Analysis of dynamic performance data for the assessment of cognitive states – Results from aviation, assembly tasks and maritime transportation

Marcus ARENIUS[1], Jürgen KLIPPERT[2], Jorgos ATHANASSIOU, Oliver STRÄTER[1]

[1] *FG Arbeits- und Organisationspsychologie, Universität Kassel*
Heinrich-Plett Str. 40, D-34132 Kassel
[2] *IG Metall Vorstand, Wilhelm-Leuschner-Str. 79, D-60329 Frankfurt a. M.*

Kurzfassung: This article will demonstrate how dynamic performance measurements can be used for inferring cognitive states in a variety of field settings. This type of data analysis can be used complementary to traditional data collection methods in order to bridge some of the known weaknesses of the methods. The authors will present results from several studies demonstrating how data from portable motion tracking and eye-tracking devices can be used to assess three main aspects: the onset and time-based development of worker fatigue in assembly tasks, the time-based impact of situation-induced affective states on visual attention and decision-making in maritime transportation and the transition between safety-relevant cognitive coping strategies by pilots in aviation. The benefits of this kind of approach will be demonstrated by showing how work systems can be improved by providing a better fit between cognition and workplace design.

Schlüsselwörter: mental fatigue, coping mechanisms, affective states, assembly tasks, aviation, maritime transportation

1. Introduction

While the optimization of physical work demands, particularly in the wake of the demographic change, has been a priority issue across various industry sectors, the optimization of work for cognitive demands is still very much work-in-progress. One of the main reasons is that the assessment of cognitive states remains a difficult issue. Traditional subjective, self-report measurements (e.g. questionnaires, interviews) are poorly suited for addressing the dynamic and fluid nature of human cognition, as they mostly provide “snapshot” representations of cognitive states (questionnaires). Advances in the technology of portable performance measurement provide a promising approach for addressing these types of issues and provide a complementary approach to traditional methods. Several studies will be presented how dynamic performance data can be used to address different issues related to cognitive states.

2. Analysis of eye-tracking dynamics in aviation – cognitive coping strategies

In order to assess cognitive coping strategies in aviation, we have to move from traditional eye-tracking data to the analysis of eye-tracking dynamics (Arenius, 2014). A short summary on what is meant by cognitive coping will be given.

In any work environment, human operators have to cope with uncertainty and unexpected events based on an (by necessity) incomplete picture of the overall situation. Hence, (human) performance is variable, as the adaptive and approximate process of coping with uncertainty forces people to "fill in the gaps" by relying on trade-offs, that mostly produce the desired results, but in very rare cases may lead to adverse outcomes (Hollnagel, 2009). Conversely, if human performance is approximate, that is, characterized by necessary short-cuts and heuristics, this should be reflected in the way human operators interact with their working environment and react to changes and disturbances.
Based on this reasoning, this article presents the results of a study conducted to assess the impact of an air-traffic control (ATC) information display for airborne separation the performance of commercial pilots. Pilots from a major German airline (n=10) flew a safety-critical flight scenario in a full flight simulator using the ATC display. Their performance was evaluated by expert ratings of professional trainers from the airline and by means of eye-tracking performance data. The pilots start the scenario on route from Athens to Heraklion. After ~60 seconds of flight, the pilots receive a traffic warning of a potential head on collision followed by an engine failure.
The pilots have to notice the engine failure and initiate recovery actions while at the same time avoiding the other flight on collision course. Thus, the pilots are facing conflicting goals: Should they focus on avoiding the collision with the other aircraft or primarily focus on the engine failure? Following the principle of "fly the aircraft first" the engine failure should be treated as the priority issue over the traffic warning (Arenius, Schneider, & Straeter, 2014). The results of the study demonstrated that some pilots focused on the ATC display and thus did not notice the engine failure in time. Using the ordination technique non-metrical multidimensional scaling (NMDS), a representation of the correlation between the visual focus on the information sources can be calculated (Straeter & Arenius, 2014). The closer the points are in the NMDS representation, the stronger the correlation between the gazes on the two information sources, that is, the stronger the common visual focusing of them. This representation can thus reveal the division of visual attention associated with this coping strategy of prioritizing the traffic warning over the engine failure. In order to create the representation, the eye-tracking data of a pilot that did not notice the engine failure in time has to be separated into meaningful segments that provide insight into the coping process associated with the onset of the traffic warning and the engine failure. The NMDS representation of the first segment shows that, with the onset of the collision warning, the strategy of the pilot changes. The ATC-display showing the warning, the central console and the autopilot as input device for a new heading to avoid collision, are in strong joint visual focus, while the other displays are completely discarded.

The second NMDS representation shows the visual attention after the pilot has noticed engine failure, ca. 20 seconds after its onset. Visual focus is placed on displays important for handling the engine failure - the primary flight display (PFD) and the engine and warnings display (E/WD) - however also on the ATC-display which is not required if the principle "fly the aircraft first" is followed. The other displays are disregarded completely. This means, that the pilot did not abandon the handling of the traffic warning although the engine failure has been identified. Thus visual attention is directed away from the most pressing task - the handling of the aircraft - to managing the traffic. Therefore, the NMDS representation shows the different coping strategies associated with safety-critical changes (traffic warning and engine failure) in the work environment.

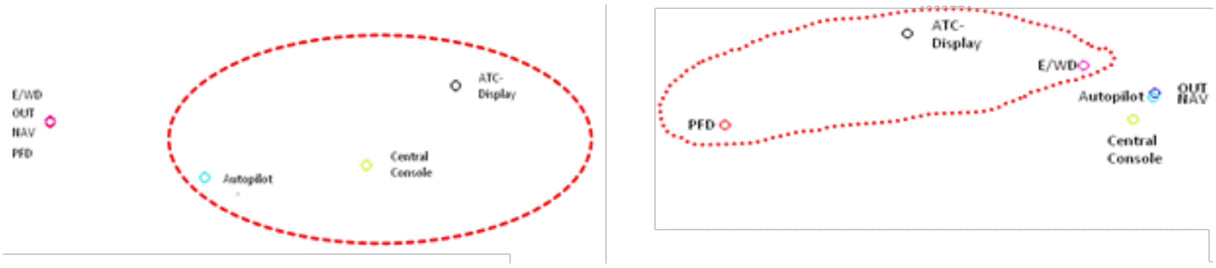


Figure 1. *Left-hand: After the traffic warning, information sources relevant for collision avoidance are focused (red circle). Right-hand: The pilot notices the engine failure occurring directly after the traffic warning late; however still also focuses on ATC-display (irregular red shape) [Note: the NMDS representation is dimensionless]*

3. Portable motion tracking and eye-tracking for assembly tasks

Simple manual tasks are characterized by recognition-action-coupling in the way of „object related actions“(ORAs, Land et al. 1999): Objects are grasped and placed on defined positions. During performance, ORAs require spatial orientation of the whole body and the recognition of task-relevant objects. Therefore, cognitive processing of visual information is required. Before performing an action, the action relevant object is firstly fixated with the eyes and secondly grasped with the hand.

The mental workload associated with simple manual tasks tends to be underload because the demands for performing a simple manual task are comparatively low. Thus, the source for mental fatigue of the operators is predominantly mental underload in combination with the repetitive character of the tasks.

To measure mental fatigue by primary task performance seems to be inadequate because operators can compensate for low workload by increasing their effort to maintain vigilance. Therefore, mental fatigue may not be determined by using performance outcomes of simple manual tasks. The measurement of mental fatigue by subjective response (e.g. questionnaires) may fail as well because performance and perception tend to deviate. Thus, the use of physiological parameters to assess mental fatigue seems as a promising approach (Hancock, Meshkati, & Robertson 1985). The time between fixating and grasping is the time required for cognitive processing of the information for controlling the body's motor functions. This time span is called eye-hand-latency (EHL, Pelz und Canosa 2001). The EHL in milliseconds (ms) was tested as an indicator for mental fatigue in simple manual tasks (Klippert 2014a, 2014b). Both studies demonstrate that there is a fatigue-related pattern of the EHL over time. At the beginning of an eight hour working shift the EHL is shorter than the time at the end of the shift. In both cases, the tasks studied were conducted in u-shaped production cells with several machines served by one operator. The production cells were organized following the chaku-chaku approach, i. e. the operator's task was to step from machine to machine and to load and unload objects which have been manufactured automatically by the machines. The two working systems in the studies (Klippert 2014a and 2014b) require different actions to be performed. In one case the work consisted of simply to loading and unloading objects. If a technical disturbance occurs during work, it was prohibited to take appropriate action to eliminate the interference by the operator himself. In the other case, the elimination of technical disturbances was explicitly defined as part of the operator's task.

For both cases a t-test for independent samples was conducted to compare the EHL values for the beginning and the end of the shift. In the work system with very simple tasks, the test confirmed an effect on α-level of 0.1 (p=0,09) that the EHLs at the end are longer than in the beginning. Cohen's d (0,35) and the Bravais-Pearson Correlation (Effect r = 0,19) confirm a slight effect (Klippert 2014a). For the second case the t-test is significant on α-level of 0.05 (p=0,011). The Bravais-Pearson Correlation (Effect r = 0,28) confirms a slight to medium sized effect and Cohen's d (0,59) confirms a medium sized effect (Klippert 2014b).

The difference in the effect strength may be caused by the differences in the work content. The first task leaves the operators very little freedom for action which means that there is a very low variation in task performance for the operator. The operator is forced to follow the machines step by step and to load and unload the objects. The person performing the task does not necessarily need to know anything about the production process and the technology at all. In the case of a technical disturbance the operator has to call a member the maintenance staff who will eliminate the interference. Afterwards the process continues. The second task is more diverse compared to the first one. The elimination of disturbances is part of the work task and the operators had the appropriate qualification to do so. Therefore this additional task may contribute to task diversity which may be the cause of the differences in effect strength measured in the two cases. However, two studies are insufficient to verify this conclusion and we need to study far more cases.

In order to gain a better understanding of the effects, the concepts used in the study have to be differentiated further. This is the difference between states of mental fatigue and fatigue-like-states (Hacker & Richter, 1984).

4. Emotions and cognition, time-bound effects in Eye-tracking

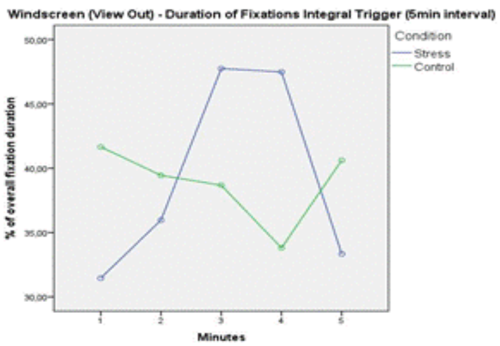
The impact of affective states on behavioral responses has been rather marginally investigated. The reason for this may lay in lack of appropriate methods and the inherent difficulty of modeling and integrating the potential impact of affect-related factors on – cognitive – performance but also in the normative view of the role of affective and emotional states in the behavior of rational actors in the working context. Affects are considered as mostly negative interferences at the cost of rational decision-making and behavior, and hence are treated as unwelcome and unacceptable behavioral by-products that should not be there at the first place.

Despite this persistent view, affective phenomena constitute integral features of everyday life and everyday work (Hollnagel, 2002; Sträter, 2005). As such they should not be treated in a normative but rather in a pragmatic, descriptive manner. Research should focus on the assessment and analysis of the specific ways that such phenomena drive behavior especially in safety-critical situations.

Athanassiou (in press) addresses this issue within the scope of a PhD activity which was driven by the assumption that cognitive and affective functions are tightly interrelated and underlie similar principles of activation that drive information processing and behavioral adaptations but additional overly affective load may induce distinct – observable – adaptation responses (Sträter, 2005). This fundamental research question was tested empirically in the context of simulated ship management tasks.

The empirical study was conducted in the full-mission ship handling simulator facility at Jade University of Applied Studies. Two task-based scenarios were constructed, one designed to employ situation-induced affect stemming from task occurrences (triggers). The trigger considered here was associated with immediate operational challenges (crossing situation with another vessel in shallow and narrow waters and restricted available time for evasive measures). The behavior of the traffic “opponent” in the affective treatment condition was characterized by uncooperative and face-loss communication.

A mixed 2x5 factorial design was employed with one between factor (affective treatment with two levels) and one within factor (time). N = 13 students of nautical studies participated in the study. The visual attention has been operationalized through measures of fixation duration and fixation dispersion towards crucial information sources on bridge, and overall fixation shifts for five minutes immediately after the onset of the affect-inducing trigger event. The results concerning the view out / windscreen (due to the criticality of the information source for unexpected situations) and the aspect of overall fixation shift number will be briefly showcased. The results of the multivariate analysis of variance (Pillai’s Trace) showed a significant interaction between affective conspicuousness of the event and duration of fixation over time (F (4, 8) = 4.581, p = .032, partial η^2 = .696) in terms of fixation duration rates towards the view out of the windscreen. The result suggests that individuals made distinct use of the information source when faced with explicit affective load in conditions of severe time constraints. Visual information search and consultation were directed more intensively towards the outside world and not its representations (i.e. Radar screen) during the 3rd and 4th minutes after the trigger onset (when the situation developed in a manner that posed severe threats for vessel safety). Lower rates for the rest of the time suggest an increased ability to distribute visual attention towards available information in a less fixated manner. The behavior was inversed (and somehow with more moderate “deflections” during time interval) for individuals facing a similar situation with only the respective cognitive load, hence suggesting differences in necessary control efforts in order to adjust to the situational demands due to the affect inducement.



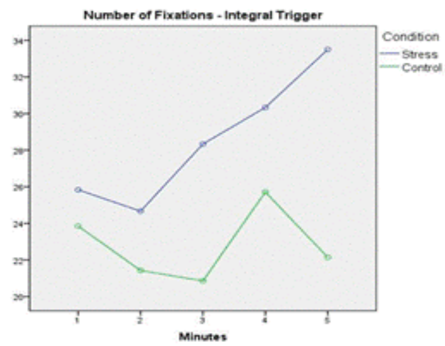


Figure 2. *Mean values for ´fixation duration rates towards the view out (left) and number of fixations towards information sources on the bridge (right) within the five minutes interval at the encounter with the affect-inducing trigger.*

In addition, the mixed ANOVA – nearly significant – result of the overall fixation shifts (F (1, 11) = 3.878, p = .075, partial η^2 = .261) suggested an increased need for fixation shifts. The derived visual responses imply different control and adaptation efforts as a result of the specific impact of additional situation-induced affective responses during dynamic time-bound task performance.

5. Discussion

This article demonstrated how the ordination technique NMDS can be used to capture dynamics of eye-tracking behavior in aviation and provide the basis for assessing cognitive coping strategies when facing adverse conditions (traffic warnings/engine failure). Furthermore, the analysis of eye-hand latency for assembly tasks demonstrated how fatigue-related effects can be identified in assembly tasks in field settings. Finally, the maritime study showed how eye-tracking performance can be linked to affect-induced cognitive states. When uncovered this information can serve as a basis for

- Fitting the work place to coping strategies of pilots. The different strategies should be accepted as plausible courses of action and supported rather than eliminated.
- Task redesign to counter the effects of mental fatigue over time in assembly tasks
- Rule out the effect of affect-induced activation in maritime transport in order to ensure that the work system stays resilient even when facing emotionally challenging situations

6. Literature

Arenius, M. (2014). Function Allocation in Human-Automation Interaction by means of Dynamic Performance Data. In GFA (Ed.), *Proceedings of the 60. GFA-Conference: Gestaltung der Arbeitswelt der Zukunft*. Dortmund: GFA Press.

Arenius, M., Schneider, M., & Straeter, O. (2014). Evaluation of a display for airborne separation in safety-critical situations. Gehalten auf der 6th International Conference on Research in Air Transportation, Istanbul, Turkey.

Athanassiou, G. (in press) Mariner's adaptive behaviour under stress: Effects of workload and affective arousal on the cognitive behaviour of novice bridge officers of the watch and on safety performance of bridge teams during a complex ship management task. PhD Thesis. University of Kassel

Hancock, P. A., Meshkati, N., & Robertson, M. M. (1985). Physiological reflections of mental workload. Aviation, Space, and Environmental Medicine, 56, 1110-1114.

Hacker, W. & Richter, P. (1984). Psychische Fehlbeanspruchung: psychische Ermüdung, Monotonie, Sättigung und Stress. Band 2. Berlin, Heidelberg, New York, Tokio: Springer.

Hollnagel, E. (2002). Time and time again. Theoretical Issues in Ergonomics Science, 3(2), 143–158.

Hollnagel, E. (2009). *The ETTO principle: efficiency-thoroughness trade-off: why things that go right sometimes go wrong*. Surrey: Ashgate Publishing, Ltd.

Klippert, J. (2014a). Blickbewegungen als Indikator für psychische Beanspruchung? Zur Bedeutung der Wahrnehmungs-Handlungs-Kopplung am Beispiel manueller Tätigkeiten in der Industrie. Paper presented at the Frühjahrskonferenz 2014, München.

Klippert, J. (2014b). Wahrnehmungs-Handlungs-Kopplung bei teilautomatisierten Montagetätigkeiten Paper presented at the Psychologie der Arbeitssicherheit und Gesundheit, Dresden.

Land, M. F., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. Perception, 28, 1311-1328.

Straeter, O., & Arenius, M. (2014). Assessing the Complex Interaction and Variations in Human Performance using Non-metrical Scaling Methods. In S. Kadry & A. El Hami (Eds.), *Numerical Methods for Reliability and Safety Assessment - Multiscale and Multiphysics Systems* (pp. 579–598). Springer Press.

Sträter, O. (2005). Cognition and safety: An integrated approach to systems design and assessment. Aldershot: Ashgate.