Applied Ergonomics 59 (2017) 104–113

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Latent error detection: A golden two hours for detection

Justin R.E. Saward*, Neville A. Stanton

Transportation Research Group, Civil, Maritime and Environmental Engineering and Science Unit, Faculty of Engineering and the Environment, Southampton University, Southampton, SO17 1BJ, UK

ARTICLE INFO

Article history:
Received 4 February 2016
Received in revised form
9 August 2016
Accepted 26 August 2016

Keywords:
Latent error detection
Aircraft maintenance
System cues
Human performance variability

ABSTRACT

Undetected error in safety critical contexts generates a latent condition that can contribute to a future safety failure. The detection of latent errors post-task completion is observed in naval air engineers using a diary to record work-related latent error detection (LED) events. A systems view is combined with multi-process theories to explore sociotechnical factors associated with LED. Perception of cues in different environments facilitates successful LED, for which the deliberate review of past tasks within two hours of the error occurring and whilst remaining in the same or similar sociotechnical environment to that which the error occurred appears most effective. Identified ergonomic interventions offer potential mitigation for latent errors; particularly in simple everyday habitual tasks. It is thought safety critical organisations should look to engineer further resilience through the application of LED techniques that engage with system cues across the entire sociotechnical environment, rather than relying on consistent human performance.

1. Introduction

Undetected error in any safety critical system generates a latent condition that can contribute to a future system failure thus the detection of latent error is an essential element of effective safety management (Rasmussen and Pedersen, 1984; Reason, 1997; Shorrock and Kirwan, 2002; Wiegmann and Shappell, 2003; Flin et al., 2008; Aini and Fakhru'l-Razi, 2013). A systems failure occurs when there is inadequate control of the sociotechnical factors (across multiple environments) within a defined operating context, which impacts human performance (Leveson, 2004; Woods et al., 2010). In a safety critical system, this can lead to a hazardous condition due to design-induced errors (Stanton et al., 2009a).

Errors or erroneous actions, that can result in a latent condition impacting safety, are a normal by-product of human performance variability induced by the sociotechnical environment thus a systems approach to error research is essential in seeking out interventions prompting the successful recovery from latent error; the systems view now being widely accepted (Hutchins, 1995; Dekker, 2014; Reason, 2008; Woods et al., 2010; Cornelissen et al., 2013; Hollnagel, 2014; Chiu and Hsieh, 2016). Indeed, Reason (1990) highlighted it is only with the benefit of hindsight that it is possible to label behaviour as erroneous (in that an action led to an undesirable outcome or safety failure). In adopting the systems view, some research has rejected the label 'human error' in favour of 'human performance variability', which includes both normative and non-normative performance (Woods et al., 2010; Dekker, 2014). This approach emphasises the broad spectrum of human behaviour, rather than a dichotomy, and therefore a need to engineer resilient systems (Hollnagel et al., 2006). Arguably, 'human error' remains a dominant term to describe performance variability, provided it is used carefully as an indicator of broader sociotechnical issues within system design (Stanton and Baber, 1996; Reason and Hobbs, 2003; Reiman, 2011). Thus the term 'error' or erroneous act is argued to refer to any situation where the required performance was not enacted as expected due to system induced influences. As such, the term error refers to performance variability, which encapsulates a range of human interactions within a given sociotechnical environment that includes non-normative and normative accepted behaviours in the workplace (Reason, 1990; Woods et al., 2010; Cornelissen et al., 2013). Here, the term latent error is used simply as a signpost to residual errors (system induced erroneous acts) that were not detected prior to an action or proximal to the action event, i.e. system-induced errors that pass undetected and therefore lie hidden in the sociotechnical system rather than individual human failings/error.

Saward and Stanton (2015a) introduced the relatively

* Corresponding author.
E-mail addresses: j.saward@soton.ac.uk (J.R.E. Saward), n.stanton@soton.ac.uk (N.A. Stanton).

http://dx.doi.org/10.1016/j.apergo.2016.08.016

unexplored phenomenon of Latent Error Detection (LED) in response to the seemingly spontaneous self-detection of latent errors post-task completion. A cohort of air engineers in Royal Navy aircraft maintenance reported this phenomenon, where LED appeared to offer a defence against latent errors. Combining a systems view of human error with a multi-process approach to error research, Saward and Stanton (2015b) offered early findings from an exploratory study that considered retrospective accounts of the proposed LED phenomenon from a sample of naval air engineers. They reported LED to be prevalent and that time, location and other system cues appeared to be important. It was recognised that retrospective recall is susceptible to memory decay effects thus their exploratory study required follow-up with a real-world study. LED describes events where system failures are successfully recovered by individuals via system-triggered detections of hidden errors; in each case, post task completion. Thus LED arguably aligns with performance that goes right, before suffering an adverse effect (Reason, 2008); in other words, Safety II situations when modelling accident causation (Hollnagel, 2014).

Due to the apparent paucity of LED research, a review of literature for transferrable theories has led to a multi-process approach to LED research that combines theories on Prospective Memory (PM), Supervisory Attentional System (SAS) monitoring and schemata (Saward and Stanton, 2015a). The PM element refers to the creation of intent to carry out an action and the SAS for monitoring of the schema-action-world cycle, which is characterised by the Perceptual Cycle Model (PCM). Schema theory, embedded within the PCM, highlights the human interaction with the sociotechnical system via the bottom-up processing of external sensory data against top-down knowledge of the world (schemas) within a perceptual cycle (Niesser, 1976; Cohen et al., 1986; Plant and Stanton, 2013a). This forms the transactional relationship between a schema-action-world cycle and system cues in the external world that trigger intended actions (Niesser, 1976; Norman, 1981; Mandler, 1985; Stanton et al., 2009b; Plant and Stanton, 2013a). Studies on PM indicate that intentions (schema selection) are 'loaded' into memory to act upon later, which generates a 'to-do list' or internal marker (Sellen et al., 1997; Marsh et al., 1998; Van den Berg et al., 2004). The SAS described by Norman and Shallice (1986) is thought to be the attentional mechanism to continually monitor the external environment for external cues and monitor the perceptual cycle for correct execution of intent on the 'to-do' list (Norman, 1981; Norman and Shallice, 1986; Smith, 2003; Einstein and McDaniel, 2005; Saward and Stanton, 2015b). The SAS is also argued to regulate schema housekeeping, which is a term used to simply highlight the function of monitoring the perceptual cycle to confirm an action is completed as well as collecting feedback from the action to facilitate learning and the acquiring of experience (Saward and Stanton, 2015b). Developed schemas that form internal memory structures, which can be accessed to respond to a particular task, are known as genotype schemas whilst phenotype schemas refer to the actual response when executing a task (Niesser, 1976; Reason, 1990). Thus schema housekeeping is thought to highlight the cyclic update of genotype schema (for schema learning) by reviewing previous schema-action-world information and thereby providing the opportunity for the detection of latent errors (Saward and Stanton, 2015b).

PM research has also found that the successful recall of intentions is cue dependent and can be triggered automatically by external cues present in environmental contexts (Tulving, 1983; Bargh and Chartrand, 1999; Guynn et al., 2001; Einstein and McDaniel, 2005). Particularly, easily recognised cues in the external environment, being mostly visual or auditory, are effective triggers of internal markers where written word cues have been found to be more likely to trigger recall than picture cues (Kvavilashvili and Mandler, 2004; Mazzoni et al., 2014).

Sellen (1994) offered that an operator could fail to detect an error because the success of the action could have been imperceptible; indicating the schema-action-world cycle is dependent on cue recognition (e.g. where an oil filler cap was not quite seated correctly or a similar but wrong lubricant was used on a component). Autonomous schema housekeeping may later highlight the genotype/phenotype mismatch; accounting for why latent errors can appear to come to mind spontaneously (Reason, 1990; Stanton et al., 2009b; Einstein and McDaniel, 2005). If cue information was imperceptible at the time an activity was carried out then differences in the activity or location associated with an LED event needs consideration. Saward and Stanton (2015b) findings highlighted the link between LED and cognition distributed across different sociotechnical environments (Hutchins, 2001; Stanton et al., 2014; Grundgeiger et al., 2014). Here LED was argued to be most successful when post-task schema housekeeping takes place immersed in the same or similar environment to that which the error occurred, although latent errors occurring in the workplace were also detected in unrelated surroundings. This extended task-related cue recognition across a range of unrelated sociotechnical environments where different environments could be accessed for cue information by internally visualising/reconstructing past activity. For example, when at home, driving a car, walking, showering, etc. Further, and by extending transferable theories on memory, LED is thought more likely to occur when alone (not interacting with others) and mostly during periods of unfocused attention such as inactivity, day dreaming or engaged in largely autonomous activities that do not require high levels of concentration on the task in-hand (Kvavilashvili and Mandler, 2004; Smallwood and Schooler, 2006; Rasmussen and Berntsen, 2011).

PCM theory describes a schema-action-world cycle (Plant and Stanton, 2013a) therefore time must be associated with the frequency of the cycle. Saward and Stanton (2015b) argued that the perceptual cycle persistently reviews past task performance through schema housekeeping, for which distributed cues must remain available across sociotechnical environments for LED to occur. This could facilitate a 'golden window' for LED to occur. Initial findings indicated most latent errors were detected within two hours of occurring. It was also argued that this persistence is largely autonomous, leading to the unintentional review of a past task that perhaps accounts for seemingly spontaneously chance detections, although the intentional review of past task is expected to be more successful than the autonomous condition.

Using the multi-process approach to systems thinking described above, the researchers wanted to advance existing literature associated with LED via the real-world study of naval air engineers. To determine whether this cohort exhibits normal cognitive behaviours, literature was reviewed for an appropriate instrument to employ. The Cognitive Failures Questionnaire (CFQ) scores an individual's propensity for everyday cognitive failures using 25 questions scored 0–100 against a 5-item Likert coding (Broadbent et al., 1982). A high mean score (>51) indicates a propensity for cognitive failures (Broadbent et al., 1982). Whilst organisational safety performance has rightly moved away from focusing on individual human failings to a system induced view of erroneous acts, knowledge of individual performance variability within a defined cohort is argued to remain important in anticipating the level of resilience that must be engineered into safe systems within the workplace (Reason, 2008; Woods et al., 2010; Reiman, 2011; Cornelissen et al., 2013). Wallace et al. (2002) administered the CFQ questionnaire to US Navy personnel whilst Bridger et al. (2010) studied a large cohort of naval personnel in the Royal Navy. Both found these cohorts to exhibit normal performance variability representative of skilled workers, which is a relevant benchmark to inform the current study. Literature also indicates that those with a high CFQ score are more susceptible performance variations leading to erroneous acts, due to poor executive function, yet they are likely to have developed a personal coping strategy in the knowledge that they are prone to cognitive failures (Reason, 1990; Wallace et al., 2002; Mecacci and Righi, 2006; Day et al., 2012). Applying systems thinking, it is argued that this should be interpreted differently: that those with a high CFQ score are likely to be less receptive to external cues that trigger the necessary schema response, for which unreported behaviours may have been created that engage with the sociotechnical environment to engineer resilience.

Overall, the researchers wanted to know how individuals engage with system cues for successful LED with the aim of advancing knowledge on the nature and extent of LED so that potential ergonomic interventions can be identified to engineer resilience. It was hypothesised that the exploratory findings by Saward and Stanton (2015b) would remain applicable in a real-world study and confirm that most LED events occur within two hours of the erroneous act; significantly, when alone during periods of unfocused attention. Further, sensory data from familiar everyday cues present within the engineer's workplace are expected to facilitate successful LED through engagement with the perceptual cycle, for which physical or auditory cues are expected to dominant (Saward and Stanton, 2015b). The hypothesis drives the requirement for real-world study to understand what promotes Safety II behaviour in the workplace in terms of LED, where it is believed resilience also exists through successful latent error recovery (Flin et al., 2008; Reason, 2008; Finomore et al., 2009; Woods et al., 2010; Hollnagel, 2014). The researchers selected a diary study as an effective method to capture everyday LED events (Reason, 1990; Cassell and Symon, 2004; Robson, 2011) whilst the CFQ was administered to simply affirm normal cognitive behaviours in naval air engineers by relating to research from wider populations.

2. Methodology

2.1. Participants

A convenience sample was conducted (Robson, 2011), which comprised representative numbers of junior and senior naval air engineers from the target population in Royal Navy helicopter squadrons. Air engineers all train and operate to the same standards and practices thus significant differences do not exist between squadrons in terms of the working environment or employment, which need accounting for in the analysis of data. Six squadrons were available for the study, consisting of 695 engineers, of which 173 engineers participated (mean age = 29.99 years, sd = 6.81, range 18–48). This represents 25% of the population and includes both males (n = 164) and females (n = 9). Female participants accounted for 5.2% of the sample, which is representative of the population. As the low count of females is not statically significant, no separate analysis of female responses could be conducted within the scope of the current study. Flanagan (1954) argued that the number of events was more important that number of participants, for which Twelker (2003) recommended no less than 50 events were needed for data to be meaningful. Thus 173 participants were considered acceptable to yield sufficient events for analysis within the resources available for the study.

Trainees were not selected to avoid variance due to undeveloped skills (Fitts and Posner, 1967). 60% attrition was anticipated due to participant dropout or unusable diary entries thus the minimum number of returned diaries was expected to be 70 and therefore sufficient for analysis.

Ethics approval was received from Southampton University (Ethics No. 13496) and a Participant Information Sheet (PIS) was created in accordance with Ministry of Defence research ethics committee guidelines.

2.2. Diary design

A self-report diary was used to capture everyday LED events observed in the workplace, thereby avoiding intrusion but with adjacency and detail (Reason, 1990; Cassell and Symon, 2004; Robson, 2011). The diary was constructed according to Flanagan's (1954) Critical Incident Technique (CIT) where the term critical simply refers to a significant LED event reported by the participant. Neutrally worded questions were generated according to multi-process theories shown in Table 1. Intentionally, the diary was not designed against questions from the CFQ as the diary was designed to capture system factors. Free text descriptions of LED events were avoided since Schluter et al. (2008) found experienced nurses found it hard to describe their error behaviours. Questions were designed to give largely quantitative responses as the exploratory study by Saward and Stanton (2015b) was qualitative. To help further ensure construction validity, diary methodologies were reviewed (Oppenheim, 1992; Sellen et al., 1997; Cassell and Symon, 2004; Johannessen and Berntsen, 2010; Mace et al., 2011; Robson, 2011).

2.3. Piloting

A squadron not involved in the main study was approached for 10 air engineers to practice administration and test the diary booklet. A small group of university research staff also tried the diary for general usability and question comprehension. Feedback from the pilot was provided via a follow-up interview. Based on Wiegmann and Shappell's (2001) guide for an effective taxonomy, participants were asked to comment on: the comprehensiveness of the diary questions; whether the questions were sufficiently wide-ranging and captured everything they wanted to record about their LED event; and how usable they found the diary booklet. The general readability of the PIS was also assessed against the Flesch reading ease score and amended to achieve a score of 60.1 (standard readability). Based on piloting, changes were also made to the administration and diary booklet to remove repetition, typographical errors and ambiguity in some questions.

2.4. Data collection procedure

Approval for the study was received from local engineering management prior to participants receiving a standardised verbal brief, which included an explanation of each diary question. Naval air engineers receive flight safety briefs and training on error types (GEMS: Reason, 1990) as part of the UK MoD aviation error management system. However, the researchers confirmed participant understanding of error types during the verbal brief and instructions were printed in the diaries, which included examples. The CFQ, participant register and consent forms were then completed prior to issuing the diary booklet. Participants were asked to record each LED event as near to the occurrence as possible to counter memory decay effects. To avoid the completion of the diary causing an unsafe distraction, a notepad was included in the booklet for participants to make quick notes for later completion. The notepad also allowed participants to record any additional comments they wanted to record to avoid limiting any important data not considered in the design of the study. After 2 months, the researchers personally collected completed diaries to preserve anonymity from line managers.

Table 1
Diary questions.

Factor	Question	Response options (additional comments in brackets not published in diary)
PM	Q1. Please give a brief description of the error event.	General narrative (to understand context for LED event)
Time	Q2. At what time did the error event occur?	Time of day
PCM	Q3. What type of task was it?	Complex/Simple/Don't Know (looking for task complexity)
Cue	Q4. What was the cue to do this task?	Event/Time/Both
PCM	Q5. What was the error type?	Slip/Lapse/Mistake/NK (according to Reason's (1990) GEMS)
Location	Q6. Where were you when the error occurred?	AMCO/Hangar/Line/Maintenance office/Issue centre/Storeroom/ Aircraft/Workshop/Flight Deck/Other (At Work locations)
Time	Q7. At what time did you recall the error (post task completion)?	Time of day (to calculate time between the error occurring and detection)
Location	Q8. Where were you when you recalled the error?	AMCO/Hangar/Line/Maintenance office/Crew room/Issue centre/ Storeroom/Aircraft/Workshop/Flight Deck/Home or Mess/Bed/Vehicle/ Gym/Other (At Work and Not At Work locations)
SAS	Q9. What were you doing when you recalled the error?	Planning/preparing maintenance activity/Conducting similar maintenance activity/Conducting dissimilar maintenance activity/ Walking/Driving a vehicle/Exercising (e.g. cycling, jogging)/Showering/ Eating/Socialising (e.g. in a bar)/General work-related discussion/ Daydreaming/Resting/Entertainment (i.e. reading, TV, internet, etc)/ Sleeping/Other
SAS	Q10. Did you intentionally review your past tasks/activities?	Yes/No
SAS	Q11. (If Q10 'yes') Was this part of your personal routine?	Yes/No
PCM	Q12. On checking your work, was the error:	Real/False alarm (looking for successful detection of an latent error)
Cue	Q13. Did anything in your immediate location appear to trigger the error recall?	Sound/Equipment/Document/Smell/Taste/General vista/Other (looking to identify system cues)
Cue	Q14. What were you thinking about at the time of the error recall?	Work-related thoughts/Non work-related thoughts
SAS	Q15. Were you alone when the error was recalled?	Yes/No
PCM	Q16. The specific error was very clear to me.	Likert coding: Strongly Agree = 1, Strongly Disagree = 5 (Clarity of error)
PCM	Q17. I was very confident that my past task was in error.	Likert coding: Strongly Agree = 1, Strongly Disagree = 5 (Error confidence)
SAS	Q18. The error recall occurred when I was highly focused on the activity at Q9.	Likert coding: Strongly Agree = 1, Strongly Disagree = 5 (Task focus)

3. Results

3.1. Description of sample

37% (n = 64) of engineers returned their diaries, which is close to the anticipated maximum of 40% determined from piloting. The mean age of those who returned their diary was 30.70 years (sd = 7.41, range 19–48). 38% (n = 24) of returned diaries were blank, as the participant had not experienced an LED event during the 2 months of the study. The 40 completed dairies contained 51 usable entries, after 13 entries were dismissed due to conflicting responses or the recorded error example was not an LED event. Overall, the minimum number of CIT events recommended by Twelker (2003) was achieved. A Kolmogorov-Smirnov test on age within the sample of 64 engineers who returned their diaries showed D(64) = 0.13, $p < 0.01$; indicating the distribution of sampled mean ages deviate from a normal distribution, positive skews towards a mode of 28. However, this is representative of the population in naval aircraft squadrons where there are approximately 2.5 times more junior engineers of a younger age than older senior engineers.

3.2. Analysis

Category variables from the diary questions were mapped against each other to construct a simple 87 × 87 matrix. The matrix was not used for analysis accept to facilitate a targeted approach to data analysis. The matrix provided a general indication that LED events were particularly associated with simple event-based tasks involving lapses. These events were mostly detected accurately (few false alarms) in the workplace without the intentional review of a past task and whilst attending to an on-going task working alone. Here, thinking about work and the presence of physical objects (cues) appear related. Thus the following analysis focuses any these areas.

3.2.1. CFQ scores

The mean CFQ score for the 64 engineers who returned their diaries was M = 38.00 (n = 64, sd = 10.77, se = 1.34), for which a Kolmogorov-Smirnov test showed D(64) = 0.10, p = n.s; indicating the distribution of sampled mean CFQ scores do not deviate significantly from a normal distribution. The mean CFQ score for air engineers who returned their diary but reported no LED events was M = 35.71 (n = 24, sd = 10.56) whilst M = 39.37 (n = 40, sd = 10.78) for those reporting a LED event. Whilst the mean CFQ score was slightly higher for those who reported an attentional failure, the mean is still low within normal range and a *t*-test showed no significance between group means (t = 1.32, df = 62, p = n.s), although a small effect exists (r = 0.2). Participants were also asked whether they intentionally reviewed (IR) past tasks for errors or if recall appeared to be spontaneous and therefore an unintentional review (UR) occurred. Those with a high mean CFQ score (≥51) reported slightly more UR (57%, n = 8) than IR (43%, n = 6) and those with a low mean CFQ score reported more URs (70%, n = 26) than IRs (30%, n = 11). A 2 × 2 contingency table was constructed using categories for high and low mean CFQ scores against UR and IR, for which a Chi-square test showed no significant association ($\chi^2_{(1)} = 0.79$, p = n.s).

3.2.2. Diary responses

Question 1 provided context to confirm no significant difference in operating environment existed compared to the initial study conducted by Saward and Stanton (2015b). Thus, intentionally, no qualitative analysis was attempted.

Questions 2 & 7 were used to calculate the time (T) between the error (e) and latent detection (d), recorded as *T*(e-d), for which

Table 2 provides a summary of mean times for T(e-d) against location. A Kolmogorov-Smirnov test on timing data showed D(51) = 0.36, $p < 0.05$; indicating the distribution of the times for T(e-d) deviated from a normal distribution (positive skew = 2.58). The distribution gave a mean time for T(e-d) = 120 min (n = 51, sd = 216, se = 30, range 2–1020 min) with a mode of 30 min and median of 30 min. Notably, 78% (n = 40) LEDs occurred within 120 min.

Question 3 recorded latent errors associated with either complex or simple tasks. For example, a complex task included maintenance activities such as rigging flying controls and in-depth fault diagnosis. Examples for simple tasks include checking oil levels, returning tools, basic data entry tasks and logistics activities such as sending and receiving stores. Participants reported 14%(n = 7) complex tasks and 86%(n = 44) simple tasks.

Responses to Question 4 indicated 92%(n = 47) were event-based, 2%(n = 1) task-based and 6%(n = 3) were recorded as both. For event-based activities, 64%(n = 30) were associated with UR and 36%(n = 17) were associated with IR, for which 71% (n = 12) of this group reported against Question 11 that their IR was part of a personal routine (not part of a mandated procedure). A 2 × 2 contingency table was constructed using the categories for review type against the main cue for the task carried out in error. Due to low frequencies a Fisher's Exact Test was carried out, which showed no significant association (P = 1.0, p = n.s). Counts for IR and UR are shown against each activity in Table 3.

Reason's (1990) GEMS taxonomy was used in Question 5, which expands on Norman's (1981) research that described error types based upon the incorrect use of schemas. Significantly more lapses were reported than mistakes and slips: 90% (n = 46); 6% (n = 3); and 4%(n = 2), respectively. Thus the post-task detection of latent execution errors (slips and lapses) represented 94% of all maintenance tasks reported by participants with planning errors (mistakes) accounting for 6%. Table 4 provides example narratives against the GEMS taxonomy and also Norman's (1981) schema-related error types for completeness.

Derived from Questions 6 & 8, Table 2 also provides a count for locations where the error occurred and LED event. 73%(n = 37) error events were detected in the same or similar environment to that which the error occurred. Note that the AMCO is the Air Maintenance Coordination Office where most aircraft paperwork is controlled and maintenance organised. The Issue Centre is where ground equipment and tools are issued. The Line and Flight Deck are where aircraft operations are conducted, which is similar to a Ramp in civilian contexts. These locations could be grouped as environments At Work and Not At Work as shown in Table 3. This shows 80% (n = 41) LED events occurred whilst At Work and 20%(n = 10) whilst at Not At Work. A 2 × 2 contingency table was constructed for environment against review type. Fisher's Exact Test showed no significant association (P = 1.0, p = n.s), although UR was dominant At Work (n = 27) and Not At Work (n = 7). Table 3 also shows activity at the time of recall, reported against Question 9. This indicates 25%(n = 13) participants were engaged with similar maintenance task, 20%(n = 10) were planning or preparing to conduct maintenance task or 14%(n = 7) were simply walking (between activities). The remaining activities (n = 21) are highlighted in Table 3, noting that activities covering dissimilar maintenance, exercising and entertainment are not included, as participants reported none. Table 3 also shows 33% (n = 17) of participants intentionally reviewed past tasks.

Question 13 asked participants to record anything in the immediate physical environment that they believed might have triggered their error recall. Responses are shown in Table 3, for which aircraft documentation accounted for 35%(n = 14) and aviation equipment 30%(n = 12). A 2 × 4 contingency table was constructed for environment against triggers for Documentation (n = 14), Vista (n = 5), Sound (n = 4) and Equipment (n = 12). Other (n = 5) was not used, as participants did not provide examples. A Fisher's Exact Test showed no significant association (P = 0.42, p = n.s); however, 30% (n = 12) of all reported triggers were At Work and related to aircraft documentation followed by 28%(n = 11) aviation equipment.

Table 3 also highlights responses to Question 14, which asked participants to record what they were thinking at the time their LED event. 78% (n = 40) reported work-related thoughts and 22% (n = 11) non work-related thoughts. Worked-related thoughts

Table 2
Error locations.

Location (n = 51)	Response	Count	Mean time T(e-d) (min)	Example
Q6: Location of Error Occurrence	AMCO	19		
	Hangar	8		
	Line	3		
	Maintenance Office	2		
	Issue Centre	8		
	Storeroom	3		
	Aircraft	4		
	Workshop	0		
	Flight Deck	2		
	Other	2		Crew room, Head Office
Q8: Location of LED (ungrouped)	AMCO	15	106	
	Hangar	9	24	
	Line	3	19	
	Maintenance Office	1	15	
	Issue Centre	3	45	
	Storeroom	0	–	
	Aircraft	3	28	
	Workshop	0	–	
	Flight Deck	3	23	
	Crew room	2	45	
	Home/Mess	5	273	
	Bed	1	420	
	Vehicle	2	392	
	Gym	0	–	
	Other	4	324	Locker room, briefing room, bar(×2)

Table 3
General environment against associated factors.

Environment (Q8)	Activity (Q9)	Review (Q10)		Work Thoughts (Q14a)			Non Work Thoughts (Q14b)			Physical Trigger (Q13)	
		IR	UR	Past	In-hand	Future	Past	Moment	Future	Item	Example
At Work (n = 41)	Planning/preparing maintenance (n = 10)	6	4	4	4	2				Document = 5	*Aircraft documentation*
										Vista = 1	*Scenery inside building*
										Other = 1	*None specified*
	Conducting similar maintenance (n = 13)	5	8	2	9	2				Equipment = 4	*Computer, rotor blades*
										Document = 6	*Aircraft documentation*
										Vista = 1	*None specified*
										Sound = 2	*Aircraft noise, headset volume*
	Walking (n = 7)	2	5	1	2	3			1	Equipment = 5	*Aircraft, tools, toolbox*
										Other = 1	*Felt keys in pocket*
	Eating (n = 1)		1						1	–	*None specified*
	General work discussion (n = 3)	1	2	1		2				Other = 1	*None specified*
	Daydreaming (n = 3)		3	1		1		1		Vista = 1	*None specified*
	Resting (n = 1)		1					1		–	*None specified*
	Other (n = 3), i.e. changing clothes, paperwork & auditing		3		2			1		Equipment = 2	*Screw bag*
										Document = 1	*Aircraft documentation*
Not At Work (n = 10)	Driving vehicle (n = 1)		1					1		Sound = 1	*Work-related topic on car radio*
	Showering (n = 2)		2				1	1		Equipment = 1	*Keys*
										Other = 1	*None specified*
	Socialising (n = 2)	1	1	1				1		Vista = 2	*None specified*
	General work discussion (n = 1)		1	1						Sound = 1	*Colleague's voice*
	Resting (n = 2)	1	1	1				1		Document = 2	*Aircraft documentation*
	Sleeping (n = 1)	1		1						–	*None specified*
	Other (n = 1), i.e. readying for work		1					1		Other = 1	*None specified*
	Totals	17	34	13	17	10	1	8	2	40	

Table 4
Example narratives against the GEMS taxonomy (Reason, 1990) and Norman's (1981) schema-related error types.

Example narrative	Erroneous act (error) classification	
	GEMS (Reason, 1990)	Schema action (Norman, 1981)
'I did not replace the oil filler cap correctly … had to go back and check'. *'Walking from the hangar to the flight deck I dropped a tool in my pocket'.*	Slip	Correct intention selected but faulty schema(s) activation.
'Having prepared a Lynx [helicopter] *for flight … the book* [aircraft documentation] *was completed. Just prior to launch, I realised that I hadn't cleared a Pt1 entry* [statement of required maintenance]*'.* *'Forgot to fit main rotor spectacles* [rotor blade securing device]*'.*	Lapse	Correct intention selected but schema(s) not triggered.
'Card raised [maintenance paperwork] *for a maintenance task required post flying serial … incorrect aircraft annotated on paperwork'.* *'Failed to co-ordinate* [complete] *a maintenance work order correctly'.*	Mistake	Incorrect formation of intent. Wrong schema selected based on incorrect perception of external sensory data.

ranged from thinking about past maintenance activities, to the task in-hand through to maintenance to be carried out at a later time. Non work-related thoughts ranged from past personal errands, simply existing within the 'moment' through to thinking about a personal task to be done later. Examples include a previous social event, in the moment watching TV or thinking about what PC game to play later. When an LED event occurred, 59% (n = 30) of participants were alone according to the responses to Question 15 (i.e. not actively engaged with another person such as talking or working on a task together), during which 87% (n = 26) occurred when carrying out largely autonomous tasks requiring little focused attention. Aviation related examples include planning and conducting simple maintenance tasks (n = 8) such as basic aircraft servicing, tool checks and simple logistic tasks. Non-aviation related included walking (n = 7), driving a car (n = 1), showering (n = 2), daydreaming (n = 3), resting (n = 2), sleeping (n = 1) and other (n = 2) such as changing clothes and getting ready for work.

Question 12 indicated that all participants checked their work when a latent error came to mind, for which 92% (n = 47) found the error to be real whilst 8% (n = 4) experienced a false alarm. Fig. 1 describes the distribution of responses (n = 50) to the Likert coding specified for Question 16, 17 & 18. Participants generally agreed 46%(n = 23) or strongly agreed 36%(n = 18) that their specific error was very clear to them. Participants were uncertain 26%(n = 13) or agreed 26%(n = 13) that they were very confident in their past task being in error whilst 28%(n = 14) strongly agreed. 30%(n = 15) strongly agreed that they were highly focused on the activity they reported at Question 9. This was closely followed by 28%(n = 14) who agreed although 20%(n = 10) strongly disagreed.

4. Discussion

4.1. Individual factors

4.1.1. CFQ scores

CFQ scores were not used to predict LED events or propensity for human failings. The mean CFQ score for participants who returned their dairy was similar to other studies of naval personnel (Wallace et al., 2002; Bridger et al., 2010). A low mean CFQ score indicated the sample of air engineers possessed good executive function found in skilled workers. Studies have shown this helps cope with high workloads (Finomore et al., 2009; Bridger et al., 2010). However, participants described routine habitual tasks thus workload did not necessarily influence LED. The majority of LED events were

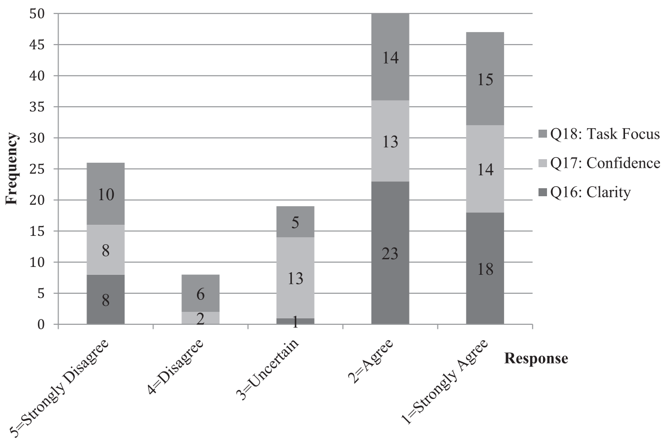

Fig. 1. Responses to questions on task focus, confidence and clarity.

associated with simple tasks. Arguably, the routine nature of everyday tasks infers low arousal. This may account for the high number of simple tasks although good executive function should result in increased cognitive awareness. Thus it can only be surmised that complex tasks tend to be more safety critical. Expectedly, an organisation's safety management system attempts to defend critical tasks with more layers error protection (Safety I: Hollnagel, 2014) than simple tasks; perhaps resulting in fewer examples of LED involving complex tasks.

A high CFQ score may indicate someone who is likely to be less receptive to external cues, against which behaviours may have been created that promote resilience. However, the sample possessed a low mean CFQ score and since more URs than IRs were reported, results appear to support literature that reports cohorts of skilled workers are less likely to need to deliberately check their work (compared to less skilled workers with high mean CFQ score) (Reason, 1990; Mecacci and Righi, 2006; Day et al., 2012). However, the researchers expected a greater number of IR than UR events as air engineers operate in a safety critical environment. A possible explanation is that the reported errors were of a potentially minor nature and everyday experience (e.g. forgetting to return equipment, keys left in pocket, basic data entry errors, etc). Thus general behaviours may exist within the cohort of naval air engineers where everyday minor tasks are not considered to represent sufficient concern to warrant deliberate checking. Here, the systems approach looks to offer mitigation either by promoting cues to achieve LED or through organisational resilience to undetected latent errors.

4.1.2. Error type

Lapses were associated with most LED events. Error research often reports the dominance of lapses, which offers an account for the overall high number of execution errors (Woods, 1984; Reason, 2008; Kontogiannis and Malakis, 2009; Blavier et al., 2005; Saward and Stanton, 2015b). The low count of LED events involving slips was expected as action errors tend to be detected proximal to the error event due to the transparency of the action (Reason, 2008), for which the action-world element of perceptual cycle also supports this expectation, as action errors should be readily apparent during performance monitoring. Arguably, mistakes tend to be less evident to attentional mechanisms due to the implicit lack of awareness of the genotype/phenotype schema mismatch for a given task, which may account for the low count in LED events associated with planning errors. Therefore, the post-task detection of latent errors is likely to be problematical for the PCM since important cues are simply not recognised (Norman, 1981; Plant and Stanton, 2013b). This incorrect perception of the world element of the PCM creates a lack of situational awareness. When situational awareness is absent during the perceptual cycle, it is argued that schema housekeeping is ineffective unless an important cue is recognised later. Thus reviewing past tasks can lead to a re-gain of PCM situation awareness as the genotype/phenotype schema mismatch between intent and the external world is identified from system cues that trigger error recall. The fact that this activity occurs after a task is completed may add to PM theories such that the internal marker or ‘to-do’ list is not deleted from memory upon completion; otherwise there would be no reference against which to conduct schema housekeeping. Also the detection of latent errors without a conscious attempt to review past tasks appears to confirm the presence of the SAS and an autonomous capability in schema housekeeping over time, although it is unclear whether this autonomy is systematic or simply dependent on cue recognition and is therefore indiscriminate.

4.2. Sociotechnical factors

4.2.1. Time

Most LED events occurred within two hours of the error event, which is a similar finding to Saward and Stanton (2015b). This is an important finding as Plant and Stanton (2013a) highlighted PCM theory involves a schema-action-world cycle, for which Saward and Stanton (2015b) argued that the perceptual cycle persistently reviews past task performance and thus time must be associated with the frequency of this cycle; facilitating a ‘golden window’ in which most work-related LEDs occur. Although the standard deviation for the current study was large, indicating the distribution of mean times would benefit from a larger sample to improve statistical confidence, timing data represent real-world evidence. Additionally, the correlation with a separate sample of air engineers offers confidence that there appears to be a relationship between time and the perceptual cycle, which can be persistent. T(e-d) for IR was found to be almost 3 times quicker than UR; indicating deliberately focused PCM attention can identify a latent error more quickly than the autonomous condition. In each case, LED events demonstrate information about a completed task remains accessible to the perceptual cycle, even though the intended task may have been removed from the internal ‘to-do’ list. Here it is perhaps autonomous schema housekeeping that is responsible for continued engagement with the perceptual cycle over time although deliberate intervention via an intentional review is also effective. If there is indeed a delayed cycle time associated with LED then accommodating this delay in maintenance planning may reduce the likelihood of latent errors transitioning to flight. However, this would be a challenging intervention in a real-world operating context.

4.2.2. Environment

LED is thought to be most successful when schema housekeeping takes place in the same or similar environment to that which the error occurred (Saward and Stanton, 2015b). In the current study, most LED events occurred At Work in the same or similar location to that which the error occurred (e.g. an error occurring in the hangar and was later recalled in the AMCO). This seems reasonable as most LEDs occurred within two hours whilst the engineers were still working. Work-related LED events also occurred when Not At Work although there were fewer reported but this does provide evidence that the PCM is able to remotely link to workplace cues to detect past errors, despite being physically present in a different sociotechnical environment. However, there are significantly greater delays to LED when Not At Work as highlighted in Table 2. In most cases though, error recall was associated with a cue. This demonstrates the importance of system cues distributed across environments to trigger recall.

4.2.3. Cues triggering recall

Participants reported familiar and recognisable workplace cues that they perceived to trigger schema recall. At Work, participants reported physical work-related cues such as sounds, aircraft paperwork, ground support equipment, toolbox, etc. Not At Work, participants reported physical items such as keys or a screw bag but also indirect perceptions of work-related cues. For example, a participant reported thinking about aircraft paperwork at home before realising an entry was made in error, another participant reported holding a general discussion about work with a colleague when a past error came to mind and one via an aircraft related news item on the radio. This appears to highlight the perceptual cycle's dependency on cues to trigger LED, which can involve the internal visualisation or phonological review of work-related cues in addition to physical cues away from the workplace. If the perceptual cycle can access representations of work-related cues when not in the workplace, either by visualising or reconstructing past activity internally, it is therefore argued important cues must remain available through residual memories and be sufficiently meaningful to trigger any genotype/phenotype schema mismatch; leading to successful LED (Reason, 1990; Stanton et al., 2009b; Saward and Stanton, 2015b). Kvavilashvili and Mandler (2004) found memories of past events where mostly triggered by easily recognised physical cues in the external environment, which are mostly visual or auditory. In the current study, At Work aircraft documentation and aviation equipment were prevalent whilst only 2 occasions of sound-related triggers were reported whilst At Work (aircraft noise and headset volume). Further, Mazzoni et al. (2014) argued that written word cues are more likely to trigger past memories (and thus LED) than picture cues. Although theirs was a laboratory-controlled experiment, current data appear to support this position since written words are implicit with aircraft documentation. Of note, the picture cues they employed were card cues and not physical objects encountered in the workplace. The only work-related pictures that the researchers found in the workplace (apart from technical diagrams in maintenance manuals) are those on Flight Safety posters but participants did not report LED events involving posters thus no analysis could be attempted.

For LED events where no physical trigger was perceived, this could indicate that the participant was simply not aware of the trigger or confirms autonomous schema housekeeping; perhaps giving rise to chance detections. In this situation, the general environment may be the trigger as opposed to specific cues. For example, Sellen (1994) found that intentions are often recalled due to contextual factors rather than spontaneous retrievals. Thus simply being immersed in an associated environment or recreating associations internally may aid LED.

Few LED events were associated with a false alarm, which suggests schema housekeeping often identifies latent errors correctly. Here participants mostly agreed or strongly agreed that their error was clear to them, although they were slightly less confident that the past task was actually in error. For example, “I don't think I replaced the oil filler cap correctly, but did I?” Arguably, this may imply weaknesses can lie in the transactional relationship between schemas and the action-world element of the PCM rather than the actual schema being faulty. LED events often occurred working alone on another task, which requires little focused attention. This seems to suggest the PCM has capacity to attend to schema housekeeping, even when attending to an on-going task, which may limit false alarms provided other people do not distract the engineer. Few participants thought about work-related topics when Not At Work whilst most reported thoughts ‘in the moment’ and ‘task in-hand’. Both are argued to be largely unfocused activities associated with autonomous behaviour and therefore may offer further evidence that cognitive capacity is afforded to schema housekeeping during these periods of activity. This should be expected according to Kvavilashvili and Mandler (2004) who found memory recall is most likely during periods of unfocused attention either during periods of inactivity or engaged in largely autonomous activities that do not require high levels of concentration on the on-going task. This risks a lack of cue recognition needed to assure situation awareness of the perceptual cycle proximal to the error event but also seems to suggest the PCM has in-built capacity to attend to an on-going activity whilst conducting schema housekeeping. This argument appears to receive support from Wilkinson et al. (2011) who similarly found expert operators exhibit the capacity to respond to cues related to the on-going task whilst also engaging with the external environment for error checking. That said, Fig. 1 shows participants also reported they were highly task focused when the LED occurred. Since all reported tasks are argued to be routine (see Table 3), it is considered more likely that they were not highly focused in terms of cognitive demands but more that the participant simply means it was the only task they were engaged with.

5. Interventions

The significant number of LED events associated with lapses may indicate weaknesses in a cues ability to trigger the intended schema action (Norman, 1981). Thus safety-focused organisations should consider the ‘strength’ of existing cues as an intervention to help enhance an error-detecting environment through assured engagement with the perceptual cycle. If physical objects (aviation equipment) and word cues (aircraft paperwork) mostly influence accurate LED, this may offer other avenues of engagement with the perceptual cycle. For example, specific word cues strategically placed alongside data entry areas in aircraft paperwork (i.e. the maintenance process or signature sheet) such as ‘check units’, ‘panels’, ‘keys’, ‘tools’, etc. This deliberately targets the triggers reported by participants. Aircraft paperwork (hard copy or electronic paperwork) is often completed in a separate location to where actual maintenance is conducted, therefore placing related paperwork near to where maintenance or supporting activities are carried out may improve transactional relationships; potentially reducing error initiation and/or enhancing LED. For example, signing for aircraft work as near to the aircraft as possible by moving necessary paperwork to the aircraft or issuing the air engineer with a portable e-tablet so specific task elements can be intentionally reviewed whilst remaining immersed in the same physical environment; thereby avoiding dissociation of socio-technical context. Further, messages on Flight Safety posters could be replaced with simple images of relevant objects or perhaps use small display stands positioned in locations such as the AMCO or Line. Displaying actual physical objects associated with common errors could enhance cue recognition. For example: an oil dipstick; padlock and key; fuel filler cap; or indeed a scaled model of the aircraft. A formal ‘stop and check’ of simple tasks, even for minor tasks, may offer effective intervention. This is thought to be most effective if conducted within two hours of a completed task and alone to avoid distraction, and if conducted in the vicinity of where the task was executed. Interventions are considered especially important for simple everyday habitual tasks carried out alone. Here the ‘Stop, Look and Listen’ strategy used in UK road safety campaigns could be applied to air safety. ‘Stop’ refers to the PCM cycle-time, ‘Look’ refers to sensing physical cues or the internal visualisation of past tasks and ‘Listen’ refers to phonological cues that could simply include the internal voicing of the ‘to-do’ list. Importantly, the golden two hours window in which most LED events occurred is an intervention on its own but clearly any of the interventions described above are likely to shorten detection times (depending on when the intervention is enacted).

6. Study limitations

Real-world research is clearly challenging and limited the data that could be collected but ecological data was essential to gain the necessary insight into erroneous acts with successful recoveries in the workplace, for which it is believed the current study has advanced understanding on LED. Analysis was limited to 51 usable LED events, which may mean LED is not as prevalent as thought or that the 2 months given to complete the diary was not long enough for more LED events to occur. Feedback from squadron engineering management is thought to provide a further explanation for the low return. It was highlighted that participant workload was very high and so they were not always able to make diary entries whilst a number of engineers were re-employed away from the squadron or on a short notice course. Additionally, the self-reporting diary approach may have biased the count of LED events due to increased vigilance. Thus future LED research would benefit from study of a larger sample over a longer period in a cohort that is able to commit fully to completing diary entries. Since the study was limited to a population of highly skilled engineers, a sample of unskilled workers should be considered as well as scenarios involving less familiar tasks and/or high workloads. In the study of human performance variability, only the context changes (Robson, 2011; Cheng and Hwang, 2015) thus it would be advantageous to conduct LED research in other workplace contexts where there are clearly more sociotechnical factors of interest than could be covered in the current study.

It was beyond the scope of the current study to report the safety risk of not detecting the latent errors highlighted in the current study. To make this assessment requires hierarchical task analysis and accident causation modelling to explore how particular latent errors or erroneous acts might contribute to a safety occurrence. Similarly, modelling of the entire At Work and Not at Work environments, to report frequencies of all complex/simple tasks carried out by the engineers in all locations where erroneous acts occurred and later recalled, was also beyond the scope of the current study. Participant high workloads and the safety critical nature of the observed squadrons did not permit this additional data collection as more extensive diary questions and/or separate observations would have been necessary, i.e. participants would need to report all complex/simple tasks carried out each day, in addition to 24/7 tracking of participant movements to report location frequencies. Thus, the risk/benefit of LED interventions, modelled against frequencies for all possible maintenance-related activities would warrant separate research.

7. Summary

The aim of the current study has been to advance knowledge on the nature and extent of LED through diary accounts of LED events so that potential ergonomic interventions can be identified. A multi-process approach to systems research was used that combines theories on PM, SAS and the PCM. Additionally, the CFQ was administered to simply affirm that the sample exhibited normal cognitive behaviours associated with skilled workers thus the current findings are likely to be transferrable to other populations of skilled workers. Previously unreported LED events appear to show successful human performance (Safety II) to be effective upon the deliberate review of past tasks within a golden window of two hours of the erroneous act occurring; notably during periods of unfocused attention and whilst working alone in the same or similar sociotechnical environment to that which the error occurred. Several sociotechnical factors associated with LED were studied so that ergonomic interventions could be identified, which are anticipated to enhance LED. Application of these ergonomic LED interventions using a systems approach is considered especially important for simple everyday habitual tasks carried out alone where perhaps individual performance variability is most likely to pass unchecked if there are deficiencies in an organisations system defences. It has been argued that LED interventions are likely to offer further resilience against human performance variability by helping to re-gain situation awareness within the perceptual cycle through deliberate engagement with system cues; particularly physical objects such as equipment or written words. However, it is recognised that the interventions identified in the current study need to be deployed within real-world contexts to operationalize and test their true benefit in terms of risk mitigation, frequency and effectiveness over time. By definition, LED compliments Safety II scenarios yet any LED intervention is also likely to support Safety I safety strategies. Thus it is believed safety critical organisations should look for further resilience using LED intervention techniques that deliberately engage with system cues across the entire sociotechnical environment and full range of normal workplace behaviours; thereby providing opportunities for enhanced organisational mitigation for human performance variability.

References

Aini, M.S., Fakhru'l-Razi, A., 2013. Latent errors of socio-technical disasters: a Malaysian case study. Saf. Sci. 51, 284–292.

Bargh, J.A., Chartrand, T.L., 1999. The unbearable automaticity of being. Am. Psychol. 54, 462–479.

Blavier, A., Rouy, E., Nyssen, A.S., De Keyser, V., 2005. Prospective issues for error detection. Ergonomics 48 (7), 758–781.

Broadbent, D.E., Cooper, P.F., FitzGerald, P., Parkes, K.R., 1982. The cognitive failures questionnaire (CFQ) and its correlates. Br. J. Clin. Psychol. 21 (1), 1–16.

Bridger, R.S., Brasher, K., Dew, A., Sparshott, K., Kilminster, S., 2010. Job strain related to cognitive failure in naval personnel. Ergonomics 53, 739–747.

Cassell, C., Symon, G., 2004. Essential Guide to Qualitative Methods in Organizational Research. Sage, London.

Cheng, C.-N., Hwang, S.-L., 2015. Applications of integrated human error identification techniques on the chemical cylinder change task. Appl. Ergon. 47, 274–284.

Chiu, M.-C., Hsieh, M.-C., 2016. Latent human error analysis and efficient improvement strategies by fuzzy TOPSIS in aviation maintenance tasks. Appl. Ergon. 54, 136–147.

Cohen, G., Eysenck, M.W., LeVoi, M.E., 1986. Memory. Open University Press, Milton Keynes.

Cornelissen, M., Salmon, P.M., Jenkins, D.P., Lenné, M.G., 2013. A structured approach to the strategies analysis phase of cognitive work analysis. Theor. Issues Ergon. Sci. 14 (6), 546–564.

Day, A.J., Brasher, A., Bridger, R.S., 2012. Accident proneness revisited: the role of psychological stress and cognitive failure. Accid. Anal. Prev. 49, 532–535.

Dekker, S., 2014. The Field Guide to Understanding Human Error, third ed. Ashgate, Aldershot.

Einstein, G.O., McDaniel, M.A., 2005. Prospective memory: multiple retrieval processes. Curr. Dir. Psychol. Sci. 14 (6), 286–290.

Fitts, P.M., Posner, M.I., 1967. Human Performance. Brooks/Cole, Belmont, CA.

Finomore, V., Matthews, G., Shaw, T., Warm, J., 2009. Predicting vigilance: a fresh look at an old problem. Ergonomics 52 (7), 791–808.

Flanagan, J., 1954. The critical incident technique. Psychol. Bull. 51, 327–358.

Flin, R., O'Connor, P., Crichton, M., 2008. Safety at the Sharp End: a Guide to Non-technical Skills. Ashgate, Aldershot.

Grundgeiger, T., Sanderson, P.M., Dismukes, R.K., 2014. Prospective memory in complex sociotechnical systems. Z. Psychol. 222 (2), 100–109.

Guynn, M.J., McDaniel, M.A., Einstein, G.O., 2001. Remembering to perform actions: a different type of memory? In: Zimmer, H.D., Cohen, R.L., Guynn, M.J., Engelkamp, J., Kormi-Nouri, R., Foley, M.A. (Eds.), Memory for Action: a Distinct Form of Episodic Memory. Oxford University Press, New York, pp. 25–48.

Hutchins, E., 1995. How a cockpit remembers its speeds. Cogn. Sci. 19, 265–288.

Hutchins, E., 2001. Distributed cognition. In: Neil, J.S., Paul, B.B. (Eds.), The International Encyclopaedia of the Social and Behavioral Sciences. Pergamon, Oxford, UK, pp. 2068–2072.

Hollnagel, E., 2014. Safety-I and Safety-II: The Past and Future of Safety Management. Ashgate, Farnham, UK.

Hollnagel, E., Woods, D.D., Leveson, N., 2006. Resilience Engineering: Concepts and Precepts. Ashgate Publishing, Aldershot.
Johannessen, K.B., Berntsen, D., 2010. Current concerns in involuntary and voluntary autobiographical memories. Conscious. Cogn. 19, 847–860.
Kontogiannis, T., Malakis, S., 2009. A proactive approach to human error detection and identification in aviation and air traffic control. Saf. Sci. 47 (5), 693–706.
Kvavilashvili, L., Mandler, G., 2004. Out of one's mind: a study of involuntary semantic memories. Cogn. Psychol. 48, 47–94.
Leveson, N., 2004. A new accident model for engineering safer systems. Saf. Sci. 42, 237–270.
Mace, J.H., Atkinson, E., Moeckel, C.H., Torres, V., 2011. Accuracy and perspective in involuntary autobiographical memory. Appl. Cogn. Psychol. 25 (1), 20–28.
Mandler, G., 1985. Cognitive Psychology: an Essay in Cognitive Science. Erlbaum, Hillsdale, NJ.
Marsh, R.L., Hicks, J.L., Bink, M.L., 1998. Activation of completed, uncompleted and partially completed intentions. J. Exp. Psychol. Learn. Mem. Cogn. 24, 350–361.
Mazzoni, G., Vannucci, M., Batool, I., 2014. Manipulating cues in involuntary autobiographical memory: verbal cues are more effective than pictorial cues. Mem. Cogn. 42, 1076–1085.
Mecacci, L., Righi, S., 2006. Cognitive failures, metacognitive beliefs and aging. Personal. Individ. Differ. 40, 1453–1459.
Niesser, U., 1976. Cognition and Reality: Principles and Implications of Cognitive Psychology. Freeman, San Francisco.
Norman, D., 1981. Categorization of action slips. Psychol. Rev. 88, 1–15.
Norman, D., Shallice, T., 1986. Attention to action: willed and automatic control of behavior. In: Davidson, R., Schwartz, G., Shapiro, D. Plenum (Eds.), Consciousness and Self Regulation: Advances in Research. New York Press.
Oppenheim, A.N., 1992. Questionnaire Design, Interviewing and Attitude Measurement. Pinter, London.
Plant, K.L., Stanton, N.A., 2013a. What is on your mind? Using the perceptual cycle model and critical decision method to understand the decision-making process in the cockpit. Ergonomics 56 (8), 1232–1250.
Plant, K.L., Stanton, N.A., 2013b. The explanatory power of schema theory: theoretical foundations and future applications in ergonomics. Ergonomics 56 (1), 1–15.
Rasmussen, A.S., Berntsen, D., 2011. The unpredictable past: spontaneous autobiographical memories outnumber autobiographical memories retrieved strategically. Conscious. Cogn. 20, 1842–1846.
Rasmussen, J., Pedersen, O.M., 1984. Human factors in probabilistic risk analysis and risk management, 1990. In: Reason, J. (Ed.), Human Error. Cambridge University Press, Cambridge, pp. 173–174.
Reason, J., 1990. Human Error. Cambridge University Press, Cambridge.
Reason, J., 1997. Managing the Risks of Organizational Accidents. Ashgate, Aldershot.
Reason, J., 2008. The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries. Ashgate, Aldershot.
Reason, J., Hobbs, A., 2003. Managing Maintenance Error: a Practical Guide. Ashgate, Aldershot.
Reiman, R., 2011. Understanding maintenance work in safety-critical organisations – managing the performance variability. Theor. Issues Ergon. Sci. 12 (4), 339–366.
Robson, C., 2011. Real World Research, third ed. Wiley, Chichester.
Saward, J.R.E., Stanton, N.A., 2015a. Individual latent error detection and recovery in naval aircraft maintenance: introducing a proposal linking schema theory with a multi-process approach to human error research. Theor. Issues Ergon. Sci. 16 (3), 255–272.
Saward, J.R.E., Stanton, N.A., 2015b. Individual latent error detection: is there a time and a place for the recall of past errors? Theor. Issues Ergon. Sci. 16 (5), 533–552.
Schluter, J., Seaton, P., Chaboyer, W., 2008. Critical incident technique: a user's guide for nurse researchers. J. Adv. Nurs. 61 (1), 107–114.
Sellen, A.J., 1994. Detection of everyday errors. Appl. Psychol. Int. Rev. 43, 475–498.
Sellen, A.J., Louie, G., Harris, J.E., Wilkins, A.J., 1997. What brings intentions to mind? An in situ study of prospective memory. Memory 5, 483–507.
Shorrock, S.T., Kirwan, B., 2002. Development and application of a human error identification tool for air traffic control. Appl. Ergon. 33, 319–336.
Smallwood, J., Schooler, J.W., 2006. The restless mind. Psychol. Bull. 132, 946–958.
Smith, R.E., 2003. The cost of remembering to remember in event based prospective memory: investigating the capacity demands of delayed intention performance. J. Exp. Psychol. Learn. Mem. Cogn. 29, 347–361.
Stanton, N., Baber, C., 1996. A systems approach to human error identification. Saf. Sci. 22 (1–5), 215–228.
Stanton, N.A., Salmon, P., Harris, D., Marshall, A., Demagalski, J., Young, M.S., Waldmann, T., Dekker, S., 2009a. Predicting pilot error: testing a new methodology and a multi-methods and analysts approach. Appl. Ergon. 40 (3), 464–471.
Stanton, N.A., Salmon, P.M., Walker, G.H., Jenkins, D., 2009b. Genotype and phenotype schemata and their role in distributed awareness in collaborative systems. Theor. Issues Ergon. Sci. 10 (1), 43–68.
Stanton, N.A., Salmon, P.M., Walker, G.H., 2014. Let the reader decide: a paradigm shift for situation awareness in sociotechnical systems. J. Cogn. Eng. Decis. Mak. 9 (1), 44–50.
Tulving, E., 1983. Elements of Episodic Memory. Oxford University Press, New York.
Twelker, P.A., 2003. The critical incident technique: a manual for its planning and implementation, 2008. In: Schluter, J., Seaton, P., Chaboyer, W. (Eds.), Critical Incident Technique: A User's Guide for Nurse Researchers, Journal of Advanced Nursing, vol. 61(1), pp. pp. 107–114.
Van den Berg, S.M., Aarts, H., Midden, C., Verplanken, B., 2004. The role of executive processes in prospective memory tasks. Eur. J. Psychol. 16 (4), 511–533.
Wallace, J.C., Kass, S.J., Stanny, C.J., 2002. The cognitive failures questionnaire revisited: dimensions and correlates. J. Gen. Psychol. 129 (3), 238–256.
Wiegmann, D.A., Shappell, S.A., 2001. Human error perspectives in aviation. Int. J. Aviat. Psychol. 11 (4), 341–357.
Wiegmann, D.A., Shappell, S.A., 2003. A Human Error Approach to Aviation Accident Analysis: the Human Factors Analysis and Classification System. Ashgate, Aldershot.
Wilkinson, W.E., Cauble, L.A., Patel, V.L., 2011. Error detection and recovery in dialysis nursing. J. Patient Safe. 7 (4), 213–223.
Woods, D.D., 1984. Some results on operator performance in emergency events. In: Whitfield, D. (Ed.), Ergonomic Problems in Process Operations, vol. 90. Institute of Chemical Engineering Symposium, Ser., pp. pp. 21–31
Woods, D.D., Dekker, S., Cook, R., Johannesen, L., Sarter, N., 2010. Behind Human Error, second ed. Ashgate, Farnham; Burlington, VT.