# Eye movements and verbal communication as indicators for the detection of system failures in a control room task

*Carmen Bruder, Carolina Barzantny, & Dirk Schulze Kissing*
*German Aerospace Center, Department of Aviation and Space Psychology*
*Germany*

**Abstract**

In modern control rooms, operators need to monitor visual information representing large technical systems. Operators usually monitor together in teams in order to detect abnormal system behaviour in time. It remains an open question which performance indicators are valuable for assessing a team member's capabilities of detecting abnormal system behaviour. The present study investigates the value of monitoring behaviour and communication behaviour for predicting the performance results of subjects attempting to detect system failures while executing a control room task. A simulation of a generic control room was implemented in order to enable synchronized measurement of monitoring processes in teams. The monitoring behaviour was measured by tracking the eye movements of the team members while they were monitoring for system failures. Simultaneously, the communication behaviour between team members was recorded. Eye-tracking data and communication data were analysed including the interaction with team members' performance in detecting system failures in time. Data from 21 three-member teams indicate that there are significant differences in communication and to some extent in eye-movement, between operators who detect system failures in time and those who fail to do so. The findings are discussed in the context of personnel selection and training team members in control rooms.

**Introduction**

This paper presents an eye-tracking study that investigates the monitoring and communication behaviour of operators while collaboratively supervising the dynamic processes of a control room simulation. In this study, monitoring behaviour was measured using eye tracking. By tracking the operator's eye movements, the visual attention processes while gathering relevant information as well as detecting abnormal system behaviour could be visualized. Furthermore, recording verbal communication behaviour between team members makes it possible to indicate the coordinative processes while monitoring together. By specifically investigating how monitoring and communication behaviour can be used to predict the performance of operators attempting to detect system failures, the goal is to provide initial indications for selecting and training operators in control room teams.

In D. de Waard, F. Di Nocera, D. Coelho, J. Edworthy, K. Brookhuis, F. Ferlazzo, T. Franke, and A. Toffetti (Eds.) (2018). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2017 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

*Collaborative monitoring in control room teams*

The control room is an example of a working environment where operators supervise complex and dynamic processes together. Control rooms can be found particularly in domains where safety is of critical importance, such as airport operational centers, air traffic control centers, nuclear power plant control and military control centers, where human error can have severe consequences (Hauland, 2008; Salas et al. 2008). As monitoring is one of the core tasks in control rooms, teams of operators are required to monitor the system appropriately (Sharma et al., 2016). In control rooms, not only is the individual situation awareness relevant, but also the situation awareness of the team. Through interactions, operators in a team can dynamically modify each other's perceptual and active capabilities (Gorman et al., 2006). However, when monitoring a system, it is essential that team members work together effectively and cooperatively (Cooke et al., 2000; Salas et al., 2008). In order to coordinate their activities in such "centers of coordination," not only do individuals have to be aware of their own situation, but they must also be aware of their team members' situation (e.g. Suchman, 1997).

The importance of communication in control operations has been stressed by Carvalho et al. (2007). Communication as a "meta-teamwork process that enables the other processes" (Papenfuss, 2013, p. 319) provides indications for the coordinative activities while monitoring. Cooke et al. (2013) stressed that, especially in critical situations, "it is not only critical that teams correctly assess the state of the environment and take action, but how this is accomplished (p. 279)". As a consequence, recording the quality and degree of a team's communication provides insight into how the group deals with critical situations.

*Measuring collaborative monitoring*

A variety of studies support the idea that eye movements offer an appropriate means for measuring the efficient and timely acquisition of visual information (e.g. Findlay & Gilchrist, 2003; Underwood et al., 2003; for an overview see Holmqvist et al., 2011). Based on this research, eye movement parameters that reflect the human monitoring performance have been identified (Grasshoff et al., 2015; Hasse & Bruder, 2015). Bruder et al. (2014) investigated the link between these eye movement parameters and the monitoring behaviour of experts, compared the monitoring behaviour of experts with novices (Bruder et al., 2013), and used eye movements to research differences in monitoring behaviour resulting in detected automation failures and behaviour resulting in missed failures (Bruder & Hasse, 2016).

While the results of previous studies give valuable insight into eye movements during the process of monitoring individually, the present study focuses on collaborative monitoring behaviour in a team task. In this context, monitoring behaviour leading accurate failure detection will be compared with monitoring behaviour that leads to missed failures. Additionally, the communication behaviour while monitoring will be taken into account. The following research questions will be addressed: What are valuable performance indicators in a team task with respect to communication quality and monitoring behaviour that differentiate between accurate failure-detection and missed failures?

## Method

An empirical study was undertaken requiring collaborative monitoring while performing a control room team task.

*Simulation of a generic control room*

In the present study, the simulation of a generic control room, called ConCenT (Generic Control Center Task Environment), was used to enable synchronized measurement of monitoring processes in teams (Schulze-Kissing & Bruder, 2016). ConCenT replicates different control room tasks by simulating the production processes of several technical facilities spread over three locations, which are supervised by a team of three human operators. It simulates four different tasks: monitoring the distributed production processes, reporting system deviations (failures), diagnosing the sources of deviations and remedying the deviating processes by deciding between two alternative choices. These four tasks have to be managed within a team of three operators. Since this paper presents findings concerning the monitoring task and the reporting task, these two tasks are described in more detail. Figure 1 shows a screenshot of the monitoring screen of ConCenT.


*Figure 1. Monitoring screen of ConCenT containing the displays of nine production facilities and three power stations, which are distributed over three locations*

In the monitoring and reporting task, each team member had to observe nine of 27 gauges in total and three joint power station gauges with the objective of reporting deviations from standard processes within a time span of four seconds. Each of the 27 gauges represented the production processes of a single production line. Deviations could be recognized when one of the black arrows, indicating the current value on each of the 27 gauges, exceeded or fell below the tolerance range (marked green). Before a deviation occurred, a specific constellation of production processes indicated this kind of critical situation. Critical situations could only be identified when the distributed information on the production processes was communicated between team members. As a consequence, the team was able to anticipate deviations in the production. Sharing all relevant information on the production processes therefore helped identify critical situations and anticipate as well as helped report any system deviations.

*Eye tracking system*

Each participant was seated in front of a 24-inch LCD computer display at a distance of approximately 60 cm. Eye movements were recorded remotely by using the Eye Follower System manufactured by LC Technologies, Inc. The system operated at 120 Hz and was combined with the simulation tool ConCenT to ensure that both systems used the same timestamp. The fixation-detection algorithm was set with a minimum sample for fixation detection of six gazes on a particular screen point – within the deviation threshold of 25 pixels.

*Sample*

The study was conducted with a sample size of $N$ = 63. Of this total, 41 individuals were applicants for air traffic control training (ATC) at DFS (German Air Traffic Control), while the remaining 22 individuals were students and graduates from different universities. All participants were between 18 and 34 years old ($M$ = 21.57, $SD$ = 3.39) and 47.6% were female (52.4% male). ATC participants were recruited with a personal call from DLR (German Aerospace Center), Hamburg, and compensated €25 for their participation in the 2.5hrs experiment. Students were recruited via social media and with flyers posted on the campus of the University of Hamburg.

*Procedure*

The three participants in each team performed the experiment at the same time, each with a separate computer and eye tracking system. A room divider was installed between the participants to prevent direct communication and eye contact. Written instructions introduced participants to their general tasks as operators working in a control center, and explained their specific responsibilities while monitoring the system, diagnosing errors and solving problems. Following this, each team was guided through a practice scenario that lasted about ten minutes. Throughout the practice scenario, participants familiarized themselves with how to anticipate, detect and report deviations from standard processes in time. After the practice scenario, participants confirmed their understanding of the monitoring procedure and the other required tasks. The test scenario began with the ramp-up of the gauges and ended after 72 minutes. A manipulation check was done, and participants were required to complete a questionnaire regarding their attitudes towards teamwork. Finally, participants were asked to give their impressions of the study.

*Design and measurements*

The present study investigates the relationship between team members monitoring as well as communication behaviour and their capabilities of detecting system deviations. The dependent variables included the monitoring behaviour (tracking eye movements) and the quality of communication. The quasi-independent variable was the performance level (deviation reported successfully vs. deviation missed). These two groups (cases of successful detection of deviations and cases of missed deviations) were created post-hoc. A deviation was successfully reported if a participant clicked on the button "Diagnose" next to the gauge within the corresponding time frame (4s). Each of the six deviations could either be detected (= successful detection of deviation) or not detected (= missed).

*Measuring monitoring behaviour and communication quality*

Eye movements were recorded while monitoring the distributed production processes as well as reporting system deviations. Afterwards, they were synchronized with the logged simulation events before and during the occurrence of deviations. At first, twelve areas of interest were defined for each team partner (A, B, C): nine gauges for the production processes and three gauges for the power stations. For each of the six deviations in the test scenario, AOIs were predefined according to where an operator's attention should be allocated within the interval before and while a deviation occurred. It was defined in advance, which gauges must be monitored to anticipate system deviations and this decision was based on the information necessary for detecting critical situations.

Regarding the timely allocation of attention on relevant AOIs when detecting deviations, four successive monitoring phases were defined (1. identification phase, 2. verification phase, 3. anticipation phase, 4. detection phase). Within each of these four monitoring phases, the team member had the opportunity to share their information in order to allocate their attention in an ideal way. In the first two phases, identification and verification, the team member had to share their information to find out whether or not there was a critical situation. In the third phase (anticipation), they had to anticipate the gauge where the deviation could happen. In the last phase (detection), the deviation could occur and had to be reported. The eye tracking parameters on the relevant AOIs were analysed for each monitoring phase, team partner and deviation.

The relative fixation count (rfc) was calculated in terms of the predefined, relevant AOIs for each of the four monitoring phases. The rfc is defined as the ratio between the number of fixations on relevant AOIs and all fixations within a given time span. Relative parameters ranged from 0 to 1, with 0 indicating that no eye movements fell on predefined AOIs within a time period, and with 1 indicating that all eye movements fell on the predefined AOIs within that time period.

During the test scenario, the verbal communication of each team member was recorded. An audio file logged the identities of each speaker, the content of the information exchanged, and the duration of this communication. Each audio file was analysed with respect to the necessary communication in all six intervals before a system deviation. This analysis provided the basis for determining the quality of communication. For each of the four monitoring phases, participants could score on a scale from 0 (no communication or wrong communication of necessary information) to 1 (right communication/no communication needed) in each of the 6 intervals before a system deviation.

**Results**

Data from 52 subjects were reported, each of whom experienced six deviations within the test scenario. Data were excluded from the reported results when a scenario was not completed due to technological problems (18.1%), if they failed the manipulation check the manipulation check was not passed (4.8%), and when eye movement data were missing or showed major inconsistencies (3.2%). For communication analyses, additional data were excluded when no communication was recorded by the system (14.8%). In sum, eye-tracking data, communication data and deviation-detection data from 212 deviations were included in the statistical analyses. On a scale from 0 to 6, an average of 4.33 (SD = 1.37) deviations were reported with an average response time of 2.17 seconds (SD = 0.56; see Table 1 for a detailed overview).

*Table 1. Descriptive performance data (N = 52)*

| | *Deviation detected* | | *Response time* | |
|---|---|---|---|---|
| *Deviation* | *n* | *%* | *M* | *SD* |
| *1* | 25 | 48.1 | 2.79 | 0.74 |
| *2* | 25 | 48.1 | 2.68 | 0.76 |
| *3* | 44 | 84.6 | 1.95 | 0.70 |
| *4* | 42 | 80.8 | 2.01 | 0.87 |
| *5* | 42 | 80.8 | 2.00 | 0.67 |
| *6* | 47 | 90.4 | 1.76 | 0.70 |
| *All* | | | 2.17 | 0.56 |

Looking at the eye tracking data, the attention allocation of the test subjects implies that in the case of successful detection of deviations, relevant AOIs were focused on more intensively if the deviation was detected successfully (see Figure 2, which shows the second deviation in the test scenario as an example).

*Figure 2. Comparison of attention allocation in a case of successful detection of deviation (left) and a missed deviation (right), illustrated by the eye tracking data (N = 52) during the anticipation phase of the second deviation in the test scenario (marked yellow)*

A variance analysis with repeated measurement was conducted to compare the main effects of monitoring phase and the interaction effect between monitoring phase and performance in detecting deviations on the relative fixation count. The factors PHASE (four levels: identification, verification, anticipation, detection) and DETECTION (two levels: detected, not detected) were defined and analysed. See Table 2 for descriptive data of the eye tracking parameter. Multivariate tests showed a significant effect for PHASE [$F(3, 240) = 5.038$, $p < .005$; Wilk's $\lambda = .94$, partial $\varepsilon^2 = .059$]. It could be shown that subjects fixated relevant AOIs most frequently within the identification phase (1) and the verification phase (2). No significant effect of the interaction between PHASE and DETECTION [$F(3, 240) = 2.297$, $p = .078$; Wilk's $\lambda = .97$, partial $\varepsilon^2 = .028$] on eye tracking parameter was found. Post hoc tests indicated that accurate deviation detection is only related to a higher frequency of fixations on relevant AOIs during the anticipation phase [$t(305)=-2.22$, $p<.05$)]. Concerning the identification phase, verification phase and detection phase, differences between cases of accurate and missed deviation detection were not significant [$p > .05$]. The interaction of DETECTION and PHASE on relative fixation counts on relevant AOIs is shown in Figure 3 (left).

*Table 2. Descriptive data for the eye tracking parameter (relative fixation count) and communication quality parameter in the four monitoring phases (rows), separately for deviations detected and deviations NOT detected (columns).*

| | *Deviation detected* | | *Deviation NOT detected* | |
|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* |
| *Relative fixation counts* | | | | |
| *Identification (1)* | 0.47 | 0.25 | 0.51 | 0.20 |
| *Verification (2)* | 0.48 | 0.26 | 0.44 | 0.28 |
| *Anticipation (3)* | 0.45 | 0.23 | 0.37 | 0.19 |
| *Detection (4)* | 0.44 | 0.20 | 0.38 | 0.18 |
| *Communication quality* | | | | |
| *Identification (1)* | 0.99 | 0.11 | 0.86 | 0.38 |
| *Verification (2)* | 0.47 | 0.50 | 0.31 | 0.47 |
| *Anticipation (3)* | 0.74 | 0.44 | 0.55 | 0.50 |
| *Detection (4)* | 0.15 | 0.36 | 0.25 | 0.44 |

Following, a variance analysis with repeated measurement was conducted to compare the main effects of monitoring phase and the interaction effect between monitoring phase and performance in detecting deviations on communication quality. The factors PHASE (four levels: identification, verification, anticipation, detection) and DETECTION (two levels: detected, not detected) were defined and analysed. See Table 2 for descriptive data of the communication quality. Multivariate tests showed a significant effect for PHASE [$F(3, 320) = 336.142$, $p < .001$; Wilk's $\lambda = .24$, partial $\varepsilon^2 = .759$]. It could be shown that subjects communicated accurate information most frequently during the identification phase (1) and anticipation phase (3). The interaction between PHASE and DETECTION [$F(3, 320) = 5.457$, $p < .005$; Wilk's $\lambda = .95$, partial $\varepsilon^2 = .049$] on communication quality was found. Post hoc tests showed that accurate deviation detection is related to higher communication quality during the identification phase [$t(98.11)=-3.20$, $p<.05$)], verification phase [$t(182.43)=-2.42$, $p<.05$)] and anticipation phase [$t(152.92)=-3.61$, $p<.05$)], but not during the detection phase [$p > .05$]. The interaction of DETECTION and PHASE on communication quality is shown in Figure 3 (right).

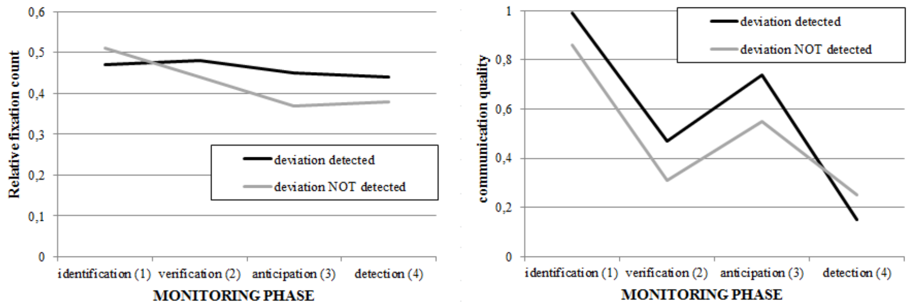


*Figure 3. Interaction effects of detection * time unit (estimated mean values) on the communication quality as the relative frequency of correctly communicated information (left) and on the relative fixation counts on relevant AOI (right)*

## Discussion and further research

The present study investigated the role of monitoring behaviour and communication behaviour as performance predictors for the detection of failures (=deviations) in a control room team task. To the subjects were given the task of monitoring dynamic processes in a team of three operators with the objective of anticipating and detecting deviations from standard processes by communicating relevant information adequately. To summarize the results, data from 21 three-member teams indicate that there are significant differences in communication and to some extent in eye-movement, between operators who detect system deviations in time and those who miss the deviations. This is shown by the fact that successful failure detection is related to a higher frequency of communication and focusing attention on relevant information during the anticipation phase.

Comparing the predictive value of communication quality and monitoring behaviour, the relationship between the frequency of monitoring relevant information and the detection of system deviations is clearly weaker than the relationship between the frequency of communicating relevant information and the detection of system deviations. However, in the case of successful failure detection, relevant information is monitored more frequently shortly before the deviation occurs when the automation failure should be anticipated. This is quite understandable, because monitoring relevant information within the anticipation phase is only possible if the subject has identified the critical production system together with the team partners, thus leading to successful detection of system deviations in time.

Contrary to prior expectations, no substantial relationship between successful deviation detection and monitoring behaviour within the identification phase, information phase and detection phase was found. Besides this, the effect sizes on eye tracking parameters are small. These may be due to the fact that technical problems lead to losses of eye tracking data, but also to certain methodological shortcomings of predicting deviation detection by means of the eye movements of human operators. Further research will improve the reliability of eye-movement indicators by adjusting the definition of information that is relevant for detecting deviations.

With respect to communication behaviour, the differences between detected and missed automation failures were highest when the system deviation could be verified, which happened in the second monitoring phase. This result implies that successful failure detection is highly related to adequate communication of relevant information at the beginning of an upcoming situation. A deviation can only be detected in time if the team members communicate the relevant information and identify the critical production system together with the team partners.

Predicting the detection of system failures in a team task within a dynamic setting using eye tracking and communication quality is an innovative strategy that enables the development of new approaches for personnel selection and training. Learning from the differences in monitoring and communication behaviour between successful and unsuccessful failure detection will be helpful in selecting successful trainees and providing them with appropriate training. Especially the monitoring and communication patterns related to successful detection may be useful in order to give trainees direct feedback on their own monitoring behaviour or to demonstrate “correct” monitoring behaviour.

Further research is replicating this study with a larger sample of 48 teams and prior technical problems are being reduced, which will lead to a significant gain in the volume of data. In contrast to the study reported here, in further research the effect of team coordination within a monitoring task is systematically investigated by comparing the monitoring behaviour of communicating teams to a control condition where all channels for oral communication are blocked.

**References**

Bruder, C., Eißfeldt, H., Maschke, P., & Hasse, C. (2013). Differences in monitoring between experts and novices. In *Proceedings of the HFES 57th Annual Meeting, 2013* (pp. 295-298). Sage, Thousand Oaks, CA.

Bruder, C., Eißfeldt, H, Maschke, P., & Hasse, C. (2014). A model for future aviation: Operators monitoring appropriately. *Aviation Psychology and Applied Human Factors*, *4*, 13-22.

Bruder, C., Weber, P., & Hasse, C. (2016). To Look and (Not) See: Predicting the Detection of Automation Failures Based on the Eye Movements of Human Operators. In *Proceeding of the HCI-Aero '16 Proceedings of the International Conference on Human-Computer Interaction in Aerospace*. New York, NY, USA: ACM Press.

Carvalho, P.V.R., Vidal, M.C.R., & de Carvalho, E.F. (2007). Nuclear power plant communications in normative and actual practice: A field study of control room operators' communications. *Human Factors in Ergonomics and Manufacturing*, *17*, 43–78.

Cooke, N. J., Salas, E., Cannon-Bowers, J. A. & Stout, R. (2000). Measuring team knowledge. *Human Factors*, *42*, 151-173.

Cooke, N.J., Gorman, J.C., Myers, C.W. & Duran, J.L. (2013). Interactive team cognition. *Cognitive Science*, *37*, 255–285.

Findlay, J.M. & Gilchrist, I.D. (2003). *Active Vision*. Oxford (UK): Oxford University Press.

Gorman, J.C., Cooke, N.J., & Winner, J.L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics, 49*, 1312-1325.

Grasshoff, D., Hasse, C., Bruder, C., & Eißfeldt, H. (2015). On the development of a monitoring test for the selection of aviation operators. In D. Harris (Ed.), *Proceedings of Engineering Psychology and Cognitive Ergonomics, 12th International Conference EPCE 2015, held as Part of HCI International 2015 (*pp. 537-546). Berlin, Heidelberg: Springer.

Hasse, C. & Bruder, C. (2015). Eye Tracking Measurements and their Link to a Normative Model of Monitoring Behaviour. *Ergonomics*, *58*, 355-367.

Hauland, G. (2008). Measuring individual and team situation awareness during planning tasks in training of en route air traffic control. *The International Journal of Aviation Psychology*, *18*, 290-304

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: Oxford University Press.

Papenfuss, A. (2013). Phenotypes of teamwork - an exploratory study of tower controller teams. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp.319-323).

Salas, E., Cooke, N.J., & Rosen, M.A. (2008). On Teams, Teamwork, and Team Performance: Discoveries and Developments. *Human Factors*, *50*, 540-547.

Sharma, C., Bhavsar, P., Srinivasan, B. & Srinivasan, R. (2016). Eye gaze movement studies of control room operators: A novel approach to improve process safety. *Computers & Chemical Engineering*, *85*, 43-57.

Suchman, L. (1997). Centers of coordination: A case and some themes. In L.B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools and reasoning: Essays on situated cognition* (pp. 41–62). Berlin: Springer.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics, 46*, 629-646.