

Workflows and individual differences during visually guided routine tasks in a road traffic management control room



Sandra D. Starke^{a,*}, Chris Baber^a, Neil J. Cooke^a, Andrew Howes^b

^a School of Engineering, University of Birmingham, B15 2TT, UK

^b School of Computer Science, University of Birmingham, B15 2TT, UK

ARTICLE INFO

Article history:

Received 13 July 2016

Received in revised form

22 December 2016

Accepted 10 January 2017

Available online 24 January 2017

Keywords:

Control room operations

Eye tracking

Information search

ABSTRACT

Road traffic control rooms rely on human operators to monitor and interact with information presented on multiple displays. Past studies have found inconsistent use of available visual information sources in such settings across different domains. In this study, we aimed to broaden the understanding of observer behaviour in control rooms by analysing a case study in road traffic control. We conducted a field study in a live road traffic control room where five operators responded to incidents while wearing a mobile eye tracker. Using qualitative and quantitative approaches, we investigated the operators' workflow using ergonomics methods and quantified visual information sampling. We found that individuals showed differing preferences for viewing modalities and weighting of task components, with a strong coupling between eye and head movement. For the quantitative analysis of the eye tracking data, we propose a number of metrics which may prove useful to compare visual sampling behaviour across domains in future.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Control room design: fundamentals and challenges

Ergonomic control room design requires consideration of many factors such as personnel, systems design or equipment layout (Wood, 2004), is subject to standards such as ISO 11064 (part 1–7) or BS EN ISO 6385 and is discussed by substantial literature (Noyes and Bransby, 2001; Ivergard and Hunt, 2008). An important ergonomic design consideration for control rooms is that the available technology has to be aligned with human behaviour, requirements and limitations (Hughes and Kornowa-Weichel, 2004); this helps to maximize the efficiency and effectiveness of operators. Human information processing in control rooms has received scientific interest in domains as varied as air traffic control (Stein, 1992; Endsley and Rodgers, 1996), airplane cockpit design (Steelman et al., 2011), nuclear power plant control (Chang Hoon et al., 2006; Kim et al., 2013) or monitoring of CCTV (Howard et al., 2011; Stainer et al., 2013). Human error as a consequence of insufficient ergonomic workspace design has arguably led to

catastrophic accidents such as Three Mile Island, Chernobyl and Bhopal (Meshkati, 1991). The causes for human error are manifold (Kirwan, 1992; Reason, 2000; Dekker, 2014). A possible contributing factor to human error is the incomplete use or false interpretation of visual information; this has for example received substantial interest in medical image analysis (Krupinski, 2010), but also in air traffic control (Stein, 1992) or cockpit design (Hanson, 2004). Furthermore, people might not use available visual resources as expected by designers; an example is the discrepancy between expected and observed behaviour in CCTV control rooms (Smith, 2004) or the selective use of available information sources by air traffic controllers (Stein, 1992; Seok et al., 2006). Analysis of operators' visual workflows and preferences allows understanding of such discrepancies and helps designing towards reliable resource usage.

Most control rooms are built around the presentation of visual information through a multitude of display screens (Hénique et al., 2008; Ivergard and Hunt, 2008; Stanton et al., 2009). In air traffic control, visual information perception has been described as most crucial next to voice communication (Meyer et al., 2013), and this extrapolates to many other control room domains. Activities of operators in control rooms range from making predictions about criminal activity in CCTV footage (Troschianko et al., 2004), to monitoring and control of process control plants (Kim et al., 2013)

* Corresponding author.

E-mail address: s.d.starke@bham.ac.uk (S.D. Starke).

or evaluating risks for collision in air traffic control (Landry, 2011). In each domain, a core activity of the operator involves interpreting the visual information that is displayed to them in order to infer the state of the system. In order to understand how this information is used, studies from the domain of psychology and ergonomics can apply a methodology called 'eye tracking', which allows to measure where operators are looking, what information sources they use and how they combine information.

1.2. Understanding user behaviour and cognitive processes through eye tracking

Humans have to move their eyes to attend to relevant information sources because only central vision provides a high resolution and sharp rendering of a scene (Henderson, 2003; Land, 2006; Borji and Itti, 2014). This is due to the distribution of photoreceptors within the retina and the non-linear representation of this information in the visual cortex (Snowden et al., 2012). Eye tracking allows recording where someone is looking with central vision, which is called the 'point of gaze'; it has been used as part of study design in thousands, if not tens of thousands, of research papers (Tien et al., 2014) and has long been used to inform user interface design (Jacob and Karn, 2003; Poole and Ball, 2006). For example, scan patterns (the sequence of attended regions of interest) and fixation duration can be used to identify sub-optimal layout of interfaces or the perceived importance of individual user interface elements (Pohl et al., 2009; Burch et al., 2011).

The use of eye-tracking to study operator activity and decision making has been employed in a range of control room domains (Moray and Rotenberg, 1989; Lin et al., 2003; Shepley et al., 2009; Moore and Gugerty, 2010). Eye tracking has also been used with a view to detecting operator fatigue and impairment of visual vigilance (McIntire et al., 2014a,b). Moray and Rotenberg (1989) for example demonstrated that, when dealing with incidents, operators tend to increase the frequency of fixations on the failed system component, rather than increasing the duration of fixations, and that information processing becomes restricted to one information source at the expense of attending to subsequent or parallel incidents. This suggests that operators might not optimally use available resources (Smith, 2004). Analysis of operators' workflows and preferences could aid the understanding of such discrepancies. To date, reports of visual scanning behaviour (the act of moving the point of gaze across a scene) across a wide range of domains is lacking and hence conclusions are often drawn on a case-by-case basis.

Gaze shifts can be executed by eye movement alone or accompanied by head movement (Wollaston, 1824; Bizzi et al., 1972; Morasso et al., 1977; Zangemeister and Stark, 1982; Guitton and Volle, 1987; Goossens and Opstal, 1997; Oommen and Stahl, 2005). Gaze shifts larger than 45°–50° visual angle have to be executed by head movement simply because the eyes do not rotate further within the head (Proudlock and Gottlob, 2007; Freedman, 2008). Gaze shifts larger than 75°–90° additionally need rotation of the upper body due to the functional limits of head/neck rotation (Proudlock and Gottlob, 2007). For example, in a study of visual search in a mail room, 80% of search time was spent moving head and body rather than only the eyes (Foulsham et al., 2014). Recent work has highlighted that aligning eyes and head, rather than diverting gaze laterally, results in better performance during visual search tasks (Nakashima and Shioiri, 2014). Hence, aligning eye and head orientation is likely beneficial for cognitive information processing. The interaction between eye-, head and arm movements has previously been investigated in air traffic control (Boyer, 1995) and so it would be interesting to determine the relationship between eye- and head-movement in other control room

environments. For this paper, the control room environment under consideration concerns road traffic management.

1.3. Road traffic control rooms: purpose and goals

Road traffic management involves the monitoring of traffic, responding to incidents and influencing road user behaviour. Given that incidents can contribute to some 25% of the overall congestion levels on major roads (UK Highways Agency, 2009), it is important that any incident is resolved as quickly as possible. Regional Control Centres, such as the 'Direction Interdépartementale des Routes Centre-Est' (DIR-CE) in Grenoble, France, are the central focus of communications regarding major roads. They monitor traffic flow (through CCTV, through verbal reports or through sensor data from the roads or vehicles) and control the Variable Message Signs on these roads. In broad terms, the goals of such a centre can be summarized as follows (Folds et al., 1993): i.) maximise the available capacity of the roadway system; ii.) minimise the impact of incidents (accidents, debris, etc.); iii.) contribute to demand regulation; iv.) assist in the provision of emergency services; and v.) maintain public confidence in the control centre operations and information provision.

1.4. Aims and scope of this study

The present study is a case study, constrained by the availability of staff and environmental factors. Aim of the study was to i) present insight into operator behaviour in a road traffic control room, both from the perspective of qualitative work analysis and quantitative visual sampling analysis, and ii) present a number of eye tracking metrics which we deem useful to compare visual sampling behaviour across domains in future. We use a Hierarchical Task Analysis (HTA) and eye tracking to study information sampling behaviour and workflow in a road traffic management control room. Visual scanning via eye- and head movement forms the sensory foundation for decision making and actions in control rooms and would benefit from further exploration, especially in context of user interface- and control room design. The present study provides a rare reference dataset on visual scanning behaviour in a fully operational road traffic management facility.

2. Study setting

2.1. Control room layout

Data for this study were collected at the road traffic management facility at DIR Centre Est, Grenoble, France. The control room under investigation for this study consisted of multiple displays (Fig. 1). In front of the operator, at arm's length, is an arced arrangement of five monitors, containing the following information sources and components:

- D1 ("Display 1") - generic display with access to internet and software packages.
- D2 ("Display 2") - user interface (UI) for incident logs. The workflow of operators is systematically guided by this UI, which contains for example dropdown menus and text entry fields to document incident details.
- D3 ("Display 3") - interactive schematic map of the road traffic network.
- D4 ("Display 4") - live CCTV feed, which the operator can select from a number of available feeds. The selected camera can be controlled by the operator through zooming, panning and rotating.
- D5 ("Display 5") contains an auxiliary schematic interface

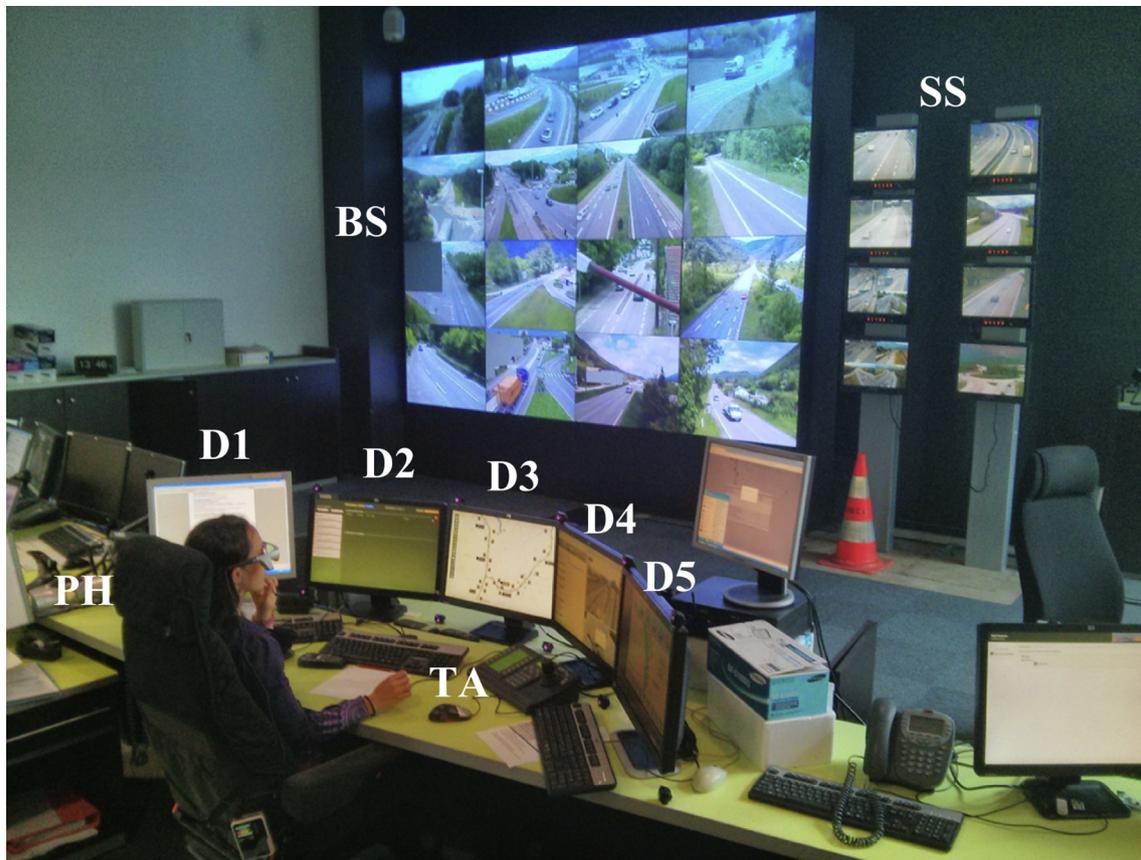


Fig. 1. Road traffic control room at DIR Centre Est, Grenoble, France. A control operator who participated in the study is wearing a mobile eye tracker (Tobii glasses) while performing a standard road traffic control task.

In the far field are two panels displaying CCTV information:

- SS (“Small Screens”) - bank of eight CCTV feeds arranged in two columns, presenting colour CCTV footage which can be interacted with from the operator's desk (D4). This display is located approximately 5 m away from the operator.
- BS (“Big Screens”) is a large projection of 16 CCTV feeds, arranged as a 4×4 array. This display is not interactive and camera views are fixed. This display is located at the same distance as SS.

On the table (TA) in front of the operator are standard PC peripherals and phones/radios (PH) for communication with stakeholders from outside the facility such as traffic patrol staff and emergency services personnel.

2.2. Scenario

DIR Centre Est is a fully operational, live traffic control centre, hence forbidding the interruption of incoming data with a pre-recorded scenario. Instead, we captured the resolution of two live incidents (one broken down lorry and one broken down motorbike) by two different observers and asked another three operators to engage in a simulated task (‘object in the road’) during a period of less intense activity. The live incident concerning the lorry presented a situation where a lorry had broken down near an exit on a main road, causing congestion (Fig. 2 a). This incident took 15 min to resolve and included liaising with authorities, managing traffic and status updates. The live incident concerning the broken down motorbike presented a situation where a motorbike presumably

had an accident and was in a road close to a slip road on a quiet road (Fig. 2 b). This incident took 6 min to resolve and included communication with the police, situation monitoring and status updates. In the simulated task of ‘object in road’, operators were asked - after a pretend call - to resolve a situation where an object was found in a road. The rationale behind this simulated approach was that operators were assumed to navigate this task in a manner representative of their ‘average’ behaviour, while visual search would be required independent of the presence or absence of an object in the road.

3. Qualitative analysis of operator activity: hierarchical task analysis

3.1. Approach

To understand visual sampling behaviour in context of a person's goal structure, a Hierarchical Task Analysis (HTA) was performed, which is a common technique for goal decomposition in the UK (Annett et al., 1971; Shepherd, 2001). Through a combination of observing operators and discussion with subject matter experts (i.e., the operators, their managers and staff in other control centres), we identified those subgoals that are important in the activity of Road Traffic Management.

3.2. Identified subgoals and plans

For the operators' main goal of ‘respond to incident’ (goal 0), the following seven subgoals were identified (the full HTA is shown in Fig. 3):

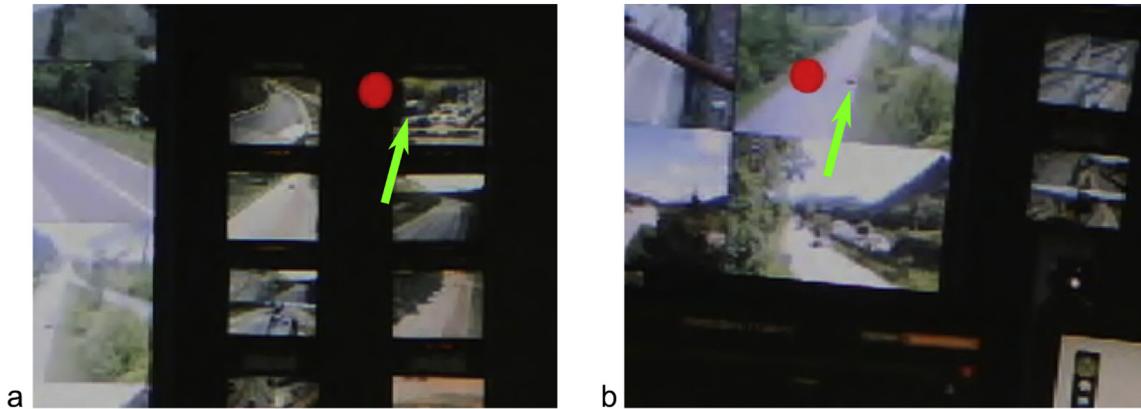


Fig. 2. Screenshots from the Tobii glasses video recording of the two live scenarios, a - broken down lorry and congestion, b - broken down motorbike. Red dot – instantaneous raw gaze point of the respective participant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

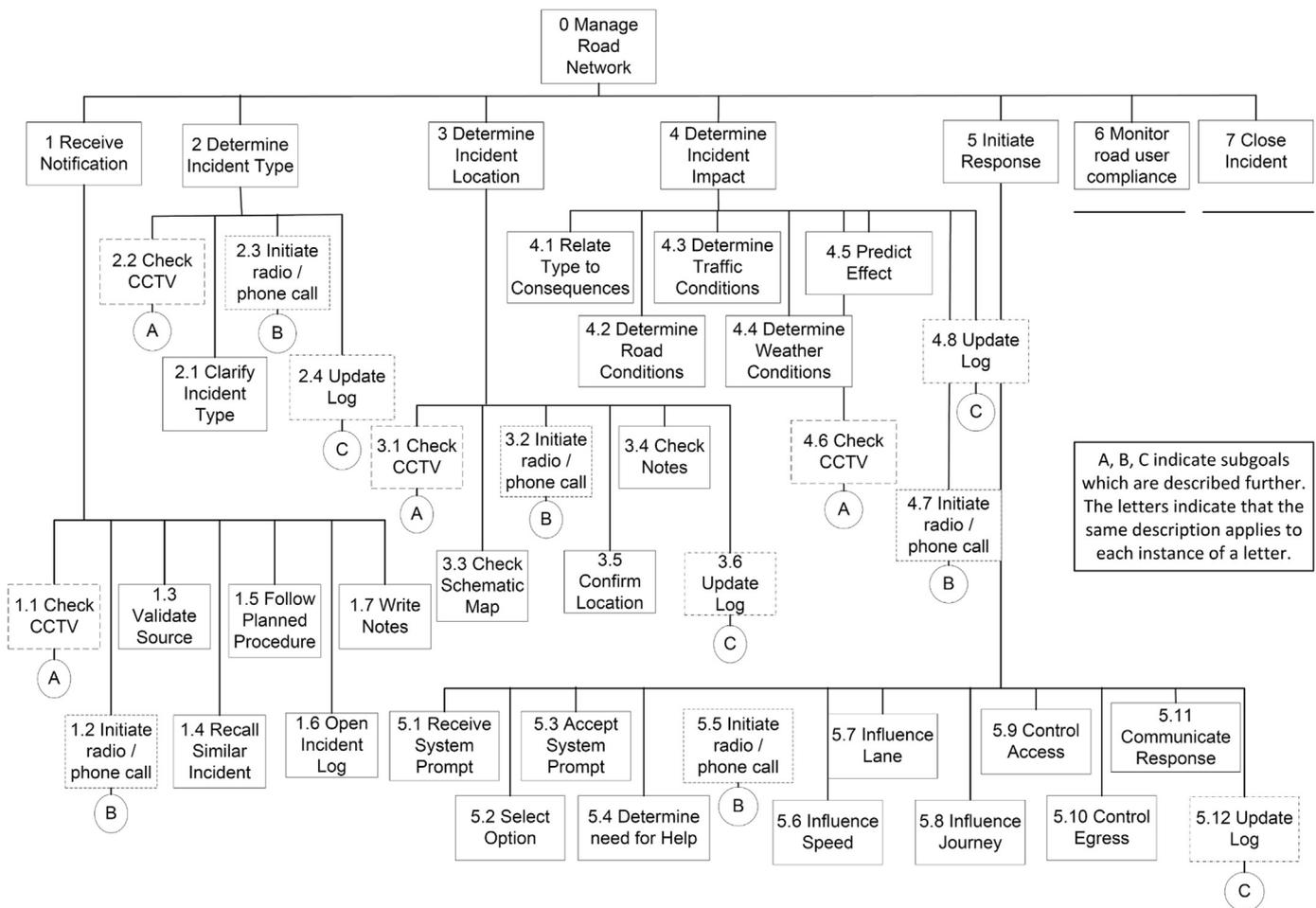


Fig. 3. Hierarchical task analysis of road traffic management based on eye tracking data and discussion with subject matter experts.

Subgoal 1: Receive Notification. The operator typically responds to an incident notification at the start of an engagement. This could arrive through different media (phone, radio, auto-detect from CCTV etc.) and the operator would make sure that the notification was from a credible source. If the incident is felt to be sufficient to require a response, an incident log is created.

Subgoal 2: Determine incident type. The operator needs to classify the incident in terms of its type. The operators spoke of

{Accidents, congestion, obstacles and incidents, road works} as examples of incident type. The initial notification would have provided some information about the type of incident but the operator might seek further clarification from CCTV or from colleagues at the scene over the phone. Having classified the incident, the operator updates the incident log.

Subgoal 3: Determine incident location. In this control room, there are two distinct strategies for determining incident location.

The operator could, on the basis of information in the initial incident report, select a CCTV camera close to the incident and then use this to look for the incident. If the CCTV shows the incident then the location of the CCTV and the nearest exit are recorded. Alternatively, the operator might refer to the schematic map screen to define the most likely location and then use this information to select a CCTV to confirm this location. While the tasks are not complicated, this shows how operators could develop different strategies for achieving the same subgoal.

Subgoal 4: Determine incident impact. Having defined a specific incident at a specific location, the next subgoal is to decide what impact this incident will have on the performance of the road network. This can involve the operator relating the type of incident to particular consequences. The operators spoke of {risk, safety, journey time, average speed, traffic density and congestion, changes in demands on the road}. An understanding of the factors which contribute to the current situation, such as weather, road conditions, traffic conditions, and how these factors are likely to change during the course of incident, influences the operators' ability to develop a coherent mental representation of the situation, respectively situation awareness (Endsley and Rodgers, 1996; Moore and Gugerty, 2010).

Subgoal 5: Initiate response. As the operator completes the incident log, the software system used by the centre shows the options available regarding the type of incident. The operator could accept the system prompt or could select another option. The response could involve changing the content of overhead signs on the road (either in terms of changing speed limits, indicating lane closures or providing advisory signs). The content of the signs is pre-defined and the operator selects from this set. If the incident cannot be dealt with by a sign from the set, the operator is not empowered to improvise or create new content; rather, new content needs to be approved by the relevant agencies. As well as modifying the signs, the operator can limit access to the road network (through control of junctions). Finally, the operator could call for road-side assistance to attend the incident. This assistance could help the vehicle or could be specialised support (paramedic or technical).

Subgoal 6: Monitor road user compliance. When a response has been initiated, the operator will check that the incident has been resolved and also that the road users affected by the response are complying with it.

Subgoal 7: Close incident log. Once the incident has been resolved, the operator closes the incident log. We noted that operators (and the control room) would have some incidents that remained open. This might be because it was taking longer to resolve than expected due to unforeseen circumstances or because the incident was open for a scheduled reason, such as road works.

Fig. 4 shows the correlation between smoothed gaze data and the different HTA stages and actions for three example participants. Gaze data were interpolated to 0–100% task time. For bins of 1% increments, the ROI is shown which was attended to most within the bin. Example plans for the identified subgoals are shown in Table 1.

4. Quantitative analysis of visual information sampling

4.1. Participants

The DIRCE control room is staffed by six trained personnel (plus two managers), working two shifts (7am to 3pm and 2.30pm–9.30pm). Over the course of a single day, the study recruited five out of six staff members. Participants ranged in age from 35 to 40 years; four participants were male and one was female. Personnel had been working for DIR-CE for at least 5 years.

Personnel had been recruited for the role by competition: the criteria for selection were rigorous people who have good analytical skills and excellent responsiveness.

4.2. Collection of eye tracking data

This study was approved by the University of Birmingham Ethics Panel (Reference Number ERN_13–0997). All participants received an explanation of the task. Each operator was equipped with a monocular mobile eye tracker (Tobii Glasses v.1, Tobii, Sweden). The sampling frequency of the system was 30 Hz. Data were collected according to the system's standard operating procedures, including a 9-point calibration and placement of infrared reference markers around the regions of interest (ROIs) within the control room. Data collection time was around 6 min with the exception of one real incident (participant 1) that lasted almost three times as long.

4.3. Data mapping

The recorded raw gaze data were automatically mapped to global space using the subset of infrared (IR) reference markers visible in each frame through custom-written code and a set of conditional pre-processing steps (please refer to [Supplementary Material 1](#) for technical details). Mapping had to be executed outside the proprietary software, since due to the room layout only 2 IR markers were captured in a substantial number of frames, while the proprietary software requires 3 visible IR markers. Head orientation was approximated by mapping the ROI in which the centre of the CMOS was located to each frame. The analysed data range started with an incoming call reporting an incident and finished with the operator indicating task completion and finishing entries on the incident log (on screen D2). Times during which operators engaged in unrelated activity, such as talking to a colleague about unrelated matters, were cropped out.

4.4. Data analysis

The scan pattern of each participant, respectively the sequence of attended regions of interest (ROIs), was analysed to quantify how long individual operators attend to specific information sources, how often they switch between sources and how information sources were combined.

4.4.1. Viewing behaviour

We calculated several metrics for analysis of the gaze data, some of which are conventional eye tracking metrics (Holmqvist et al., 2011) and some of which were borrowed from network analysis (Scott, 2013). Network analysis metrics were chosen to quantify how information sources were combined: network analysis facilitates the quantification of connections, rather than sequential or cumulative properties of a scan pattern.

Conventional metrics:

1. *Cumulative percentage viewing time per ROI.* Sum of frames for which a participant attended to an individual ROI divided by the sum of all frames for which eye tracking data was available. This metric shows how – in total – participants split their time between the available information sources.
2. *Number of visits per ROI.* Count of each dwell (continuous attention) per ROI for ≥ 5 consecutive frames (167 ms; based on the minimum viewing time of 150 ms required to process and understand a complex pattern (Thorpe et al., 1996)). This metric shows how often a participant returned to an information source.
3. *Dwell time.* Time rested per visit, here calculated across all ROIs.

Table 1
Plans associated with HTA.

Subgoal	Plan
1. Receive notification	p1a: 1.1 > 1.6 > exit p1b: 1.2 > as required and for specific information 1.7 > if appropriate 1.3 > if possible 1.4 > 1.6 > exit p1c: 1.2 > 1.1 > 1.6 > exit
2. Determine incident type	p2a: 2.4 > if unable to determine type > 2.2 OR 2.3 OR 2.1 > 2.4 > p2b: 2.1 > 2.4 > exit p2c: 2.3 > 2.1 > 2.4 > exit
3. Determine incident location	p3a: 3.3 > 3.6 > exit p3b: 3.3 > 3.2 OR 3.4 > 3.6 > p3c: 3.1 > 3.5 > 3.6 > exit
4. Determine incident impact	p4a: 4.7 > if system has option for type 4.1 > 4.8 > exit p4b: 4.1 > if familiar incident 4.4 > 4.8 > exit p4c: 4.1 > if unfamiliar incident 4.2 > 4.3 > 4.4 > 4.5 > if clarification required 4.7 > 4.8 > exit
5. Initiate response	p5a: 5.1 > 5.3 > .12 > exit p5b: 5.2 > depending on decision 5.7, 5.8, 5.10, 5.11 > 5.12 > exit p5c: 5.1 > 5.3 > long-term impact 5.9 > 5.12 > exit
6. Monitor road user compliance	No further description: this task will involve checking that road users are obeying the signs that have been set-up
7. Close incident log	No further description: once an incident has been declared terminated, the incident log will be closed.

Example plans for the seven identified subgoals. In the notation for plans, '>' signifies "followed by" to indicate sequence, numbers indicate subgoals (i.e., nodes in the HTA), text indicates 'conditions' and each plan terminates with 'exit' (which indicates that the process moves to the next top-level subgoal). Provided are three examples for each subgoal for illustrative purposes; the set of plans is not exhaustive but intended to illustrate that operators can achieve these subgoals in a variety of ways, depending on operating conditions and strategy employed by the operator.

5. (*Average*) *nodal degree*. The average number of edges connected to each node of the network and number of edges for each individual node.

Nodal degree allows exploring the function of different ROIs within the observer's work pattern: a ROI with a high nodal degree could for example be central to the work process as it may require the assimilation of information from several different ROIs, or it could indicate that the ROI holds information relevant to several other ROIs. A ROI with a low nodal degree would in contrast indicate that the information source is only attended to for very specific purposes.

6. *Leaf nodes*. Count of nodes that are only connected to one other node.

This metric identifies ROIs that serve a very specific function, such as a CCTV monitor attended to only to check entries on a schematic diagram.

Further to the quantitative metrics, we constructed what we term 'viewing networks', a schematic visualization of the frequency of connections between any two ROIs; this type of visualization is often used in network analysis and gives an intuitive insight into which nodes are combined how frequently.

4.4.2. Eye- and head movement

In order to quantify the movement frequency for eyes and head when switching between different ROIs as well as the coupling between eye and head orientation, we calculated the following metrics:

1. *ROI switch frequency*. Frequency of gaze shifts between the available ROIs, calculated as average values across the whole trial.
2. *Agreement between ROI attended to by eyes and head*. Number of all frames in which eye and head orientation were assigned to the same ROI divided by the number of all tracked frames. Further, we quantified the proportion of switches between ROIs executed by both, eye and head movement: a) by approximating head movement velocity from IR marker displacement across frames and tagging those frames as 'head moving'. If a gaze shift was made while the head was tagged as moving, this was

classified as concurrent eye and head movement; b) by checking for the match in the new ROI attended to by gaze and head orientation.

3. *Amount of head movement*. From those frames that were tagged as 'head moving' in 2 a), the percentage time that operators spent moving their head above the set threshold was calculated as the fraction of task time.

4.5. Results

4.5.1. Viewing behaviour

The selection of ROIs and the cumulative percentage viewing time varied across participants (Figs. 5 and 6). For example, when examining the traffic situation, while participant 5 attended to D4 for the majority of time with hardly any time allocated to SS, the other participants attended mainly to SS and D3. The time spent viewing D3 was $\leq 7\%$ viewing time except for participant 4, who spent 24% of viewing time on this ROI. All participants attended to the incident log (D2). According to the attention allocation above, the visit count for each ROI varied as well across participants (Table 2); the number of visits to the different ROIs was substantially higher for participant 1 due to the almost 3-fold task time.

Dwell times showed a skewed distribution as expected, median values ranging from 1.2 to 4.2 s across participants with individual dwells often lasting up to 10 s and in rare cases even 20–60 s.

A summary of the outcomes for the calculated network metrics is presented in Table 3. While values for participant 1 stood out due to the longer and more involved task, they spanned similar ranges for the remaining four participants. Number of edges and link density showed that around 20–30% of possible paths were taken while operators attended to 60–100% of ROIs. The average nodal degree was around 2–3 for undirected and 4 for directed networks and the number of leaf nodes ranged from 0 to 2. Constructed viewing networks highlighted the heterogeneity between participants, who followed different search patterns when redirecting gaze between the different displays (Fig. 6).

The median nodal degree (Table 2) was highest for D2 (incident log) and TA (table in front of operator).

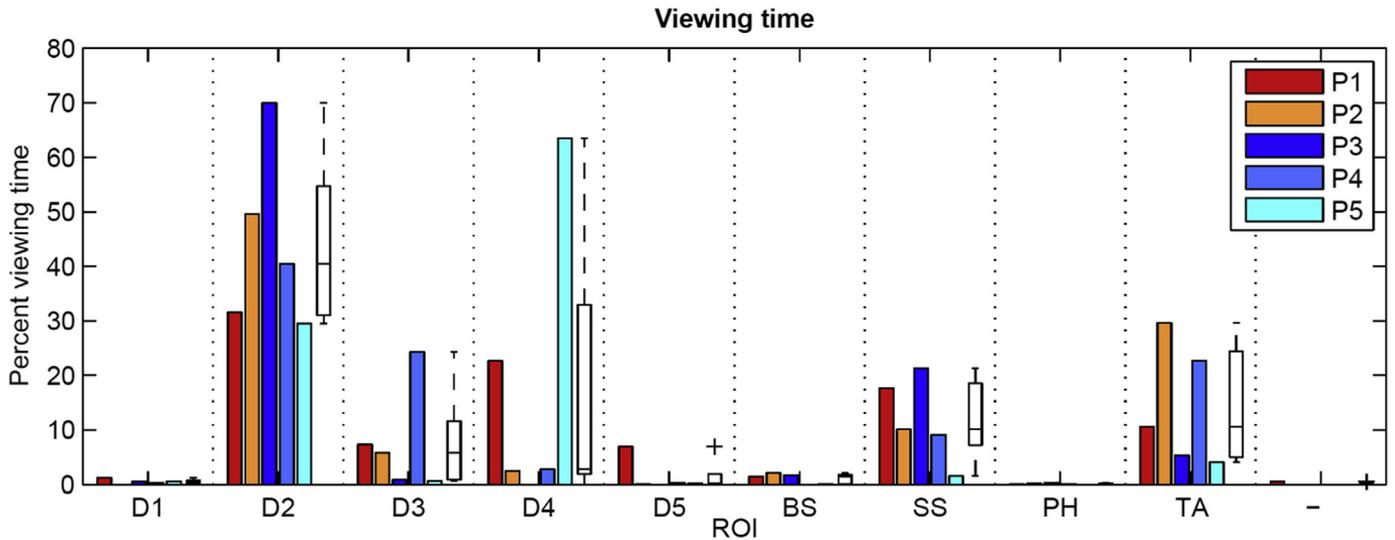


Fig. 5. Cumulative percentage viewing time for each region of interest (ROI). Data are presented for each participant (coloured) to highlight the pronounced individual differences and summarized in adjacent boxplots (showing median, interquartile range, maxima and outliers). ROIs: D1 - generic display with access to internet etc., D2 - incident log, D3 - interactive schematic map of traffic network, D4 – live, selectable and interactive CCTV feed, D5 - interactive user interface, BS – ‘big screen’: large projected 4 × 4 array of CCTV footage, SS – ‘small screens’: bank of eight CCTV feeds arranged in two columns which can be interacted with, PH: phone, TA: table/desk. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

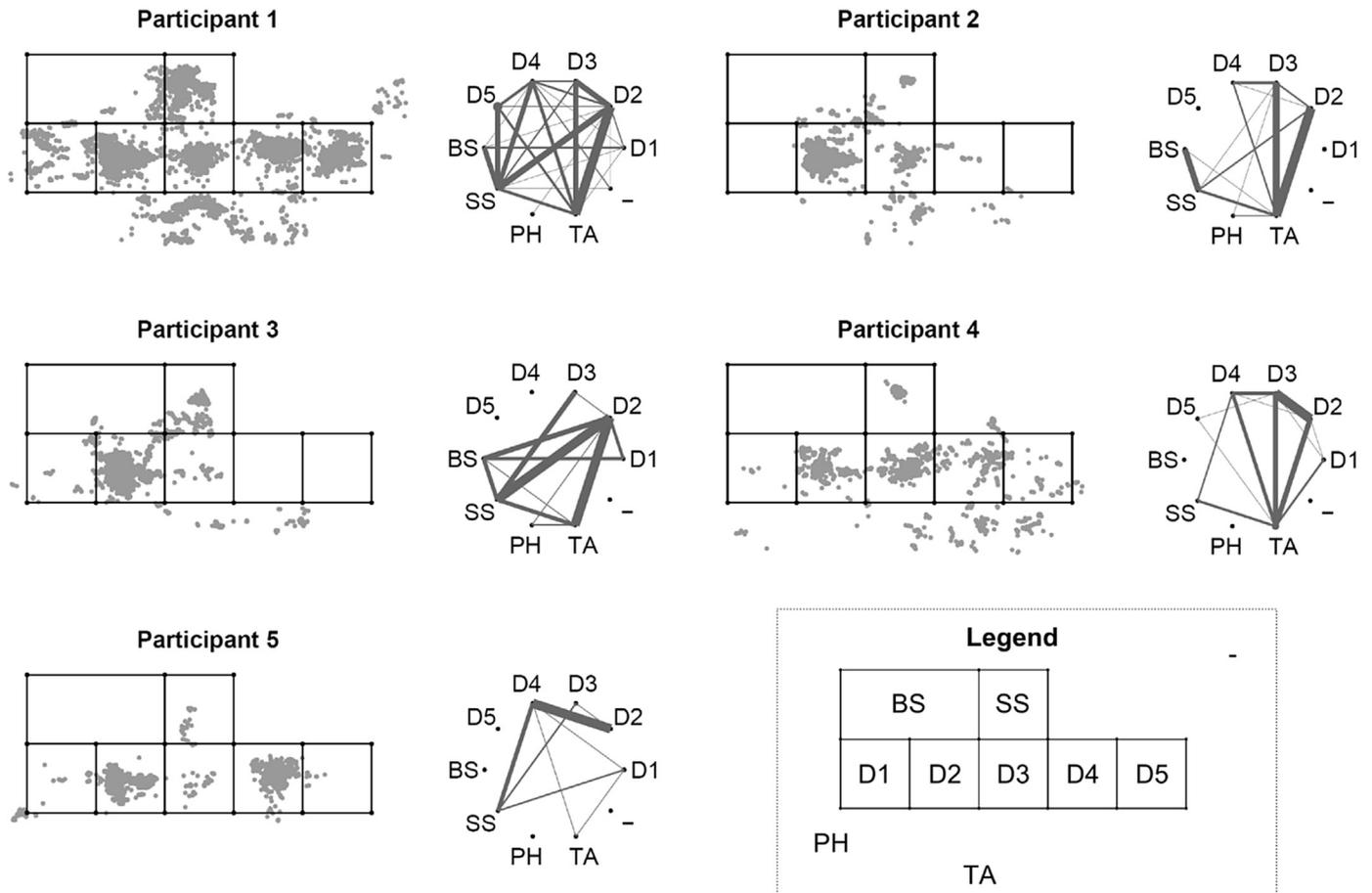


Fig. 6. Visual attention to the different regions of interest and viewing networks for each participant. On the viewing networks, the line width is representative of the number of switches between regions summed for both directions and scaled to the highest switch count.

Table 2
Region of Interest (ROI) specific metrics.

ROI	Visit count		Percentage of all visits		Nodal degree	
	Median (IQR)	Range	Median (IQR)	Range	Median (IQR)	Range
D1	3 (3)	0 to 9	4 (5)	0 to 8	3 (2)	0 to 5
D2	14 (19)	6 to 68	22 (7)	19 to 30	6 (3)	2 to 9
D3	11 (20)	2 to 35	10 (11)	6 to 23	4 (3)	2 to 5
D4	9 (15)	0 to 43	12 (13)	0 to 36	4 (3)	0 to 8
D5	0 (14)	0 to 51	0 (5)	0 to 14	0 (3)	0 to 6
BS	6 (10)	0 to 19	5 (11)	0 to 15	2 (4)	0 to 4
SS	10 (21)	4 to 69	19 (8)	5 to 23	4 (2)	2 to 8
PH	1 (1)	0 to 1	0 (2)	0 to 2	1 (2)	0 to 2
TA	20 (28)	1 to 60	17 (18)	4 to 32	6 (3)	2 to 8

This table shows median (IQR) and ranges for visit count to each ROI, the percentage of all visits allocated to each ROI and the nodal degree of each ROI based on network analysis.

4.5.2. Eye and head coupling

Switch frequency averaged across the whole task time ranged from 0.1 to 0.4 Hz across participants for both, eyes (median (IQR): 0.21 (0.13)) and head (median (IQR): 0.18 (0.16)). The median (IQR) difference between the eye- and head switch frequency was 0.02 (0.08) Hz.

The median (IQR) agreement between ROIs attended to by eyes and head on a per-frame was 91 (6) %. The number of concurrent switches between ROIs was 87 (9) % based on tagging the head as moving (method 2a) and 77 (31) % based on agreement of the new ROI attended to after the switch (method 2b). The agreement between the two methods was 70 (23) %.

The prevalence of head movement ranged from 21 to 43% task time across participants, with a median (IQR) of 30 (12) %.

5. Discussion

5.1. Operator workflow

In a situation akin to a road traffic control room, perception of the current situation primarily involves two processes. The first is the ability to collate sufficient information to define the situation. From the HTA, this information can be considered in terms of location, type and impact of an event. In routine (and planned) incidents, the definition of an 'event' can be straightforward – there is a specific and discrete incident (e.g., a collision, a broken down vehicle, an object in the road) which can be unambiguously defined in terms of the information. In this case, the ability to define the event in terms of the information is supported by the categorisation scheme applied in the options in the incident log. The second process which is relevant to perception of an event is the ability to recall previous, related examples. In our observations, this ability was supported mainly through discussion with colleagues

(although there is also the likelihood that the operator would simply remember similar events). The operators are monitoring the situation and choosing an appropriate response to make. We believe that this is not a two-stage process of comprehend and then respond, but rather an interleaving process in which both activities are performed in parallel, with one influencing the other.

5.2. Visual scanning

In this study we observed heterogeneous visual scanning behaviour, both with respect to the attended regions of interest and combination of information sources. Previous work in other domains showed comparable preferences for sources: CCTV operators primarily attended to a selected video feed on a separate monitor rather than looking at all available feeds (Stainer et al., 2013) and air traffic controllers attended to a preferred information source (Stein, 1992). In the present study, the difference in CCTV monitoring between participant 5 and all other participants was rooted in the ability to view a selected CCTV feed on the panel in the far field (SS) or on monitor D4 in the near field. Only participant 4 attended to the schematic road map on D3 for a substantial amount of time (viewing time 24%), while the other participants did not use this source much (viewing time \leq 7%). None of the operators attended to the large bank of sixteen non-interactive CCTV feeds (BS) for large amounts of time. This may be related to information overload from too much visual input; a recent review reported that an increasing number of CCTV feeds results in a drop of both operator accuracy and confidence (Stainer et al., 2013). On the other hand, the CCTV feeds were static and might not contain a view relevant to the task. Operator 2 in fact used BS to see the broken down motorbike. Task constraints are likely to affect the similarity in viewing behaviour across staff: while the completion of subgoal 1, 5 and 7 were highly constrained (requiring the incident log on D2 with access to road sign amendment and/or a phone), subgoals 2, 3, 4 and 6 allowed room for individual preferences in ROI attendance; all goals could be completed by combining a variety of information sources, such as the interactive bank of CCTV on SS, the CCTV monitor on D4 or the schematic road map on D3 or attending to a single preferred source.

The median dwell times of 1–4 s across participants compared well to previously reported values of for example 0.1–3.5 s when assessing swimming patterns (Moreno et al., 2006) or switch frequencies of around 1.1 Hz (equal to dwells of around 0.9 s) when assessing football games (North et al., 2009).

Use of the network analysis metrics revealed the following scanning behaviour: around 20–30% of possible paths were taken, which indicates that the attendance sequence was not completely random. The viewing networks highlighted preferred paths between pairwise ROIs, showing that operators had individual patterns for those information sources which they frequently attended

Table 3
Network analysis outcomes.

Metric	Matrix directionality	Median (IQR)	P1	P2	P3	P4	P5
Number of edges	Undirected	13 (7)	27	13	12	15	8
	Directed	21 (12)	46	21	18	23	12
Link density (fraction)	Undirected	0.3 (0.2)	0.60	0.29	0.27	0.33	0.18
	Directed	0.2 (0.1)	0.51	0.23	0.20	0.26	0.13
Inclusiveness (in %)	Undirected	70 (10)	100	70	70	70	60
Average nodal degree	Undirected	3 (1)	5.4	2.6	2.4	3	1.6
	Directed	4 (2)	9.2	4.2	3.6	4.6	2.4
Leaf nodes	Undirected	0 (0)	1	0	0	0	0
	Directed	1 (1)	1	2	1	1	2

This table shows median (IQR) for the network metrics as well as participant-specific data. Results are shown for undirected networks (irrespective of the direction of the gaze shift for a given edge) and directed networks.

to in a specific order. The average nodal degree of around 3 indicates that an individual ROI was attended to following attention to various other ROIs. This means that there was not one specific sequential pattern, but flexibility in how information was combined and ROIs attended to sequentially. In line with this finding, the number of leaf nodes was low, ranging from 0 to 2. The median nodal degree was highest for D2 (incident log) and TA (table in front of operator), which fits in with the function of these two ROIs: D2 serves to aggregate information from all ROIs, and TA is attended to during various points of the task to for example take notes on paper. The median nodal degree for the other ROIs was lower, indicating that they had a more specific function during sequential scanning.

The limitations to this study are the small sample size and the number of scenarios. Sample size was limited by the number of staff working at the facility, where 5 out of 6 participants chose to participate. The facility subsequently changed/updated the control room layout, so it was not possible to go back and record further scenarios or repeats. In this study we present results for two 'real' live scenarios, which differed substantially, and three 'simulated' scenarios, which we hoped would result in reasonably similar behaviour during the workflow of locating an object in the road and making the adequate responses. This case study hence may not generalize, despite giving a first insight into road traffic management behaviour. With a larger sample size and more scenarios, a future study may be able to identify and classify distinct patterns. With this case study we aimed to provide a starting point when considering the layout of a new facility and introduce metrics that may in future allow to compare operator behaviour across domains.

5.3. Eye- and head movement

The high prevalence of head movements during the task highlights that head movement should be taken into account during control room design in context of physiological requirements and possible strains. In the present study, there was wide spacing of information sources, which can require head movement *per se* when switching gaze between certain ROIs: the approximate degree visual angle between ROI centres was approximately 35°. This value is close to the physiological threshold for head movement necessity (45°–50°). A confounding factor is that the high prevalence of head movement throughout the study may have resulted in ROI switches performed by the eyes being classed as a match by chance. On the other hand, matching eye- and head switches based on the ROI attended to by the head may suffer from artefacts of the head ROI mapping, as this was performed using the simplifying assumption that the CMOS centre is a reliable indicator of head orientation.

5.4. Considerations for control room design

Having to follow a clearly defined procedure through the necessity to complete the incident log ensured that despite heterogeneous sampling approaches, subgoal completion remained structured and directed. The differences in scanning patterns between operators leave two opposing hypotheses for optimal control room or interface design: on the one hand, it may be beneficial to accommodate different individual workflows so that operators can use material which feels most intuitive and insightful for them. On the other hand, in the related domain of air traffic control, trainees are "explicitly instructed to acquire a scan pattern across the radar display and other equipment that allows coverage of the scope without being lured into any particular problem in one part of the airspace" (Durso and Dattel, 2006). This approach to constrain viewing modalities and learn scan paths may be advantageous because it encourages a more structured visual scanning protocol,

ensuring that the focus is kept on the task at hand and encouraging a more consistent behaviour between different operators.

Acknowledgements

The authors would like to thank Philippe Coutard and the staff of DIR Centre-Est for making this study possible and Natan Sorin Morar for proof-reading of, and comments on, an initial version of the paper. This work was part funded European Union FP7 project SPEEDD (619435).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.apergo.2017.01.006>.

References

- Annett, J., Duncan, K.D., Stammers, R.B., Gray, M.J., 1971. Task Analysis. (HMSO), London.
- Bizzi, E., Kalil, R.E., Morasso, P., 1972. Two modes of active eye-head coordination in monkeys. *Brain Res.* 40 (1), 45–48.
- Borji, A., Itti, L., 2014. Defending Yarbus: eye movements reveal observers' task. *J. Vis.* 14 (3).
- Boyer, D.J., 1995. The Relationship Among Eye Movements, Head Movements, and Manual Responses in a Simulated Air Traffic Control Task. DTIC Document. US Department of Transportation, https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/1990s/media/am95-23.pdf.
- Burch, M., Konevtsova, N., Heinrich, J., Hoeferlin, M., Weiskopf, D., 2011. Evaluation of traditional, orthogonal, and radial tree diagrams by an eye tracking study. *Visualization and computer graphics.* *IEEE Trans* 17 (12), 2440–2448.
- Chang Hoon, H., Jong Hyun, K., Seung Jun, L., Poong-Hyun, S., 2006. Investigation on relationship between information flow rate and mental workload of accident diagnosis tasks in NPPs. *Nucl. Sci. IEEE Trans.* 53 (3), 1450–1459.
- Dekker, S., 2014. *The Field Guide to Understanding human Error*. Ashgate Publishing, Ltd.
- Durso, F.T., Dattel, A.R., 2006. Expertise and transportation. In: Ericsson, K.A., Charness, N., Feltovich, P.J., Hoffman, R.R. (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press, New York, pp. 355–371.
- Endsley, M.R., Rodgers, M.D., 1996. Attention distribution and situation awareness in air traffic control. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 40 (2), 82–85.
- Folds, D., Brooks, J., Stocks, D., Fain, W., Courtney, T., Blankenship, S., 1993. *Functional Definition of an Ideal Traffic Management System*. Georgia Tech Research Institute, Atlanta, GA.
- Foulsham, T., Chapman, C., Nasiopoulos, E., Kingstone, A., 2014. Top-down and bottom-up aspects of active search in a real-world environment. *Can. J. Exp. Psychology/Revue Can. de Psychol. expérimentale* 68 (1), 8–19.
- Freedman, E.G., 2008. Coordination of the eyes and head during visual orienting. *Exp. Brain Res.* 190 (4), 369–387.
- Goossens, H.H.L.M., Opstal, A.J.V., 1997. Human eye-head coordination in two dimensions under different sensorimotor conditions. *Exp. Brain Res.* 114 (3), 542–560.
- Guitton, D., Volle, M., 1987. Gaze Control in Humans: Eye-head Coordination during Orienting Movements to Targets within and beyond the Oculomotor Range.
- Hanson, E.K.S., 2004. Focus of Attention and Pilot Error. *Eye Tracking Research and Applications Symposium (ETRA)*.
- Henderson, J.M., 2003. Human gaze control during real-world scene perception. *Trends Cognitive Sci.* 7 (11), 498–504.
- Hénique, E., Lindegaard, S., Hunt, B., 2008. Design of Large and Complex Display Systems. *Handbook of Control Room Design and Ergonomics*. CRC Press, pp. 83–138.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J., 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press.
- Howard, C., Gilchrist, I., Troscianko, T., Behera, A., Hogg, D., 2011. Task relevance predicts gaze in videos of real moving scenes. *Exp. Brain Res.* 214 (1), 131–137.
- Hughes, G., Kornowa-Weichel, M., 2004. Whose fault is it anyway?: A practical illustration of human factors in process safety. *J. Hazard. Mater.* 115 (1–3), 127–132.
- Ivergard, T., Hunt, B., 2008. *Handbook of Control Room Design and Ergonomics*. CRC Press, London.
- Jacob, R.J., Karn, K.S., 2003. Eye tracking in human-computer interaction and usability research: ready to deliver the promises. *Mind* 2 (3), 4.
- Kim, J.T., Shin, S.K., Kim, J.H., Seong, P.H., 2013. An experimental approach to estimate operator's information processing capacity for diagnosing tasks in NPPs. *Ann. Nucl. Energy* 59 (0), 100–110.
- Kirwan, B., 1992. Human error identification in human reliability assessment. Part

- 1: overview of approaches. *Appl. Ergon.* 23 (5), 299–318.
- Krupinski, E.A., 2010. Current perspectives in medical image perception. *Atten. Percept. Psychophys.* 72 (5), 10.3758/APP.3772.3755.1205.
- Land, M.F., 2006. Eye movements and the control of actions in everyday life. *Prog. Retin. Eye Res.* 25 (3), 296–324.
- Landry, S., 2011. Human centered design in the air traffic control system. *J. Intelligent Manuf.* 22 (1), 65–72.
- Lin, Y., Zhang, W.J., Watson, L.G., 2003. Using eye movement parameters for evaluating human–machine interface frameworks under normal control operation and fault detection situations. *Int. J. Human-Computer Stud.* 59 (6), 837–873.
- McIntire, L.K., McIntire, J.P., McKinley, R.A., Goodyear, C., 2014a. Detection of Vigilance Performance with Pupillometry. *Eye Tracking Research and Applications Symposium (ETRA)*.
- McIntire, L.K., McKinley, R.A., Goodyear, C., McIntire, J.P., 2014b. Detection of vigilance performance using eye blinks. *Appl. Ergon.* 45 (2 PB), 354–362.
- Meshkati, N., 1991. Human factors in large-scale technological systems' accidents: three mile Island, bhopal, Chernobyl. *Organ. Environ.* 5 (2), 133–154.
- Meyer, L., Schultz, M., Schmidt-Roßleben, S., Fricke, H., 2013. Understanding the safety-relevance of visual cue perception at a Surface Manager HMI. In: *SIDS 2013-Proceedings of the SESAR Innovation Days*.
- Moore, K., Gugerty, L., 2010. Development of a novel measure of situation awareness: the case for eye movement analysis. In: *Proceedings of the Human Factors and Ergonomics Society*.
- Morasso, P., Sandini, G., Tagliasco, V., Zaccaria, R., 1977. Control strategies in the eye-head coordination system. *Systems, man and cybernetics. IEEE Trans.* 7 (9), 639–651.
- Moray, N., Rotenberg, I.R.A., 1989. Fault management in process control: eye movements and action. *Ergonomics* 32 (11), 1319–1342.
- Moreno, F.J., Saavedra, J.M., Sabido, R., Luis, V., Reina, R., 2006. Visual search strategies of experienced and nonexperienced swimming coaches. *Percept. Mot. Ski.* 103 (3), 861–872.
- Nakashima, R., Shioiri, S., 2014. Why do we move our head to look at an object in our peripheral Region? Lateral viewing interferes with attentive search. *PLoS ONE* 9 (3), e92284.
- North, J.S., Williams, A.M., Hodges, N., Ward, P., Ericsson, K.A., 2009. Perceiving patterns in dynamic action sequences: investigating the processes underpinning stimulus recognition and anticipation skill. *Appl. Cogn. Psychol.* 23 (6), 878–894.
- Noyes, J., Bransby, M., 2001. *People in Control: Human Factors in Control Room Design*. IET.
- Oommen, B.S., Stahl, J.S., 2005. Overlapping gaze shifts reveal timing of an eye–head gate. *Exp. Brain Res.* 167 (2), 276–286.
- Pohl, M., Schmitt, M., Diehl, S., 2009. Comparing the Readability of Graph Layouts Using Eyetracking and Task-oriented Analysis. In: *Proceedings of the Fifth Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging*. Eurographics Association, Victoria, British Columbia, Canada, pp. 49–56.
- Poole, A., Ball, L.J., 2006. Eye tracking in HCI and usability research. *Encycl. Hum. Comput. Interact.* 1, 211–219.
- Proudlock, F.A., Gottlob, I., 2007. Physiology and pathology of eye–head coordination. *Prog. Retin. Eye Res.* 26 (5), 486–515.
- Reason, J., 2000. *Human Error: Models and Management*.
- Scott, J., 2013. *Social Network Analysis*. (Sage), London.
- Seok, H.H., Kyeong, T.K., Sun, M.S., Jong, S.Y., Jae, I.L., Sun, M.C., Young, W.S., 2006. Effects of expertise and situation complexity on visual attention and action planning for air traffic control. In: *Proceedings of the Human Factors and Ergonomics Society*.
- Shepherd, A., 2001. *Hierarchical Task Analysis*. Taylor and Francis, London.
- Shepley, J.P., Johnson, C.M., Sanchez, J., Smith, E.C., 2009. Eye-tracking Analysis of Near-term Terminal Automation for Arrival Coordination. 9th AIAA Aviation Technology, Integration and Operations (ATIO) Conference. Aircraft Noise and Emissions Reduction Symposium (ANERS).
- Smith, G.J.D., 2004. Behind the screens: examining constructions of deviance and informal practices among CCTV control room operators in the UK. *Surveillance & society. CCTV Spec.* 2, 376–395.
- Snowden, R., Snowden, R.J., Thompson, P., Troscianko, T., 2012. *Basic Vision: an Introduction to Visual Perception*. Oxford University Press.
- Stainer, M.J., Scott-Brown, K.C., Tatler, B., 2013. Looking for trouble: a description of oculomotor search strategies during live CCTV operation. *Front. Hum. Neurosci.* 7 article 615.
- Stanton, N.A., Salmon, P., Jenkins, D., Walker, G., 2009. *Human Factors in the Design and Evaluation of Central Control Room Operations*. CRC Press.
- Stelman, K.S., McCarley, J.S., Wickens, C.D., 2011. Modeling the control of attention in visual workspaces. *Hum. Factors J. Hum. Factors Ergonomics Soc.* 53 (2), 142–153.
- Stein, E.S., 1992. *Air Traffic Controller Visual Scanning*, DTIC Document.
- Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. *Nature* 381 (6582), 520–522.
- Tien, T., Pucher, P.H., Sodergren, M.H., Sriskandarajah, K., Yang, G.-Z., Darzi, A., 2014. Eye tracking for skills assessment and training: a systematic review. *J. Surg. Res.* 191 (1), 169–178.
- Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D.B., Wilson, A., 2004. What happens next? The predictability of natural behaviour viewed through CCTV cameras. *Perception* 33 (1), 87–101.
- UK Highways Agency, 2009. *Highways Agency Network Management Manual: Part 7-traffic Incident Management and Contingency Planning*. http://www.standardsforhighways.co.uk/ha/standards/nmm_rvsc/docs/nmm_part_7a.pdf.
- Wollaston, W.H., 1824. On the apparent direction of eyes in a portrait. *Philosophical Trans. R. Soc. Lond.* 114, 247–256.
- Wood, J., 2004. Control room design. In: Sandom, C., Harvey, R.S. (Eds.), *Human Factors for Engineers*. The Institution of Engineering and Technology, London, UK.
- Zangemeister, W.H., Stark, L., 1982. Types of gaze movement: Variable interactions of eye and head movements. *Exp. Neurol.* 77 (3), 563–577.