17. Salmon, P. and N. Goode. *UPLOADS: Overview of a new approach to incident reporting for the outdoor sector*. 2013. 5th Asia Oceania Camping Congress.

