Normal work and drifting systems – Using dynamic performance measurements for uncovering performance variability in complex systems

Marcus ARENIUS, Oliver STRÄTER

FG Arbeits- und Organisationspsychologie, Universität Kassel
Heinrich-Plett Str. 40, D-34132 Kassel

Abstract. Modern organizations have grown increasingly complex. This development can partly be traced back to the technological developments that have taken place, particularly the widespread application of automation to increase efficiency and reliability. This has in turn led to systems that are more interconnected and centralized, where human operators control large subsystems and where errors may lead to severe or catastrophic outcomes. This in turn has led to higher demands on reliable performance for the remaining human contribution (e.g. maintenance, oversight, human automation cooperation). Recent research has indicated that in order to understand breakdowns in performance in complex systems, we have to shift from “hunting broken components” to identifying critical interactions that contributing to events and accident causation on a system level. This can be accomplished by understanding system drift, where individually correct behavior from operators can combine in a way that produces a net negative effect on system safety. In order to understand the human contribution to system drift, it is necessary to capture individual variations in human performance that normally would go unnoticed and to assess how these variations relate to the way that operators define their work (functions of work). This will lead to a better understanding in how to model how individual variations in performance may sum up in a way that is detrimental for system safety. Thus, this paper will show how eye-tracking data from professionals in aviation, railway transportation and surface mining can be used to assess individual variations in performance and how to relate these variations to the functions of work defined by the respective human operators.

Keywords. human factors, human reliability assessment, task modeling, eye-tracking, functions of work

1. Performance Variability and Drifting Systems

Models of human work are one of the fundamental building blocks in human reliability assessment (HRA) and therefore constitute an important step in the safety assessment of any system. This holds especially true for complex work systems, where task models are needed in order to get overview of actual system functioning.

Traditionally task analysis is conducted by using procedures as input for defining the steps that the human operators have to go though in order to conduct their work. This has approach has been devised as "work-as-imagined" as humans usually deviate in from the procedures in various ways in their day to day work. Therefore, task-based HRA-methods are usually not capable of capturing or representing performance variability, as they mainly draw on procedures as a basis for describing how humans conduct their work (Sträter, Dolezal, Arenius, & Athanassiou, 2012).This performance variability of human operators is vital for keeping the system within safe operating limits. However in rare cases, performance variability may combine in such a way that it leads to adverse outcomes. This gradual transition of system behavior towards the boundaries of safe performance has been termed “drift” and originates from the propagation of performance variability in a given system (Dekker, 2011). Thus, in order to understand drift, it its necessary to understand everyday variability of work.

2. Functions of Work-As-Done

Functions describe the way in which work is actually carried out. The difference between a function and a task is that a function describes an activity as it is done in daily work (work-as-done) and not in terms of the relevant procedures or rules (work-as-imagined). One way to define functions is to use the knowledge of SMEs (subject matter experts) on how they actually conduct their work. This however can be a tedious process as “[…] functions do not work or fail, but rather vary in how they are carried out (Hollnagel, 2012, p. 53)”. That is, expert may not agree on what typical features of aspects of function are, e.g. when a function typically starts or ends (for a full description of the six aspects of a function see (Hollnagel, 2012). Some experts may conduct the function continuously other may switch the function with other functions in between or interrupt the function completely to conduct more urgent work, etc.. This does not mean that there is a "right" or "wrong" way in which a function is conducted, rather it shows that there may be considerable individual differences in how they are carried out. This is especially true for the "time" aspect of a function, that is when it starts, when it ends and whether it is done continuously or not.

The recent advancement in biometric technology seems as a promising approach for capturing these individual differences in performance. Eye-tracking has been used in a variety of domains to address or improve human factors related issues. Many of these studies have focused on different aspects of cognition of humans conducting tasks in real-life or simulated settings.

Eye-tracking has been used to assess dynamic attention allocation when searching for a target stimuli and for the assessment of monitoring strategies of pilots in the aviation domain (Günebak, Allgaier, & Sträter, 2010). Furthermore, eye-tracking has also been used in order to fit the workspace to the attention requirements of the human operators in different settings (Arenius & Buch, 2012) In safety research, studies have demonstrated that eye-tracking data can be used for cognitive task analysis (Kurland, Gertner, Bartee, Chisholm, & Mcquade, 2006; Seagull & Xiao, 2001).

The common denominator of these studies is the assumption that the eye-tracking data of human performance can be meaningfully analyzed and set into relation to performance. Thus, the high temporal resolution of eye-tracking data should allow for the identification of these spatio-temporal patterns that in turn can be used to determine e.g. onset of a function, gradual transition into other functions and the cyclical nature of the functions.

3. Method and Eye-tracking and Approach for Data Analysis

In order to illustrate the approach for data analysis, a dataset from aviation will be described in detail and used as example for the analysis steps. The same approach is applied to data from other domains (railway and surface mining).

3.1 Aviation: Flight Scenario

A total of 10 male employed pilots participated in the study (n=10). Other demographic data could not be collected due to privacy restrictions of the participants. The scenario is a flight route of the standard training program by an airline and it is flown in the full-scope simulator that is used by the airline for regular checks of pilot performance. Thus, the pilots are familiar with the route and the simulator. The pilots start on route from Athens to Heraklion. After ~60 seconds, the pilot receive a traffic warning (TW) followed by an Engine Failure at a specific location on route. The location were the pilots receive the traffic warning is known for high traffic volume and therefore a likely location for actual occurrence of traffic warnings. Following the traffic warning, the pilot may chose one of three options or ignore the display completely while assessing and handling the Engine Failure. The scenario ends when the aircraft is under control and in a safe state. During the flight, the eye-tracking data of the pilots was recorded with a portable eye-tracker. The main Areas of Interest for the study were the Primary Flight Display (PFD), the Navigation Display (NAV), the Engine and Warning Display (E/WD) and the Flight Back were the ATC-Display was located (Fig. 1).

Figure 1. *The different Areas of Interest for the scenario traffic warning and engine failure. The ATC-display was placed at the location of the Flight Back* (Schneider, 2014).

3.2 Approach to data analysis

In order to identify individual variations in human performance over time, it is necessary to analyze the dynamics eye-tracking data:

- task identification (average beginning/end),
- gradual transition into other task (duration/intensity shift),
- Rhythm of tasks: anti-cyclicality and reoccurrence
- Relative importance and timing of parallel tasks

This can be done by plotting the cumulative durations of the gaze in percent on a specific area of interest (e.g. the primary flight display, PFD). The correlation matrix of the different areas of interest represents the dependence relationship between the different areas of interest (AoI). However, in order to correctly assess the correlations between the different areas of interest, the time-based effect of the gaze distribution

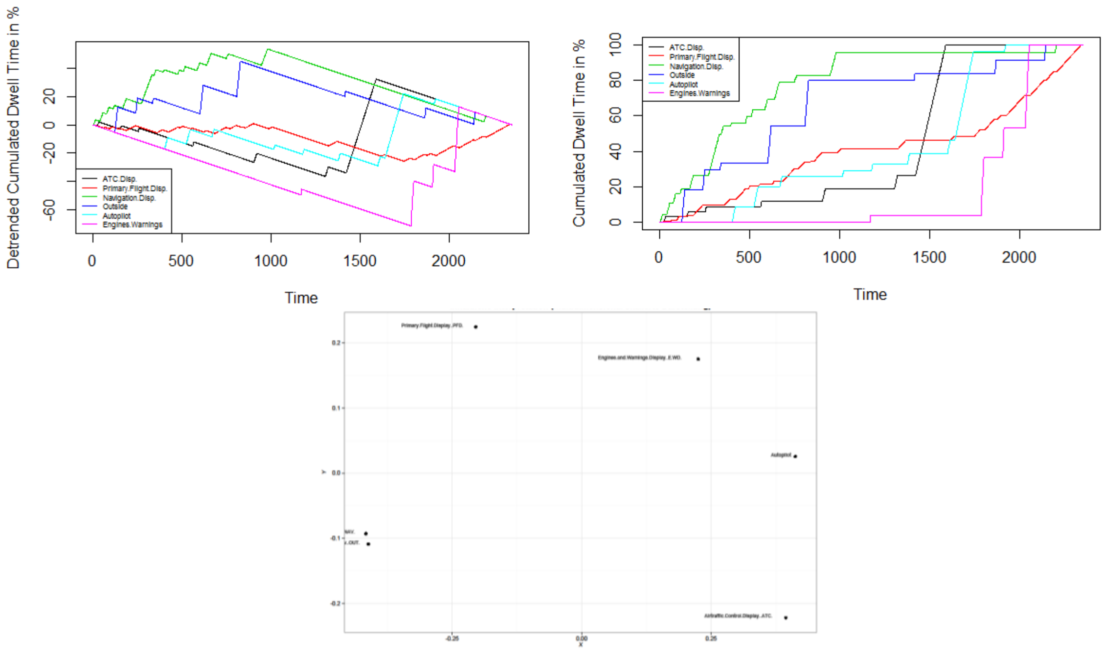

Figure 2. *Removing time-based effects of cumulative time series of fixations (from upper left to upper right). Serves as basis for calculating the MDS-Representation of the correlation matrix of the eye-tracking data (lower side). Proximity in the MDS representation indicates that the displays are highly interdependent [Note: MDS-representation is dimensionless]*

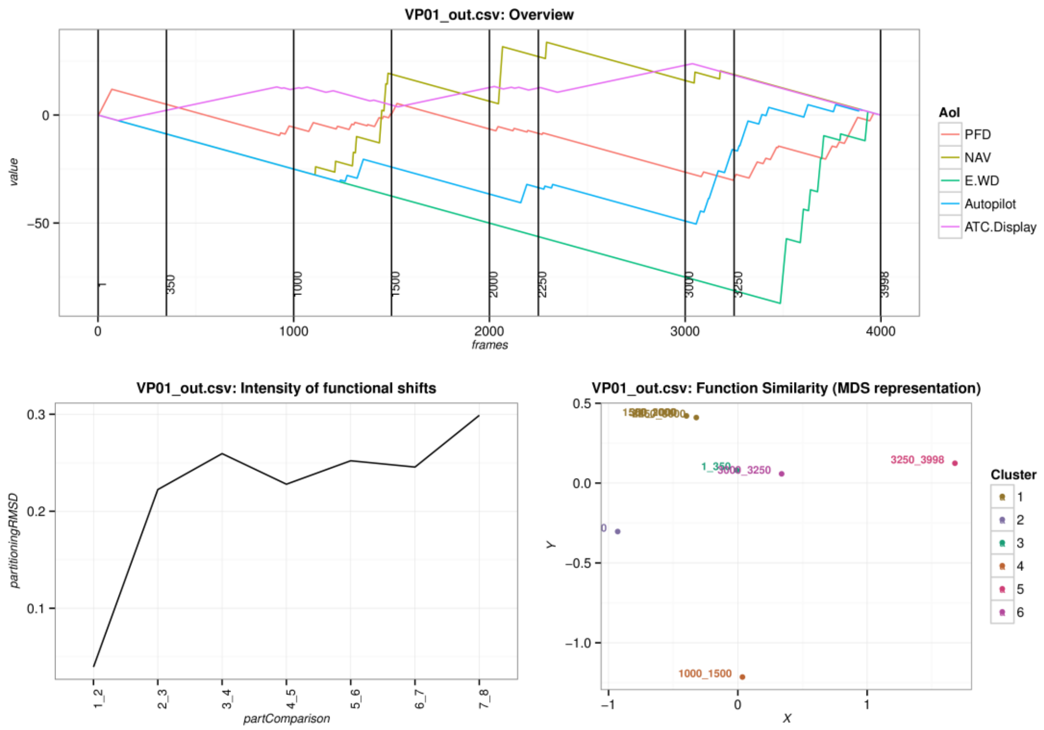

Figure 3. *The best partitioning of aviation eye-tracking data that maximizes the difference between the parts. The upper figure shows different functions as partitions of the eye-tracking data; the difference between the functions in terms of the root mean squared deviation (RMDS) of the MDS representations (lower left) and finally the degree of similarity of the different functions (clustering of the partitions obtained when comparing the MDS representations with each other, lower right).*

has to be compensated by aligning the line through origin over the x-axis, otherwise only positive correlations are obtained (ref. fig. 2). The correlation matrix between the different areas of interest can then serve as input for determining the MDS-representation, which in turn shows the interdependence of the gazes on the areas of interest (fig. 2).

However, since the functions of an human operator may vary during an eye-tracking session, there has to be a systematic approach for identifying the relevant parts of the eye-tracking data that should be converted into an MDS-representation. This can be done by maximizing the difference between the correlation matrices of the eye-tracking data parts.

4. Results: Functions in various technical domains

Fig 3 shows the best partitioning for the eye-tracking data of one pilot in the aviation study. For every part of the eye-tracking data a correlation matrix has been computed and converted into an MDS-representation (upper side of fig. 3). These MDS-representations have in turn been compared with each other to determine their degree of similarity. This similarity measure as been used to determine the different functions that the eye-tracking data parts should be associated with (clusters, lower right, fig 3).

Fig. 4 shows the difference in results obtained for task from other technical domains, especially since there are more relevant parts in the eye-tracking data (operation of bucket excavator and train driving respectively).

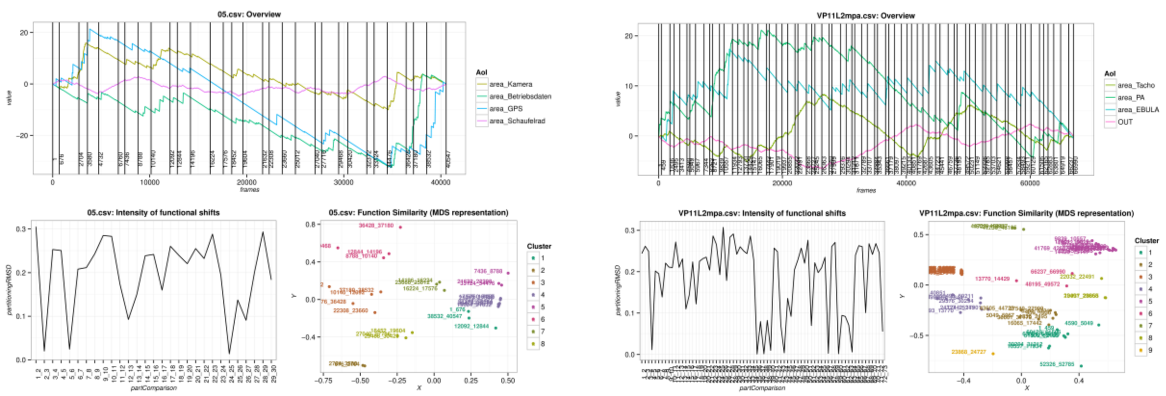

__Figure 4.__ Demonstration of the application of the approach for eye-tracking data from an operator of bucket excavator in surface mining (left-hand side) and of a train driver in a full scope simulator (right-hand side). There are more partitions that maximizes the difference between the parts for both, however, the data clusters well indicating a satisfactory basis for the identification of functions.

5. Discussion and next steps

In order to satisfy the requirements for modern human reliability assessment (HRA), functions of everyday work have to be defined. Usually, functions may vary in the way that they are implemented by the human operators.

Arenius et al (in press) demonstrated how the ordination technique (non-)metrical multidimensional scaling (MDS) can be used to capture dynamics and individual variations of eye-tracking behavior in aviation. However, since the partitioning of the eye-tracking data was done based on the onset of a specific event (engine failure during flight in a full scope simulator) there has to be a systematic approach for finding the relevant partitions of the eye-tracking data if these events are not known beforehand. This is done by identifying the partitions that maximize the difference between the correlation matrices of gazes on the areas of interest, which serve as the basis for calculating the MDS-representation. Furthermore, the necessary steps for converting the eye-tracking data into a format suitable for this partitioning approach have been shown. Results from aviation, railway and mining demonstrated how the MDS representations from the partitioned eye-tracking data can be used to identify patters of variability within an eye-tracking recording session in various technical domain.

This can serve as a basis for the identification of functions of the different activities, which the operators define during the conduct of their daily work and aid the data-driven definition of models of work-as-done, which are a necessary requirement for addressing current issues in human factors and human reliability assessment (HRA).

Several issues have to be addressed in future analysis:

- The representations of individual patterns of variability have to be rendered comparable to other individuals
- A suitable and intuitively understandable representation of the patterns of variability has to be developed
- The identified patters of variability have to be compared against known patterns of variability, to determine the feasibility of the results

6. References

Arenius, M., & Buch, M. (2012). Ergonomische Optimierung des Großgeräte-Führerstandes im Braunkohlen-Tagebau. In *GFA (Eds.): Gestaltung nachhaltiger Arbeitssysteme* (pp. 101–104). Dortmund: GFA-Press.

Dekker, S. (2011). *Drift Into Failure: From Hunting Broken Components to Understanding Complex Systems*. Farnham: Ashgate Publishing Co.

Günebak, S., Allgaier, S., & Sträter, O. (2010). Systemgestaltung durch Kombination mentaler Belastungsanalyse und Blickbewegungsaufzeichnung. In *GFA (Eds.): Neue Arbeits- und Lebenswelten gestalten* (Vol. 56, pp. 627–630). Darmstadt: GFA Press.

Hollnagel, E. (2012). *FRAM: The Functional Resonance Analysis Method: Modelling Complex Socio-technical Systems*. Farnham: Ashgate Publishing, Ltd.

Kurland, L., Gertner, A., Bartee, T., Chisholm, M., & Mcquade, S. (2006). *Using Cognitive Task Analysis and Eye Tracking to Understand Imagery Analysis*. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA456286

Schneider, M. (2014). *Entwicklung eines Zusatzinformationssystems für Piloten und dessen Untersuchung hinsichtlich Auswirkungen auf das Verhalten von Piloten in kritischen Flugsituationen [Development and Evaluation of a Display for Airborne Separation]* (Master Thesis). University of Kassel, Department of Human and Organisational Engineering & Systemic Design, Kassel.

Seagull, F. J., & Xiao, Y. (2001). Using eye-tracking video data to augment knowledge elicitation in cognitive task analysis. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 45, pp. 400–403).

Sträter, O., Dolezal, R., Arenius, A., & Athanassiou, G. (2012). Status and Needs on Human Reliability Assessment of Complex Systems. *SRESA Journal of Life Cycle Reliability and Safety Engineering*, *1*(1), 22–43.

Acknowledgement. This work is part of an phd-activity sponsored by the Eurocontrol HALA-network.