TECHNISCHE UNIVERSITÄT
CHEMNITZ

Designing a workplace in the aviation domain: The transition to a remote air traffic control workplace by analysing the human-computer interaction.

von der Fakultät für Human- und Sozialwissenschaften der Technischen Universität Chemnitz
genehmigte Dissertation zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt von: Maik Friedrich, geboren am 08.11.1981 in Schmalkalden

Gutachter: Prof. Dr. Joseph F. Krems, Technische Universität Chemnitz

Prof. Dr. Mark Vollrath, Technische Universität Braunschweig

Tag der Verteidigung: 20. Juni 2019

(http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa2-341604)

Zusammenfassung

In der heutigen Zeit ist die effiziente Nutzung aller verfügbaren Ressourcen von zentralem Interesse. Im Bereich der Flugführung hat die Fernüberwachung von Flughäfen aus diesem Grund über die letzten 10 Jahre immer mehr Bedeutung erlangt. Die größten Vorteile der Fernüberwachung liegen in der geringeren Abhängigkeit von Flughafengebäuden und deren Instandhaltung, einer vereinfachten Personalplanung (vor allem bei kleinen Flughäfen) sowie dem möglichen Hinzufügen von zusätzlichen Informationen beim Arbeitsplatz zur Fernüberwachung. Insbesondere das Designen eines Arbeitsplatzes zur Fernüberwachung hat in diesem Zusammenhang eine Schlüsselrolle eingenommen. Die größte Herausforderung bei dieser Umsetzung ist ein Mensch-Computer-Interaktionsmodell, das die Verlagerung der Arbeitsplätze unterstützt, indem es die Einflüsse auf den Operateur und dessen Aufgaben beschreibt.

Die vorliegende Dissertation fokussiert sich auf die Anwendung und Verbesserung eines Mensch-Computer-Interfacemodells zur Umwandlung eines Arbeitsplatzinterfaces ohne Beeinflussung der Aufgabe des Operateurs. Das präsentierte Modell konzentriert sich auf den Informationsfluss am Arbeitsplatz, anstatt technisch machbare Konzepte zu präsentieren. Es besteht aus 3 Teilen, welche sich separat mit dem Einfluss des veränderten Interfaces auf den Informationsfluss auseinandersetzen. Das Modell wird für die Thematik der Fernüberwachung spezifiziert und angewendet. Nur wenige Publikationen beschäftigten sich bisher mit Strategien, mit denen Towerlotsen Flugführung durchführen. Daher ist das Ziel dieser Arbeit, einen soliden Beitrag zur Entwicklung der Psychologie im Bereich Mensch-Computer-Interaktion zu leisten, welcher durch praktische Anwendungen und Erweiterungen der Methodik untermauert wird.

Der Hauptunterschied zwischen dem konventionellen Arbeitsplatz und dem Arbeitsplatz zur Fernüberwachung ist der Verlust der Außensicht und des Fernglases und deren Ersatz durch Kamerasysteme. Neue Systeme in der Flugsicherung können die menschliche Leistung beeinflussen. Daher hat die Fernüberwachung besonderen Einfluss auf die Bereiche Ausrüstung des Arbeitsplatzes, das Benutzerinterface und menschliche Leistungsfähigkeit. Die Herausforderung für die Ausrüstung des Arbeitsplatzes besteht darin, Informationen zu identifizieren, welche durch die Fernüberwachung reduziert wurden und diese durch zusätzliche Informationensysteme bzw. Assistenzsysteme zu ergänzen. Für den Faktor Benutzerinterface ist das Ermitteln und Analysieren von dynamischen Informationen besonders wichtig. Für die menschliche Leistungsfähigkeit besteht hier die Frage, wie sich Arbeitslast und Situationsbewusstsein kombiniert auf die Leistung auswirken und welche Konsequenzen das auf die Arbeit an einem Arbeitsplatz zur Fernüberwachung hat.

Alle Herausforderungen wurden im Detail analysiert. Für den Faktor Ausrüstung des Arbeitsplatzes zeigen zwei Analysen die große Vielzahl an Indikatoren, welche verwendet werden können, um die Veränderung des Informationsflusses zu bestimmen. Die detaillierte Analyse des Windsack Indikators liefert ein Beispiel, wie die verschiedenen Indikatoren angewendet werden können. Die zweite

Analyse zeigt, wie die bestehenden Indikatoren zur Lotsenaufgabe um spezielle Wetterindikatoren erweitert werden, um den Aspekt der Überwachung des Luftraums vollständig abzudecken. Für den Faktor Benutzerinterface wurde eine besondere Blickanalyse mit dem Namen Integration Guideline for Dynamic Areas of Interest (IGDAI) entwickelt. Diese erlaubt, die dynamischen Informationen innerhalb des Interfaces eines Arbeitsplatzes zu analysieren. Sie wird auf den Arbeitsplatz zur Fernüberwachung angewendet. Für den Faktor menschliche Leistungsfähigkeit zeigt eine detaillierte Analyse, wie Arbeitslast und Situationsbewusstsein die Leistung bei niedriger und hoher Aufgabenlast beeinflussen. Durch das Anwenden von IGDAI konnten zwei Kontrollstrategien in Abhängigkeit zur Aufgabenlast identifiziert werden.

Das bereitgestellte Model für die Veränderung des Interfaces ohne Beeinflussung der Operatoraufgabe stellt einen Sonderfall in Bereich der Mensch-Computer-Interaktion dar. Der Übergang vom konventionellen zum Fernüberwachungsarbeitsplatz ist ein sich immer noch fortsetzender Prozess. Weitere Entwicklungen im Bereich der Fernüberwachung von Flughäfen sind notwendig, um den zukünftigen Herausforderungen an die Flugführung zu begegnen. Deshalb stellen die in dieser Dissertation dargestellten Konstrukte, erarbeiteten Methoden sowie Ergebnisse eine solide Basis für zukünftige Forschungsarbeiten bereit.

Summary

The efficient usage of all available resources is a central interest of our time. In air traffic management, the topic of remote tower operations has increased in importance over the last 10 years. Herein, the design of a remote tower workplace plays a key role in the successful implementation of remote tower operations. Less dependency on building and maintaining airport control towers, an improved human research planning (especially for small airports) and an increase in available information to the conventional tower workplace are central advantages of remote tower operations. However, a potential challenge for this approach is an HCI model that supports the transition by describing the influence on the operator task.

This dissertation focuses on the application and improvement of an HCI approach to redesign a workplace by changing the interface without influencing the task of the operator. The presented model focuses on the flow of information rather than the presentation of technical possibilities. It consists of three parts that each individually measure and analyse the influence that a redesigned interface has on the flow in information. This model is specified and applied to remote tower operations. Prior to this dissertation, there were only a few publications connected to the strategies that air traffic control officers (ATCO) in the tower use to control traffic and virtually no publications connected to the practical implications for working at a remote tower workplace. Therefore, the goal was to provide a well-founded contribution to the development of psychology in the area of human-computer interaction by applying the psychological theories and extension of the methodology.

The main difference between the conventional and the remote tower workplaces is the replacement of out-the-window view and binoculars by camera systems. Based on what influences the human performance in connection with new systems developed in air traffic control, most changes afflict the general workstation and equipment, the user interface, and human resource management. The challenge for the factor workstation and equipment is to identify the information decrease at the remote tower workplace and its replacement with additional information whilst simultaneously ensuring that this information can be tested in a standardised manor throughout a variety of research projects and several different prototypes. The challenge for the factor user interface was the analysis of the dynamic information presented at the remote tower workplace. The challenge for the human resource management is to identify how workload and situation awareness influence performance.

In sum, all challenges are analysed in detail. For the factor workstation and equipment, two analyses showed a large variety of indicators that are applicable to evaluate the difference in the flow of information between the conventional and the remote tower workplace. The first analysis of the windsock indicator provided an example of how the different metrics can be applied. The second analysis showed that the weather remote tower metrics extend the existing remote tower metrics and thereby complete the aspects of the monitoring that an ATCO has to perform. For the factor user

interface an advanced gaze analysis, called Integration Guideline for Dynamic Areas of Interest (IGDAI) was developed. This allows for a detailed analysis of the dynamic information presented at the remote tower workplace. For the factor human resource management, a detailed analysis shows how situation awareness and workload influence performance within low and high task load phases. By applying IGDAI, the existence of two control strategies for the Air Traffic Control (ATC) environment that are each related to the task load phases could be identified as well as the extent to which these might afflict remote tower operations.

The provided model of redesigning only the interface presents a detailed approach for a special case in HCI. The transition from the conventional to the remote tower operations is an ongoing process that will be continued. The development in the domain of remote tower operations seems to be stable and necessary to keep up with the challenges of future air traffic management. Therefore, the analysed constructs, developed methodologies and presented results from this dissertation provide a seminal basis for the necessary future research.

Table of Contents

I Synopsis

1 Introduction

The process of developing and implementing workplaces always presents a challenging task driven by a variety of motivations. Especially with restricted resources and fixed technical specifications, this process needs to be straightforward. A process like this is more likely to succeed if built on research that focuses on interfaces and interaction. Human-Computer Interaction (HCI) supports different stages of each design and implementation process by combining the engineering of the interface with the psychological theories of cognitive processes and the empirical analysis of user behaviour (Issa & Isaias, 2015). One way of improving HCI is by using the feedback that is provided when applied to practical tasks, e.g. redesigning a workplace (Badham & Ehn, 2000).

Redesigning a workplace in air traffic control is a long-lasting process. Can an airport be remotely controlled from a different location than the control tower? With this research question, the ViTo (“Virtual Tower”) project started in 2002. It was the first project ever to address the topic of remote tower operations[1] for an airport. The project started with a survey analysing the state-of-the-art virtual reality technology in Europe and the United States of America and thereby shaping an initial concept (Fürstenau, Werther, Mittendorf, & Schmidt, 2004). During the following 15 years, knowledge about the procedures and the information gathering of an air traffic control officer (ATCO) was increased by several research initiatives. This moved remote tower research into the focus of HCI.

When developing an HCI model for redesigning a workplace, the following questions in connection with the conventional tower workplace arise. How can the influence of a remote workplace on the ATCO be measured? How can the process of information gathering be analysed in detail? Which factors influence the procedures and tasks of normal ATCO operations? These are a few essential research questions that have to be answered within the process of transferring the ATCO to a remote tower workplace. For safety reasons, the redesign of the workplace has to lead to equal or better performance of the operator and can therefore be seen as a special case in HCI.

Prior to this dissertation, there were only a few publications connected to the strategies that ATCOs in the tower use to control traffic and virtually no publications connected to the practical implications for working at a remote tower workplace. Therefore, the goal was to provide a well-founded HCI contribution to redesigning a workplace in the area of air traffic control (ATC) by applying the psychological theories and extension of the methodology. At the same time, approaches have been developed which allow the ATCOs to transition from the conventional tower workplace to the remote workplace in the most secure and efficient manner possible. Participants at a large-scale validation at

[1] In the context of this dissertation, remote tower is defined as transferring an air traffic control officer at the control tower of an airport to a distant workplace. For a detailed definition see section **Fehler! Verweisquelle onnte nicht gefunden werden.**.

the Erfurt-Weimar airport (Friedrich, 2016) made this possible. The validation concentrated on the feasibility of remote tower operations.

With the finalisation of this dissertation, three first authorship articles will have been published in peer-reviewed journals on the subject of "Remote tower operations and eye tracking". These form the basis for this dissertation. Furthermore, four conference contributions (chapter IV) and one book chapter (Friedrich, 2016) have also been published.

This synopsis focuses on the presentation of the research framework (see chapter 2), summarises the research in the three articles and connects them to each other. The individual articles describe in detail the research hypothesis, the methodical developments, and analyses. Therefore, chapters 3 and 4 of this synopsis only provide a short overview and refer to sections within the articles. The following three articles are appended to this synopsis:

Article 1: How to Evaluate Remote Tower Metrics in Connection to Weather Observations. An Extension of the Existing Metrics

Article 2: A Guideline for Integrating Dynamic Areas of Interests in Existing Set-up for Capturing Eye Movement: Looking at Moving Aircraft

Article 3: The Influence of Task Load on Situation Awareness and Control Strategy in the ATC Tower Environment.

2 Research framework and goals

This section provides the definition of the most important terms used in this dissertation, followed by an overview of how this dissertation is connected to the specific research area of HCI and how it is contributing to the development of this area. The focus of this dissertation is its contribution to psychological concepts and methods in HCI, applied to the specific topic of remote tower operations and its emerging issues.

2.1 Human-computer interaction

Since the mid-1980s, the research focus of workplaces connected to computers has been combined as HCI. In the early stages of HCI, psychology was seen as one of the key factors for this new field (Newell & Card, 1985). HCI "is a discipline concerned with the design, evaluation, and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them" (Preece et al., 1994, p. 7). HCI includes all processes in connection with interfaces and interaction. Head (1999, p. 4) defines interface as the "visible piece of a system that a user sees or hears or touches". Interaction combines possible operator inputs into the system, e.g. by mouse click, keyboard stroke, gesture, eye movement or voice command.

The fast technical advances that drive "an aggressive invention of novel products in HCI and interaction design domains" (Zimmerman, Forlizzi, & Evenson, 2007, p. 500) make it necessary to carefully revaluate the HCI process when applied to a new domain. The application of HCI is an interconnected process between identifying the information that is presented on the interface and which task the operator has or wants to fulfil by interacting with the system. Thereby, the methods for "effectively using input and output techniques so that computers and users have more seamless interactions" (Head, 1999, p. 12) have to be identified. On the one hand, this increases the effort for the design process in a new domain, but on the other hand, through a comprehensive analysis of user behaviour, an interaction model can be defined that allows predictions regarding the usage of a final system.

In HCI, the redesigning of an existing system has the characteristic that the task of the operator and the environment do not change throughout the process, but only the interface and how to interact with it. This task of redesigning is common and often applied for e.g. web pages (Borges, Morales, & Rodríguez, 1996; Dix, 2009; George, 2005). In the air traffic domain, the redesigning of a system is undertaken with great caution due to the safety-critical task and high degree of automation already implemented (Billings, 2018). This dissertation supports the development of a model for the redesigning of an air traffic control workplace.

2.2 Remote Tower Operations

Air traffic control (ATC) involves the monitoring and control of movements in the air or on the ground, with the goal of a safe, organised and efficient handling of traffic (Mensen, 2014).

Kontogiannis and Malakis (2013) characterise the work in ATC as simultaneous, multiple, with goal-driven tasks that are interconnected and have to be accomplished under time pressure. Thereby, errors almost always inflict enormous consequences. Pfeiffer, Valtin, Müller, and Rosenthal (2016) found that the most stressful situations in ATC have a high traffic load, are unexpected events or involve malfunctions of the equipment. ATC is separated into airport control, en-route control, and approach control (Mensen, 2014). This dissertation focuses on remote tower control; a fourth workplace deviated from airport control.

Remote tower operation of an airport is defined as any technical installation with the goal to provide ATC for that particular aerodrome from a distant location. Therefore, the conventional tower workplace is to be replaced by a remote tower workplace. The major technical challenges for the remote tower workplace are the reconstruction of the out-the-window view and an adequate replacement for the binoculars. These challenges are normally met by a high resolution camera system that transmits the images from the airport to a remote tower workplace (Fürstenau, 2016). Since all other ATC systems are generated independently from the control tower, their transmission is not considered an issue. Currently, the reconstruction of the out-the-window view is connected to a substantial technical effort concerning the camera technology and bandwidth for transmission. Looking at the current technical standard in the film industry, an almost perfect reproduction of the out-the-window view would theoretically be possible but would exceed any economical aspect of remote tower operations (Fürstenau, 2016). Therefore, the pivotal HCI question is which information from the out-the-window view is essential for the ATCO and how this information can be transferred and presented to the remote tower workplace.

The initial step for the transition from a conventional tower workplace to a remote tower workplace is to develop almost identical work environments to reduce the stress for an ATCO working in the new environment. In the medium term, the application of HCI is an essential basic, because it also enables applications that were not possible in the conventional tower workplace, for example advanced visual assistance systems within the out-the-window view. In the long term, any new applications to the remote tower workplace need to have a solid research foundation to build upon, especially with the surveillance of multiple remote airports (Papenfuss & Friedrich, 2016; Schneider, Möhlenbrink, & Kiesel, 2010).

Considering the medium and long-term development, a well-defined model for the interface and the interactions required for performing ATC safely and efficiently is needed. The final implementations of the remote tower operations depend on the necessary information and how this is presented to the ATCO. Detailed knowledge of the major influencing factors and information sources for the work of an ATCO (see especially chapter I) is needed for a comprehensive application of the remote tower workplace and future extensions.

2.3 Remote Tower Research

At the beginning of remote tower research, between 2002 and 2005, there was hardly any research on the ATCO in the conventional tower workplace with regard to the replacement of the out-the-window view. The few available studies were associated with the introduction of new assistance systems, e.g. electronic flight strips (MacKay, 1999; Truitt, 2006) or the optimisation of workflows, e.g. on the runway (Anagnostakis, Clarke, Bohme, & Volckers, 2001; Atkins & Brinton, 2002) or as contingency tower (McKeon, 2009). The first research projects concerning remote tower operations focused on general feasibility in connection with technical issues, e.g. camera optics and data transfer. In chronological order, those were ViTo (Virtual Tower) (Fürstenau et al., 2004), RapTOr (Remote Airport Tower Operation Research, 2005-2007) (Schmidt, Rudolph, Werther, & Fürstenau, 2006), and ROT (Remotely Operated Towers, 2006-2008) (Saab Security, 2008)[2]. Later projects, such as RAiCe (Remote Airport Traffic Control Center, 2008-2012) (Fürstenau et al., 2007) or SESAR 6.8.4 (Single European Sky ATM Research Project 6.8.4 2011-2015) used large-scale validation and real-time simulations to look at the impact of remote tower operations on the ATCOs' performance.

The current developments in the field of remote tower operation show a large variety of different technical implementation. For example, three different prototypes (COOPANS, INDRA and Frequentis/DLR) for remote tower operations were developed and evaluated within the SESAR 2020 (Single European Sky ATM Research Joint Undertaking) research programme. The different prototypes were developed within the same research programme but follow different approaches for composition and representation of the remote tower workplaces. The similarities between these prototypes are that they started as technical projects focussing on the camera systems and none has an accurate representation of the out-the-window view. However, as mentioned above, the research projects focused on technical solutions rather than a description of the task or the necessary requirements to perform the task itself.

For an ATCO, the information-gathering through the interface changes when applying a reconstructed out-the-window view as panorama. Available eye tracking studies within the ATC tower are connected to specific systems, e.g. weather display (Ahlstrom, 2005) or radar screen (Bruder, Leuchter, & Urbas, 2003). Only a few studies (Lange, 2014) analyse the influence of the out-the-window view by applying eye tracking as a method for measurement. This can be explained by the cost for such a field study and the extreme effort in connection with the eye data analysis in such an environment (Papenmeier & Huff, 2010). One special issue for eye tracking is the synchronised capturing of eye movement and relevant dynamic information, e.g. positions of aircraft and vehicles at the apron. By solving this issue, a comprehensive analysis of the information-gathering process can be realised and the influence of the out-the-window view evaluated.

[2] For a detailed and chronological summary of the research projects for remote tower operations see Fürstenau (2016)

The detailed evaluation of the information-gathering process allows for insight into the interface requirements of an ATCO work environment. Questions regarding increased task load and reduced situation awareness become relevant, especially with the decrease or change in information-gathering with remote tower operations. Wickens, Mavor, and McGee (1997) define the interconnection between task load and individual performance with control strategy, whereas situation awareness is seen as part of the individual performance. A change in the workplace would either influence the task load or would lead to an adapted control strategy that would then also influence the situation awareness. By accepting this dependency and its effects, the replacement of an interface seems quite simple, as long as the same information is available.

2.4 Embedding into HCI

As mentioned above, from the early years to the recent research on remote tower operations, a predominant part related only to the technical aspects of video transmission and the duplication of the workplace. Because of the following three reasons, this focus on the technical aspects altered in the direction of HCI. Firstly, the technical limitations of video reconstruction could be demonstrated in several studies, e.g. (Ellis, Fürstenau, & Mittendorf, 2011). Secondly, air traffic service providers showed increased interest in participating and accelerating the research as a potential source of savings. And thirdly, through the involvements of the end-user (ATCO) in the research, reluctance concerning remote tower operations and the participation in it could be reduced.

The main research challenge in connection with HCI is how to remotely control an airport without an equally good interface of the out-the-window view. The most important questions are which information is essential, how is the visual representation at the remote tower workplace (Friedrich, 2016), and their influence on remote tower operations. Also, the psychological processes while working at a remote tower workplace and their influence on the decision-making process have comprised a small or non-existent part of recent research. Therefore, a greater focus on the HCI enriches the entire design process by incorporating the end-user performance into the assessment instead of only designing and evaluating a specific technical system (Poulton, 1966). As mentioned above, besides implementing remote tower operations, this dissertation presents a basis for a sustainable development that might enable additional research, e.g. multiple remote tower operations.

This dissertation applies HCI to the field of remote tower operations. The focus lies on the analysis of the conventional tower workplace with regard to the acquisition of information and the evaluation of the remote tower operations in connection with the comparability between both workplaces. A model that describes the interface of the conventional tower with respect to the task could help to unlock the full potential of remote tower operations. This dissertation extends psychology in the field of HCI with the theoretical (section 3) and the methodological development (section 4).

2.5 Research goals of the dissertation

Building on the current research, the following research goals are defined:

1. The first goal, based on the results and in accordance with the transferable theories from the different research areas, was to develop a model that describes the influence that the redesign of a workplace interface only can have on the operator. Therefore, it allows for a hypothesis regarding the implications for the performance of the operator and possible influential factors. The model was developed continuously throughout articles 1 to 3.
2. The second goal was an extension to the methodology for the evaluation of the different technical remote tower prototypes, with regards to operational requirements. Based on the methodology developed by Friedrich (2016), an extension and enhancement is provided by article 1. The focus on the extended methodology was deliberately placed on the effects of the changed HCI. A synthesis of this article can be found in section 4.1.
3. The third goal was to develop a methodology to record the information-gathering process of ATCOs in relation to their tasks. A general method for recording eye tracking data in relation to the environment and the dynamic information, e.g. moving objects, is required as support and preparation for research goal 4. The method is published in article 2 and a short summary is provided in section 4.2.
4. Based on the methodology of research goal 3, the fourth goal was the detailed analysis of the interaction of the ATCOs' with their environment. The focus was on the factors of workload, situation awareness and control strategies in relation to the dynamic environment and the individual performance. The experiment and results are presented in article 3 and a short summary is provided in section 4.3.

The motivation behind these research goals was a well-founded psychological understanding of the conventional tower workplace and the implication for remote tower operations. This should facilitate the transition between the two workplaces and support future developments through a well-grounded approach.

3 The development of a new workplace

3.1 Redesign of a workplace

The user-centred design of a new workplace includes several stages, beginning with the requirements for the task that the operator has to perform, the identification of the necessary information for the interface, and the needed interaction possibilities. Figure 1 presents the primary flow of information from the environment through the workplace to the operator and back. The workplace has to present an abstract representation of the environment allowing the operator to gather all needed information for his/her task. This dissertation focuses on the case where the operator task and the environment stay the same and only the workplace changes.

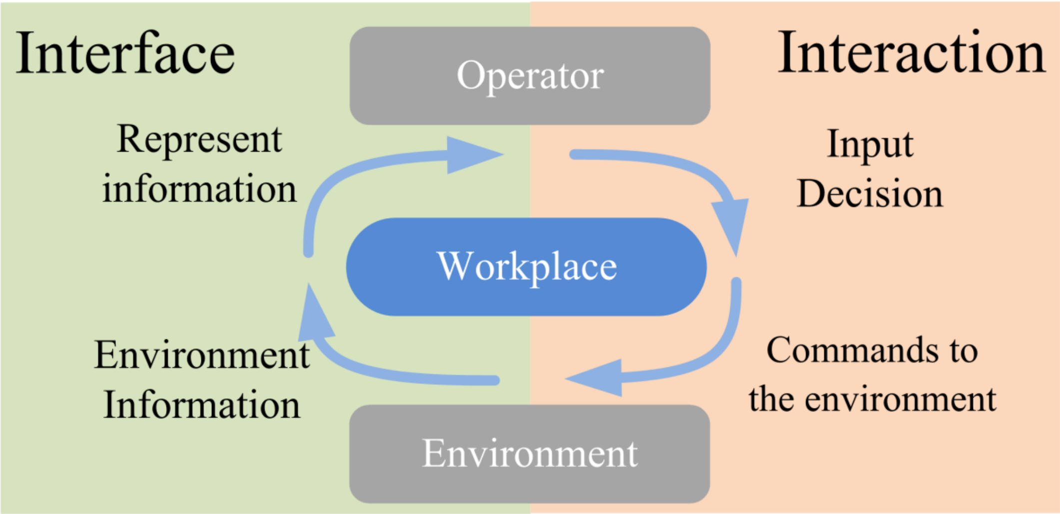

Figure 1. The cycle flow of information from the environment through the workplace to the operator.

The model described in this dissertation investigates the case of redesigning a workplace where only the workplace interface changes. This model supports the process of redesigning in case the operator has to perform the same task with a different interface to the same environment as before. The model concentrates on the redesign of the workplace interface by keeping the provided information equal throughout the process. Even if the system that provides the information changes, this design approach ensures the represented information stay the same, e.g. by suggesting an additional sensor or advanced assistance system.

The model for redesigning only the workplace interface (Goal 1, see 2.4) consists of three parts. Firstly, the design factors influencing each operator in relation to his/her task have to be evaluated. Secondly, based on the evaluated design factors, the influence of the redesign of the interface has to be analysed by selecting or defining metrics that allow measuring of the difference within the flow of information. Thirdly, also based on the evaluated design factors, the influential factors on the operator have to be evaluated to ensure that the redesign does not affect his/her decision process.

3.2 Design factors in Aviation

The design factors influencing the operators in aviation were published by EUROCONTROL and FAA (2015). They provide an overview of factors that mainly influence the human performance in connection with the development of new ATC systems. Figure 2 presents the summary of these twelve factors in a structured manner and separated into the following main areas; human resource management, selection, physical restriction, and safety culture. The application of remote tower operations has a minor to major impact on all twelve factors.

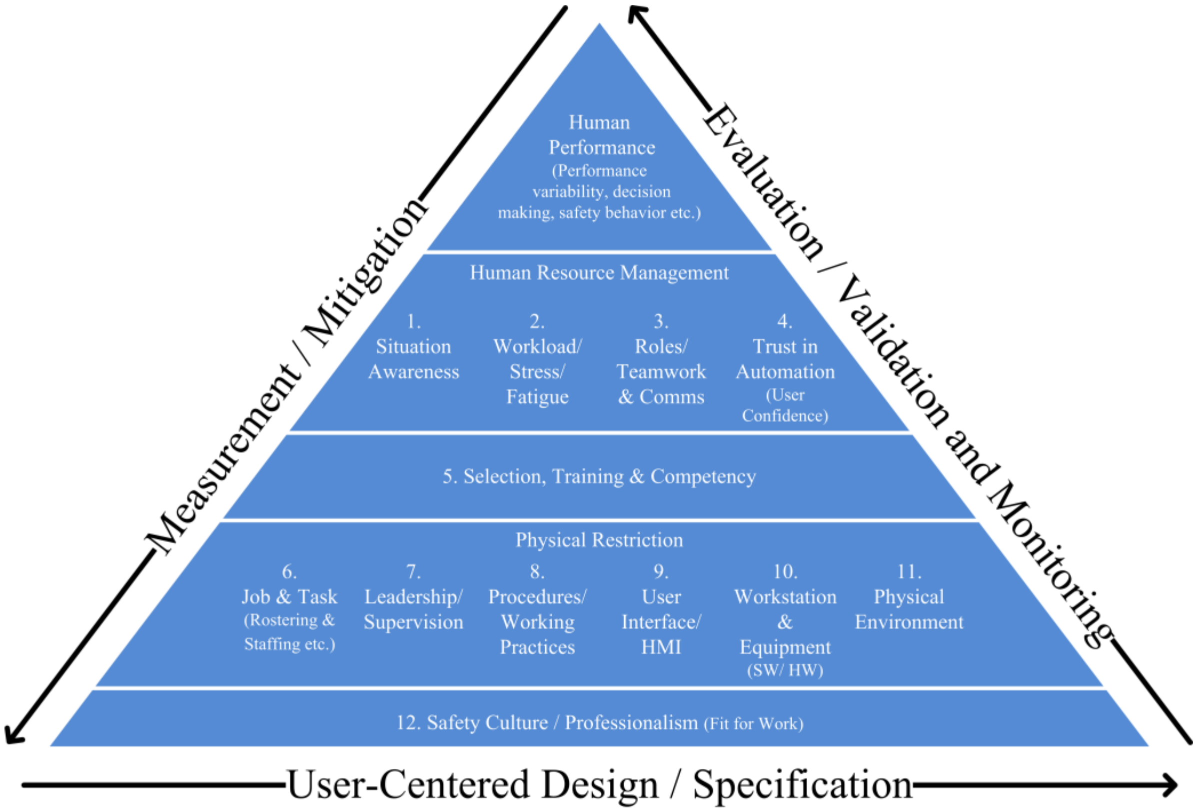

Figure 2. Influence of human performance in connection with new systems developed for ATC (EUROCONTROL & FAA, 2015)

In air traffic control, the airport environment information is presented on the workplace and the ATCO gathers a situation representation from there. He/she decides how the traffic should behave to ensure safety and efficiency and then radios each traffic participant with instructions. Ignoring the physical replacement away from the airport tower, the major redesign between the conventional and the remote tower workplaces is the replacement of the out-the-window view. In most projects this is implemented with the live stream from a camera system. The redesign from the out-the-window view to a camera system changes the interface to the airport environment. Therefore, the major impacts are inflicted onto workstation/equipment (see 10. in Figure 2), the user interface/HMI (see 9. in Figure 2) and human resource management (see 1. and 2. in Figure 2).

3.3 Remote Tower Metrics

When applying the HCI model for redesigning a workplace to the workstation and equipment (see 10. in Figure 2) for remote tower operations, the challenge is to develop a remote tower workplace with an interface that presents the same airport information as a conventional tower workplace. The environment in Figure 1 is the airport and the operator is the ATCO. As generalised in Figure 1, the flow of information is from the airport environment via the workplace to the ATCO and back. The workplace can either be a conventional or remote tower workplace. Following the model of redesign, the influence of redesigning the interface on the ATC task has to be analysed by selecting or defining remote tower metrics that allow measuring of the differences.

The model for redesign also supports the switching from the conventional tower to the remote tower workplace because it describes where the information is presented on each interface. A variety of research projects (see 3.1) have led to several different prototypes for a remote tower workplace and each has a slightly different interface. Each prototype was tested in a separate validation exercise, but the results are only valid for each particular prototype, because no combined model of redesign was used. Therefore, a set of remote tower metrics derived from the model of redesign is needed to evaluate the technical implementation in relation to the ATC task. The selection and evaluation of these indicators and the connected methodology is derived from research goal 1 (section 2.4) and resulted in article 1 (chapter II)

3.4 Dynamic Areas of Interests

Besides the workstation and equipment, the redesign of the interface applies, of course, to the user interface design (see 9. in Figure 2) of the remote tower workplace. For the process of user interface design, eye movement analysis is considered a standard method for determining the information-gathering process (Holmqvist et al., 2011). In comparison to other domains of eye movement analysis, such as reading (e.g. Castelhano & Rayner, 2008; Rayner, 1995) or image processing (e.g. McCarthy, Sasse, & Riegelsberger, 2004; Rayner & Pollatsek, 1992), the workplace of an ATCO has dynamic objects, e.g. moving aircraft and vehicles within the out-the-window view. Those dynamic objects also have to be considered for remote tower operations. Therefore, a guideline for advanced gaze analysis, called Integration Guideline for Dynamic Areas of Interest (IGDAI), was developed. The guideline is connected to research goal 3 (section 2.4) and described in detail in article 2 (chapter I).

3.5 Adaptation and strategy shifts

In its third part, the model for redesigning describes the need for identification of all the influential factors on the operator. For remote tower operations, these are the influences of the human resource management (see 1. and 2. in Figure 2) and they are still not fully understood. ATCOs do not only understand a situation by perceiving the position of all moving objects; they also have to understand their relevance in relation to safe and efficient ATC. As shown for the driving task (Baumann & Krems, 2007), ATCOs must also be able to make assumptions concerning the future states of the

airport. The research on situation awareness and workload and how they influence performance indicates that both have to be measured, analysed and interpreted together (Wickens et al., 1997). At the same time, ATC also allows for different strategies that influence situation awareness and workload.

Changing between different strategies to solve a task or operate an environment is called strategy shift and can be observed in a variety of domains. Studies that examine strategy shifts in multiple task environments are, for example, arithmetic (Staszewski, 1988), logical problem solving (Charness, 1991; Simon, 1975), spelling (Siegler, 1986), or text editing (Card, Moran, & Newell, 1983). Card et al. (1983) were also able to show that strategy shifts can depend on an environment parameter. Geary and Brown (1991) showed that the more basic knowledge about a task is available, the faster and more efficient the selection of a strategy becomes. Sperandio (1971) looked at strategy shifts within the ATC workplace in particular and found that these often occurred to compensate in high-workload situations, e.g. by spending less time on a single aircraft.

Sperandio (1978) uses control strategies to describe the interaction between workload and the dynamic environment, e.g. by prioritising in which order aircraft are handled. Therefore, the dynamic environment of an airport directly influences the workload and the control strategy, independent of the type of operation (conventional or remote tower). Situation awareness, workload and control strategy also influence each other if, for example, a changed control strategy requires different information about the environment.

Following the model of redesign, the development of the remote tower operations depends on situation awareness, workload, control strategies and their combined influence on the performance. The necessary approach to analyse the human resource management should be focusing on situation awareness and workload in connection with the control strategy applied by the ATCO. Knowledge of their influence on the individual ATCO performance will help the model to evaluate new design concepts in the early stages of their development. It should also prevent situations, e.g. with high workload, where the meaning of elements within the environment cannot be properly processed and then leads to no or incorrect actions (e.g. Krems & Baumann, 2009). A study presenting the described approach is published in article 3 (chapter I). The results of the study help to reach research goal 4 (section 2.4).

4 Methodological aspects of the dissertation

4.1 Identify and evaluate remote tower metrics

The remote tower operation requirements concentrating on visual perception though the interface were a starting point for the development of the remote tower metrics (Ellis & Liston, 2008; Hannon et al., 2008). Following research goal 1 (section 2.4) and the model of redesign, a method for extending the existing evaluation that allows measuring the difference on the flow of information between different workplaces is presented in article 1 (chapter II).

In cooperation with ATCOs as subject matter experts, the functional requirements of remote tower operations were refined and consolidated to identify remote tower metrics (Friedrich, 2016). This process focused on the visual perception through the interface that is essential for monitoring as part of the ATC task and is also most influenced by the remote tower operations. Following Friedrich (2016) and the article in chapter II, a total of 14 remote tower metrics were identified and evaluated.

The remote tower metric in chapter II focuses on six weather events (chapter II, table 1) that are one of the most influential factors for ATC. The complete process of identification was performed with experienced ATCOs in several workshops. Three steps were used to identify the remote tower metrics. Firstly, the weather information needed in connection with ATC tasks was identified. Secondly, specific tasks for each type of weather information were defined. Thirdly, indicators were defined for each of the specific tasks to allow performance measurement. By applying these three steps, the resulting remote tower metrics focus on the visual perception of static and dynamic information on the interface.

The suitability of the identified remote tower metrics was evaluated with a large-scale validation experiment. Within the evaluation process, the feasibility, comparability and effort were determined for each remote tower metric. Feasibility defines how applicable the remote tower metric is in comparison to the standard ATC task. Comparability describes the correlation between the different sources of information that are used. Effort is defined as the financial expense that is required to test the metric and helps to optimise and plan the available resources. For the presentation of the three indicators, an evaluation table was introduced to summarise the results and simplify the selection of remote tower metrics in preparation for a parallel evaluation of different prototypes.

4.2 Evaluate dynamic areas of interest

Research goal 3 (section 2.4) is connected to the analysis of the information-gathering process of an ATCO in relation to the dynamic environment of an airport. A method for measuring the change is necessary to understand the influence that the redesign has on the information-gathering process. Previous studies focussing on eye movement and dynamic environments in the domain of driving (e.g. Akamatsu et al., 2001; Palinko, Kun, Shyrokov, & Heeman, 2010), marketing (e.g. Josephson & Holmes, 2006; Nielsen, 2006) or air traffic management (e.g. Bruder et al., 2003; Möhlenbrink,

Oberheid, & Werther, 2008) rather focused on the application of individual approaches to the methodology. IGDAI (Integration Guideline for Dynamic Areas of Interest, see chapter I, section "IGDAI") presents a structured approach to connect eye movement to dynamic information within an artificial environment.

IGDAI presents an efficient way to cover the set-up, storage, and data analysis issues that arise when combining eye movement and dynamic areas of interest. The set-up issue is to configure the eye tracking environment to capture the needed information. The storage issue is used to identify the information that has to be recorded and derived into a suitable data structure. The data analysis issue is used to match eye movement and dynamic areas of interest. IGDAI does not determine a fixed solution but defines requirements that have to be fulfilled by solving each of the issues in connection with the individually available hardware and software.

IGDAI addresses the three issues by defining 8 requirements that are interconnected and thereby provide a workflow. In connection with the set-up issue, three requirements describe the necessity of the synchronised capturing of eye movement and dynamic information. For the storage issue, three requirements define the efficient and overall storage of the information. Finally, two requirements define the process of how to combine the eye movement and the dynamic information. By implementing all requirements, the recording of dynamic information is incorporated into the complete eye movement process. Therefore, the methodology simplifies and structures the process of capturing dynamic information and eye movement and allows, if applied to an ATC workplace, insights into which information the ATCO collects from the interface that is presented to him/her.

4.3 Measuring Situation Awareness

EUROCONTROL and FAA (2015) placed situation awareness at the same level as workload and summarise them as human resource management (see Figure 2). Even though there are scientists who doubt the usefulness of the situation awareness construct (Dekker, 2015; Dekker & Woods, 2002), for this dissertation it is an essential part of research goal 4 (section 2.4). Situation awareness is used to describe the interaction between workload and the control strategy of an ATCO. Salmon et al. (2009) argue that a consistent method for capturing situation awareness is difficult to identify due to the different concepts and how to measure them. In consequence, each method has to be revaluated in connection with a new environment with different preconditions. A detailed definition of situation awareness, how it is measured and its application in different domains can be found in article 3 (see chapter I, section 2.2). It also shows a lack in online measurements of situation awareness.

As an extension to the existing measurements of situation awareness in the ATC domain the SARA-T (Kraemer & Süß, 2015), an online probe measurement without interruption, was evaluated. In comparison to online freezing methods, SARA-T allows for a less intrusive, more natural, and realistic method to question the ATCOs about their current "picture" of the situation. If available, SARA-T

analyses the log files of a simulation in real-time and adapts the questions to the current traffic situation. SARA-T influences the work performance only by a minimum of additional workload, because accepting questions is only allowed if the resources are available.

5 Discussion and implications

5.1 Summarising the findings

The central findings of research goals 1 to 4 (section 2.4) are summarised in articles 1 to 3. For research goal 1, a summary is provided in the Discussion & Conclusion section of article 1. In the introduction of article 3, the connection to research goal 1 is also drawn by presenting remote tower operations as a new workplace environment that will be strongly investigated in the future. For research goal 3, the pursued methodology is described in the IGDAI section (see also section 4.2) and verified in the “Results and Discussion” section of article 2 (chapter I). As applied methodology, IGDAI is also described in the eye tracking section of article 3. Research goal 4 is summarised in the Discussion section of article 3. The final conclusions are summarised in the following three paragraphs. In the following sections, the theoretical (5.2) and practical (5.3) implications for each research goal are presented.

The two analyses in article 1 showed a large variety of metrics that are applicable to evaluate the influence of the redesign of the interface. An analysis of the windsock metric provided an example of how the different metrics can be applied to measure the difference in the flow of information. By applying this particular metric, it could be shown that the different workplaces have no impact on the correctness of the answer but on the reaction times of the participants. The second analysis showed that the weather remote tower metrics extend the existing remote tower metrics and thereby complete the aspects of the monitoring that an ATCO has to perform. As improvement to the previous selection methodology, an evaluation table was introduced to support future validation with a simple tool to select the correct remote tower metrics.

The method, IGDAI, used to evaluate the dynamic areas of interest is as good as a manual categorisation prepared with state-of-the-art software (e.g. BeGaze). The advantage of IGDAI over the manual categorisation is that the analysis time is fixed during the experimental set-up rather than a linear increase depending on the number of participants. In an exemplary validation study, it could be shown that IGDAI reached the same results in 6 hours, whereby a manual categorisation of the 20 data sets would have taken approximately 11 hours, increasing by 35 minutes with every additional data set. Also, the results of the study showed IGDAI’s capability to categorise eye movement in relation to dynamic areas of interest. IGDAI helps in identifying the individual information-gathering processes of each participant while they are performing their exemplary task.

Article 3 presents insights into the application of a probe technique for the assessment of situation awareness in connection with the influence of task load on the control strategy. SARA-T proved to be as useful as SASHA (Dehn, 2008), an after-run self-rating questionnaire specialised for aviation, in evaluating the situation awareness of ATCOs. As an online probe method, SARA-T also provided insights on the situation awareness during the high and low task load phases. The situation awareness

and performance did not vary significantly between the task load phases. The eye tracking, which implements the IGDAI method, implicates through a change in information-gathering the existence of two control strategies for the ATC environment that are each related to the task load phases. Both strategies use the same information, e.g. the position and speed of aircraft within the aerodrome, to update the situation awareness and use it to plan future traffic, as suggested by the situation awareness model by Baumann and Krems (2007). For high task load phases, the information is preferably gathered from the radar screen, because the task-relevant information (e.g. callsign and position) is combined and available at one glance. In the low task load phases, more attention is given to the aircraft themselves on the video panorama, probably because it is connected to additional information, e.g. status of the landing gear that is provided if requested.

5.2 Theoretical implications

The final conclusions on the theoretical implications are restricted by the partly indirect indicators and the correlative approach of the field studies. Also, articles 1 and 2 concentrate strongly on improving the methodology to lay the foundation for article 3. The conclusion should be evaluated under these restrictions.

5.2.1 The influence of the interface

The model for redesigning a workplace (section 3.1) that only changes the interface provides valuable insights into the general approach of interface redesign. This model implies that within HCI, a separation between the engineering and the redesign process is possible, depending on detailed task requirements and the implications for the operators' capacities to collect the information from different sources, if provided. The three parts of the model of redesign are detached from the engineering implementation and concentrate on the flow of information in relation to the task and operator capabilities. Therefore, this model reduces the influence of engineering within redesign to a minimum and places emphasis on a detailed task analysis.

5.2.2 Different redesigns

Looking at the overall approach and design factors (section 3.1), a wide range of factors are influenced when developing a remote tower workplace. The defined weather metrics allow measuring of the difference in the flow of information. With the variety of developed prototypes for remote tower operations, a structured approach for evaluating each of them and comparing them against each other is necessary. ATC within Europe is similar in procedures and task descriptions; a combined solution would therefore be applicable all over Europe. A procedure (section 4.1) that focuses on the task itself and can be transferred to remote tower metrics is an additional step towards remote tower operation. The validation shows that each remote tower metric can be judged differently depending on the capabilities of each prototype, but always has to be evaluated against the baseline, the existing controller tower workplace.

5.2.3 Eye movement with dynamic environments

As mentioned above, eye movement analysis is essential for capturing and consequently understanding the information-gathering process while performing a task. Of course, analysing the eye movement in relation to the dynamics changing in an environment is only necessary if the source of information changes position in relation to the operator, but due to the nature of many different domains, e.g. driving (e.g. Krems & Baumann, 2009; Palinko et al., 2010), marketing (e.g. Nielsen, 2006), or air traffic management (e.g. Möhlenbrink et al., 2008), the movement of information within an environment is common. In the “Existing approaches” section article 2 (chapter I), the standard methodologies are summarised and their pros and cons are evaluated against each other. Following the current application of eye tracking and the theoretical descriptions (e.g. Bruder et al., 2003), the application of eye movement in dynamic environment increases the significance of an eye tracking experiment.

In addition to the increase in significance, the development of efficient eye tracking methodology is as important as the development of a new eye tracking hardware itself. New eye tracking hardware is developed and distributed every year, especially the application of eye tracking glasses and monitor bars, e.g. Tobii (Tobii Technology, 2018) or SMI (SensoMotoric Instruments, 2018), but extended methods for data analysis and integration of the eye trackers into the task environment have still to be accomplished by the eye tracking researcher. Therefore, research such as in article 2 (chapter I) is not only important for this dissertation (see section 3.4) but also to the community of eye tracking researchers. These articles extend the existing methodologies and make additional data analysis approaches accessible.

5.2.4 Compensation strategies for Remote Tower

The major issue that ATCOs have with remote tower operations is connected to the replacement of the out-the-window view with a video reconstruction (see chapter IV, Papenfuss & Friedrich, 2016). ATCOs often note that the redesign comes with a decrease in visibility that leads to an increased workload (Fürstenau, 2016). As presented for different domains (section 3.5), this dissertation also presents results in the domain of ATC on how ATCOs cope with situations that have high workload. In article 3 (chapter I), ATCOs changed their strategy by changing their source of information. They switch from the out-the-window view to a radar screen that provides the same information but presented differently. Friedrich and Möhlenbrink (2013) (see chapter I) demonstrated the same behaviour for an exemplary monitoring task in a field study with ATCOs by comparing conventional and remote tower workplaces directly. Given that the same information is available on the interface of the remote tower workplace only presented in different formats, the theoretical implication is that ATCOs are likely to cope with the redesigned environment.

5.3 Implications for the application

5.3.1 Redesigns and predictions

The three parts of the model for redesigning only the interface of a workplace support the general approach of HCI. Firstly, it should encourage mainly engineering driven redesign approaches to consider not only the operator but also the task and the interdependency of both. This ensures that the operator can perform his/her task as safely and efficiently as before the redesign. Secondly, by looking at the needed information for the task, the redesign of the interface can be specific about the changes of information representation. This supports the process of switching from one workplace to the other, because the training could concentrate on the changed flow of information. Thirdly, identification of the necessary information for the operator task provides the basis for possible assistant systems and how they might influence the task. Therefore, the model also opens up the design process for additional assistance systems. For example, if missing information after the redesign is identified throughout the process, or the representation of critical information has to be more salient.

5.3.2 Simplified metric selection

The process of defining and selecting metrics that allow measuring of the differences in the flow of information is supported by the weather indicator evaluation table, as proposed by article 1 (section II, table 7). Its application would have positive implications, especially for the coming research initiatives in SESAR2020. The application of the evaluation table allows a compact overview of all monitoring tasks connected to weather at the redesigned interface. This reduces resources spent on identifying the metrics in every single research initiative by only concentrating on the specifics that might add to or alter the indicator evaluation table. The provision of effort as an evaluation for metrics supports the selection of indicators in connection with the available resources.

Considering that all research initiatives testing remote tower operation prototypes would use the same remote tower metrics, the consolidation of results would provide insights into the advantages and disadvantages of each single prototype. If technically possible, this could lead to a combined prototype sharing the advantages.

5.3.3 Effort in preparation

The main implications for the application of IGDAI are the detailed insights into dynamic areas of interests and the independence of the number of participants in relation to the time of data analysis. More insight into dynamic information could encourage researchers working with dynamic task environments, e.g. car, cockpit, or ATCO workplace, to re-evaluate previous studies to improve their results significantly. In the past, some researchers decided to replace the dynamic areas of interest with static areas of interest to simplify the data analysis. Now, dynamic areas of interest can be described as they are and do not have to be substituted by static areas of interest that might only partly fit the needs

of the study. The reduction of effort in data analysis could also lead to additional studies in dynamic task environments.

Since the effort for the data analysis is no longer connected to the number of participants and generally carried out in preparation for the experiment, researchers can focus on the data analysis right away rather than preparing their data. This increases the attractiveness of eye tracking in total and also motivates researchers who are new to the field to apply it. Eye tracking might already be a standard tool in the tool box of HCI, but not every generation of researchers should have to start with the development of their own analysis methodology. As mentioned above, for this dissertation IGDAI is the basis for article 3 (Chapter I). It was a fast and efficient way of obtaining the detailed insight into the information-gathering process of an ATCO to evaluate the change of the used control strategy.

5.3.1 Evaluation situation awareness

As described in article 3 (section I), the evaluation of situation awareness is a process that is still under development. Even in preparation of the specific sample, there is a huge field of studies which indicates that the situation awareness is affected by age and experience (Bolstad, 2001; Gawron, 2008; Patten, Kircher, Ostlund, Nilsson, & Svenson, 2006). Additional analyses testing the influence of these factors, e.g. by using correlation analysis between age, experience, and situation awareness, should become a standard for studies looking at the construct and especially before testing new methods.

The SARA-T method, developed by Kraemer and Süß (2015), that allows an actual real-time assessment of the situation awareness, was evaluated in the ATC environment. It provided the same results as the standardised SASHA_Q (Jeannot, Kelly, & Thompson, 2003) that was developed specifically for the context of measuring situation awareness in the ATC domain. Due to the situation dependency in connection with the amount of traffic, SARA-T allows for further insights into the situation awareness during normal work procedures. These insights can be compared later on in remote tower operations to identify an increase or decrease in situation awareness in connection with a specific situation or in general. Therefore, SARA-T not only helps to understand the situation awareness (see 2. in Figure 2); it also allows monitoring of possible changes inflicted by the redesign of the workplace.

5.4 Critical reflection of the methodology

For a correct interpretation of all results presented in this dissertation, a critical reflection of the used samples and the methodology is necessary.

5.4.1 Specifications of the samples

The recruitment for the field study presented in article 1 (chapter II) was performed by the German Air Navigation Service Provider (DFS). The first issue is derived from the fact that each ATCO has to be trained and licenced for a particular aerodrome and working position. Only ATCOs from the airport Erfurt-Weimar were allowed to participate; this reduces the generalisability of the study. A second

issue is that the regulations of air traffic services are provided on an international level by ICAO (International Civil Aviation Organization, 2001, 2007), but are implemented and sometimes extended or changed by national ATC regulations. Therefore, the task analysis that led to the resulting monitoring task might not be applicable for ATCOs from countries other than Germany. This also reduces the generalisability of the results in article 1 (chapter II).

In general, the topic of remote tower operation is still controversially discussed among ATCOs and therefore the recruitment of participants is difficult. Both studies working with experts, article 1 (chapter II) and article 3 (chapter I), had particularly small sample sizes. As explained above, for article 1 this was the maximum number of participants that could be provided by the DFS. For article 3, all ATCOs from the airport Braunschweig-Wolfsburg and two additional ATCOs were willing to participate and were recruited via an email mailing list. Since article 1 concentrated on descriptive data analysis and selecting and defining weather metrics, the sample size had only a minor influence on the results. However, within article 3, the smaller sample size was accounted for by the selection of statistical analysis methods that are classified as more robust against this issue, e.g. for the correlation analysis, Kendall's tau was used instead of the more common Spearman rank.

5.4.2 Specification of the methodologies

The diversity of research goals (section 2.4) made the application of three research methodologies necessary. A field study at an airport was conducted in article 1 (chapter II), to capture the input for the remote tower metrics directly at the workplace. The evaluation of IGDAI was performed using a microworld to simulate a dynamic environment (chapter I) that would show the advantages and the correctness of the approach. Article 3 (chapter I) presents a real-time simulation study to analyse workload, situation awareness and control strategies in a realistic but controlled environment. The selection of methodologies depended on the research goals and resources available to conduct each study. Whereas there were no specific restrictions with the microworld and the real-time simulation, except reduced external validity, the field study had some minor restrictions.

The reduced availability of ATCOs made a formal scale development including pre-test impossible in advance of the field test. The selected metrics regarding the monitoring of weather could only be answered by ATCOs. This could have influenced some of the results (Moosbrugger & Kelava, 2011; Stade, Von Bernstorff, Niestroj, & Nachtwei, 2011). The specifications of field studies also made the control of some factors impossible, e.g. the experience of the ATCOs and their knowledge of remote tower operation prior to the study. Also, only monitoring as a major part of the ATCOs task was part of the study, because using any additional equipment to perform ATC is under strict regulation by the Federal Office of Civil Aeronautics.

5.5 Revenue for psychological research

The model of redesign for a workplace interface only contributes to the field of HCI, by defining three important parts that need to be considered. This approach is based on the psychological aspects of the operators looking only at the information needed to perform the task. Even though redesigning of interfaces is performed in many domains (e.g. Dix, 2009), the safety-critical issues in some domains justify and demand a detailed HCI theory and model-based approach to describe the operator and performed task. The model presented in this dissertation supports HCI by focusing on the flow of information and how this information might be affected by a redesign of the interface. The model is tested through its application in the special case of redesigning an interface for remote tower operations.

The development in the domain of remote tower operations seems to be stable and necessary to keep up with the challenges of future air traffic management. The involvement of ATCOs in the research process, e.g. in workshops as in article 1 (chapter II) or as participants (chapters II and I), increases their acceptance towards the technology (Fürstenau, 2016). In addition, the air navigation service provider's attempt at keeping the same level of safety in combination with a more efficient use of available resources also facilitates the development of technology for remote tower operations. The transition from the conventional tower workplace to remote tower operations is an ongoing process that will be continued. Therefore, the analysed constructs, developed methodologies and presented results from this dissertation provide a seminal basis for future research.

The application of new technical systems and assistance systems (Bengler et al., 2014; Rudin-Brown & Jamson, 2013; Stevens, Brusque, & Krems, 2013) is an ongoing topic within traffic psychology. In particular, the understanding of the influential factors that support or reject the process of behaviour adaptation need to be carefully analysed, to support any transition process. This dissertation looked at the changes introduced by a new system on a task level (article 1) and evaluated the difference by looking at the information gathered. The results indicated the need for additional assistance systems that might close the gap between the different information provided. The next logical step of developing visual assistant systems has already been started by Friedrich, Pichelmann, Papenfuß, and Jakobi (2017) (see also chapter I) with 3 proprietary implementations. Therefore, this dissertation can be judged as an asset to the field of traffic psychology.

As described in section 3.4, the application of dynamic areas of interest within the domain of HCI is still connected to increased effort for analysing the data gathered (Papenmeier & Huff, 2010). This dissertation introduced a methodology that reduced the effort for data analysis, not only for the domain of air traffic management but also for researchers who experiment with dynamic environments and are interested in eye movement. The evaluation of dynamic areas of interest in connection with newly developed prototypes seems necessary and helpful in influencing and optimising the ease of use. In the

coming years, the methodological approach presented in this dissertation is expected to support research in other areas of the HCI.

Finally, the contribution of this dissertation towards explaining the connection between workload, situation awareness, and control strategy has to be mentioned. Vidulich (2000) stated that since the introduction of situation awareness, there have been a great many discussions about the relation between workload and situation awareness as well as their link to performance, e.g. Ma and Kaber (2005) or Lee, Jeon, and Choi (2012). Based on article 3, this dissertation showed the influence of workload on situation awareness and the change in control strategy. Further research on these constructs has to consider their connection and their reciprocal influence combined with the effect of changing control strategies.

Concluding all sections, this dissertation could provide input for traffic psychology and HCI with additional influence and support to other domains and research topics in connection with psychology. The same applies to the methodologies that are not restricted to one particular domain and therefore could be transferred if needed. The results could be valuable for a better understanding of workload, situation awareness and control strategy in a large number of different domains, especially wherever an operator has to interact with an environment to perform a task.

6 Literature

Ahlstrom, U. (2005). Work domain analysis for air traffic controller weather displays. *Journal of Safety Research, 36*(2), 159-169. doi:https://doi.org/10.1016/j.jsr.2005.03.001

Akamatsu, M., Okuwa, M., Funakawa, M., Kitajima, M., Onuki, M., Life, Q., & Co, M. P. (2001). *Whole-View Driving Simulator For Measuring Driving Behavior In Town Environment*. Paper presented at the Driving simulation conference, Valbonne.

Anagnostakis, I., Clarke, J.-P., Bohme, D., & Volckers, U. (2001). *Runway operations planning and control sequencing and scheduling.* Paper presented at the System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on.

Atkins, S., & Brinton, C. (2002). Concept description and development plan for the surface management system. *Journal of Air Traffic Control, 44*(1), 1-8.

Badham, R., & Ehn, P. (2000). Tinkering with technology: Human factors, work redesign, and professionals in workplace innovation. *Human Factors and Ergonomics in Manufacturing & Service Industries, 10*(1), 61-82. doi:doi:10.1002/(SICI)1520-6564(200024)10:1<61::AID-HFM4>3.0.CO;2-O

Baumann, M. R. K., & Krems, J. F. (2007). Situation awareness and driving: A cognitive model. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments* (pp. 253-265). London: Springer.

Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C., & Winner, H. (2014). Three Decades of Driver Assistance Systems: Review and Future Perspectives. *IEEE Intelligent Transportation Systems Magazine, 6*(4), 6-22. doi:10.1109/MITS.2014.2336271

Billings, C. E. (2018). *Aviation automation: The search for a human-centered approach*: CRC Press.

Bolstad, C. A. (2001). *Situation awareness: does it change with age?* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Borges, J. A., Morales, I., & Rodríguez, N. J. (1996). *Guidelines for designing usable World Wide Web pages.* Paper presented at the Conference Companion on Human Factors in Computing Systems.

Bruder, C., Leuchter, S., & Urbas, L. (2003). Konzept für automatische Auswertungen von Blickbewegungen bei dynamischen Informationen. *MMI interaktiv, 6*, 25-34.

Card, S. K., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.

Castelhano, M. S., & Rayner, K. E. (2008). Eye movements during reading, visual search, and scene perception: An overview. *Cognitive and Cultural Influences on Eye Movements, 2175*, 3-33.

Charness, N. (1991). *Expertise in chess: The balance between knowledge and search*. Cambridge (UK): Cambridge University Press.

Dehn, D. M. (2008). Assessing the Impact of Automation on the Air Traffic Controller: The SHAPE Questionnaires. *Air Traffic Control Quarterly, 16*(2), 127-146. doi:10.2514/atcq.16.2.127

Dekker, S. W. (2015). The danger of losing situation awareness. *Cognition, technology & work, 17*(2), 159-161.

Dekker, S. W., & Woods, D. D. (2002). MABA-MABA or abracadabra? Progress on human–automation co-ordination. *Cognition, technology & work, 4*(4), 240-244.

Dix, A. (2009). Human-computer interaction. In *Encyclopedia of database systems* (pp. 1327-1331): Springer.

Ellis, S. R., Fürstenau, N., & Mittendorf, M. (2011). Frame Rate Effects on Visual Discrimination of Landing Aircraft Deceleration: Implications for Virtual Tower Design and Speed Perception. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 55*, 71-75. doi:10.1177/1071181311551015

Ellis, S. R., & Liston, D. (2008). Visual Features Involving Motion Seen from Airport Control Towers. In E. Witrant, E. Fridman, O. Sename, & L. Dugard (Eds.), *11th IFAC Proceedings* (Vol. 43, pp. 152-157): Elsevier.

EUROCONTROL, & FAA. (2015). *A Human Performance Standard of Excellence*. Retrieved from www.skybrary.aero

Friedrich, M. (2016). Which metrics provide the insight needed? A selection of remote tower evaluation metrics to support a remote tower operation concept validation. In N. Fürstenau

(Ed.), *Virtual and Remote Control Tower* (1 ed., pp. 221-240). Berlin, Heidelberg, New York: Springer Publishers.

Friedrich, M., & Möhlenbrink, C. (2013). Which data provide the best insight? A field trial for validating a remote tower operation concept. In FAA/EUROCONTROL (Ed.), *Proceedings of the USA/FAA Air Traffic Management R&D Seminar*: http://www.atmseminarus.org/10th-seminar/papers/.

Friedrich, M., Pichelmann, S., Papenfuß, A., & Jakobi, J. (2017). The Evaluation of Remote Tower Visual Assistance System in Preparation of Two Design Concepts. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics: Performance, Emotion and Situation Awareness* (1 ed., pp. 285). Vancouver: Springer International Publishing.

Fürstenau, N. (2016). *Virtual and Remote Control Tower: Research, Design, Development and Validation*: Springer International Publishing.

Fürstenau, N., Schmidt, M., Rudolph, M., Möhlenbrink, C., Friedrich, M., Papenfuss, A., & Kaltenhäuser, S. (2007). *Steps towards the Virtual Tower: Remote Airport Traffic Control Center (RAiCe)*. Paper presented at the ENRI International workshop on ATM/CNS.

Fürstenau, N., Werther, B., Mittendorf, M., & Schmidt, M. (2004). Virtual Tower. *Sonderheft "Wettbewerb der Visionen 2001 – 2004"*, 16-21.

Gawron, V. J. (2008). *Human performance, workload, and situational awareness measures handbook*: CRC Press.

Geary, D. C., & Brown, S. C. (1991). Cognitive Addition: Strategy Choice and Speed-of-Processing Differences in Gifted, Normal, and Mathematically Disabled Children. *Developmental Psychology, 27*.

George, C. A. (2005). Usability testing and design of a library website: an iterative approach. *OCLC Systems & Services: International digital library perspectives, 21*(3), 167-180.

Hannon, D., Lee, J., Geyer, M., Mackey, S., Sheridan, T., Francis, M., . . . Malonson, M. (2008). Feasibility Evaluation of a Staffed Virtual Tower. *Journal of the Air Traffic Control, 50*(1), 27-39.

Head, A. J. (1999). *Design wise: A guide for evaluating the interface design of information resources*: Information Today, Inc.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*: OUP Oxford.

International Civil Aviation Organization. (2001). *Annex 11 to the convention on International Civil Aviation - Air Traffic Services* (Vol. 13): ICAO.

International Civil Aviation Organization. (2007). *Procedures for Air Traffic Services - Air Traffic Managemnet* (Vol. 15): ICAO.

Issa, T., & Isaias, P. (2015). *Sustainable Design*. London: Springer-Verlag.

Jeannot, E., Kelly, C., & Thompson, D. (2003). *The Development of Situation Awareness Measures in ATM Systems*, Brussels, Belgium.

Josephson, S., & Holmes, M. E. (2006). Clutter or content? How on-screen enhancements affect how TV viewers scan and what they learn. *Main, 38*, 155-162. doi:10.1145/1117309.1117361

Kontogiannis, T., & Malakis, S. (2013). Strategies in controlling, coordinating and adapting performance in air traffic control: modelling 'loss of control'events. *Cognition, technology & work, 15*(2), 153-169.

Kraemer, J., & Süß, H.-M. (2015). Real Time Validation of Online Situation Awareness Questionnaires in Simulated Approach Air Traffic Control. *Procedia Manufacturing*, 3152-3159.

Krems, J., & Baumann, M. (2009). Driving and Situation Awareness: A Cognitive Model of Memory-Update Processes. In M. Greenlee (Ed.), *Human Centered Design* (pp. 986-994): Pabst Science Publishers.

Lange, M. (2014). *Information Search for Decision-Making in Tower Air Traffic Control: A Field Study*. (Master), Tu Chemnitz,

Lee, Y. H., Jeon, J.-D., & Choi, Y.-C. (2012). Air Traffic Controllers' Situation Awareness and Workload under Dynamic Air Traffic Situations. *Transportation Journal, 51*(3), 338-352. doi:10.5325/transportationj.51.3.0338

Ma, R., & Kaber, D. B. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics, 35*(10), 939-953.

MacKay, W. E. (1999). Is paper safer? The role of paper flight strips in air traffic control. *ACM Trans. Comput.-Hum. Interact., 6*(4), 311-340. doi:10.1145/331490.331491

McCarthy, J. D., Sasse, M. A., & Riegelsberger, J. (2004). Could I have the menu please? An eye tracking study of design conventions. *People and Computer*, *17*, 401-414.

McKeon, D. (2009). NATS Delivers World: First in Virtual Control Towers. *International Airport Review*, *13*(5), 44-46.

Mensen, H. (2014). *Moderne Flugsicherung: Organisation, Verfahren, Technik[Modern Air Traffic Management: Organisation, Procedures, Methodes]*: Springer-Verlag.

Möhlenbrink, C., Oberheid, H., & Werther, B. (2008). *A Model Based Approach to Cognitive Work Analysis and Work Process Design in Air Traffic Control*. Paper presented at the Annual Meeting of the Human Factors and Ergonomics Society Europe Chapter, Braunschweig.

Moosbrugger, H., & Kelava, A. (2011). *Testtheorie und Fragebogenkonstruktion [Test theory and questionair construction]*: Springer.

Newell, A., & Card, S. K. (1985). The prospects for psychological science in human-computer interaction. *Hum.-Comput. Interact., 1*(3), 209-242. doi:10.1207/s15327051hci0103_1

Nielsen, J. (2006). F-Shaped Pattern For Reading Web Content. *Search, 2009*, 2007.

Palinko, O., Kun, A. L., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. *Proceedings of the 2010 Symposium on EyeTracking Research Applications ETRA 10*, 141. doi:10.1145/1743666.1743701

Papenfuss, A., & Friedrich, M. (2016). Head Up Only – A design concept to enable multiple remote tower operations. *35th Digital Avionics Systems Conference*.

Papenmeier, F., & Huff, M. (2010). DynAOI: A tool for matching eye-movement data with dynamic areas of interest in animations and movies. *Behavior research methods, 41*, 179-187. doi:10.3758/BRM.42.1.179

Patten, C. J., Kircher, A., Ostlund, J., Nilsson, L., & Svenson, O. (2006). Driver experience and cognitive workload in different traffic environments. *Accid Anal Prev, 38*(5), 887-894. doi:10.1016/j.aap.2006.02.014

Pfeiffer, L., Valtin, G., Müller, N., & Rosenthal, P. (2016). *The Mental Organization of Air Traffic and its Implications to an Emotion Sensitive Assistance System* (Vol. 8).

Poulton, E. C. (1966). Engineering Psychology. *Annual Review of Psychology, 17*(1), 177-200. doi:10.1146/annurev.ps.17.020166.001141

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-Computer Interaction*. Wokingham, England: Addison-Wesley Longman Ltd.

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. *Studies in Visual Information Processing, 6*, 3-22. doi:10.1016/S0926-907X(05)80003-0

Rayner, K., & Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal Of Psychology, 46*, 342-376.

Rudin-Brown, C., & Jamson, S. (2013). *Behavioural adaptation and road safety: Theory, evidence and action*: CRC Press.

Saab Security. (2008). Remotely Operated Tower -The Future of Being Present [Press release]

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics, 39*(3), 490-500.

Schmidt, M., Rudolph, M., Werther, B., & Fürstenau, N. (2006). *Remote Airport Tower Operation with Augmented Vision Video Panorama HMI*. Paper presented at the Proc. 2nd Int Conf. Res. in Air Transportation ICRAT, Belgrade.

Schneider, N., Möhlenbrink, C., & Kiesel, A. (2010). *Doubling the task - Effects of task switching during simultaneous control of multiple airports.* Paper presented at the HFES Europe Chapter conference, Berlin.

SensoMotoric Instruments. (2018). Eye Tracking Solutions by SMI. Retrieved from https://www.smivision.com/

Siegler, R. S. (1986). *Children's thinking*. Englewood Cliffs, NJ: Prentice-Hall.

Simon, H. A. (1975). The Functional Equivalence of Problem Solving Skills. *Cognitive Psychology, 7*, 88.

Sperandio, J.-C. (1971). Variation of Operator's Strategies and Regulating Effects on Workload. *Ergonomics, 14*(5), 571-577. doi:10.1080/00140137108931277

Sperandio, J.-C. (1978). The Regulation of Working Methods as a Function of Work-load among Air Traffic Controllers. *Ergonomics, 21*(3), 195-202. doi:10.1080/00140137808931713

Stade, M., Von Bernstorff, C., Niestroj, N., & Nachtwei, J. (2011). *(Not) Everybody's Darling: Value and Prospects of Multiple Linear Regression Analysis and Assumption Checking* (Vol. 15).

Staszewski, J. J. (1988). Skilled memory and expert mental calculation. *The nature of expertise*.

Stevens, A., Brusque, C., & Krems, J. (2013). *Driver adaptation to information and assistance systems.* (Vol. 2). Stevenage, UK: IET.

Tobii Technology. (2018). Tobii.com - Tobii is the world leader in eye tracking. Retrieved from https://www.tobii.com/

Truitt, T. (2006). *Concept development and design description of electronic flight data interfaces for airport traffic control towers* (DOT/FAA/TC-TN-06/17). Retrieved from

Vidulich, M. A. (2000). *The relationship between mental workload and situation awareness.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Wickens, C. D., Mavor, A. S., & McGee, J. P. (1997). *Flight to the future: Human factors in air traffic control*. Washington, D.C.: National Academies Press.

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). *Research through design as a method for interaction design research in HCI.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

II Article 1: How to Evaluate Remote Tower Metrics in Connection to Weather Observations. An Extension of the Existing Metrics

The following version of the article is not equal to the article published in Aviation Psychology and Applied Human Factors with doi: http://dx.doi.org/10.1027/2192-0923/a000142. This is not the original version of the article and therefore cannot be uses for citations. Please do not distribute or publish this article without the permission of the authors.

Citation: Friedrich, M., & Möhlenbrink, C. (2018). How to evaluate remote tower metrics in connection to weather observations. An extension of the existing metrics. Aviation Psychology and Applied Human Factors, 8(2). doi: http://dx.doi.org/10.1027/2192-0923/a000142

Journal: Impact Factor is expected for the year 2018.

How to evaluate remote tower metrics in connection to weather observations. An extension of the existing metrics.

Maik Friedrich Christoph Möhlenbrink

6058 words (body)

German Aerospace Center, Braunschweig

Author:

Maik Friedrich, maik.friedrich@dlr.de
+49 (531) 295 - 2598
German Aerospace Center
Lilienthalplatz 7
Braunschweig, NI D-38108

Acknowledgments

The activities developed to achieve the results presented in this article were created by SESAR Lot 1 activity (SJU/D/12-446 study) for the SESAR Joint Undertaking within the frame of the SESAR Program co-financed by the EU and EUROCONTROL. The opinions expressed herein reflect the authors' view only. The SESAR Joint Undertaking is not liable for the use of any of the information included herein. The experimental set-up was contributed by German Aerospace Center and DFS.

Abstract

Due to the different approaches for remote tower operation, a standardized set of indicators is needed to evaluate the technical implementations on a task performance level. One of the most influential factors for air traffic control is weather. This article describes the influence on remote tower operations by weather metrics and how to validate them against each other. Weather metrics are essential to the evaluation of different remote controller working positions. Therefore weather metrics have been identified as part of a validation at the airport Erfurt-Weimar. Air traffic control officers have observed weather events at the tower control working position and the remote control working position. The eight participating air traffic control officers have answered time-synchronized questionnaires at both workplaces. The questionnaires addressed operationally relevant weather events within the aerodrome. The validation experiment has targeted the air traffic control officer's ability to categorize and judge the same weather event at different workplaces. The results show the potential of standardized indicators for the evaluation of performance and the importance of weather metrics in relation to other evaluation metrics.

Keywords: Remote Tower Operations; Weather; Video-panorama; detectability; Validation;

Introduction

The development of the last years indicates that the German air traffic control transitions from the conventional air traffic controller workplace (CWP-tower) to a new controller working position remote (CWP-remote). The German Aerospace Center project RapTOr (Remote Airport Tower Operation Research, 2005-2007) focused on the feasibility of remote tower operations (RTO), while the follow-up project RAiCe (Remote Airport traffic control center, 2008-2012) focused on controlling multiple small airports from a remote center (Ellis, Fürstenau, & Mittendorf, 2011; Fürstenau, Schmidt, Rudolph, Möhlenbrink, & Halle, 2008; Möhlenbrink, Friedrich, & Papenfuss, 2009).

Besides the German Aerospace Center research, remote tower operation were also part of the project ROT (Remotely Operated Towers, 2006-2008; Saab Security, 2008) and ART (Advanced Remote Tower, 2007-2009; van Schaik, Lindqvist, & Rössingh, 2010). Such as RapTOr, both projects focused on the aspect of an Air Traffic Control Officer (ATCO) controlling one remote tower. Further, Frequentis developed the system smartVISION in cooperation with Rheinmetall Defence and tested its feasibility at Saarbruecken airport (FREQUENTIS, 2015). Remote Tower Operation was also addressed in an Operational Focus Area (06.03.01) (Committee Sesar Program, 2010) funded by the Single European Sky ATM Research Program.

The research projects above demonstrate the interests of different stakeholders in the development and application of RTO. These projects completed field trials with human-in-the-loop experiments to validate their research focuses. Möhlenbrink, Rudolph, Schmidt, and Fürstenau (2007) completed the first field trials at the Airport Braunschweig-Wolfsburg. They used a self-developed prototype equipped with four cameras to generate the video-panorama. Fürstenau et al. (2009) quantified the resolution of the video-panorama that suited best to the needs of an ATCO. In addition, the importance of visual cues and their detection and recognition within the video-panorama was presented by van Schaik et al. (2010). Friedman-Berg (2012) also explored alternative surveillance systems for the out-the-window view for their concepts on staffed NextGen Tower.

Figure 1 presents the primary flow of information from the airport environment to the ATCO via the controller working position. The controller working position works as a filter between ATCO and Airport environment and can either be the CWP-tower or the CWP-remote. The CWP-remote has the same systems as the CWP-tower, except the out-the-window view and the binoculars. The out-the-window view on the CWP-remote is a digitally reconstructed video-panorama using high resolution cameras. The binoculars are replaced by a pan-tilt-zoom camera. The replacement of the CWP-tower with a CWP-remote therefore has the most influence on the combined situation representation of the out-the-window view and the binoculars. The challenge with RTO is to develop a CWP-remote that takes different types of airport information, especially for out-the-window view and binoculars and combines them into a situation representation that allows the ATCO to come to the same conclusions and make the same decisions as with the CWP-tower.

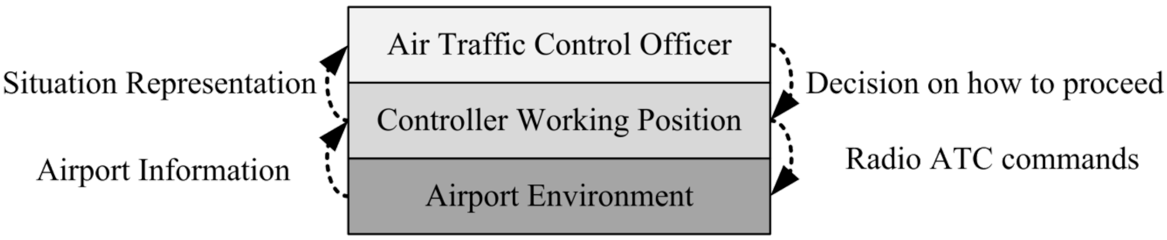

Figure 1 Flow of information from the airport environment to the air traffic control officer

Since the direct transfer from CWP-tower to CWP-remote is not possible, each RTO research initiative developed a prototype that has a unique set of specifications, e.g. size and arrangement of displays and information, the presentation of radar information, or the integration of multiple sensors. These specifications are derived from technical limitations, e.g. the amount of visual information that can be streamed from the airport to the CWP-remote. In contrast to that, the authors believe that system development should not be technical driven, but depending on functional requirements derived directly from the air traffic control task. The functional requirements describe the technical functionalities the system must or should have, to perform the task it was developed for. In general the ATCO, working the CWP-remote, has to reach the same conclusion on how to proceed as if he was working the CWP-tower.

Each different RTO system can be validated with the functional requirements, to ensure that it provides the necessary functionalities. The process of evaluating functional requirements is supported

by the definition of metrics and the specification of indicators. Metric is therefore defined as quantifiable measure that is used to track and assess the quality of the monitoring weather task at different workplaces. Considering the different RTO research initiatives, functional requirements with a fixed set of Remote Tower Metrics (RTMs) for each task area allow for a structured comparison and validation of RTO systems. Each RTM set contains indicators that represent the measurable values throughout the validation process. Therefore, RTMs allow measuring the discrepancy between CWP-tower and CWP-remote independent from the technical implementation. Friedrich and Möhlenbrink (2013) showed the influence of RTMs on the validation of remote tower system in detail. The evaluation process has to be simple and transparent to ensure the system developers can provide the information with their RTO systems. These information help the ATCO, that have to work with a specific CWP-remote system, to judge if they can work as before or with specific limitations or enhancements.

Each RTO project contributes to an increase in feasibility and the preparation of an industrial integration. Therefore, the RTO concept and functional requirements have been extended with every validation (Mullan, Lindqvist, Svensson, Abel, & Ankartun, 2012). Friedrich and Möhlenbrink (2013) proposed a system to evaluate the quality of different RTMs covering the different requirements of an air traffic control task. Friedrich (2016) described initial RTMs that were tested. The RTO system developers need RTM that are derived from functional requirements, that are reasonable and allow a qualitative evaluation of their RTO system. This article extends the existing RTMs with weather RTM derived from weather tasks and also enhances the process of evaluation, by simplifying the result summary. First, a method for improving the metrics evaluation is presented. Second, the weather indicators and their connection to the CWP-remote are described in detail. Third, the experimental setup is presented. Fourth, the results are presented and the contribution of the weather RTM to the concept of evaluating different research prototypes will be discussed. Fifth, data and methods are summarized as appropriate to judge which indicators cover the important parts of the ATCOs work.

Weather in Air Traffic Control

"There are always two systems out there — air traffic control and weather — and the responsibility and the challenge come from fitting the airplane into both of them as smoothly as possible" (Collins, 1991). Weather effects are one of the most influential issues in air traffic control with respect to visual perception, coordination, and delay. A major air traffic control task is to mitigate its influence on safety and efficiency (Krozel, D. Andre, & Smith, 2003). The ATCO is responsible for guiding the traffic within the aerodrome and has to consider several dependencies before meteorological information is used (Wickens, Mavor, & McGee, 1997). The CWP-tower is mostly influenced by weather events within the aerodrome. For example, the flight path of an aircraft is strongly connected to the speed of wind and its direction (Mensen, 2014), therefore thunder cells close to the airport have to be monitored and if necessary flight paths adapted. On the other hand, detailed knowledge on the current wind situation at the final approach can be used to reduce required stagger distances and increase the capacity without effecting safety. A third example are the three categories of visibility CAT I (>550m), CAT II (>350m), and CAT III (>200m) that define visibility ranges (Mensen, 2014). Different air traffic control procedures are applied depending on the CAT category. Even though the CAT categories are normally derived from a validated external weather source, the ATCO has to have the ability to check their correctness. These examples show the strong connection between weather information and the air traffic control task.

At the CWP-tower the two sources of weather information are the national weather provider and the out-the-window view. The primary source of weather information is the national weather provider. It is normally presented on a weather information display (Deutscher Wetter Dienst, 2017) and therefore can be duplicated at the CWP-remote. The secondary source of weather information is the out-the-window view. The monitoring of weather by the ATCO has a different significance depending on the laws of a country and the air navigation service provider. In most cases, ATCOs are only required to make tower visibility observations that might be inflicted by weather. In the United States of America, for example, if no national weather provider is located at the airport, the ATCO are likely to be responsible for performing all of the necessary weather observations (Nolan, 2010).

Because the out-the-window view is replaced by the video-panorama the transferability has to be tested in connection to the weather observation task that an ATCO has actually to perform while controlling the aerodrome. Therefore, for the definition of weather RTM it is necessary to consider that the monitoring of weather through the out-the-window view can either be “nice to have” or an essential part of the ATCO task.

Remote Tower Metrics

The earliest approaches for defining RTO requirements were concentrating on the visual perception (Ellis & Liston, 2008; Hannon et al., 2008). Friedrich (2016) describes the process of RTM identification as refinement and consolidation of the functional requirements of a CWP-remote and in cooperation with ATCOs that serve as subject matter experts. The main focus for Remote Tower Metrics (RTM) was the visual perception in relation to the surveillance task. He defined 8 RTMs divided into Aircraft (5) and Objects at the Apron (3). This article extends this work by defining RTM that allow the validation of monitoring weather events.

The following three steps were used to identify the weather RTMs. The first step was the identification of a monitoring task in relation to weather that ATCOs have to perform. The monitoring of weather is connected to the task of planning routes and providing weather information for pilots. The second step was the definition of tasks for each type of weather information. Therefore a workshop with ATCOs was performed and their input was used to specify the task in detail. The third step was the definition of indicators for the specified RTMs to allow a performance measurement.

As mentioned before, the RTMs are not tasks every ATCO has to perform. They focus on the visual perception of static and dynamic information in connection to the ATC task presented via out-the-window view. This also applies for the weather RTMs. Within this article we concentrate on the evaluation of the defined RTMs in terms of usability to evaluate a CWP-remote and comparability to existing RTMs (Friedrich, 2016). A large scale validation experiment was performed to evaluate the seven identified RTMs (Table 1) for the task of weather surveillance. These RTMs are made measurable by transforming them into indicators.

Table 1 Specified RTMs for the task weather surveillance.

Remote Tower Metric	Description
Windsock	The ATCO has to identify the status of the windsocks at the airport.
Range of sight	The ATCO has to identify the range of sight on the airport.
Speed of wind	The ATCO has to identify the speed of wind on the airport.
Intensity of rainfall	The ATCO has to identify the rainfall intensity on the airport.
Percentage of clouds	The ATCO has to identify the cover of clouds in percentage.
Type of clouds	The ATCO has to identify the type of clouds visible at the airport.

Friedrich (2016) evaluated the RTMs using effort, comparability and feasibility. Effort defines the financial expense to set-up the validation environment for testing a particular RTM. This helps plan the available and restricted resources. The RTMs are therefore categorized into low, medium, or high. Because the RTMs discriminate between CWP-tower and CWP-remote, their validity to describe the difference is derived from the sources of information. Comparability is defined as the correlation between the sources of information used on both workplaces. Feasibility of each RTM is their operational value, a task an ATCO has to perform. Feasibility is defined as the extent to which the RTM is applicable to normal work situations. Friedrich (2016) used a debriefing questionnaire to determine the feasibility. All three aspects were standardized and summarized equally to calculate the RTMs Score.

Method

Participants

Eight ATCOs, from local or regional airports (age M = 35, SD = 12.9; work experience M = 14, SD = 12), participated in this validation experiment. They were all employed by the German Air Navigation Service Provider. Four participants had known the RTO project in advance and none of them received additional payment. They participated during their typical working hours.

Apparatus

As mentioned before the most important difference between the CWP-remote and the CWP-tower is the visual reproduction of the out-the-window view (Fürstenau et al., 2008; Fürstenau et al., 2009;

Schmidt, Rudolph, Werther, Möhlenbrink, & Fürstenau, 2007). Therefore, the apparatus for the validation experiment consisted of a CWP-tower, a CWP-remote prototype, and synchronized questioning software.

The experimental set-up was installed at the airport Erfurt-Weimar, a regional airport with a local tower in central Germany. The local tower contained all systems to provide air traffic service and therefore was used as CWP-tower. The sources of information used for the experiment were limited to the out-the-window view, binoculars, air situational display, flight plan data and the weather information system.

The CWP-remote prototype is presented in Figure 2. The reproduction of the out-the-window view was realized by five high definition cameras equipped with a 2/3" – charge-coupled device sensor and a focal length of 8mm set up on a camera platform on top of the local tower. The resolution of the camera system reached 2 arcmin angular. The video stream of the high definition cameras was displayed on five 40" monitors arranged in a quadrant. This set-up allowed a 200° field of view on the aerodrome. A pan-tilt zoom camera was mounted on top of the camera platform to provide a technical substitute for the binoculars. The pan-tilt zoom camera could be moved freely within the full 360° viewing range and had 12 pre-sets for hot spots. It was displayed on the 27" Wacom display. The remaining displays contained the weather information system, the air situation display, and flight plan data. These systems are mirrored information from the CWP-tower system. The participants were placed at a distance of 1.8m to the video-panorama.

Figure 2 CWP-remote at the airport Erfurt – Weimar (Friedrich, 2016)

Table 2 contains the indicators derived from the weather Remote Tower Metrics (RTMs, see Table 1) that were used in this validation. Each RTM is connected to an indicator of the same name and made measurable by defining a question that has to be answered at both workplaces. Each indicator is defined at least by one question, except for the windsock status. The four windsocks on the airfield were numbered for individual questioning. This summarizes to a total of 6 indicators and 9 questions to measure them. The source of information was asked for in addition to each question. This is necessary to differentiate and categorize the influences that the CWP-remote has on the information gathering. The source of information could be indicated as panorama (out-the-window view or video-panorama), magnification (binoculars or pan-tilt zoom camera), weather information system, or a combination of these systems.

Table 2 Connection between the weather RTMs and their implementation as a question

Indicator	Questions to capture the weather indicator	Possible Answers
Windsock	Which direction is wind sock X pointing? (X with 1,2,3, or 4)	South, South West, West, North West, North, North East, East, and South East
Range of sight	What is the current range of sight in meters?	Meters in numbers
Speed of wind	What is the current speed of wind in knots?	Knots in numbers
Intensity of rainfall	Which intensity has the current rainfall?	None, moderate, high
Percentage of clouds	What is the current percentage of clouds?	Percent in 10% steps between 0% and 100%
Type of clouds	Which type of cloud is visible?	Text answer

The survey software ControlSurvey was used to question the participants during the trials. ControlSurvey was developed by the German Aerospace Center. It allows a synchronized questioning at multiple workplaces with questions that are prepared in advance. The software allows the preparation of a question catalog and setting time intervals for these types of questions. It also allows flexible changes to comprehend minor deviations from planned scenarios. The questions can be presented at multiple workplaces at the same time.

Design

The main aspect of the validation experiment is the comparison of the CWP-tower and CWP-remote to test the indicators. The experimental environment was not certified to perform actual air traffic control; therefore a passive shadow-mode field trial with synchronized questioning for both workplaces was conducted. An ATCO on each workplace monitored the weather within the aerodrome and answered questions (Table 2) at the same time. The independent variable within is the workplace of the ATCOs.

Figure 3 shows the relation of all components of the experimental set-up. At each workplace an experimenter read the questions to the participants and collected the answers afterwards. ControlSurvey activated the same questions simultaneously for both workplaces. The scenario was not predefined and the order and timing of questions was randomized. Only the type of question and

the frequency within the scenario was determined in advance. The minimum period of time between two questions was 30 seconds, which could be prolonged if the question was not answered within this timeframe. The randomization reduces the influences of the changing weather in every run and the synchronized questioning ensures that the same weather situation prevails on both workplaces.

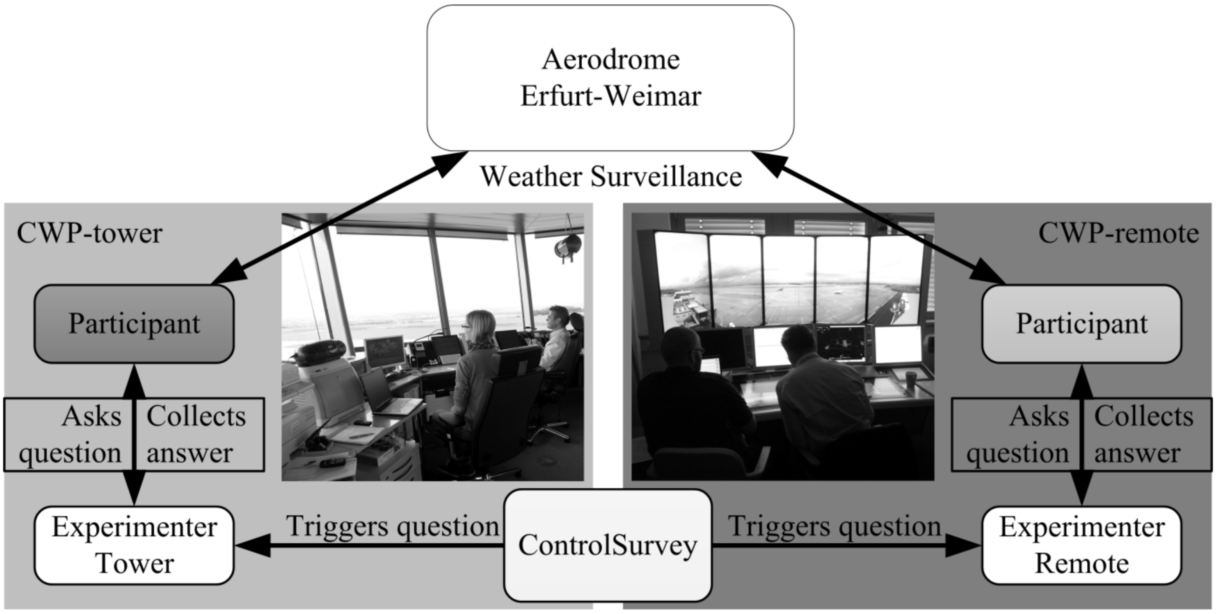

Figure 3 CWP-tower, CWP-remote, and ControlSurvey as experimental set-up

The design enables a direct comparison between the two workplaces depending on the measured difference in answers given by the participants. This increases the objectivity of the RTMs (Table 1). The influence of the CWP-remote on the monitoring of weather is quantified by the difference in indicators (Table 2). The differences are measured as the correctness of the given answer, their response times, and their change in used source of information.

Procedure

Following the study design, two participants were tested simultaneously. Therefore the eight participants were randomly divided into groups of two. The four groups were evenly distributed onto the validation days between the 17th and the 20th of July in 2012. Two runs were conducted each day. Before the first run, the participants were randomly assigned to either one of the workplaces. They switched workplaces before the second run. Throughout the experiment the participants did not interact with any traffic within the aerodrome or performed other air navigation services.

The daily schedule was identical for all validation days. Each morning, the new participants were briefed about remote tower operation (RTO) and the validation experiment. Then the groups

were split up and positioned at their assigned workplaces. The participant at the CWP-remote had to complete a 30 minute pan-tilt zoom camera training. Afterwards the first run was conducted. The participants switched workplaces and the second pan-tilt zoom camera training was completed before the second run. The runs had an equal length of 140 minutes. At the end of each day, the participants completed a final debriefing questionnaire with an average duration of 60 minutes.

After the start of each validation run, ControlSurvey randomly triggered one question after another to the experimenters of the workplaces. Comparable sets are generated by the synchronised occurrence of the questions on each workplace, independent of the RTM. The experimenter asked the question provided by ControlSurvey and typed in the participants' answers. Response time was measured as the time between the presentation of the question to the experimenter and the input of the answer. The influence of the experimenter on response times was minimized by a reading and typing training in advance to the validation experiment. A set of comparable questions was only generated if both participants answered. In addition, Fürstenau, Mittendorf, and Friedrich (2014) conducted a conservative analysis including sets that do not have both answers. Each question from Table 2 was followed by the question concerning the source of information.

Results

Before going into detail with the results, it has to be stated that the process of identifying RTM validating the indicators was performed with German ATCOs. Also, the airport Erfurt-Weimar is controlled by the German air navigation service provider. In the first part, we present a detailed analysis of the indicator Windsock, to give an example of the possibilities for basic data analysis. In the second part, we evaluate the proposed indicators in terms of expense for realization, comparability, and feasibility. This is necessary to validate the weather RTMs for future experiments. As mentioned above, this article focuses on the validation of weather RTMs that can be used in the future to simplify the process of validating a CWP-remote. Throughout the complete validation the visual range was always CAT I.

A total of 81 comparable sets of questions were collected during the validation. Each question concerning the weather RTMs (Table 1) was asked at least once per run. An average of 15 questions

sets were collected per trial. Within all runs, 18 sets were not completed and therefore rejected. Table 3 presents a detailed overview of the amount of sets collected that are comparable.

Table 3 The amount of collected sets that are comparable sets per indicator

Indicator		Amount of collected sets	Amount of comparable sets
Windsock	Windsock 1	5	3
	Windsock 2	7	6
	Windsock 3	10	9
	Windsock 4	6	6
Range of sight		14	13
Speed of wind		17	14
Intensity of rainfall		9	6
Percentage of clouds		15	12
Type of clouds		16	12

Basic analysis of Windsock indicator

The following section gives a sample analysis that can be performed using a single Windsock indicator. The analysis concentrates on the correctness of the answers, the response time of the participants and the used sources of information. Four windsocks with different distances to the tower were used. The distances are 200m (windsock 1), 750m (windsock 2), 850m (windsock 3), and 1750m (windsock 4). The windsock positions were dependent on the layout of the airport. In total 24 comparable sets of questions were collected throughout the validation experiment (Table 3).

Table 4 summarizes the results in regards to the correctness of answers and the response times on both workplaces. The amount of comparable sets for each of the windsocks is available in Table 3. The sets were used as a direct comparison between the CWP-tower and CWP-remote. For the purpose of this analysis we consider the CWP-tower as status quo for monitoring weather. Therefore, CWP-tower was considered as always correct, because only the deviation between the workplaces is relevant for the RTMs. Therefore, the CWP-remote answer was correct if it was equal to the CWP-tower answer.

The synchronized questioning also facilitates a valid analysis of the response times. The response time within the validation depended on the time an experimenter needed to read the question

out loud and the participant to answer it. The response times for the Windsock indicators are also analyzed separately for each windsock. The deviation between response times was calculated by using a paired t-Test. The difference between both workplaces is significant $t(22)=-4.19$, $p<.001$, whereas the CWP-tower participants are faster with an average of 9.6 seconds ± 2.2 (SEM). This analysis allows a detailed view on the response times and the influence CWP-remote has on them. Cohen's effect size value ($|d| = 1.23$) suggested a moderate to high practical significance. This analysis is possible for all indicators, if synchronized capturing is used.

Table 4 Correctness of answer and response time for the windsock indicator

	Workplace	Windsock 1 (200m)	Windsock 2 (750m)	Windsock 3 (850m)	Windsock 4 (1750m)
Correctness of answer	CWP-tower	100% (SD=0)	100% (SD=0)	100% (SD=0)	100% (SD=0)
	CWP-remote	100% (SD=0)	83% (SD=22)	100% (SD=0)	100% (SD=0)
Response time	CWP-tower	11.4s (SD=2.3)	14.8s (SD=8.9)	9.0s (SD=4.5)	11.6s (SD=9.1)
	CWP-remote	22.3s (SD=6.9)	22.9s (SD=13.5)	19.8s (SD=8.8)	23.3s (SD=10.3)

Figure 4 presents the sources of information used to identify the windsocks. The sources of information were categorized as panorama (out-the-window or video-panorama), magnification (binocular or pan-tilt zoom camera), weather information system, or a combination of those systems. At the CWP-remote the main source of information is the magnification independent from the distance. In contrast, the sources of information at the CWP-tower are a mixture of panorama and magnification. Comments from both workplaces indicated that the speed of wind is also important because an inflated windsock is easier to identify. These comments are used to evaluate the indicator for its feasibility in the next part of the results section.

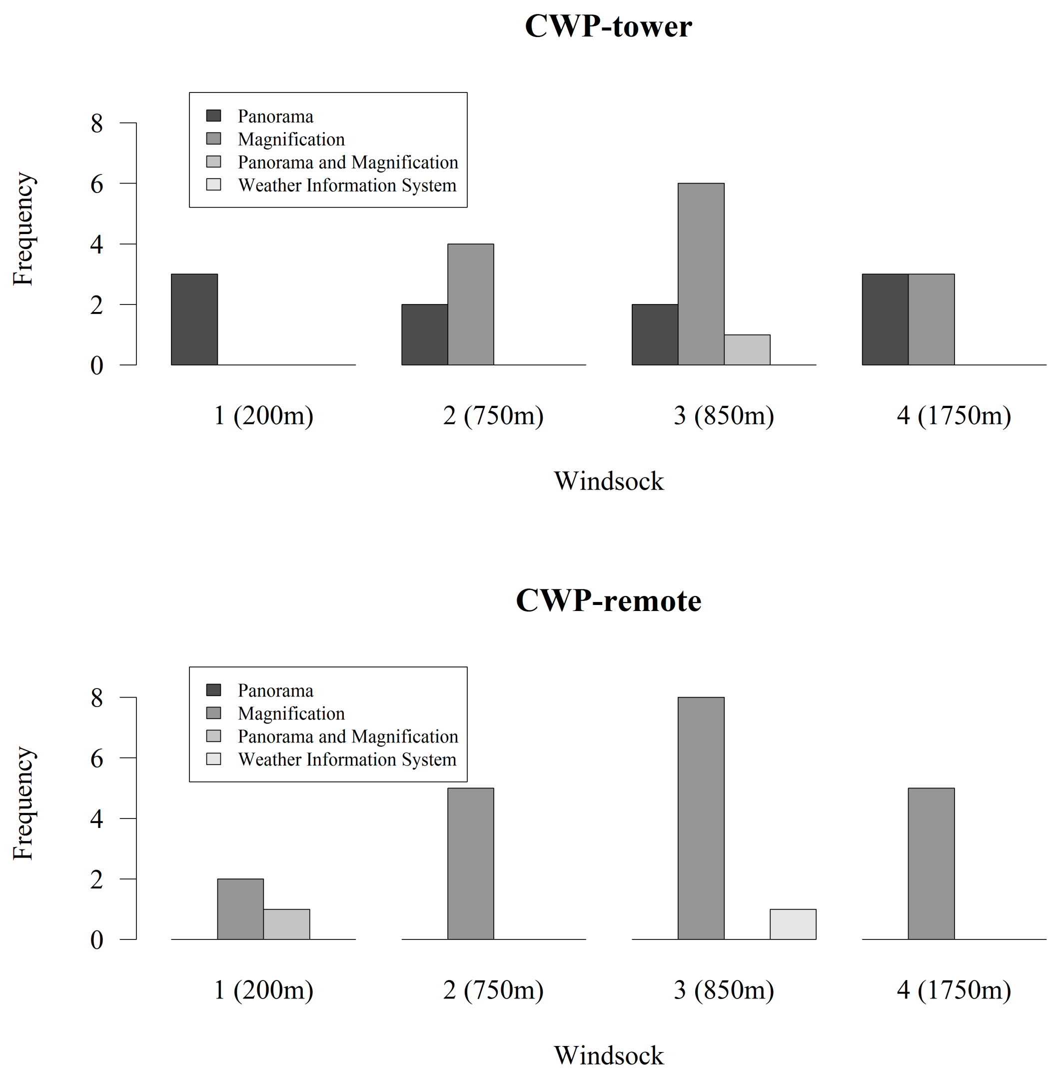

Figure 4 Sources of information used to answer the windsock indicator.

Evaluation of the weather indicators

The weather RTMs are all relevant to the work of an ATCO and therefore the evaluation of the indicators (Table 3) builds upon the aspects feasibility, comparability, and effort. Those aspects were already applied to the existing RTMs (Friedrich, 2016). The first method for evaluating the different aspects against each other is the standardization of effort, comparability, and feasibility and summarized them into the RTM score (Friedrich, 2016). The second method is an evaluation table for weather indicators.

Feasibility describes how feasible the weather indicator is for the ATCO during the validation. The participants rated feasibility within the debriefing, using a 6-point Likert Scale (1=

totally not feasible; 6= totally feasible). This evaluation is extended by comments from the experimenters concerning the practical feasibility during the validation trails. A higher feasibility value means that the indicator is easy to implement and applies to each validation environment. Table 5 shows the summarized feasibility values and the comments provided by the experiments concerning the feasibility. All indicators are above the scale average of 3.5.

Table 5 Feasibility values from the debriefing questionnaire and the experimenter comments.

Indicators	Feasibility (N=8)	Comments by the experimenter
Windsock	5.50 (SD=0.53)	Windsock 4 was not visible in the video-panorama.
Range of sight	5.25 (SD=0.71)	
Speed of wind	5.00 (SD=0.93)	Indicator was always answered with the weather information system.
Intensity of rainfall	4.50 (SD=0.53)	Rainfall was not available at every run.
Percentage of clouds	5.75 (SD=0.46)	
Type of clouds	5.63 (SD=0.52)	

The technical implementation is judged best when the same sources of information are used in both workplaces. Comparability is defined as the correlation between sources of information used at both workplaces to answer the indicators. The sources of information taken into account are panorama, magnification, and weather information system. The correlation is high if the participants use similar sources of information and low if they switch sources of information to answer the questions. Only the pairs with correct answers were considered. As additional information for a low correlation the major switching tendency ($r > 0.5$) was distinguished by correlating the three sources of information pairwise against each other (from CWP-tower to CWP-remote). Table 6 shows Comparability and the major switching tendency for the indicators. The results show that speed of wind, intensity of rainfall, and type of clouds have a strong comparability. In cases with low comparability either a major switching tendency could be identified or the dispersion between used sources of information was too high. An overall result is the tendency of the CWP-remote to switch away from the video-panorama to the magnification or weather information system.

Table 6 Comparability results for the indicators with major switching tendency

Indicators	Comparability	Major switching tendency
Windsock	0.28	From panorama and weather information system to magnification and weather information system
Range of sight	-0.05	From panorama to weather information system
Speed of wind	1	
Intensity of rainfall	0.79	
Percentage of clouds	0.28	
Type of clouds	0.87	

Effort describes the effort needed to realize the indicators. Friedrich (2016) needed an experimental vehicle and an aircraft to perform manoeuvres. Therefore, the weather indicators were rated depending on their financial effort and categorized in 3 (cost < €100 / per run), 2 (cost < €1000 / per run), or 1 (cost > €1000 /per run). All six indicators were graded with 3, because they relate to external influences that cannot be created on purpose. Also, no additional equipment is needed to generate situations to capture them which lead to no effort to use them as comparison between CWP-tower and CWP-remote.

The RTM score was calculated using e.q.1, to rate and compare the indicators against each other. At first feasibility and comparability were standardized with the sample mean and sample standard deviation to allow norm-referenced interpretation of all values. All three are defined in a way that a higher value means a positive influence on the validation. Since effort was equal for all indicators its influence for the RTM score is 0 and therefore it was not taken into account.

$$RTM\ score = \frac{(Z_{Feasibility} + Z_{Comparability} + Z_{Effort})}{3} \quad \text{e.q.1}$$

Figure 5 presents the RTM score for each of the six indicators. The results show that each indicator has a different RTM score to rank their quality of each metric.

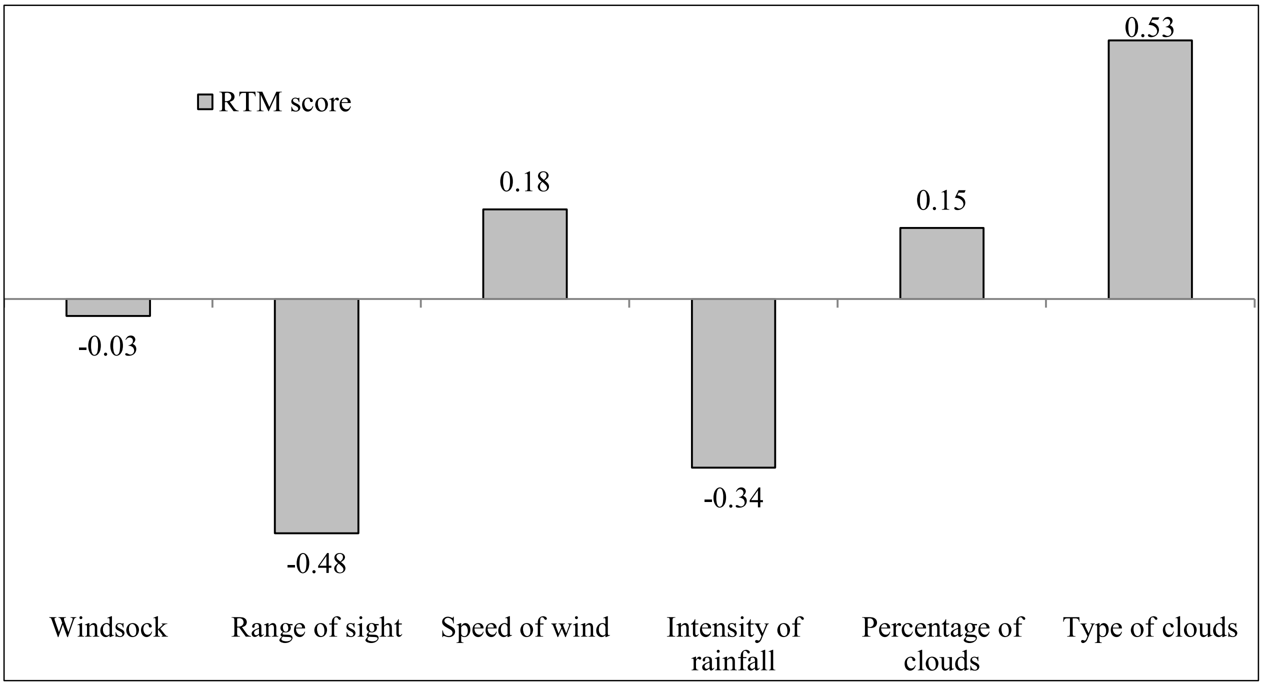

Figure 5 The ranking of all RTMs in relation to their RTM score.

The final analysis summarizes feasibility, comparability, effort, and the source of information to simplify the evaluation of different indicators against each other. Table 7 presents the indicators' evaluation table. The results for feasibility, comparability, effort are summarized from Table 5, Table 6, and the effort evaluation. Feasibility is considered a YES if it was above the scale average. Comparability is considered YES if the correlation was strong or a major switch tendency could be identified. Effort is considered YES if the effort is low or medium.

Table 7 Evaluate your RTO-System (with regard to weather RTM)

Indicator	Feasibility (Yes;No)	Comparability (Yes;No)	Effort (Yes;No)	Source of Information	Use this metric (if all are yes)
Windsock	Yes	Yes	Yes	weather information system and magnification	Yes
Range of sight	Yes	Yes	Yes	weather information system	Yes
Speed of wind	Yes	Yes	Yes		Yes
Intensity of rainfall	Yes	Yes	Yes		Yes
Percentage of clouds	Yes	No	Yes		No
Type of clouds	Yes	Yes	Yes		Yes

Discussion & Conclusion

The results provided in the previous section have to be interpreted with a strong connection to German Air Traffic Control. The two analyses showed the variety of indicators and their quality to evaluate a CWP-remote against a CWP-tower. The windsock indicator explained the possibility to distinguish the Air Traffic Control Officers' (ATCO) performances with regards to a single air traffic control task. The evaluation of the weather indicators showed that independent from their relative importance to the task, they can vary in quality to discriminate between workplaces. The question stands, how they can significantly support the process of evaluating the CWP-remote.

Basic analysis of Windsock indicator

The indicator represents the air traffic control task to determine the wind direction performed at CWP-tower and CWP-remote and is an example of the opportunity provided by each indicator. The analysis of the windsock indicator provides an overview how the weather indicators can be used to measure performance on different workplaces.

The results show no significant change in correctness of the answers, but a prolonging of reaction times at the CWP-remote. Since air traffic control tasks can also be time critical in some

situations this result indicates a decrease in overall performance. The main reason for the decrease in reaction times might be found on the different sources of information used to gather the answer. This result is a small insight into the effect that the CWP-remote has on a single air traffic control task and needs to be incorporated into a broader spectrum of tasks. Also the overall correctness of CWP-tower could elevate in future analyses with improved technical assistant systems. As mentioned above this article does not focus on the implications for RTO. Instead we were able to show that the basic analysis of the windsock indicator can help to show the differences between both workplaces for one selected task of monitoring. This approach needs to be extended for all weather related tasks.

Evaluation of the weather indicators

As indicated in the previous section, a valid conclusion about the difference between the two workplaces can only be reached with a number of different RTMs and indicators that cover all aspects of the air traffic control task. The weather RTMs extent the existing RTMs to cover all aspects of weather that are relevant for the air traffic control task. Therefore, the evaluation of the weather RTMs and their indicators is important to increase and sort their validity. This does not only apply to the CWP-remote presented in this article but might hopefully be applied for CWP-remote systems that will be designed and tested in the future.

Feasibility quantifies how well the indicator fits into the procedure of the validation. This is shown by the subjective rating of the ATCOs but also by the comments of the experimenters. We consider the feasibility classification also as a learning indictor for further changes to the indicators before finalization. The results show that all weather indicators can be applied to a validation without great effort. The major problem is seen in the comment that some weather phenomena (e.g. rain or clouds) are not always available and therefore the indicator cannot always be applied.

The results for comparability show that the comparison between the two workplaces is not possible based on single systems. One example result is that the participants moved away from the panorama and to the weather information system. This can be used to rate different systems against each other. Comparability is only above 0.75 for three weather indicators. The other indicators show a transition in systems used to gather the information. Of course, this is connected to some of the

comments from the experimenters evaluating feasibility. Alterations in the used sources of information are considered to reduce comparability.

The effort classification of indicators is not relevant for weather indicators, because their effort is equal. The weather itself is unpredictable and therefore the effort is mostly dependent on the experimenter to initialize the correct weather indicator if the real situation fits.

The RTM score as a summarized value between the weather indicator show that there are differences in quality and validity. Some weather indicators are better than others, and if the system developer has to choose between some of them, Friedrich (2016) proposed the selection should remain on the one with the highest RTM scores. In general this leads to a proposal of a set of indicators that need to be defined for the evaluation of RTO. This process of selection among the highest RTM scores does not incorporate the airport and its dependency to weather influences and is therefore inflexible to the adaptation to a new airport. Therefore, the authors present a RTM scores table (Table 7) to simplify the selection of weather metrics and indicators. This way, it is more flexible to the demands of different environments. The table summarizes the evaluation of the weather indicators into one single overview that is easy to understand; and also flexible enough to be extended for all RTMs that are available at the moment (Friedrich, 2016) and might come in the future.

Outlook

In line with the results of the accomplished validation exercises under the operational focus area “Remote Tower”, the evaluation of weather RTMs within this validation exercise provides an additional step for the RTO validation. The article focused on the weather RTMs and their indicators rather than the results of the validation exercise itself. The validation shows that indicators can be judged differently depending on their quality to distinguish between different systems. This article extends the existing RTM with the concept of weather RTMs and therefore lays the foundation for a systematic validation of RTO in the future. Taking the strong connection to German ATC into account the process of identifying different RTM is not completed and has to be updated and extended to different procedures that exist throughout Europe’s different Air Navigation Service providers. The necessity of this process is increased by the permanent transformation of air traffic control concepts,

e.g. the introductions of Resilient Synthetic Vision to the tower (Masotti, Bagassi, & De Crescenzio, 2016).

The next step is to bring the weather RTMs and the existing RTMs together into one single overview and use them to evaluate different remote tower workplaces and compare their results. This helps to identify the pros and cons of every system in relation to the real work environment and therefore supports the task of developing an adequate workplace replacement at a remote location. Therefore the authors propose to use the RTMs and indicators presented in this article for testing them with different prototypes and determine the agreement of the quantitative results.

References

Collins, R. L. (1991). *Mastering the systems: Air Traffic Control and weather*: Wiley.

Committee Sesar Program. (2010). *Overview and Release 1 Plan v1.0*, Brussel.

Deutscher Wetter Dienst. (2017). Deutscher Wetterdienst - Airlines, airports. Retrieved from https://www.dwd.de/EN/specialusers/aviation/cg_airport/airport_node.html

Ellis, S. R., Fürstenau, N., & Mittendorf, M. (2011). Frame Rate Effects on Visual Discrimination of Landing Aircraft Deceleration: Implications for Virtual Tower Design and Speed Perception. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 55*, 71-75. doi:10.1177/1071181311551015

Ellis, S. R., & Liston, D. (2008). Visual Features Involving Motion Seen from Airport Control Towers. *11th IFAC/IFIP/IFORS/IEA*.

FREQUENTIS. (2015). *DFS vertraut bei der Umsetzung des revolutionären Remote Tower-Systems an drei deutschen Flughäfen auf die FREQUENTIS-Lösung smartVISION*, Wien.

Friedman-Berg, F. (2012). *Staffed Next GenTower Human-in-the-Loop Camera Integration Evaluation*.

Friedrich, M. (2016). Which metrics provide the insight needed? A selection of remote tower evaluation metrics to support a remote tower operation concept validation. In N. Fürstenau (Ed.), *Virtual and Remote Control Tower* (1 ed.). Berlin, Heidelberg, New York: Springer Publishers.

Friedrich, M., & Möhlenbrink, C. (2013). Which data provide the best insight? A field trial for validating a remote tower operation concept. *Tenth USA/Europe Air Traffic Management Research and Development Seminar*.

Fürstenau, N., Schmidt, M., Rudolph, M., Möhlenbrink, C., & Halle, W. (2008). Augmented vision videopanorama system for remote airport tower operation. *Proc. ICAS 2008, 26th Int. Congress of the Aeronautical Sciences.* Anchorage.

Fürstenau, N., Schmidt, M., Rudolph, M., Möhlenbrink, C., Papenfuss, A., & Kaltenhäuser, S. (2009). Steps Towards the Virtual Tower: Remote Airport Traffic Control Center (RAiCe). *ENRI Int. Workshop*. Tokyo.

Hannon, D., Lee, J., Geyer, M., Mackey, S., Sheridan, T., Francis, M., . . . Malonson, M. (2008). Feasibility Evaluation of a Staffed Virtual Tower. *Journal of the Air Traffic Control Association*, 27-39.

Krozel, J., D. Andre, A., & Smith, P. (2003). *The future national airspace system: design requirements imposed by weather constraints*. Paper presented at the Guidance, Navigation, and Control, Austin, Texas.

Masotti, N., Bagassi, S., & De Crescenzio, F. (2016). Augmented Reality for the Control Tower: The RETINA Concept. In L. T. De Paolis & A. Mongelli (Eds.), *Augmented Reality, Virtual Reality, and Computer Graphics: Third International Conference, AVR 2016, Lecce, Italy, June 15-18, 2016. Proceedings, Part I* (pp. 444-452). Cham: Springer International Publishing.

Mensen, H. (2014). *Moderne Flugsicherung: Organisation, Verfahren, Technik (VDI-Buch)*: Springer Vieweg.

Möhlenbrink, C., Friedrich, M., & Papenfuss, A. (2009). RemoteCenter: Eine Mikrowelt zur Analyse der mentalen Repräsentation von zwei Flughäfen während einer Lotsentätigkeitsaufgabe. *Tagungsband 8. Berliner Werkstatt Mensch-Maschine-Systeme*. Berlin: Berlin-Brandenburgische Akademie der Wissenschaften.

Möhlenbrink, C., Rudolph, M., Schmidt, M., & Fürstenau, N. (2007). Wahrnehmungsexperimente im RApTOr Demonstrator. *2. RTO-Workshop, Poster presentation*. Braunschweig.

Mullan, C., Lindqvist, G., Svensson, T., Abel, M., & Ankartun, P. (2012). *OSED for Remote Provision of Air Traffic Services to Aerodromes, including Functional Specification for Single & Multiple aerodromes*, Brussel.

Nolan, M. S. (2010). *Fundamentals of Air Traffic Control* (Fifth Edition ed.): Wadsworth Publishing Company, Belmont, CA.

Saab Security. (2008). *Remotely Operated Tower -The Future of Being Present*.

Schmidt, M., Rudolph, M., Werther, B., Möhlenbrink, C., & Fürstenau, N. (2007). Development of an augmented vision video panorama human-machine interface for remote airport tower operation. Paper presented at the Human Interface and the Management of Information. Interacting in Information Environments, Berlin, Heidelberg.

van Schaik, F. J., Lindqvist, G., & Rössingh, H. J. M. (2010). Assessment of visual cues by tower controllers. *11th IFAC/IFIP/IFORS/IEA*. Valenciennes, France.

Wickens, C. D., Mavor, A. S., & McGee, J. P. (1997). *Flight to the future: Human factors in air traffic control*: National Academies Press.

III Article 2: A Guideline for Integrating Dynamic Areas of Interests in Existing Set-up for Capturing Eye Movement: Looking at Moving Aircraft

The following version is a pre-copyedit version of an article published in Behavior Research Methods. The final authenticated version is available online at: http://dx.doi.org/10.3758/s13428-016-0745-x

Reprinted by permission from Copyright Clearance Center: Springer Nature, Behavior Research Methods: **Citation:** Friedrich, M., Rußwinkel, N., & Möhlenbrink, C. (2016). A guideline for integrating dynamic areas of interests in existing set-up for capturing eye movement: Looking at moving aircraft. Behavior Research Methods, 49(3). doi: http://dx.doi.org/10.3758/s13428-016-0745-x

Journal: Impact Factor by the time of online publication (JCR Social Science Edition 2016) was 3.62.

A Guideline for Integrating Dynamic Areas of Interests in Existing Set-up for Capturing Eye Movement. - Looking at Moving Aircraft

Maik Friedrich[1] Nele Rußwinkel[2] Christoph Möhlenbrink[1]

7277 words (body)

1 - German Aerospace Center, Braunschweig

2 – Technische Universität Berlin

Author:

Maik Friedrich, maik.friedrich@dlr.de
+49 (531) 295 - 2598
German Aerospace Center
Lilienthalplatz 7
Braunschweig, NI D-38108

Abstract

Today, capturing the behavior of a human eye is considered a standard method for measuring the information gathering process and thereby gaining insights into cognitive processes. Due to the dynamic character of most task environments there is still a lack of a structured and automated approach for analyzing eye movement in combination with moving objects. In this article, we present a guideline for advanced gaze analysis, called IGDAI (Integration Guideline for Dynamic Areas of Interest). The application of IGDAI allows gathering dynamic areas of interest and simplifies its combination with eye movement. The first step of IGDAI defines the basic requirements for the experimental setup including the embedding of an eye tracker. The second step covers the issue of storing the information of task environments for the dynamic AOI analysis. Implementation examples in XML are presented fulfilling the requirements for most dynamic task environments. The last step includes algorithms to combine the captured eye movement and the dynamic areas of interest. A verification study was conducted, presenting an air traffic controller environment to participants. The participants had to distinguish between different types of dynamic objects. The results show that in comparison to static areas of interest, IGDAI allows a faster and more detailed view on the distribution of eye movement.

Keywords: eye movement, dynamic area of interest, experimental setup, requirements

Introduction

The interaction and observation of dynamic objects is one of the most common actions of intelligent life. Humans routinely analyse their surroundings to find and react to dynamic objects. Example tasks are manoeuvring in traffic or playing basketball. Within those task environments, analysing the movement of dynamic objects, like a car or a ball, helps us to evaluate our current situation. This enables us to select an action to enhance the outcome of our nearest future, such as braking in front of a traffic light or catching the ball. The knowledge on the information gathered to perform a task therefore increases the insight into cognitive processes. The motion of the eye in combination with the used sources of information is defined as eye movement (Seifert, Rötting, & Jung, 2001). This article describes an Integration Guideline for Dynamic Areas of Interest (IGDAI) for capturing and integrating dynamic objects into an experimental setup, and combining them with eye movement.

Over the past decades, human computer interaction analysis has gained in attractiveness due to the fact that the best product does not only have the best functionality but also an intuitive control concept. This also led to an increase in popularity for eye movement analysis, as it is an objective method for measuring the information gathering process of humans interacting with interfaces. Eye movement analysis has been used in a variety of domains, such as reading (e.g., Castelhano & Rayner, 2008; Granka, Joachims, & Gay, 2004; Rayner, 1995) or image processing (e.g., Henderson, Weeks, & Hollingworth, 1999; McCarthy, Sasse, & Riegelsberger, 2004; Rayner & Pollatsek, 1992). The first level analysis of eye movement data usually focuses on the separation between fixations and saccades. Second level analysis concentrates on summarizing first level results in relation to the task environment, such as the combination of an area of interest (AOI) with distribution of eye movement, scan paths or transition matrices. AOIs are defined regions where a certain information is presented within the task environment (e.g. screen) (Duchowski, 2007).

The first level analysis requires an algorithm for separation of raw eye movement data into fixations and saccades. For that matter numerous algorithms are already well documented and can be divided by real-time versus post analysis (Duchowski, 2007) and velocity-, area- or dispersion-based

(Salvucci & Goldenberg, 2000). However, in a dynamic task environments smooth pursuits (Reimer & Sodhi, 2006) restrains a clear separation between fixations and saccades.

The second level analysis usually includes information about the task or the environment and combines these with the eye movement. A common approach is the application of AOI. Combining AOI with eye movement allows insights into the information gathering process within the environment, such as the meaningfulness and visibility of a status bar integrated into a website. The maximum amount of information gathering is reached by combining all visible information and all actions performed by the human. Gathering this kind of data normally takes a lot of planning and consideration of the task and the environment, especially in the case of a dynamic environment with moving objects.

This article focuses on the combination of eye movement with dynamic AOI within artificial environments. Studies combining complex dynamic environments and eye movement occur frequently in the domains of driving (e.g., Akamatsu et al., 2001; Palinko, Kun, Shyrokov, & Heeman, 2010), marketing (Bulling & Gellersen, 2010; Josephson & Holmes, 2006; Nielsen, 2006) or air traffic management (Bruder, Leuchter, & Urbas, 2003; Landry, Sheridan, & Yufik, 2001; Möhlenbrink, Oberheid, & Werther, 2008). These studies use eye movement within artificial environments that comprise moving objects that are partially or completely controlled by participant actions, such as the view out of a moving car, interaction with web pages or air traffic controller interfaces. The moving objects within these different environments are of importance to the tasks. However, these studies concentrate on the application of dynamic AOI and do not aim to improve the method itself, some of them even do not publish their method. Comments in the conclusions of those studies state that combining eye movement data with dynamic objects is extremely time consuming and specialized only for the described environment (e.g., Papenmeier & Huff, 2010).

As described above, regions within the task environment that are a source of information and useful to the eye movement analysis are considered as AOI. Attributes to describe an AOI within a task environment are identifier, shape and position. The identifier is unique for each AOI within a task environment. Shape defines the area of the AOI. Position defines the centre of the shape in relation to

the task environment. The static and dynamic AOI can both be defined using identifier, position and shape. The distinction between static and dynamic AOI can be explained using a coordination system as reference for the areas over time. Static AOI have the same position, shape and size in relation to the reference system at all times. Dynamic AOI change their position, form, orientation, or size in relation to the reference system over time. An example for a dynamic task environment is a radar screen because it contains both types of AOIs. A radar screen normally displays a two dimensional top view of airspace sectors. Assuming that the radar screen is fixed over a particular sector, it provides a two dimensional coordination system as reference for AOIs. The radar screen sometimes shows motor highways, rivers or forests, to increase the orientation for the air traffic controller. These areas never move in relation to the screen and therefore represent static AOIs. The radar screen also shows aircraft that move through the sector. The moving aircraft are AOIs that change their position in relation to the screen and are therefore dynamic.

Considering the differences between static and dynamic AOIs, an analysis of eye movement data in combination with only static AOIs would not provide a comprehensive view on the used sources of information. The solution is to define dynamic AOIs for moving objects and combine them with eye movement data in post processing. Thereby the separation of the dynamic information and the static background becomes possible and increases the significance of the data analysis.

The radar screen example above demonstrates the three main issues for applying dynamic AOIs. The first issue (Set-up requirements) is used to configure the environment to capture the needed information. The second issue (Storage requirements) is used to identify the information that have to be recorded and derive a suitable data structure. The third issue (Data Analysis Methods) is used to match eye movement and AOIs. All three issues depend on each other; this means that the solution to the first issue inflicts the solution to the second and third issue. An overall approach that addresses all issues and is universal for the application to any task is needed for applying dynamic AOI’s to any environment. Below an overview of existing approaches for combining eye movement and dynamic environments is provided.

Existing approaches

Research connected to dynamic AOIs ranges from theoretical descriptions (e.g., Bruder et al., 2003) to the application of methods that are customized to the studied task or environment (e.g., Papenmeier & Huff, 2010). The three issues identified above are usually covered in the method section by shortly describing the individual solutions. Only a few publications try to give an overall view covering all three issues connected to dynamic AOIs. To assess the different approaches we have to investigate how the three issues are solved.

The approaches can be categorized into online and offline. Applying dynamic AOIs online means to match the gaze vector with the dynamic objects within the environment at the time the eye movement is recorded. In contrast, recording dynamic AOIs offline means the matching is made in a post processing step. One example for an online matching of eye movement and dynamic AOIs is presented by Sennersten et al. (2007). They combined an eye tracking system (Tobii) and a game engine (e.g. CryEngine (Crytek GmbH, 2016)) and transferred the gaze position and direction from the eye tracker online into the virtual environment generated by the game engine. Thereby they cover the Set-up requirements and the Storage requirements. Then they calculate where the eye vector hits the first surface within their scene. The name of this surface is then stored for further analysis. This covers the Data Analysis Method issue. Avoiding the post processing of eye movement and dynamic AOI's by online matching is time efficient. However, playback of the analysis is sometimes difficult because the complete virtual environment, all participant actions and eye vectors have to be stored. This problem is increased if the analysis has to be rerun, in case e.g. that some objects (e.g. windows) are to be removed from the analysis (Papenmeier & Huff, 2010) or the calibration of the eye tracker had a small divergent. The online matching of dynamic AOIs also has the disadvantage of not applying any method for correcting the eye movement data (Hornof & Halverson, 2002; Zhang & Hornof, 2011). Due to the research presented above, the authors believe that at the moment, methods of replay and post correction are necessary in the field of eye movement analysis. Therefore, the offline matching of dynamic AOIs is pursued in the following research.

A straightforward approach for realizing offline matching of dynamic AOIs is to use an eye tracking system and a scene camera. Those combined generate videos that include the scene and the calculated eye position, also covering Set-up requirements and Storage requirements. The Data Analysis Method is covered while post processing the dynamic objects. They are identified offline and manually (e.g., Land & McLeod, 2000; Sennersten, 2004), by separated rating of every frame of the combined video. This approach is time consuming and subjective since the position of the eye in relation to dynamic AOIs can be interpreted differently.

BeGaze 3.0, an analysis software developed by SMI, also supports offline definition of dynamic AOI. Here the analyst can define key frames with AOI and the software interpolates the positions of the AOIs depending on the number of frames between the key frames. The dynamic AOI can even be stored for repetitive use. However, the interaction between operator and environment changes the most scenarios and creates an individual set of moving objects for each run (and therefore also AOI), therefore the repetitive use of BeGaze is limited. An automated procedure for synchronising video files with eye movement data is presented by Papenmeier and Huff (2010). They present DynAOI which matches the gaze position with three-dimensional models underlying the presented video files. The approach converts the captured eye movement data on a two dimensional screen into a three-dimensional representation of the events within the presented video files. The gaze vector within the three-dimensional representation then is used to determine all objects on its path, with the object then serving directly as AOIs. Huff, Papenmeier, Jahn, and Hesse (2010) use this approach with real time generated videos.

Bruder et al. (2003) propose an approach by presenting a set of technical requirements for capturing the dynamic information, covering the Set-up Requirements. The requirements include that eye movement is captured in relation to the screen on which the dynamic objects are presented. They cover the Storage Requirements by suggesting the application of markers, generated by the environment and stored with the eye movement data indicating a change of information on a screen.

Another approach for the Data Analysis Method, depending on a synchronized and separate recording of eye movement and dynamic AOIs, is proposed by Zelinsky and Neider (2008) using the

shortest distant rule for assigning the eye movement data on a screen to the presented dynamic AOIs. They calculate the distances between each measured eye movement point on the screen to each centre of the visible dynamic AOI. The dynamic AOI that is closest to the eye movement is then classified as "looked at". This approach can be implemented quickly and provides a good estimation of the results (Gross, Friedrich, & Möhlenbrink, 2010) if bigger dynamic AOIs further away do not cover smaller ones that are nearer to the eye movement. Fehd and Seiffert (2008) extended the shortest distant rule with a location competition analysis that calculates a weight for each moving target in relation to the nearest eye movement. The weights are accumulated over all frames to identify the targets. This approach will lead to a distribution of eye movement data on to dynamic objects but has sometimes difficulties if the participants do not look at any object.

IGDAI

IGDAI was developed as an efficient way to cover all three issues (Set-up, Storage, and Data Analysis) in a structured and flexible manner that is need for each individual eye tracking task environment. Because of the wide variety of different eye tracking tasks and environments, IGDAI does not determine a fixed solution. IGDAI addresses the main issues of integrating dynamic AOIs into an environment for capturing eye movement by defining requirements that need to be fulfilled. Therefore the implementation of the requirements are open and can be adapted to the individual environment but are at the same time interconnected by following IGDAI. IGDAI also includes an example implementation that fulfils the requirements, to raise understanding for the process.

The guideline begins by defining Set-up requirements (Figure 1 A, B, and C) that concern the process of data capturing to ensure the needed output information. Requirement D (Figure 1) covers the storage of the eye movement data with all necessary information. Then Requirement E (Figure 1) accounts for overlaid visual information and thereby supports the process of post analysis. IGDAI provides a data structure that can store dynamic information as dynamic AOIs (Figure 1 F). An example of how to store the necessary information for the post analysis is provided below. Taking all the previous steps into account, IGDAI describes the combination of eye movement and dynamic

AOIs (Figure 1 G and H) with the depth information. Practical solutions are provided for matching eye movement data with dynamic AOIs.

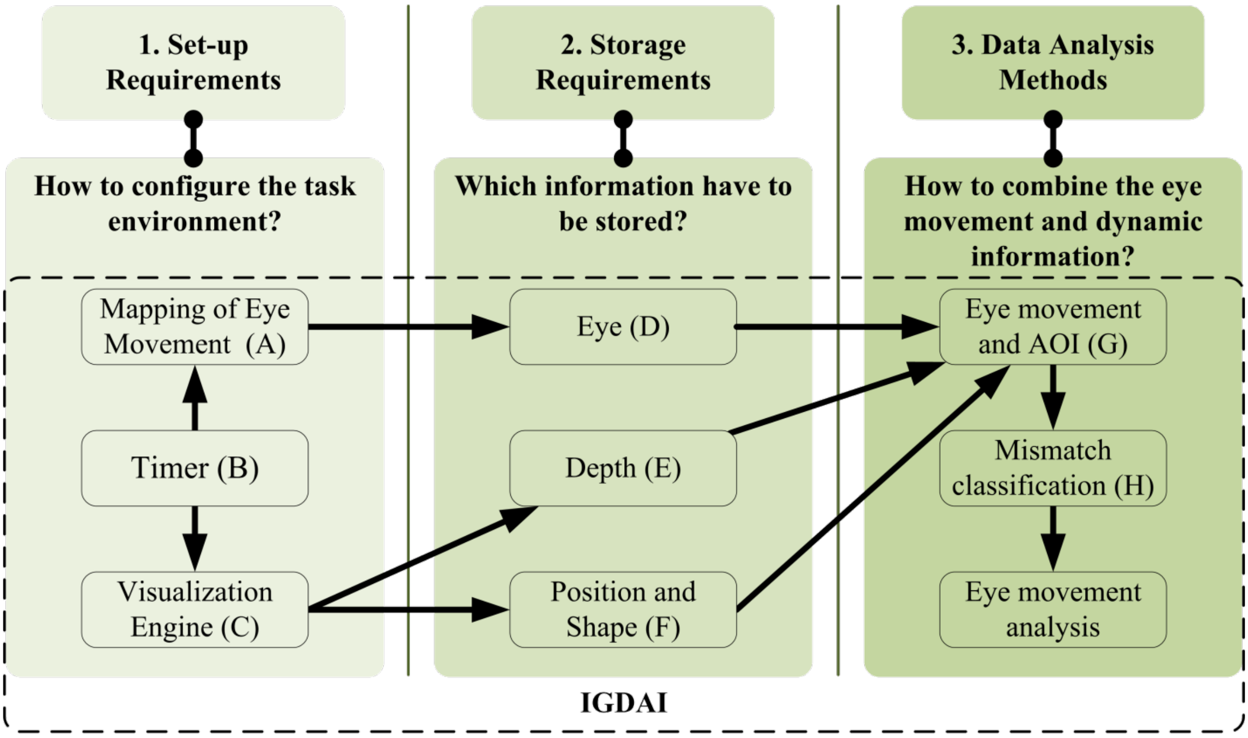

Figure 1 IGDAI separated into three issues and seven requirements.

Set-up requirements

The Set-up requirements (Figure 1 and Figure 2) are Mapping of Eye Movement (A), Timer (B) and Visualization Engine (C). Requirement A: The eye tracker shall have the capability to map the eye movement onto a reference system (e.g. a surface representing the computer monitor). Requirement B: The capturing of eye movement and dynamic information shall be time synchronized. Requirement C: The visualization engine shall log the dynamic information on the same reference system as used by the eye tracker. The main purpose of the Set-up requirements is to harmonize the reference systems between the eye tracker and visualization engine, in preparation for combining the eye movement data and the dynamic AOIs.

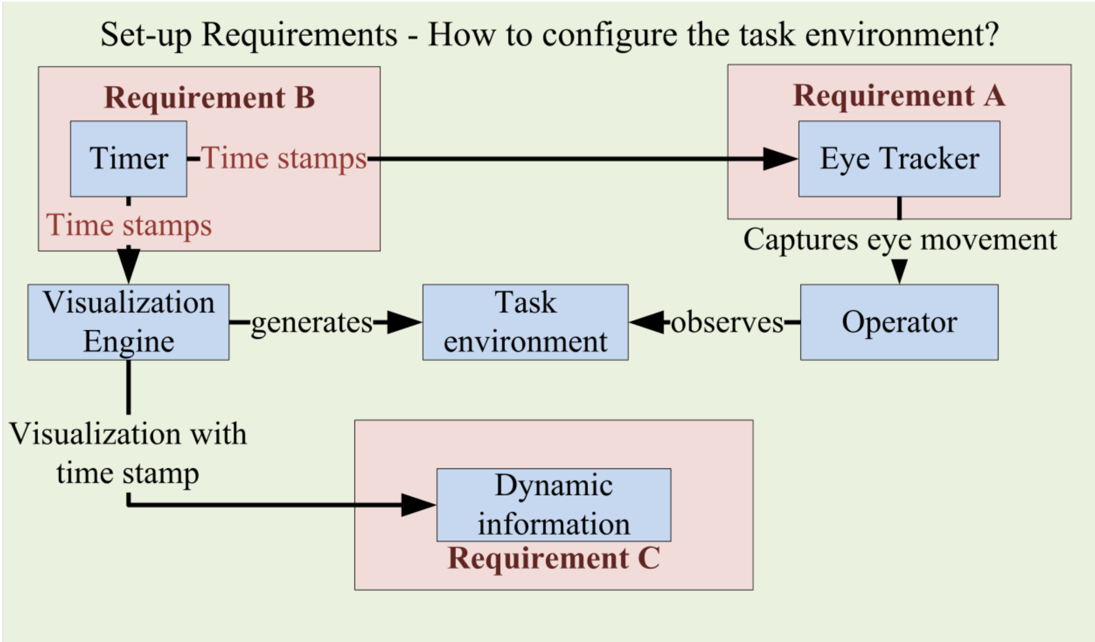

Figure 2 Set-up Requirements in relation to capturing of eye movement data and dynamic AOIs.

Figure 2 shows the components of a simplified set-up in relation to the requirements A, B and C, which we will now discuss in detail. As mentioned above, Requirement A concerns the capturing of the eye movement data in relation to a reference system. A solution to fulfil the requirement is to use virtual surfaces as reference systems and map the eye movement onto virtual surfaces that are for example equal to the screens in the environment. Therefore the eye tracker needs to have the capability of defining virtual surfaces for mapping the eye movement onto them. These virtual surfaces should be coherent with the actual monitors that present visual information. E.g. Figure 3 shows the virtual surfaces that cover a multi monitor workplace. The eye movement is determined as the intersection of gaze vector (vector with the eye position as origin) with the first virtual surface on its path. The result of this requirement is x, y and screen information of the eye movement.

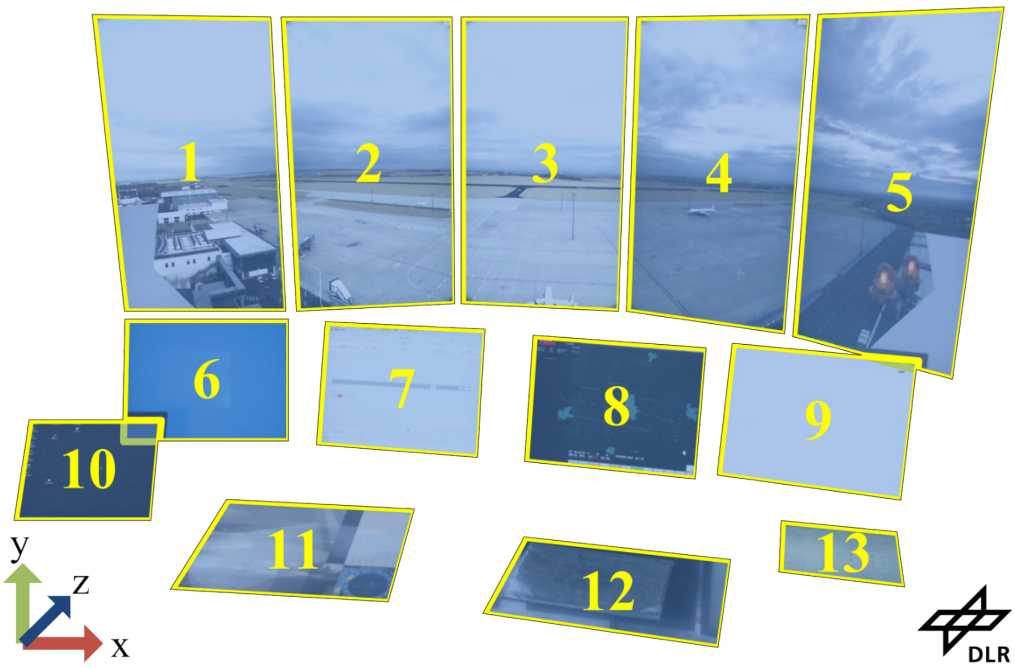

Figure 3 Multiple virtual surface setting for a multi monitor workplace (Friedrich, Möhlenbrink, & Carstengerdes, 2012)

Requirement B concerns the time synchronized capturing of eye movement and dynamic information. If the eye tracker and the visualization engine are separate systems, which they often are, their processes of capturing and storing the data need to synchronize. This requirement can be fulfilled through the application of a timestamp that is used by both systems. The introduction of a timestamp can be realized in several different ways. Figure 2 shows an example solution where an external timer provides the timestamp for both systems. An alternative implementation could be that either the eye tracker or the visualisation engine provides timestamps and distributes them to the respective other system. Dependent on the implementation the synchronized timestamps there are also different sources for time lags that also have to be taken in consideration. The three major sources are the processing time of the eye tracker, the time between logging and showing the AOIs on the screen and the duration to synchronise eye tracker and visualisation engine. The resolution of the timestamp should be in the range of milliseconds, otherwise the accuracy of the eye tracker (modern eye tracker reach up to 2000 Hz) is decreased. A rule of thumb, from the author's experience, for two systems is that the time deviation is less than 15 milliseconds for one minute. Therefore, synchronizing every 10

seconds should be enough for a 200 Hz eye tracking system. Independent from the implementation, the result of this requirement is that the presented visual information and the captured eye movement are synchronized by timestamps.

As Figure 2 shows, Requirement C concerns the generating of dynamic information in relation to the same reference system as the eye movement and with a logging rate. If the virtual surfaces (Requirement A) cover the screens in the task environment, as proposed, the reference systems for the dynamic AOIs should be the same screens. In addition to their identifier, position, shape, orientation and size, dynamic AOIs also have timestamps (Requirement B) depending on the logging rate. The logging rate depends on the context change of the task environment (e.g. 1Hz for an ATC radar workplace or 60 Hz for a driving simulator). For example if the task environment changes with 1 Hz, a logging rate of 1 Hz is enough because a higher logging rate would not generate more information but only duplicate the already logged one.

Figure 4 shows an example of the information needed to fulfil Requirement C. The two objects on the screen change their position once per second and are therefore logged with 1 Hz. The objects move for a total of three seconds. The time is synchronized with the eye movement. The name of the object is a unique identifier that is consistent over time. Each set of X and Y represent a corner of an object. The X and Y values are percentages of their particular axis (e.g. on a screen with 1920 pixel horizontal resolution X1 = 0.10 stands for pixel 192).

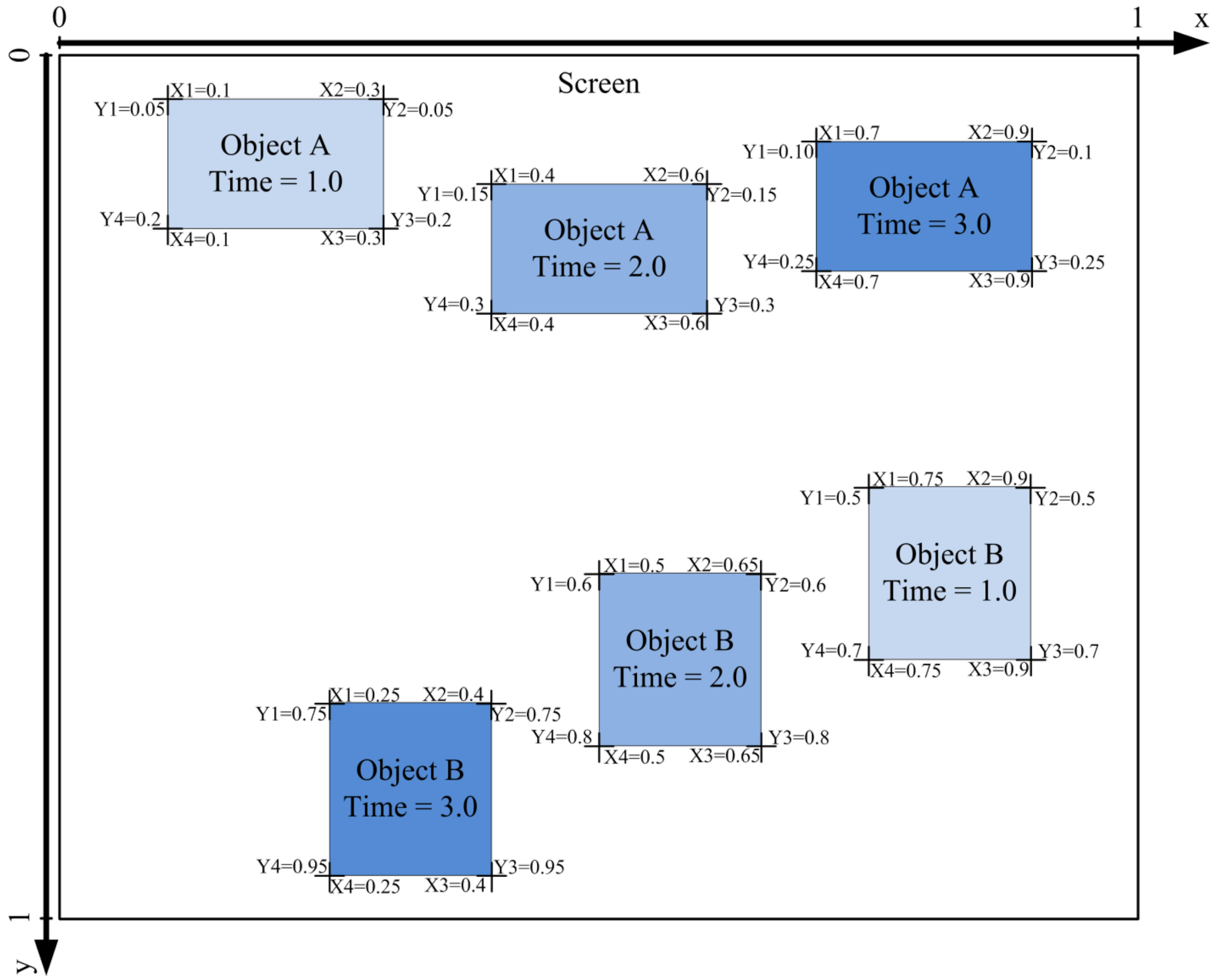

Figure 4 Two dynamic AOIs moving for three seconds in different directions

Requirements A, B and C are the necessary extensions to an existing eye movement environment. If fulfilled, the set-up captures the eye movement and the dynamic AOIs synchronized and in relation to the same reference system. This completes the first step (Figure 1) of IGDAI.

Storage requirements

IGDAI sets-up the task environments to record dynamic information, but also prepares the analysis process by structuring the storage of the dynamic information. The requirements for the storage of dynamic information are Eye (Figure 1, D), Depth (Figure 1, E) and Position and Shape (Figure 1, F). Requirement D: The eye movement data shall be stored with the timestamp. Requirement E: The dynamic information shall be stored with depth information to distinguish

between visible and hidden objects. Requirement F: The position and shape of dynamic information shall be stored as AOIs with a timestamp.

Requirement D concerns the storage of the eye movement data with the appropriate timestamp and in relation to the reference system (Requirement A). The result of Requirement A is the information of x, y and the screen of each eye movement point. Requirement D is fulfilled if this information are logged together with the timestamp of the capturing. The information can be logged as a new line into a log file. If stored correctly, this file can be loaded automatically and used for the data analysis.

Requirement E concerns the storage of AOIs that overlay each other within a task environment. This requirement can be fulfilled by storing the depth information of every AOI in relation to the reference system (Requirement A and C). In 3D task environments the depth information could be described as distance from the dynamic object to the reference system. The dynamic object with the least depth would be the visual object in the foreground unless it is transparent or semi-transparent. In 2D task environments no depth information is available. IGDAI suggests the usage of a layer model for a 2D task environment to put the dynamic objects in order in relation to the reference system.

The layer model orders the dynamic objects by high (nearest to the reference system) and low (furthest away from the reference system), but is independent from the task relevance of the visual information. The lowest layer should contain an AOI for the complete reference system (screen) to secure that each eye movement point can be assigned to at least one AOI. The amount and order of applied layers depends on the task environment and can be redefined for each different task environment.

Figure 5 shows an example implementation of the layer model adapted to a 2D air traffic control environment with four different layers. The base layer contains the screen where the visualization is presented. The window layer is reserved for dialog windows that are on top of the other layers. The object layer contains the dynamic objects that change their position in relation to the observer, for example the moving aircraft on a radar screen. The background layer contains static

information that does not change their position in relation to the reference system, for example routing structure of an air space or the speed indicator within a driving environment. The base layer contains the screen.

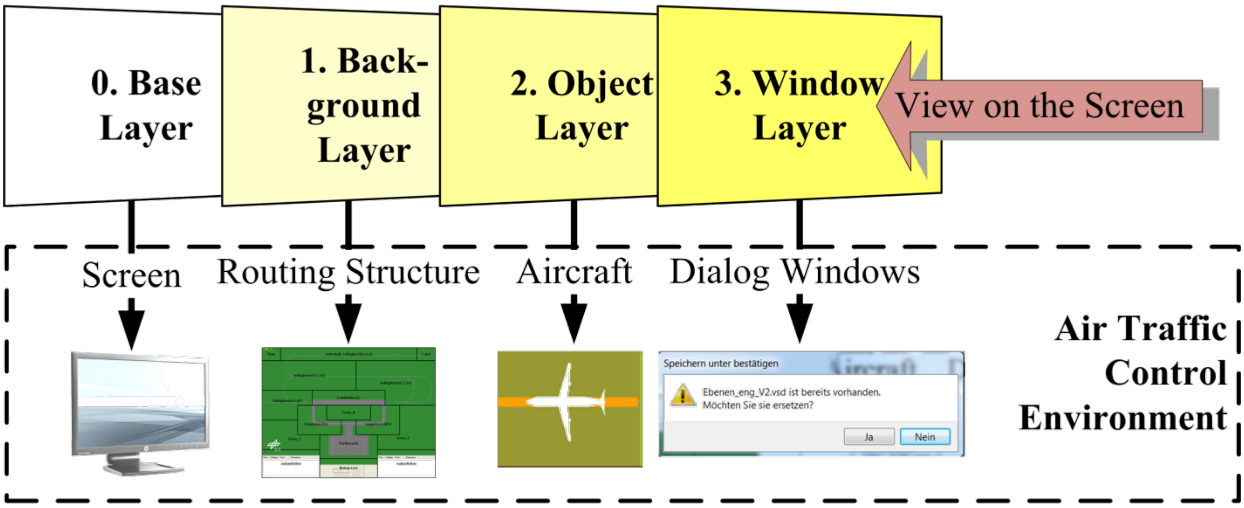

Figure 5 The layer model used for an air traffic controller task environment

Requirement F is fulfilled by storing the shape and position for every timestamp as a new line into a log file. The shape of the AOI could be defined by points along a path defining a convex or concave polygon. As preparation for the data analysis the path should be equally (clockwise or counterclockwise) for all logged AOIs, because some algorithms for solving the point-in-polygon problem have one or the other order as restriction, e.g. the Winding number algorithm (Hormann & Agathos, 2001). If a complex polygon (the path intersects itself) is used special methods have to be applied for the analysis (Galetzka & Glauner, 2012). The position of the AOI represents the information where to place the shape in relation to the screen. IGDAI supports Requirement F by providing a XML data structure that accounts for these assumptions that dynamic information do not change their position or shape over the majority of time and also integrates the depth information. The XML data structure is optimized for one screen only, which means that for a multi monitor environment only the screen with actual dynamic AOIs have to be logged.

The IGDAI XML data structure separates (Table 1) the shape of an AOI from its movement, size, or orientation and from the time it is visible. IGDAI XML has the main node types ShapeTemplateList, UpdateList, and AoiList. ShapeTemplateList contains shape template nodes that

describe the shapes (e.g. list of points) that occur during the recording. The UpdateList contains an Update node for each time interval describing the current position and depth (3D: distance; 2D: layer model) of a dynamic AOI. The nodes in UpdateList can also contain size or orientation changes. The AoiList contains the link between the dynamic AOIs and their shapes. The reuse of equal shapes reduces the amount of data stored. Table 1 shows the XML example implementation of an aircraft shape template in combination with AoiList using the shapes and the update of the screen at second 2.

Table 1 ShapeTemplateList, UpdateList, and AoiList implemented in XML.

ShapeTemplateList node for an object with 4 points	**AoiList**
<ShapeTemplate id="Object_01"> <Outline NumberPoints="4"> <Polygon> <Point x_rel="-0.05" y_rel="-0.05"/> <Point x_rel=" 0.05" y_rel="-0.05"/> <Point x_rel=" 0.05" y_rel=" 0.05"/> <Point x_rel="-0.05" y_rel=" 0.05"/> </Polygon> </Outline> </ShapeTemplate>	<DynamicAoi firstSample="22" AOI_ID="Object_A" template="Object_01" /> <DynamicAoi AOI_ID="Object_B" template="Object_01" />
UpdateList node: second 62 from the storing	
<Update Second="2.0"> <dAOI AOI_ID ="Object A" > <Position x="0.25" y="0.65" Depth="2"/> <Size s="0.25"/> </dAOI> <dAOI AOI_ID ="Object B" > <Position x="0.775" y="0.55" Depth="2"/> <Orientation r="270"/> </dAOI> </Update>	

The following description explains the process of storing dynamic AOIs within the visualisation engine. For each timestamp the dynamic objects have to be transformed to the reference systems (Requirement A and C). Then the shapes of the previous timestamp are compared with the current shapes based on their AOI_IDs. The comparison identifies if the position, orientation or size has changed. If one of the values has changed the update contains this information. If a new AOI_ID is introduced its shape is compared to all existing nodes in the ShapeTemplateList. If no match is found a new node containing the shape of the new AOI_ID is added to ShapeTemplateList. The

AOI_ID is then added to the AoiList and Update node. With Requirement D, E, and F we are now able to combine the dynamic AOIs with the eye movement data.

Data Analysis Methods

In preparation of the eye movement analysis IGDAI (Figure 1) defines a requirement for matching eye movement and AOIs (G). Requirement G: For each timestamp with eye movement data shall a set of AOIs separated by depth be available. The eye movement uses the same timestamp (Requirement A and D) as the stored AOIs. The goal is to assign dynamic AOIs to the eye movement.

Algorithm 1 describes pseudocode for automatic matching of eye movement and AOIs. By taking the layer model into account the matching algorithm would start by checking the AOIs in the highest layer (Figure 5, window layer). Checking whether an eye movement point is inside an AOI is a point-in-polygon problem (for complex polygones Galetzka & Glauner, 2012; for simple polygones Taylor, 1994). If the eye movement point does not match with the AOIs on the highest layer, the matching process is continued using the AOIs from the layer directly below (Figure 5, object layer). This process is continued for all layers until an AOI is identified that contains the eye movement point. The result is that each eye movement point is assigned to only one AOI.

Algorithm 1 Merge algorithm to identify the correct AOI for every captured eye movement point.

```
function Mergealgorithm
allEyePoints = LoadEyePoints()
DynamicAOIList = LoadAOIList()
foreach (eyePoint in allEyePoints){
	time = get_Time (eyePoint) # Determine timestamp for the current eye movement
point
	allLayers = get_Layers_From (DynamicAOIList; time) # Determine the current AOIs
	foreach (layer in allLayers){ # Start with the highest layer first
		foreach (polygon in layer){
			if Check_Point_in_Polygon (Polygon, eyePoint) == TRUE
			then {
				polygonIDs.add(Polygon,
Distance_to_PolygonCenter(eyePoint))
			}
		}
		if count(polygonIDs) > 0
		then {
			eyePoint.setAOI (Shortest_Distance(polygonIDs))
			break # jump to the next eyePoint
		}
	}
}
```

IGDAI also accounts for two special cases that arise by matching eye movement to AOIs, independent from the implementation. Requirement H: The special cases shall be addressed to conduct the eye movement analysis. The first special case occurs if an eye movement point can be assigned to multiple AOIs on the same layer. This inhibits a singular classification between eye movement and the gathered information. This special case will occur for example if a pedestrian covers a traffic sign in a driving simulation and both AOIs are classified in the same layer. In a 3D task environment the eye movement point should be assigned to the AOI with the least depth (Requirement E). In a 2D task environment IGDAI recommends to determine the distance between the eye movement point and the centre of all matching AOIs. The eye movement point is then classified based on the shortest distance (Algorithm 1, Shortest_Distance function), assuming that the AOIs surround the visual information equally centred.

The second special case is the mismatch of eye movement and AOI due to the accuracy of eye tracking systems. Figure 6 (1x) shows an example of eye movement aggregated for one second (blue dots) and the nearest AOI (square). The eye movement in Figure 6 (1x) would be classified incorrectly on a lower layer, probably the lowest one, because of the accuracy of the eye tracker system. IGDAI recommends an enlargement of the AOIs if it is possible to determine the degree of accuracy. The accuracy consists of the accuracy of the gaze vector and the accuracy of calibration. The accuracy of the gaze vector should be provided by the eye tracking manufacturer. The accuracy of calibration depends on the calibration of the participant and the stability of the calibration throughout the experiment (e.g. a head mounted eye tracking system could dislocate). Therefore, the degree of accuracy should to be checked in small intervals throughout the capturing of the eye movement, e.g. by calibration points during the experiment. A scale factor for enlargement can be calculated using the tangents of the degree of accuracy (normally expressed in angle), the distance of the participants to the screen and the original size of the AOI on the screen. Figure 6 (1x to 4x) shows the offsetting of the AOI with three different scale factors. If the degree of accuracy cannot be checked during the experiment due to experimental design or task environment constrains, IGDAI recommends using no scale factor (1x) for the eye movement analysis. Even if no scale factor is used, a set of self-selected scale factors can help to determine the quality of the recorded eye movement. This analysis is based on the assumption that each participant must dedicate an average amount of eye movement to the dynamic AOIs to solve the task. A calibration error (e.g. a dislocated eye tracking system) could lead to a strong increase of eye movement data in the perimeter of the dynamic AOIs. A systematically increasing of the dynamic AOIs helps to identify participants with an abnormal increase in assigned eye movement. An individual examination of the eye movement data of these participants should help to decide if their data is valid for further analysis. An example analysis is provided in the result section.

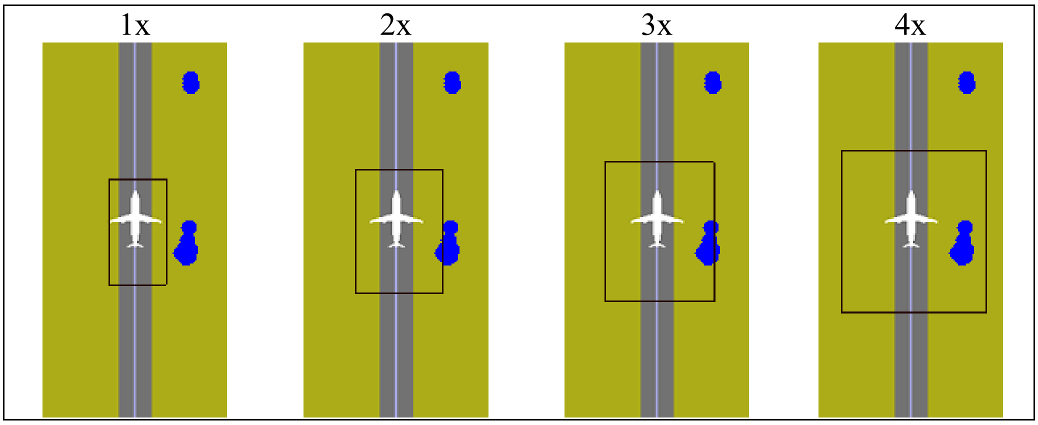

Figure 6 Eye movement data (dots) and dynamic AOIs on the object layer (square) with different area sizes.

IGDAI defines requirements for the task environment, provides a structure for recording the dynamic AOIs and then simplifies the post processing by taking into account the previous steps. IGDAI is kept generic to be applicable to a large variety of task environments. An IGDAI implementation was done by the German Aerospace Center (DLR) using a SMI head mounted eye tracker, an air traffic controller simulation (FAirControl) as task environment and the analysing software Eye Tracking Analyzer (EyeTA). The EyeTA was also developed by the DLR and represents the first software that implements IGDAI for the purpose of standardized eye movement analysis with dynamic AOIs[1].

Verification of IGDAI

The feasibility of IGDAI is shown by implementing it into a task environment and showing the effect for the analysis. We selected FAirControl (Möhlenbrink, Werther, & Rudolph, 2007) as a task environment. For the verification, we assume the focus of the study is to identify if the participants gather information by monitoring fixed points or pursuing dynamic objects to generate a decision.

[1] This software has to be requested from the author due to DLR restrictions. A short video presentation is available at dlr.de/fl/Portaldata/14/Resources/dokumente/A_Guideline_for_Integrating_Dynamic_Areas_of_Interests.mp4

Participants

The group of participants consisted of 9 female and 11 male (N = 20) students with an average age of 24 years (*SD* = 2.12 years). The participants studied at the Technical University of Berlin. They participated in the experiment in separate sessions and were compensated with 10 €.

Material

The task environment to capture the eye movement was set up as proposed by IGDAI and therefore similar to Figure 2. FAirControl represents an airfield from a top view (Figure 7). The airfield consists of a runway and two routes to reach the runway entry point (target circle). FAirControl simulates an aircraft (LandA) landing on Route Land and an aircraft (StartA) taking off on Route Start. The aircraft begin simultaneously on both routes. The time for an aircraft to fly its particular route varies between 9.5 and 11 seconds. The distance between the aircraft to reach the entry point vary between 1.5 seconds (LandA first on Target) to -1.5s (StartA first on Target), but was never zero. The speed from the aircraft on Route Land varied between 6.8 and 7.9 cm per second. The AOI of LandA was a square with sides of 2.5 cm. The speed from the aircraft on Route Start varied between 0.9 and 1.1 cm per second. The AOI of StartA was a square with sides of 1.5 cm.

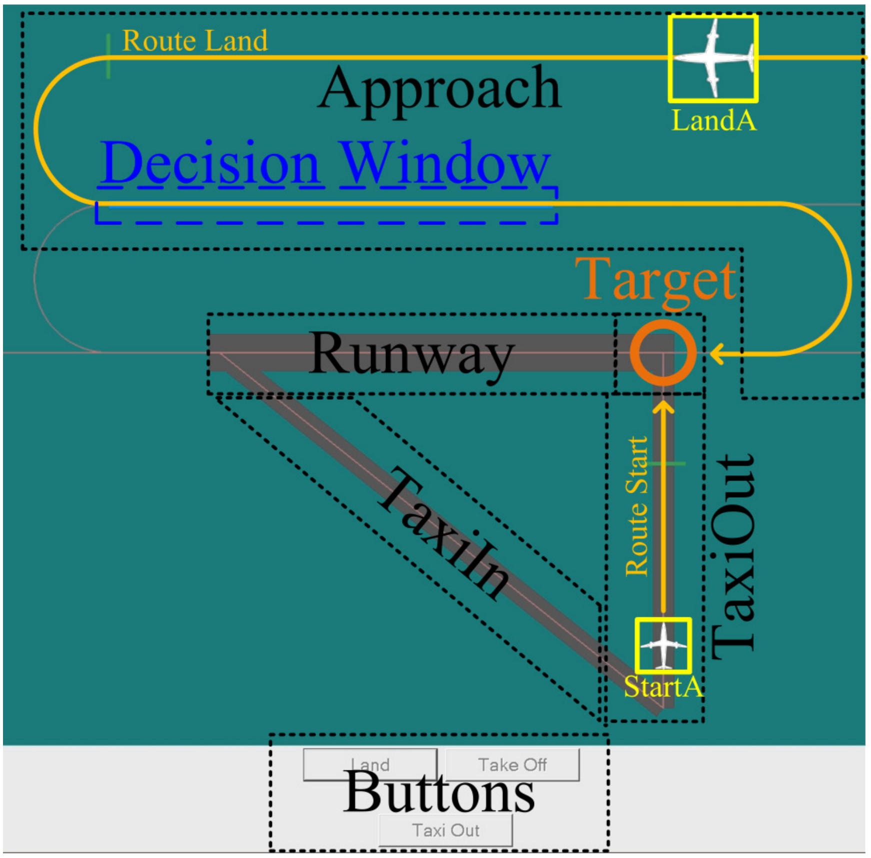

Figure 7 FAirControl, the task environment that was used for verification of IGDAI

FAirControl was presented in full screen mode on a 22-inch display (17.6-inch wide by 13.7-inch high) with a resolution of 1280 x 1024 pixels. A SMI RED eye tracker system with a 60 Hz sampling rate was used to capture the eye movement. The eye tracker warning system made sure that the participants kept a distance of 65cm and stayed within an area of 50x30 cm to the centre of the screen.

The eye tracker system fulfils Requirement A. The timer (Requirement B) was implemented into FAirControl. Therefore the eye movement data was sent to FAirControl and a timestamp was added. If no eye movement data was sent or the eye movement point was not on the screen, FAirControl added zero values to the x and y coordinates in the eye log file (Requirement D). The logging of the dynamic AOIs were also set to 60 Hz to match the eye movement. FAirControl stored the dynamic AOIs using the layer model (Requirement E) and the data structure (Requirement F). The chosen layer names and depths of the layers are equal to Figure 5, except for the window layer that was not used, because FAirControl can be handled without extra dialog windows. The base layer

contains the screen. The background layer contains AOIs for approach, runway, both taxiways (in and out), the target, and the buttons (Figure 7). The object layer contains the bounding boxes surrounding the aircraft that moved via both routs.

Procedure

After reading the instructions and prior to the experimental session, the eye tracking system was calibrated to each participant. In a 10 minutes training session, participants were familiarized with the FAirControl. The participants were instructed to monitor FAirControl and decide which aircraft would reach the runway entry point (Figure 7, Target) first. They answered by pressing one of the buttons and while the aircraft on Route Land moved through the decision window (Figure 7). After the decision was made each aircraft continued their route and the participants got feedback if they were correct or not. When the aircraft on Route Land reached the start position of Route Start the next cycle began. Every participant assessed 60 randomized cycles with an equally distributed number of aircraft from both routes reaching the runway entry point first. The experiment took approximately 17 minutes to complete.

Results and Discussion

As described above, the IGDAI layer model was used to classify the different types of AOIs used within FAirControl. Due to the problem of the clear distinction between fixation, saccades and smooth pursues in dynamic task environments the data analysis was based on the raw eye movement and the classification by the IGDAI layer model. To verify the impact of IGDAI to the analysis of eye movement, we performed three analyses. The first analysis quantifies the acceleration of data preparation by IGDAI in relation to a manual description. The second analysis quantifies the quality of the recorded eye movement. The third analysis shows the impact of the layer model to quantify the eye movement distribution.

The first analysis shows the influence of IGDAI on the AOI preparation time (amount of time needed for generating AOIs) and the quality of the AOIs. The analysis was performed using BeGaze to transcript 5 randomly selected data sets without using IGDAI. The 5 datasets had an average length of 17.2 minutes (SD = 0.95) with 2 dynamic (LandA and StartA) and 7 static (Screen, Approach,

TaxiOut, Buttons, Target, Runway, and TaxiIn) AOIs. The AOI preparation time took 35.5 minutes in average (SD = 4.9). Therefore, the transcription of all 20 data sets would have taken approximately 11 hours. The AOI preparation time without IGDAI increases linear with the number of participants. The AOI preparation time with IGDAI was zero, because the capturing of AOIs took place while conducting the study. The additional effort for implementing IGDAI into the task environment was approximately 6 hours. The implementation time for IGDAI is a fixed value independent from the number of participants. The influence of IGDAI on the quality of AOIs was determined by comparing the different dwell times in percentage. Figure 8 shows that the results from BeGaze and IGDAI are similar. The difference between the dynamic AOIs is 0.02 percent (BeGaze 31.49%; IGDAI 31.51%). Either dynamic [t(4) = -0.10807, p<.91] nor static [t(4) = 0,14887, p<.88] showed significant differences.

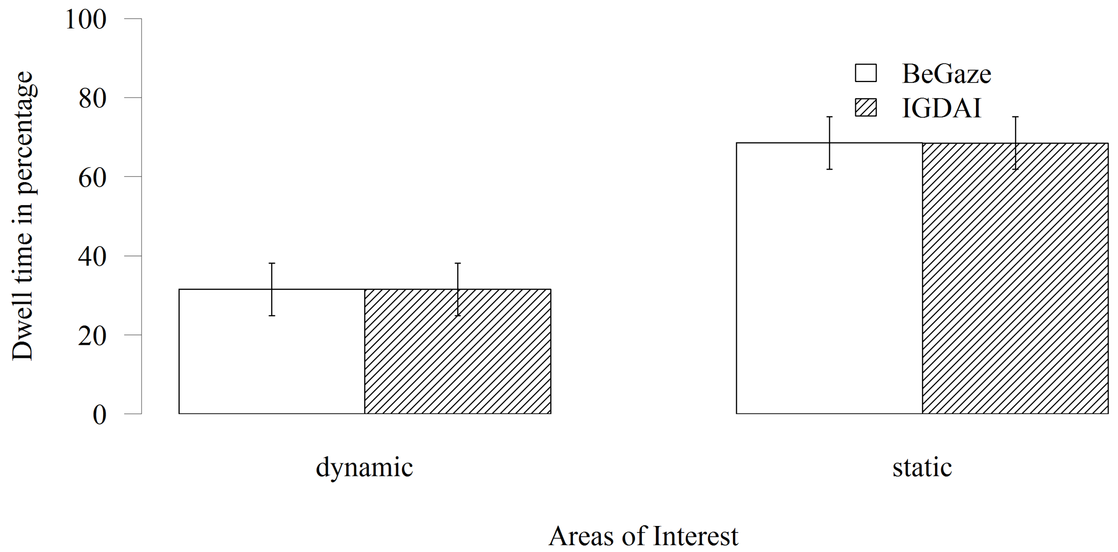

Figure 8 Comparison between dwell time in percentage for 5 participants and static or dynamic AOIs using BeGaze and IGDAI

The second analysis concentrated on the quality of the eye movement data recorded for the individual participants. First, the amount of valid eye movement per participant (valid eye movement divided by total number of captured eye movement) was calculated and compared. Eye movement was invalid if the eye tracker could not detect the eyes or the gaze vector was not on the experimental screen. For each participant the validity of the eye movement data reached more than 94% (*M* = 97%,

SD=1.8). Second, the quality of the initial calibration was determined by a 9 point verification method. After the calibration of the eye tracker, 91% (*SD*=4.5) of the eye movement data of each participant had an error of 1.76° around the actual reference points. Third, because the experimental design did not allow checking the degree of accuracy during the experiment, the alternative analysis proposed by IGDAI is to use a self-selected scale factor (2x, 3x, and 4x) to increase the dynamic AOIs. This leads to an increase in eye data assigned to the dynamic AOIs in relation to the original size (1x). If the increase shows outliers, these participant should be checked individually for calibration errors. Figure 9 shows the increase between 1x and the scale factors in boxplots to identify outliers. An individual examination of both outliers (VP4 and VP18) showed that their percentage of eye movement on dynamic AOIs did not increase at one particular factor, but was above average even for 1x. The analysis with self-selected factors could show that the increase of assigned eye data is similar for all participants, and therefore all participants are comparable on the following analysis.

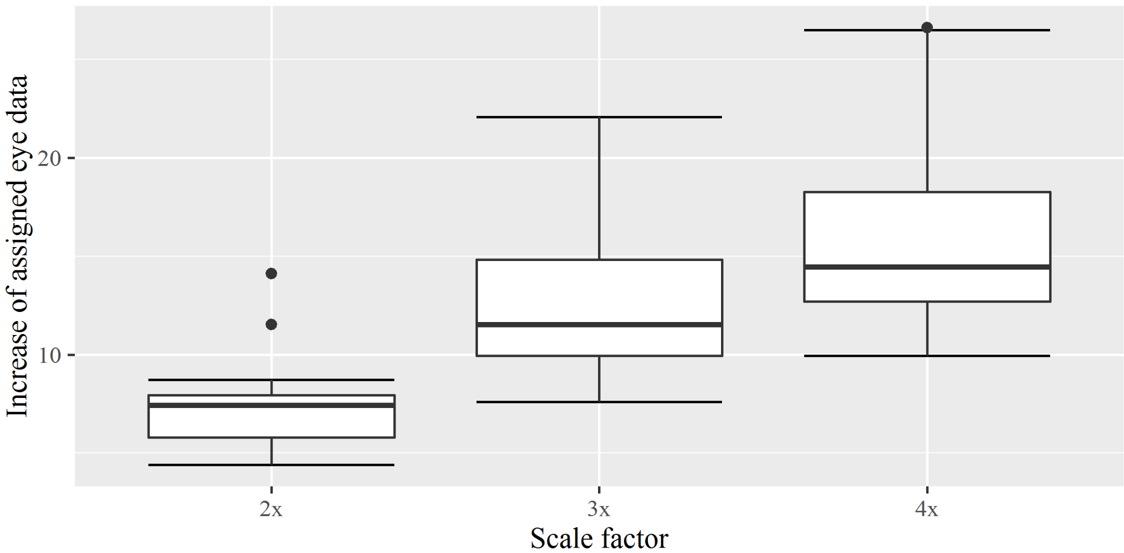

Figure 9 Increase of assigned eye data to dynamic AOIs in relation to the scale factor 1x for three self-selected factors.

The third analysis demonstrates the impact of the IGDAI layer model to the eye movement analysis. As a baseline for this analysis the base and background layer were analysed first and then contrasted with the object layer. The dynamic AOIs on the object layer had the original size (1x). Figure 10 shows the results using IGDAI layers (base and background layer – BackgroundL;

BackgroundL plus object layer – ObjectL). The bars represent the percentage of eye movement per FAirControl AOI (Figure 7). The bars for each condition summarize to 100 percent. By using only the BackgroundL the data is divided into seven AOIs. When extending the analysis with the object layer the data is divided into nine AOIs.

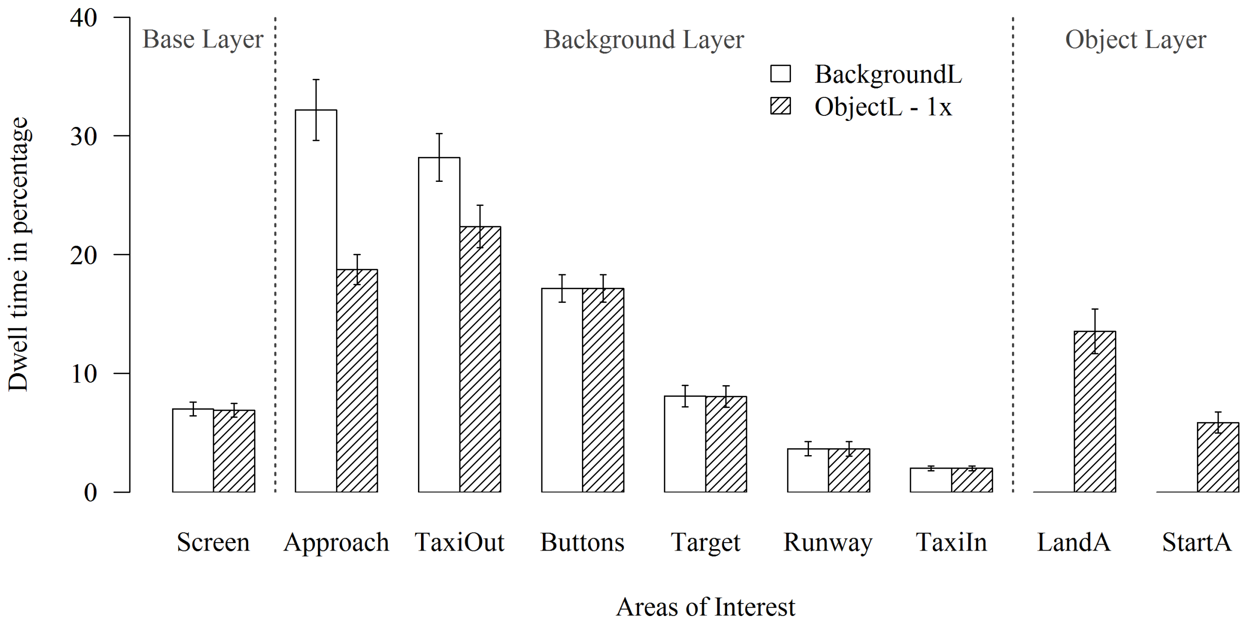

Figure 10 Mean plot for percentage of dwell times with standard error (n= 20) separated by application of different layers.

Figure 10 shows that the base layer always contains less than 7 percent of the eye movement, independent from the analysis. BackgroundL contains 91.2 percent (Runway + Target + TaxiIn + Buttons = 30.8%; Approach + TaxiOut = 60.362%) of the eye movement. Including ObjectL into the analysis has a significant decreasing effect to the percentage of eye movement data on Approach and TaxiOut [$t(19) = 9.833$, $p < .001$]. No other AOIs in the background layer showed a significant increase or decrease when ObjectL was included.

The results of the three analyses show an overall view of the possibilities of structured data analysis that IGDAI facilitates. The analysis of distributed eye movement data was increased in detail through the application of the IGDAI layer model. By the use of the new method of structuring the AOIs in a different layer we were able to see the extended distribution of eye movement for the example data. In relation to only concentrating on static AOIs this leads to a detailed insight.

Conclusion

We introduced the guideline IGDAI, which allows an automated classification of eye movement dependent on static or dynamic AOIs. IGDAI enhances the eye movement analysis by simplifying the integration of capturing dynamic AOIs by a step by step approach of integration. IGDAI is also independent from the used eye tracking technology and the task environment because it structures the flow of information that are necessary for eye movement analysis. It can be applied to simple one screen task environments, but also to complex multiple monitor task environments (e.g. Figure 3) or even aircraft cockpits (Sarter, Mumaw, & Wickens, 2007). The requirements (Figure 1) to configure the task environment build the basis for an efficient capturing of the dynamic information and consequently accelerate the data analysis.

An important step to understand how participants solve a task is to identify the information gathering sequence. This sequence has to be as accurate as possible because approximations could lead to misinterpretation. With IGDAI the information gathering sequence can objectively be identified. The verification study demonstrated the proper functioning of IGDAI and its influences to various levels of eye movement analysis. It showed that the actual position and size of a dynamic object allows a more detailed analysis than static AOIs alone.

Future IGDAI developments should extend the automatic generation of dynamic AOIs to non-artificial task environments. The generic structure of IGDAI allows the adaptation onto those task environments, such as a view through the windscreen while driving in a real environment. Methods (Hsu & Huang, 2001) and tools to support the reconstruction of video movements by post processing already exist and only need small adaptation to fulfil the IGDAI requirements, such as the DynAOI generation tool (Papenmeier & Huff, 2010).

The results showed that IGDAI lowers the effort of using static and dynamic AOIs as a method of eye movement analysis. The application of dynamic AOI is often associated with effort

that is not justified by the results. IGDAI changes that and might lead to dynamic AOIs as a standard tool for eye movement analysis.

Acknowledgments

We thank Hamilton Gross for his help with the first implementation of the IGDAI during his internship at DLR. The authors also thank all the reviewers for their helpful comments on the manuscript. Correspondence concerning this article should be addressed to M. Friedrich, DLR, Lilienthalplatz 7, 38108 Braunschweig, Germany (e-mail: Maik.Friedrich@dlr.de).

References

Akamatsu, M., Okuwa, M., Funakawa, M., Kitajima, M., Onuki, M., Life, Q., & Co, M. P. (2001). Whole-View Driving Simulator For Measuring Driving Behavior In Town Environment. *Driving simulation conference*. Valbonne.

Bruder, C., Leuchter, S., & Urbas, L. (2003). Konzept für automatische Auswertungen von Blickbewegungen bei dynamischen Informationen. *MMI interaktiv, 6*, 25-34.

Bulling, A., & Gellersen, H. (2010). Toward Mobile Eye-Based Human-Computer Interaction. *Ieee Pervasive Computing, 9*, 8-12. doi:10.1109/MPRV.2010.86

Castelhano, M. S., & Rayner, K. E. (2008). Eye movements during reading, visual search, and scene perception: An overview. *Cognitive and Cultural Influences on Eye Movements, 2175*, 3-33.

Crytek GmbH. (2016). CRYENGINE | The complete solution for next generation game development by Crytek. Retrieved from www.cryengine.com

Duchowski, A. T. (2007). Eye Tracking Methodology: Theory and Practice. 334.

Fehd, H. M., & Seiffert, A. E. (2008). Eye movements during multiple object tracking: Where do participants look? *Cognition, 108*, 201-209. doi:10.1016/j.cognition.2007.11.008

Friedrich, M., Möhlenbrink, C., & Carstengerdes, N. (2012). SESAR- JU D86 Project 06.08.04 - Single Remote TWR Ph1 V2 - Validation Report. 65.

Galetzka, M., & Glauner, P. O. (2012). A correct even-odd algorithm for the point-in-polygon {(PIP)} problem for complex polygons. *CoRR, abs/1207.3*.

Granka, L. A., Joachims, T., & Gay, G. (2004). Eye-tracking analysis of user behavior in WWW search. *Proceedings of the 27th annual international conference on Research and development in information retrieval SIGIR 04, 5*, 478. doi:10.1145/1008992.1009079

Development of a process for the evaluation of eye-fixation data using dynamic areas of interest, DLR 37 (2010).

Henderson, J. M., Weeks, P. A. J., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology Human Perception and Performance, 25*, 210-228. doi:10.1037/0096-1523.25.1.210

Hormann, K., & Agathos, A. (2001). The point in polygon problem for arbitrary polygons. *Computational Geometry: Theory and Applications, 20*, 131-144. doi:10.1016/S0925-7721(01)00012-8

Hornof, A. J., & Halverson, T. (2002). Cleaning up systematic error in eye-tracking data by using required fixation locations. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc, 34*, 592-604. doi:10.3758/BF03195487

Hsu, S.-H., & Huang, C.-L. (2001). Road sign detection and recognition using matching pursuit method. *Image and Vision Computing, 19*, 119-129.

Huff, M., Papenmeier, F., Jahn, G., & Hesse, F. W. (2010). Eye movements across viewpoint changes in multiple object tracking. *Visual Cognition, 18*, 1368-1391. doi:10.1080/13506285.2010.495878

Josephson, S., & Holmes, M. E. (2006). Clutter or content? How on-screen enhancements affect how TV viewers scan and what they learn. *Main, 38*, 155-162. doi:10.1145/1117309.1117361

Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature neuroscience, 3*, 1340-1345.

Landry, S. J., Sheridan, T. B., & Yufik, Y. M. (2001). A methodology for studying cognitive groupings in a target-tracking task. *Intelligent Transportation Systems, IEEE Transactions on, 2*, 92-100. doi:10.1109/6979.928720

McCarthy, J. D., Sasse, M. A., & Riegelsberger, J. (2004). Could I have the menu please? An eye tracking study of design conventions. *People and Computer, 17*, 401-414.

A Model Based Approach to Cognitive Work Analysis and Work Process Design in Air Traffic Control, (2008).

Möhlenbrink, C., Werther, B., & Rudolph, M. (2007). FAirControl: Ein Werkzeug zur Modellierung und Analyse von Verhaltensweisen am Lotsenarbeitsplatz: VDI Verlag.

Nielsen, J. (2006). F-Shaped Pattern For Reading Web Content. *Search, 2009*, 2007.

Palinko, O., Kun, A. L., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. *Proceedings of the 2010 Symposium on EyeTracking Research Applications ETRA 10*, 141. doi:10.1145/1743666.1743701

Papenmeier, F., & Huff, M. (2010). DynAOI: A tool for matching eye-movement data with dynamic areas of interest in animations and movies. *Behavior research methods, 41*, 179-187. doi:10.3758/BRM.42.1.179

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. *Studies in Visual Information Processing, 6*, 3-22. doi:10.1016/S0926-907X(05)80003-0

Rayner, K., & Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal Of Psychology, 46*, 342-376.

Reimer, B., & Sodhi, M. (2006). Detecting eye movements in dynamic environments. *Behavior research methods, 38*, 667-682. doi:10.3758/BF03193900

Salvucci, D. D., & Goldenberg, J. H. (2000). Identifying Fixations and Saccades in Eye-Tracking Protocols. *Eye Tracking Research and Applications Symposium*. New York: ACM Press.

Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: an empirical study combining behavioral and eye-tracking data. *Human Factors, 49*, 347-357. doi:10.1518/001872007X196685

Seifert, K., Rötting, M., & Jung, R. (2001). Registrierung von Blickbewegungen im Kraftfahrzeug. In T. Jürgensohn & K.-P. Timpe (Eds.), *Kraftfahrzeugführung SE - 14* (pp. 207-228): Springer Berlin Heidelberg.

Sennersten, C. (2004). Eye movements in an Action Game Tutorial.

Sennersten, C., Alfredson, J., Castor, M., Hedström, J., Lindahl, B., Lindley, C., & Svensson, E. (2007). *Verification of an Experimental Platform Integrating a Tobii Eytracking System with the Hifi Game Engine*.

Zelinsky, G. J., & Neider, M. B. (2008). An eye movement analysis of multiple object tracking in a realistic environment. *Visual Cognition, 16*, 553-566. doi:10.1080/13506280802000752

Zhang, Y., & Hornof, A. J. (2011). Mode-of-disparities error correction of eye-tracking data. *Behavior research methods, 43*, 834-842. doi:10.3758/s13428-011-0073-0

IV Article 3: The Influence of Task Load on Situation Awareness and Control Strategy in the ATC Tower Environment.

The following version is a pre-copyedit version of an article published in Cognition, technology & work. The final authenticated version is available online at: http://dx.doi.org/10.1007/s10111-018-0464-4.

Reprinted by permission from Copyright Clearance Center: Springer Nature, Cognition, Technology & Work: **Citation:** Friedrich, M., Biermann, M., Gontar, P., Biella, M., & Bengler, K. (2018). The influence of task load on situation awareness and control strategy in the ATC tower environment. Cognition, technology & work. doi: http://dx.doi.org/10.1007/s10111-018-0464-4

Journal: Impact Factor by the time of the publication (JCR Social Science Edition 2016) was 1.105.

The influence of task load on situation awareness and control strategy in the ATC tower environment.

Maik Friedrich[1] Maresa Biermann[2] Patrick Gontar[2] Marcus Biella[1] Klaus Bengler[2]

Words 6659 (body)

[1] German Aerospace Center, Braunschweig

[2] Technische Universität München

Author:

Maik Friedrich, maik.friedrich@dlr.de
+49 (531) 295 – 2598
ORCID 0000-0003-3742-2322
German Aerospace Center
Lilienthalplatz 7
Braunschweig, NI D-38108

Acknowledgments

This research received funding from the German Aerospace Center institutional funding mechanism dedicated to the human center approach for automation, which is sponsored by the Federal Ministry for Economic Affairs and Energy.

Abstract

The safe and efficient operation of air traffic is highly dependent on the performance of the Air Traffic Control Officer (ATCO). The Air Traffic Control Officers control the traffic within defined areas by monitoring the traffic and granting clearances. A key element in analyzing the ATCOs is their interaction with the environment through their workplace. Especially the influence of task load on their situation awareness (SA) and applied control strategy provides information on the quality of the workplace. As task load increases, controllers are able to maintain performance by using different management or compensation strategies. This article supports the evaluation of ATCO's workplaces by focusing on whether probe techniques for assessing SA are applicable for tower control operation and for measuring the influences of increased task load on the control strategy. An experiment with nine ATCOs was conducted in a simulated real-time air traffic control environment. Different measurements for SA were applied and compared regarding their efficiency and validity. The manipulation of task load and visibility influenced the SA and control strategy at the same time. Performance metrics were selected in advance to evaluate the participant's efficiency. SA was measured with a probe technique and an offline self-assessment method. Findings suggest that probe techniques increase the insight into the understanding of SA in comparison to self-assessment and that they are applicable to the air traffic control environment. Control strategies were derived from the information-gathering process via the eye-movement behavior and connected to task load. The results imply that SA is part of the individual performance and that increasing demand through task load is handled with an adaptation of the control strategy.

Keywords

Task load; Situation Awareness; Strategy identification; Eye Tracking; Air Traffic Control

1 INTRODUCTION

Air traffic control (ATC) is a system that relies highly on the capabilities of human operators for the safe processing of flights. With European air traffic estimated to double by the year 2020 (EUROCONTROL/FAA, 2010), the tasks of Air Traffic Control Officers (ATCOs) are becoming more challenging. ATC covers the surveillance and control of the movements in airspace and on the runways of airports to ensure a safe, organized, and fluent processing of air traffic (Mensen, 2014). Kontogiannis and Malakis (2013) characterize the ATC task with "time pressure, multiple goals, interconnected tasks and high consequences of error". The task takes place in a dynamic environment, in which information from different navigation, communication, and surveillance systems must be continuously processed. New technologies are currently being developed to support operational personnel, such as assistant systems in the case of low visibility (Teutsch & Postma-Kurlanc, 2014) or completely new workplaces in the case of remote tower operations (Fürstenau, 2016). The introduction of new systems, if performed without the proper knowledge of related effects (e.g. situational awareness and workload), always increases the probability of errors (Baumann & Krems, 2007; Mica R Endsley, 2000). When incidents or accidents in ATC occur, the consequences are often severe, especially in civil aviation. These events are frequently attributed to human error (Redding, 1992). This is why a thorough understanding of human performance and influencing factors on ATC is essential for further development in this research area.

The ATC task is divided into en route control, approach control, and tower control (Mensen, 2014). The majority of studies focus on the en route and approach control workplace (e.g. Durso et al., 1998; Edwards, 2013; Karikawa et al., 2014; Lee, Jeon, & Choi, 2012), but only a few have focused on ATCOs performing tower control operations (e.g. Fürstenau, Friedrich, Mittendorf, Schmidt, & Rudolph, 2013; Lange, 2014; Papenfuss et al., 2010). The ATCOs who are responsible for tower control use air/ground communications,

radar, direct view on the airfield and binoculars/zoom cameras (Hadley, Guttman, & Stringer, 1999) to regulate the speed, heading, and flight level of aircraft on the ground and in the direct proximity of the airport. Considering the expected increase in traffic, this article addresses the shortcoming in understanding the influence of workload onto the tower control operations by conducting research with operational experts in a high-fidelity simulator. This will prepare the foundation for introducing new technologies with the purpose of increasing safety and efficiency.

According to EUROCONTROL/FAA (2010), "human performance […] refers to the adequate performance of jobs, tasks and activities by operational personnel – individually and together" (p. 8). This performance depends on the person (capabilities, skills, knowledge, motivation, etc.) and the context of work (systems, organization, and environment). Consequently, human performance of ATCOs is important for the overall system safety and effectiveness of aviation operations (Hadley et al., 1999). Human performance is influenced by performance-shaping factors determining human error probability (Ambroggi & Trucco, 2011). Performance-shaping factors have been mainly investigated separately in research on human performance in ATC (Cox-Fuenzalida, 2007; Edwards, 2013). The lack of research on the interaction of these factors was addressed at an OPTICS[1] Expert Workshop in 2014 as the most important goal in human factors research (Kirwan, Bettignies-Thiebaux, & Scholte, 2014). Based on a literature review of over 400 incident reports[2] of Eurocontrol, Edwards (2013) identified the following nine performance-shaping factors as influential: Communication, teamwork, trust, fatigue, stress, vigilance, attention, mental workload, and

[1] Observation Platform for the Technical and Institutional Consolidation of Safety Research; OPTICS, www.optics-project.eu/

[2] According to EUROCONTROL Safety Regulatory Requirements (ESARRs), incidents (potentially) affecting safety must be reported to the Accident/Incident Data Reporting (ADREP) of the International Civil Aviation Organization (ICAO), www.skybrary.aero/index.php/Reportable_Incidents

situation awareness (SA). Especially the factors mental workload and SA are often used as performance measurement and frequently considered in relation to each other (e.g. Mica R Endsley, 1995b; Vidulich, 2000; Wickens, 2002), but the results of their interaction differ (Edwards, 2013; Lee et al., 2012). Performance as an objective measurement for all participants can be interpreted in terms of safety, efficiency and orderliness (Griffin, Neal, & Neale, 2000).

2 WORKLOAD, SITUATION AWARENESS, AND CONTROL STRATEGIES

2.1 Workload

When engaging in a cognitive task, the active information processing requires cognitive resources which are limited for mental operations (Kahneman, 1973). With increasing task difficulty, more resources are needed, so the operator's remaining capacity diminishes. High mental workload will occur when the performance of a task needs the majority of the available resources, while low workload indicates unused capacities. A distinction must be made between task load and workload. Task load is the external stress or demand through the task or the system, e.g. complexity or time pressure. Mental workload, in contrast, is the resulting strain or impact of these stressors on the person, depending on their individual constitution, resources, abilities, etc. (International Organization for Standardization, 1991). The relationship between task load and workload is not straightforward, as a change in task load can lead to different levels of perceived workload in different individuals (Kerkau, 2005). However, in ATC, the number of aircraft in the control traffic region (CTR) is, in experiments, frequently used to manipulate task load, which has been shown to affect mental workload: The more aircraft's to control, the higher the mental workload (Ahlstrom & Friedman-Berg, 2006; Brookings, Wilson, & Swain, 1996; Lee et al., 2012).Young, Brookhuis, Wickens, and Hancock (2015) provide an overview on studies investigating the relationship between workload and performance. In total, there is strong evidence that high

workload leads to poor performance and a higher number of errors. It can be assumed that "if demands begin to exceed capacity, skilled operators can either adjust their strategy to compensate or else performance necessarily degrades" (Young et al., 2015).

2.2 Situation Awareness

Mica R Endsley (1995b) defines SA as a hierarchical construct comprising the perception of elements in the environment together with their states and attributes, the understanding of these elements and their significance, and the projection of future states of the elements and the situation. Durso et al. (1998, pp. 1-2) consider SA as a process which is not necessarily present in consciousness but acting implicitly: "SA may sometimes involve simply knowing where in the environment to find a particular piece of information, rather than remembering what that piece of information is". D. L. Chiappe, Strybel, and Vu (2012) extend this idea by stating that individuals tend to rely on external props instead of forming a detailed mental representation: This spares the limited capacity of working memory and attention. Thus, information is not only stored internally in memory but operators also make use of external storage since "it is not necessary to have all the information stored in the mind to keep track of and perform well in complex situations" (Kraemer & Süß, 2015, p. 3153).

SA is an often-investigated factor in high-risk-environments such as the medical operating room (Shelton, Kinston, Molyneux, & Ambrose, 2012), submarine track management (Loft et al., 2015), or nuclear power plants (Yang, Yang, Cheng, Jou, & Chiou, 2012). There are fields of study that indicate age and experience have an influence on SA (e.g. Bolstad, 2001; Kass, Cole, & Stanny, 2007; Patten, Kircher, Ostlund, Nilsson, & Svenson, 2006), that makes it necessary to analyze the sample in accordance with those results. Especially considering the cascade model (Fry & Hale, 1996, 2000; Gregory, Nettelbeck, Howard, & Wilson, 2009) that tries to explain this influence and also indicates that continuous training reduces the influence of age on SA. Yet, there are critical views on the

concept of SA, doubting its usefulness as a scientific construct (Dekker, 2015; Dekker & Woods, 2002). However, several researchers address this criticism and conclude that the outlines of SA have been thoroughly drawn over the past years of research (Mica R. Endsley, 2015; Parasuraman, Sheridan, & Wickens, 2008; Vidulich, 2000). Some literature defines what SA is, by which means it can be measured, which constructs are linked to it, and which are not. It is true that the concept of SA may also be tangible for non-scientists, as interviews with ATCOs show (Edwards, 2013). The authors consider SA as a necessary and measurable construct for understanding human computer interaction, especially in relation to workload.

There are three approaches to measuring SA: (1) subjective measures in terms of self-rating or assessment by experts, (2) implicit performance measures, and (3) probe techniques (Bacon & Strybel, 2013; Durso et al., 1998). Because SA is linked to a mental picture only accessible by the particular person, the first approach, self-rating, is the one most often used. However, a difficulty is “that people are not always aware of what they don’t know” (Jeannot, 2000, p. 22). So self-rating leads to a statement of how certain the person feels about his or her SA (Mica R Endsley, 1995a), rather than to a valid estimation of the real degree of SA. Still, some subjective rating tools for SA are often used, such as the 3D SART (Situational Awareness Rating Technique, Taylor, 1990), or for the context of ATC, the SASHA (Dehn, 2008) which was derived from SASHA_Q (Jeannot, Kelly, & Thompson, 2003).

The second approach, assessing implicit performance measures focusing on observable actions and errors, may be recommendable, especially when considering their non-intrusiveness (Jeannot, 2000). The difficulty is finding the adequate performance indicator, because the relation between SA and performance is not entirely clear (Salmon et al., 2009). Additionally, the resulting performance is not the product of SA alone but several other influences (Mica R Endsley, 1995a).

The third approach is the probe technique, which consists of asking the subject questions about the current situation (Salmon et al., 2009). Probe measures can be divided into freeze techniques and online techniques. Freeze techniques interrupt the current task and blank the displays. As all sources of information are removed, the subject has to recall information from memory to answer questions about the situation, e.g. Situation Awareness Global Assessment Tool (Mica R Endsley, 1995a). In contrast to that, online techniques pose questions about the situation without interruption, e.g. SPAM (Durso et al., 1998). Compared to freeze techniques, online techniques possess a less intrusive nature and allow for a more realistic performance in simulation settings (Salmon et al., 2009) but are, however, not applicable for operational environments. Based on the SPAM method, Kraemer and Süß (2015) developed SARA-T which analyses the log files of simulation in real-time, allowing for an actual real-time assessment of SA. However, online probes have the disadvantage of representing extra workload for the operator, who has to deal with a further task demand (Jeannot, 2000). Salmon et al. (2009) emphasize the difficulty of proposing a consistent measurement method, since several concepts of SA exist which are often linked to a certain suggestion of how to measure it. In consequence, selecting the best-fitting procedure depends on the particular testing assumptions and preconditions. Given the advantages and disadvantages of each approach, the best strategy is to combine different approaches.

Since the 1990s, many studies on SA in the context of ATC have been conducted, focusing on different aspects of the concept of SA or its measurement (Jeannot, 2000). Mica R Endsley and Rodgers (1996) investigated the attention distribution of ATCOs and concluded that trade-offs must be made in deciding which elements are taken into account and in which detail. They traced the observed difficulties in identifying loss of separation between aircraft back to difficulties in passive monitoring. In his analysis of FAA reports on occupational error, Redding (1992) identified lacking maintenance of SA as the most

probable cause for most errors. In more recent studies, SA is frequently investigated in the context of task and system design in light of the future increase in air traffic (Durso & Sethumadhavan, 2008). In connection to tower control operations, findings by Teutsch and Postma-Kurlanc (2014) indicate that sight and communication (Corradini & Cacciari, 2002) are an important factor for SA.

2.3 Control Strategies

The ATC workplace allows for different strategies to compensate for increased workload (Sperandio, 1971, 1978), e.g. by decreasing the amount of time spent on one single aircraft. According to Karikawa et al. (2014), there are several possible control strategies for a certain traffic situation in en route ATC that differ in safety, efficiency, or fuel economy and thus lead to different performance. They suggest that tolerance for the variability of situations could be a main factor when selecting a strategy. Edwards (2013, p. 81) also stressed the significance of management strategies in ATC: “Controlling aircraft is a creative process, and controllers differ on styles of control strategy”. In her study, she suggested certain compensation strategies used by controllers which also serve as performance indicators. Facing the extremes of workload or situation awareness, different compensation strategies are possible. Papenfuss and Friedrich (2016) used eye tracking to identify the different control strategies depending on the dwell duration time participants spent on a specific source of information. They showed that different sets of workplace configuration lead to a change in the information-gathering process and subsequently in control strategy.

2.4 Relation of Workload, Situation Awareness, Control Strategies, and Performance

Processing always appears within a certain situation, which renders knowledge of the situation or mental models – and by this also SA – crucial for perception and action (Hendy, 1995). Concerning the relationship of workload and SA, Hendy (1995) states that good SA can reduce the processing capacity of future decisions and thus time pressure. At the same

time, though, good SA requires capacity for its formation in the time span until a decision is made. According to Lee et al. (2012), workload and SA are differently linked to performance and “these two constructs are more closely related at higher task load conditions than at lower ones.“ (p. 349). Therefore, before analyzing the influence of workload on SA and performance, different behavior in different demanding situations has to be taken into account.

The current state of research on workload, SA, control strategy, and performance indicates that none of the constructs can be measured and analyzed on its own (Wickens, Mavor, & McGee, 1997). Figure 1 presents the connections between the different constructs, beginning with the dynamic environment and its influence on workload and the selected control strategy. Workload and control strategy influence each other directly, whereas Situation Awareness can be affected if the strategy demands change. Research also suggests that the interaction is stronger when the dynamic environment generates high workload (Flin et al., 2003). If workload reaches a certain level, the control strategy is changed to keep workload on a manageable level.

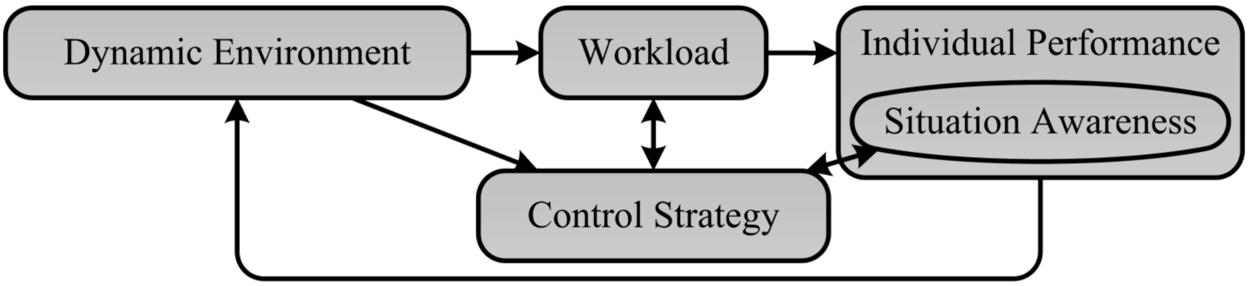

Figure 1 Connection between workload, SA, control strategy, and performance (Wickens et al., 1997)

Looking at the challenges for future tower control operations, the relation between workload, SA, control strategy, and performance still needs to be defined in more detail. Within this article, mental workload resulting from external task load will be used to influence the SA and the control strategy within the ATC tower environment. Knowledge of

their influence as performance-shaping factors is necessary to evaluate new design concepts in early stages of their development. This article supports evaluation of how ATCOs interact with their environment by focusing on two aspects: whether probe techniques for SA testing are applicable for tower control operation and if it is possible to measure the influences of increased task load on the control strategy.

3 METHOD

3.1 Sample

Participants (N = 9) were recruited from the tower Braunschweig-Wolfsburg and via an internal German Aerospace Center mailing list. They were all male, with an average age of 44.22 years (SD = 11.29), ranging from 28 to 62 years. The participants' work experience ranged from 3 to 33 years with a mean of 18.56 years (SD = 10.36). One participant had been retired for 4 years; another had been in partial retirement for 6 months. Each participant obtained 100 € and compensation for travel expenses. The ethical approval was granted by the Braunschweig University of Technology.

3.2 Design

For this study, the working position of an ATCO was simulated in a high-fidelity setting using a within-subject-design 2x2 with repeated measures for one factor to investigate the dependent variables SA, eye movement, and performance under a) varying task load conditions (low vs. high; repeated measures) and b) different visual conditions (high (A) vs. low (B) visibility). The conditions were implemented in two scenarios (Table 1), each consisting of the same sequences of altering low and high task load phases (Figure 2) and differing in visibility. Both scenarios began with low task load to provide a simple start into the simulation. Each scenario was designed to take 40 minutes. In the phases of low task load, only one moving aircraft was present, while the demand increased in the phases of high task load to a maximum of 8 aircraft. The same task load phases (low vs. high) were designed to

contain approximately the same amount and complexity of air traffic, without repeating exactly the same traffic flow to avoid training effects. Phases of high workload consisted of 5 moving aircraft and 3 additional requests for take-off. Traffic patterns used in this study were incoming scheduled flights, departing unscheduled flights, crossing unscheduled flights, and unscheduled flights following the traffic circuit. The scenarios were previously tested for feasibility by a former ATCO.

Table 1 Scenario overview

		Scenario A	Scenario B
a)	Task load conditions	Equal for each scenario, following Figure 2	
b)	Visibility conditions	high	low

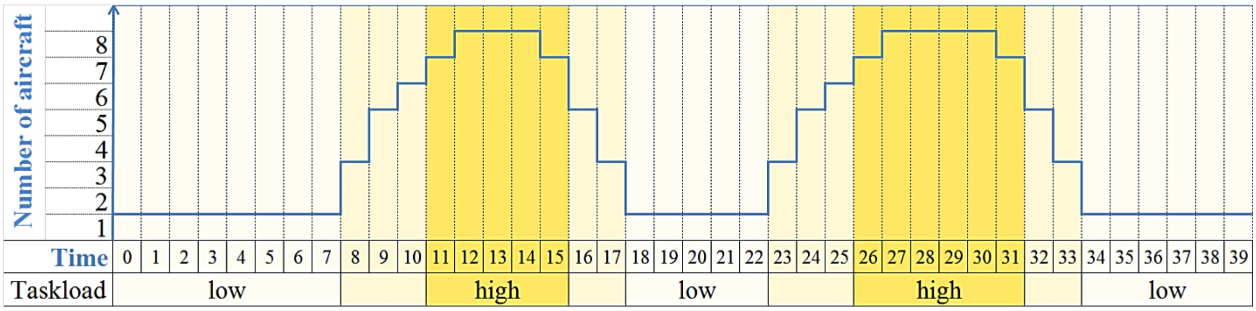

Figure 2 Overview of the scenarios in relation to task load

3.2.1 Apparatus

Figure 3 shows the experimental set-up installed in the Tower Lab of the Institute of Flight Guidance at the German Aerospace Center. The participants were seated in the center of the workspace, wearing headset and eye-tracking glasses. Five 40-inch LC displays arranged in a semi-cycle formed a compressed 180° outside view of the tower (Video Panorama). The displays in front of the participant presented Radar, Radio, digitally depicted flight strips (Flight-Strips), and Weather and Time (WaT) information. Data entry into the flight strip system was possible via mouse and keyboard. Radio communication took place with two

pseudo-pilots managing the scenario from the room next door. The SARA-T probe technique was presented on a separate display.

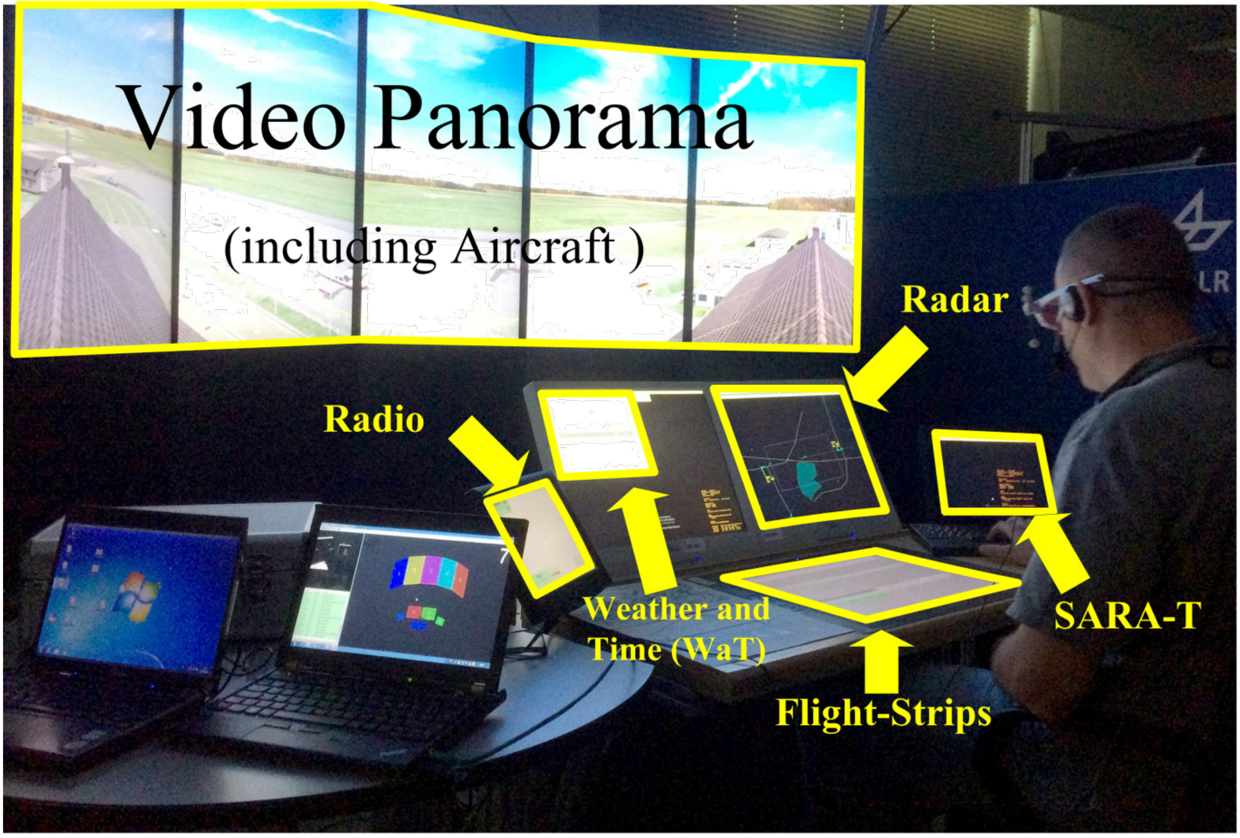

Figure 3 Tower Lab: experimental setting and AoIs

3.2.2 Performance measurement

Table 2 presents the selected performance indicators that are directly recorded by the simulation to establish a high degree of reliability and objectivity. The measurements in Table 2 were collected for each run.

Table 2 Performance measures used in the study

Performance Category	Measure	Description
Safety	Number of separation losses	A prescribed safety distance between aircraft (or aircraft and objects) must always be maintained to avoid conflict situations.
Efficiency	Number of take-offs	The scenario requires several clearances for take-off through the ATCO. The number of conducted take-offs therefore reflects the capability of the ATCO to efficiently regulate outgoing traffic.
	Aircraft taxi time	The taxi time for parking aircraft to the runway and from landing aircraft to the gate varies depending on the control style of the particular ATCO.
	Duration of radio communication	The ATCO has to communicate with the pilots via radio, giving precise instructions in a time-efficient manner to minimize the usage of the radio frequency.

3.2.3 Situation Awareness Measurement

For measuring SA, two methods were used: self-rating questionnaires after each run and probe technique during runs. In the first method, SA was evaluated through the SASHA (Dehn, 2008) consisting of six questions (Table 3) about the previous workload period(s). Answers are given on a seven-point scale ranging from “never” to “always” (item values 0 to 6). The SASHA score for SA is the total sum of the single item values, inverting the scores of items 2, 3, 5, and 6, and dividing it by six. Therefore, the SASHA score ranges between 0 (low SA) and 6 (high SA).

Table 3 SASHA items with information about the scale inverting (Dehn, 2008).

Item ID	Items	Scale inverted
	In the previous working period(s) …	
1	... I was ahead of the traffic.	
2	... I started to focus on a single problem or a specific area of the sector.	*
3	... there was a risk of forgetting something important (like transferring an a/c on time or communicating a change to an adjacent sector).	*
4	... I was able to plan and organise my work as I wanted.	
5	... I was surprised by an event I did not expect (like an a/c call).	*
6	... I had to search for an item of information.	*

For the second approach, the probe technique SARA-T was chosen. Derived from the Tower Traffic Management goal hierarchy by Strater, Tinsley, Costello, Colombo, and Endsley (2010) and by SA requirements described in Jeannot (2000), items for SA were selected. In cooperation with subject matter experts, the items were refined for tower operations. In a three-minute interval, an item is randomly selected and announced as “Incoming questionnaire” on the SARA-T display. The questions covered perception (e.g. “Please estimate the number of in-air flights.”) as well as comprehension (e.g. “Please state the call sign of the last aircraft that landed on the runway.”). While the item was being presented, the log file of the simulation was analyzed and the correct answer was made available in real-time. The correct answers and 3 distractors were presented as multiple-choice test. If the item was answered correctly, the response time was used as measurement for SA (D. Chiappe, Strybel, & Vu, 2015).

3.2.4 Eye tracking

An eye-tracking system was used to identify the source of information uses to interact with the ATC workplace. Eye data was collected with the SMI Eye Tracking Glasses 2 Wireless (SensoMotoric Instruments, Germany). This head-mounted system collects eye direction, gaze, and head position with a sampling rate of 60 Hz binocular, a gaze-tracking accuracy of 0.5° over all distances, and a gaze-tracking range of 80° horizontal and 60° vertical. A one-

point calibration was carried out to make sure that the eyes were correctly tracked by the iView System. The Integration Guideline for Dynamic Areas of Interest (IGDAI) was applied to ensure the capturing of eye data and objects (dynamic and static) with a synchronized time stamp and in the same coordination system (Friedrich, Rußwinkel, & Möhlenbrink, 2016). Due to the problem of the clear distinction between fixation, saccades, and smooth pursuits in dynamic task environments (Papenmeier & Huff, 2010), only the raw eye data dwell time on a defined area of interest (AoI) was used. The statistical software EyeTrackingAnalyser by DLR (Friedrich et al., 2016) was used to analyze the eye-tracking data. For every run, the areas of interest were used to classify the eye movement. Due to the application of IGDAI, dynamic areas of interest for the moving objects (aircraft) presented on the video panorama were defined and automatically classified. The sources of information were used to identify possible control strategies that interact with SA and task load.

3.3 Procedure

The participants were welcomed and the experimenter explained the procedure of the experiment. The participants were advised to perform ATC in their usual way. It was emphasized that the participation was voluntary and could be abandoned at any point. The participants signed a consent form in which the anonymity of data collection and analysis was guaranteed. The demographic information was collected via a preliminary questionnaire. The participants were then trained on the electronic flight strip system by creating flight strips, changing their status and deleting them. The eye-tracking glasses were put on and calibrated. The radio headset was adjusted and checked for proper function.

The participants were randomly assigned to begin with scenario A or B. 40 minutes into the run, the simulation was stopped. SARA-T was performed in parallel throughout the run. Immediately afterwards, the SASHA was presented on an additional laptop. Before the second run, the participants took a 5-10-minute break. After another 40 minutes, the second

run was stopped and SASHA was answered once more. At the end, the final questionnaire was answered concerning the feasibility of the SA probe question technique, followed by a debriefing.

4 RESULTS

The data from the first low task load phase (Figure 1) was removed from the analysis, because the participants used the first phase as training to get acquainted with the task environment and to equalize the total duration of high and low task load phases. Also, two phases of low task load from participant 6 were removed, due to a mistake in simulation procedure resulting in a high task load during these phases.

4.1 Performance

The performance metrics (Table 2) were classified in relation to safety and efficiency. The safety-critical performance separation loss happened twice (number of separation losses) throughout the study and was caused by the pseudo-pilots. Therefore, the safety performance was the same in all conditions. The number of take-offs was analyzed for scenarios A (M = 8.0, SD = 1.32) and B (M = 8.67, SD = 0.71). Table 4 (aircraft taxi time) and Table 5 (duration of radio communication) show the descriptive efficiency performance data separated by scenario and task load. Due to technical problems, radio communication times are not available for participant 7, scenario B and participant 9, scenario A. The descriptive performance data indicate that neither visibility nor task load had an influence on safety or efficiency.

Table 4 Average taxi time in seconds, separated by scenario and task load phase.

Scenario	Task load	Average taxi time (M)	SD	MIN	MAX
A	high	269.81	95.65	104.44	559.37
A	low	260.25	89.3	130.00	559.37
B	high	266.82	69.95	153.38	435.57
B	low	238.7	58.31	132.00	362.00

Table 5 Average radio duration in seconds for the participants separated by scenario and task load phase.

Scenario	Task load	Average radio communication (M)	SD	MIN	MAX
A	high	5.04	1.18	3.41	7.28
A	low	4.15	0.76	3.06	6.25
B	high	4.45	0.96	3.28	6.69
B	low	3.96	0.79	2.59	5.50

4.2 Effects on Situation Awareness

As measurements for SA, SARA-T captures the response time for an answer to a probe as well as the correctness of it. The SARA-T response times and correctness of answers were aggregated for participants, separated by scenarios and task load phases. The response times are only considered valid if the probe answer was correct. In connection to previous studies (e.g. Kass et al., 2007; Patten et al., 2006) that indicate a connection between SA, age, and experience, the first analysis presented the correlations between these factors. Figure 4 presents the results regarding the SARA-T response times and age (left) and experience (right). The same analysis was repeated using the SASHA scores instead of the SARA-T response times, but using the Kendall's rank test none of these correlations were significant ($.244 < p < .923$).

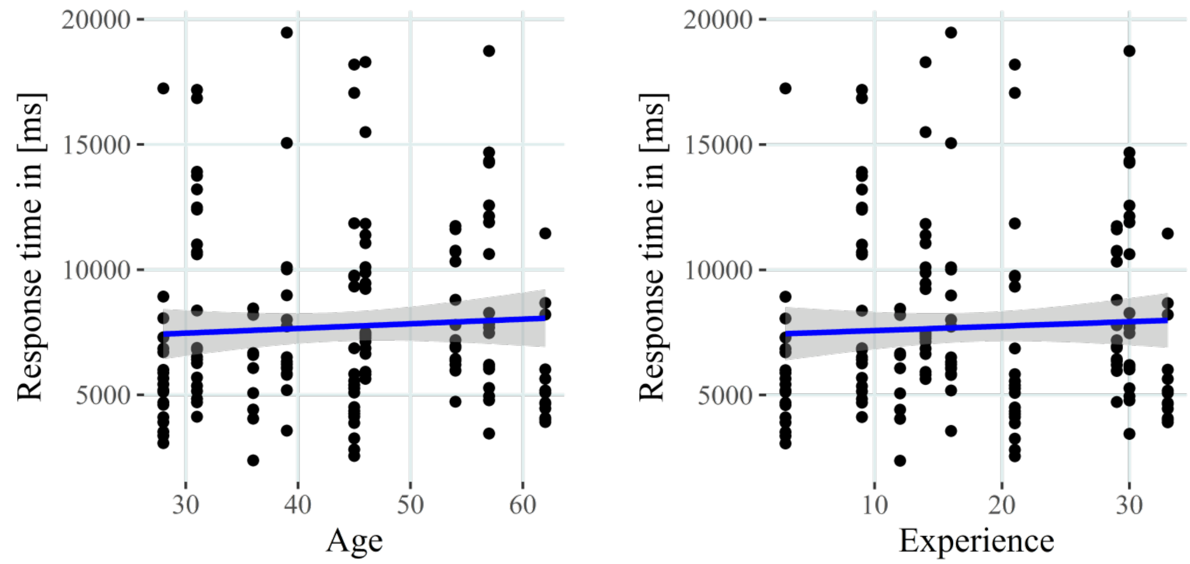

Figure 4 Correlation between SARA-T reaction time and age (left) and experience (right) for all task load phases.

Figure 5 presents the descriptive results separated by the scenario with high (A) and low (B) visibility. Although the total average of correct answers increases (Figure 5, left) from scenario A (M = 0.58, SD = 0.40) to scenario B (M = 0.80, SD = 0.25), none of the pairwise-conducted Wilcoxon Signed-Ranks test scores were statistically significantly different (.089 $< p <$ .674). Wilcoxon Signed-Ranks test was also conducted for the response times (Figure 5, right), which also led to no statistically significantly different scores (.098 $< p <$ .726).

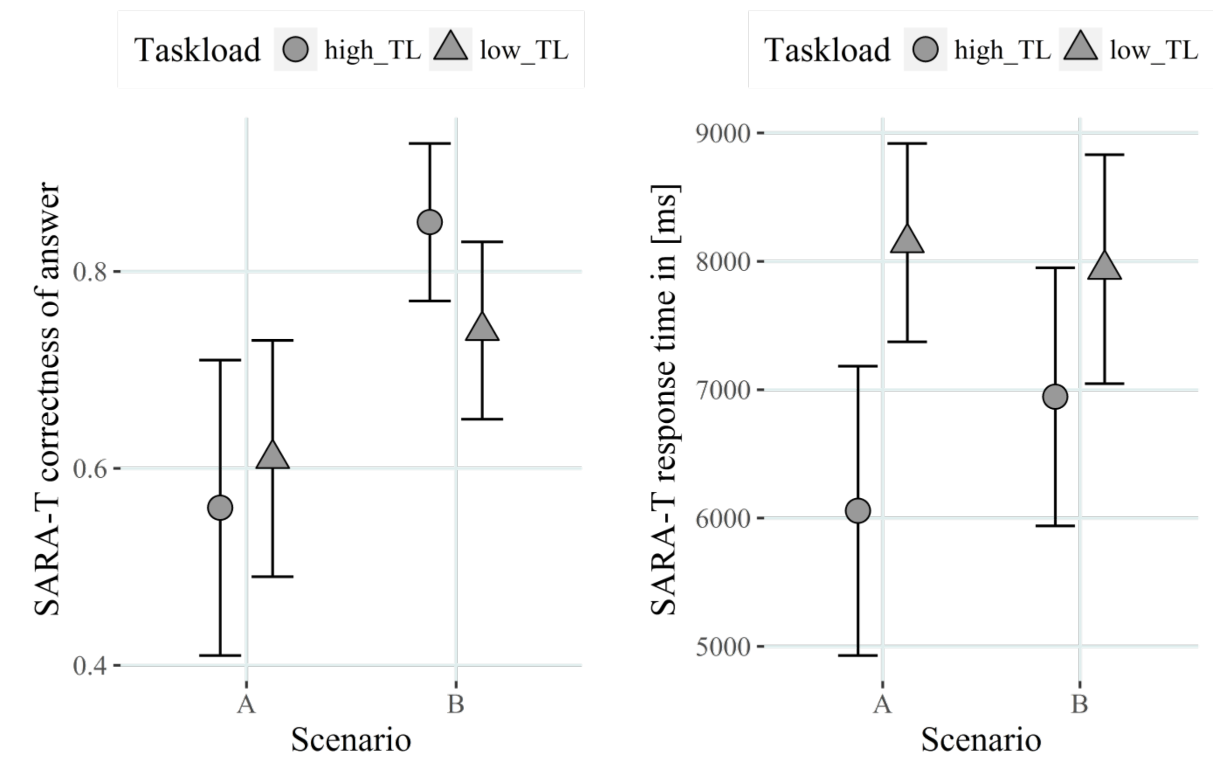

Figure 5 Answers within the scenarios (high (A) vs. low (B) visibility) and per task load phase (left, with an N=9 for each mean value) and the response times within the scenarios (high (A) vs. low (B) visibility) and per task load phase (right, with an N=9 for each mean value). The error bars in both figures represent the standard error

A Wilcoxon Signed-Ranks Test indicated that SASHA scores did not differ significantly for each scenario ($Z=6$; $p < .087$). However, a tendency can be observed: six subjects rated their SA via SASHA higher under low visibility conditions in scenario B ($M = 4.37$, $SD = 0.57$) than under high visibility in scenario A ($M = 3.95$, $SD = 0.67$). This is also supported by the SARA-T average correctness of answers.

In order to verify if there is a connection between objective probe data of particular task load phases and the retrospective self-assessment, the valid SARA-T response times were averaged per task load phase and in total and correlated separately with the SASHA scores. If the reaction times increase, the SASAH score should indicate less situation awareness. Figure 6 presents the correlation for SARA-T response times and SASHA scores

in high task load (above left), low task load (above right), and in total (below). Based on the results of the Kendall's tau test, only the response times within the high task load phases show a significant correlation with the SASHA score (r_s = -0.400, p = .08).

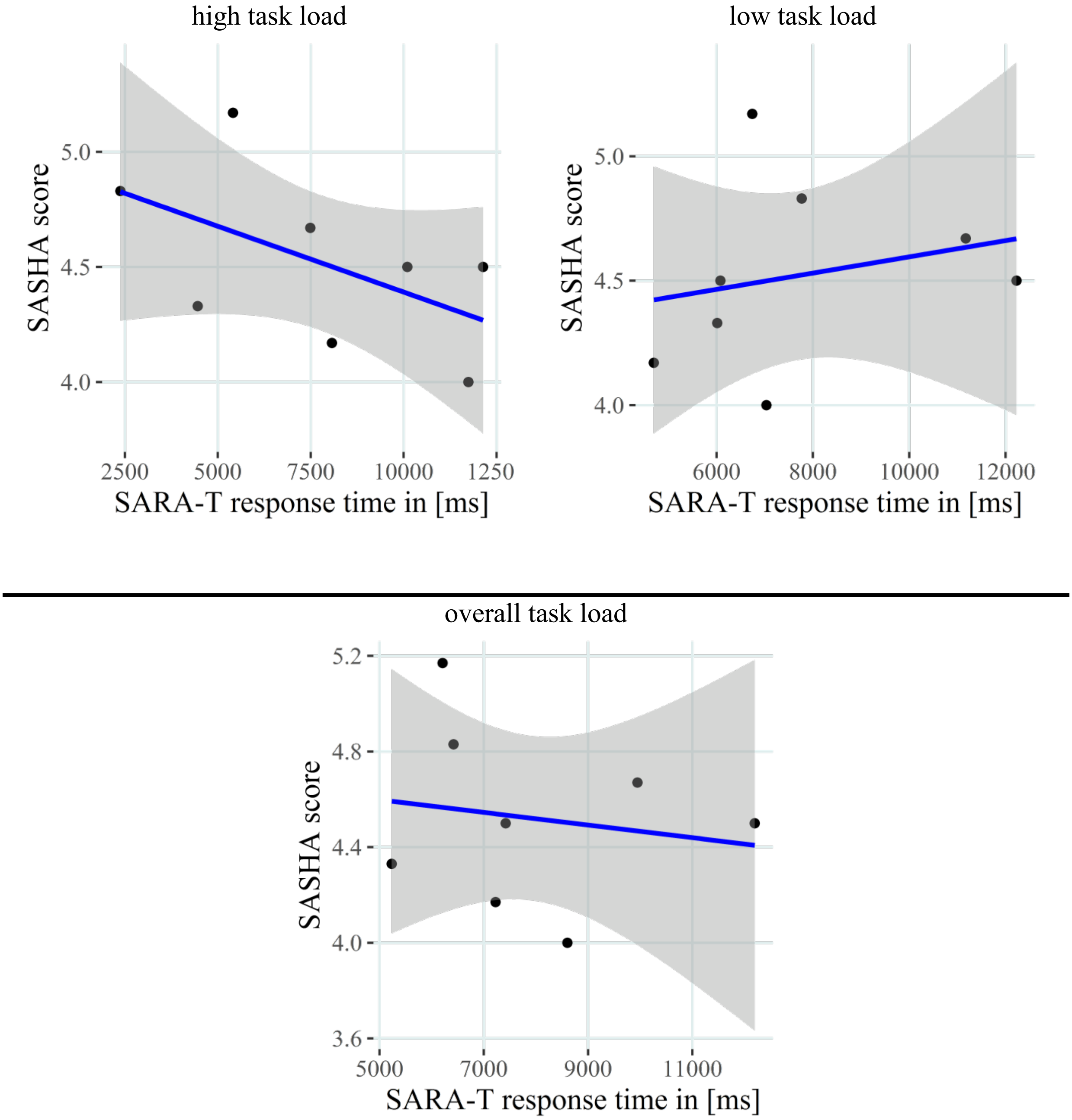

Figure 6 Correlation between SASHA and SARA-T duration for high (top left) task load, low (top right) task load, and the overall duration for scenario B

Within the debriefing, participants stated that the SARA-T questions were moderately not disruptive (M = 2.33, SD = 0.50; 1: not disruptive to 4: disruptive). The majority of the

participants deemed the SARA-T questions moderately not appropriate for assessing their situation awareness (M = 2.56, SD = 0.53; 1: appropriate to 4: not appropriate). They indicated that they felt additionally challenged by the questions (M = 2.22, SD = 0.67) and that the questions were more distracting in phases with more than four aircraft (M = 2.44, SD = 1.01; 1: distractive to 4: not distractive) than in phases with less than four aircraft (M = 3.33, SD = 0.87).

4.3 Effects on Control Strategy

The identification of the control strategy was based on the change in the information-gathering process that is measured by eye tracking. Eye movement is valid if the eye tracker detects the eyes and the gaze vector hits an area of interest (AoI, Figure 3). The amount of valid eye movement (valid eye point divided by total number of captured eye point) per participant was calculated and compared. The average of eye-data validity reached M = 72.6% (SD = 15.6), with only two participants (4 and 7) below 50%. Both participants wore glasses underneath the eye-tracking device that reduced the amount of captured eye data, but their general distribution of eye tracking data showed no sign of outliers and their data was therefore considered within this analysis. The average dwell time for all participants was spent on the AoI Flight-Strips (40%), Radar (20%), Video Panorama (17%), Aircraft (11%), WaT (5%), SARA-T (6%), and Radio (1%). For this analysis, the eye data was separated by scenario. A first inspection of data (Table 6) revealed that results from both scenarios do not differ significantly; only the results for scenario B (low visibility; highest average of correct SARA-T answers) are presented in detail.

The eye data was categorized according to task load phases to measure their influence on the control strategy. Figure 7 shows the dwell time per AoI within high and low task load. The values represent the percentage of dwell time per AoI, and the error bars show the standard error. Summarizing the values for all AoIs results in 100% and therefore represents the

complete gathering of information within the task load phases. Each participant looked at all AoIs at least once per run and task load phase, except for Radio (only 2 participants looked at it).

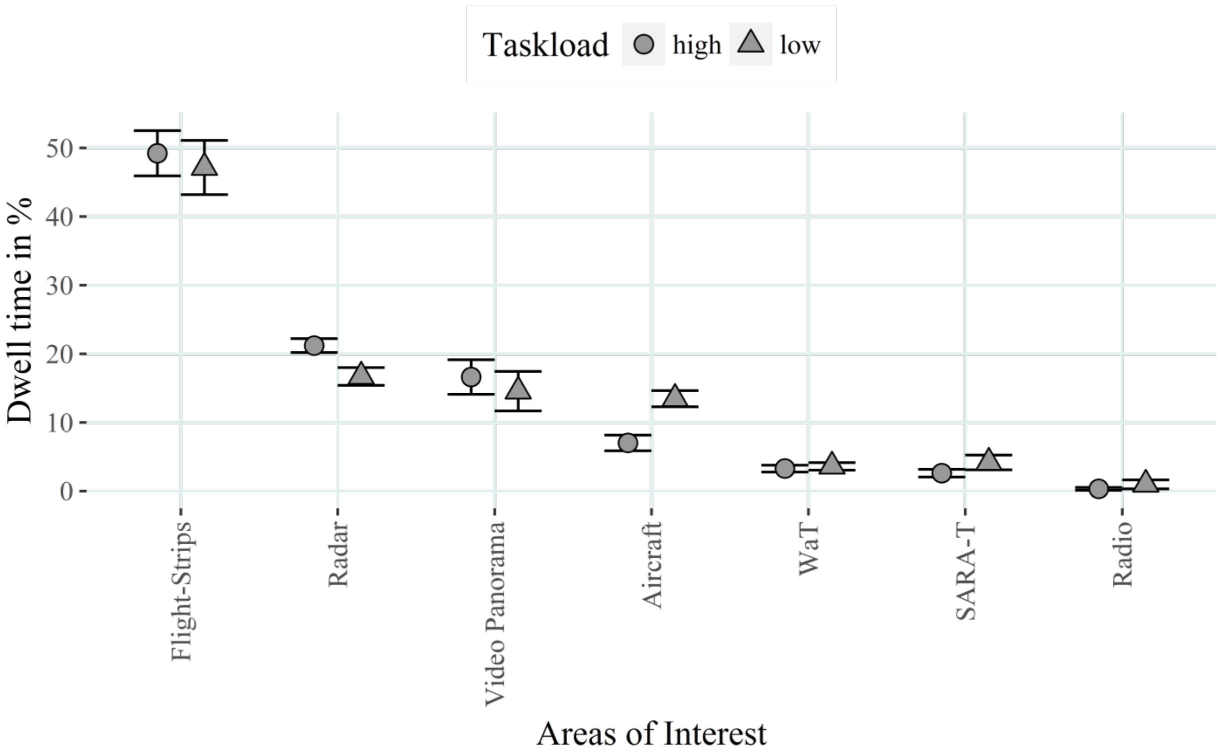

Figure 7 Dwell times for the AoIs split by task load phase for scenario B. The error bars represent the standard error

Table 6 shows the results of the Wilcoxon Signed-Ranks Tests performed for each AoI to identify significant differences between the task load phases, separated by scenario. These differences represent a change in control strategy and were identified for the AoIs Radar and Aircraft. If the task load is high, significantly more visual attention has to be directed to the Radar and less to the aircraft. This is reversed if the task load is low. No other AoI showed a significant difference depending on task load.

Table 6 F-Test results separated for each AoI and task load testing for percentage of dwell times.

AoI	Wilcoxon Signed-Ranks Tests (Scenario A)	p<0.05	Wilcoxon Signed-Ranks Tests (Scenario B)	p<0.05
Flight-Strips	V=34; p=.203		V=29; p=.499	
Radar	V=40; p=.039	*	V=40; p=.039	*
Video Panorama	V=1.88; p=.212		V=35; p=.164	
Aircraft	V=0; p=.004	*	V=0; p=.004	*
WaT	V=11; p=.203		V=18; p=.652	
SARA-T	V=7; p=.07		V=16; p=.496	
Radio	V=1; p=1		V=1; p=1	

A general criticism concerning the eye-tracking glasses was observed within the debriefing. Participants stated that wearing the eye-tracking glasses was an interference factor (M = 2.78, SD = 0.97; 1: not disturbing, to 4: disturbing).

5 DISCUSSION

Even so, in general, the sample size of this study reduces the explanatory power; the sample of nine is average for a study with ATCOs that work in the control tower (Kuk, Arnold, & Ritter, 1999; Metzger & Parasuraman, 2005; Teutsch & Postma-Kurlanc, 2014). Interesting indications were found for the application of probe techniques for SA and the measurement of the influence of increased task load on the control strategy. The focus of the experiment was set on the application of probe techniques for SA and the detection of change in control strategy in different task load phases. Therefore, the authors used a high-fidelity simulation as experimental set-up and a sample of nine ATCOs.

The manipulation of high task load for the tower ATC is more complex than for the en route or approach control, because the ATCOs can delay take-offs if they are too busy. On the contrary, with a maximum of 6 minutes in duration, the phases of low task load were probably too short to create intense monotony or boredom. However, as the performance data showed no difference between scenario and task load phases, it is indicated that the

individual performance stayed at the same level and that no overload was reached that might have impaired the results.

Concerning the influence of age and experience on SA (Kass et al., 2007; Patten et al., 2006), no correlation was found between either. This indicates that, as predicted by the cascade model (Fry & Hale, 2000; Gregory et al., 2009), constant repetition in connection to a task could decrease the effects of age and experience. It is worth mentioning that ATC is a training-intensive (3 to 4 years) profession that should guarantee the same safety and performance from each ATCO.

According to Hendy (1995), high task load at acceptable levels and good SA are linked because both involve the same internal processes. This is supported by the correlation between SARA-T and SASHA at high task load phases: when asked for a self-assessment of their SA during the previous task, participants in this study tended to refer to phases of high task load rather than to low task load, similar to Gontar and Hoermann (2015). It also indicates that retrospective self-assessment of SA is linked to the situations where a good SA is needed the most to successfully interact with the task environment. Taking into account that both performance and SA measures show no significant differences for high and low task load phases, SA itself may be an aspect of performance rather than a factor leading to it. For example, SA is considered as a cognitive skill in the NOTECHS rating scale for the assessment of pilots' Crew Resource Management skills (Flin et al., 2003). The question therefore remains as to whether SA is a factor influencing performance or if SA is the outcome of a task, thus performance itself. Considering the great number of theories, articles, and discussions on SA (e.g. Edwards, 2013; Flin et al., 2003; Hendy, 1995), this question cannot be answered unequivocally. After all, it may be a matter of definition.

Since SARA-T was classified as not disruptive during the debriefing, the extra workload for the operator, as proposed by Jeannot (2000), was kept to an absolute minimum.

However, participants of this study considered the probe questions rather not appropriate for assessing their SA, as the debriefing showed. On the one hand, this could indicate that the tower ATCOs had a different, more holistic understanding of SA which cannot be broken down to single probe questions on concrete events during simulation. On the other hand, this calls for an improvement and validation of the probe questions used in the context of tower ATC. The SASHA tendency to contradict the assumption that reduced visibility impairs SA was supported by the analysis of SARA-T. Therefore, the assumption is supported by SASHA and SARA-T, even though both tests results are not significant. We therefore believe the results of the SARA-T are reliable and valid, which is more important than the appropriateness viewed by the participants.

In contrast to the findings of Karikawa et al. (2014), the performances between the task load phases did not differ. Furthermore, through the application of the probe technique, it was shown that the SARA-T values did also not vary significantly between the task load phases. By following the model of the interconnection (Figure 1), the control strategies should change over time. As for the en route (Karikawa et al., 2014) and the approach ATC (Sperandio, 1971), the results showed the existence of different control strategies for the tower ATC environment. The analysis implies the existence of one control strategy for the high task load phases that concentrates on the radar information, and another one for low task load phases that focuses more on aircraft. Within the high task load phases, the information for managing the workplace is combined on the radar to bring all the aircraft in relative position to each other. Therefore, compared to the Video Panorama, the Radar offers the advantage of providing task-relevant information (callsign and position) at one glance. By decreasing the amount of tracking for one single aircraft, the high task load control strategies follow the same pattern as for the approach control workplace (Sperandio, 1978). The low task load phases normally had one aircraft that allowed for a use of the Video Panorama to

monitor aircraft more specifically. The authors believe that the change in information gathering is connected to the density of information at the radar display. As above, this also supports the model of the interconnection and even extends it by using the percentage of dwell time to show expected behavior.

6 CONCLUSION

In spite of the small sample size, some important indications were found, confirming the connection of the three factors task load, SA, and control strategy. The work presented in this article provided valuable insight into the human performance in tower ATC. The experimental design manipulated SA via visibility and task load via number of aircraft. The results showed that probe technique allows the online determination of SA within different task load phases. SARA-T is essential if the task load varies within a scenario, because the post-run self-assessment is not precise enough. Discriminating SA offline seems to be affected by the task load the participants have experienced during the runs. SARA-T is a helpful support for the analysis of SA, especially in a dynamic task environment such as the ATC environment. It provides a detailed view of the development of SA throughout the runs and helps in understanding its different states, depending on the situations.

Considering that performance and SA remained constant between the runs and the task load phases, the relevant factor had to be the switch between strategies. The analysis of the control strategy showed the influence of increased task load on the process of information gathering. The participants switched from monitoring the aircraft directly to monitoring the radar display when the task load increased.

Future research needs to focus further on interactions between different performance-shaping factors. For the development of a detailed model including these interconnections, overload situations should be used to identify ceiling effects. In this context, it would have been recommendable to create higher levels of task load by adding more aircraft to the

scenarios. This is needed to analyze the connections between factors more closely. In addition, this could lead to the identification of a control strategy through which the amount of traffic is actively reduced by postponing or denying clearances.

REFERENCES

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics, 36*, 623-636.

Ambroggi, M. d., & Trucco, P. (2011). Modelling and assessment of dependent performance shaping factors through Analytic Network Process. *Reliability Engineering & System Safety, 96*(7), 849-860.

Bacon, L. P., & Strybel, T. Z. (2013). Assessment of the validity and intrusiveness of online-probe questions for situation awareness in a simulated air-traffic-management task with student air-traffic controllers. *Safety science, 56*, 89-95.

Baumann, M. R. K., & Krems, J. F. (2007). Situation awareness and driving: A cognitive model. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments* (pp. 253-265). London: Springer.

Bolstad, C. A. (2001). *Situation awareness: does it change with age?* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology, 42*(3), 361-377. doi:http://dx.doi.org/10.1016/0301-0511(95)05167-8

Chiappe, D., Strybel, T. Z., & Vu, K.-P. L. (2015). A Situated Approach to the Understanding of Dynamic Situations. *Journal of Cognitive Engineering and Decision Making, 9*(1), 33-43. doi:10.1177/1555343414559053

Chiappe, D. L., Strybel, T. Z., & Vu, K.-P. L. (2012). Mechanisms for the acquisition of situation awareness in situated agents. *Theoretical Issues in Ergonomics Science, 13*(6), 625-647.

Corradini, P., & Cacciari, C. (2002). The effect of workload and workshift on air traffic control: a taxonomy of communicative problems. *Cognition, technology & work, 4*(4), 229-239.

Cox-Fuenzalida, L.-E. (2007). Effect of workload history on task performance. *Human Factors, 49*, 277-291.

Dehn, D. M. (2008). Assessing the Impact of Automation on the Air Traffic Controller: The SHAPE Questionnaires. *Air Traffic Control Quarterly, 16*(2), 127-146. doi:10.2514/atcq.16.2.127

Dekker, S. W. (2015). The danger of losing situation awareness. *Cognition, technology & work, 17*(2), 159-161.

Dekker, S. W., & Woods, D. D. (2002). MABA-MABA or abracadabra? Progress on human–automation co-ordination. *Cognition, technology & work, 4*(4), 240-244.

Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., Nikolic, D., & Manning, C. A. (1998). Situation awareness as a predictor of performance for en route air traffic controllers. *Air Traffic Control Quarterly, 6*(1), 1-20.

Durso, F. T., & Sethumadhavan, A. (2008). Situation awareness: Understanding dynamic environments. *Human Factors, 50*(3), 442-448. doi:10.1518/001872008x288448

Edwards, T. (2013). *Human performance in air traffic control.* (PhD), Dissertation, University of Nottingham,

Endsley, M. R. (1995a). Measurement of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*, 65-84.

Endsley, M. R. (1995b). Towards a Theory of Situation Awareness in Dynamic Systems. *Human Factors, 37(1)*, 32-64.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 400). Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R. (2015). Situation awareness: operationally necessary and scientifically grounded. *Cognition, technology & work, 17*(2), 163-167. doi:10.1007/s10111-015-0323-5

Endsley, M. R., & Rodgers, M. D. (1996). *Attention distribution and situation awareness in air traffic control.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

EUROCONTROL/FAA. (2010). *Human Performance in Air Traffic Management Safety: A White Paper*. Retrieved from Action Plan 15 Safety:

Flin, R., Martin, L., Goeters, K.-M., Hormann, H., Amalberti, R., Valot, C., & Nijhuis, H. (2003). Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. *Human Factors and Aerospace Safety, 3*, 97-120.

Friedrich, M., Rußwinkel, N., & Möhlenbrink, C. (2016). A guideline for integrating dynamic areas of interests in existing set-up for capturing eye movement: Looking at moving aircraft. *Behavior Research Methods, 49*(3). doi:10.3758/s13428-016-0745-x

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological science, 7*(4), 237-241.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54*(1), 1-34.

Fürstenau, N. (2016). *Virtual and Remote Control Tower: Research, Design, Development and Validation*: Springer International Publishing.

Discriminability of flight maneuvers and risk of false decisions derived from dual choice decision errors in a videopanorama-based remote tower work position., 8020 105-114 (Springer Berlin Heidelberg 2013).

Gontar, P., & Hoermann, H. J. (2015). Interrater Reliability at the Top End: Measures of Pilots' Nontechnical Performance. *International Journal of Aviation Psychology, 25*(3-4), 171-190. doi:10.1080/10508414.2015.1162636

Gregory, T., Nettelbeck, T., Howard, S., & Wilson, C. (2009). A test of the cascade model in the elderly. *Personality and Individual Differences, 46*(1), 71-73. doi:https://doi.org/10.1016/j.paid.2008.08.017

Griffin, M., Neal, A., & Neale, M. (2000). The contribution of task performance and contextual performance to effectiveness: Investigating the role of situational constraints. *Applied Psychology, 49*(3), 517-533. doi:10.1111/1464-0597.00029

Hadley, G. A., Guttman, J. A., & Stringer, P. G. (1999). Air traffic control specialist performance measurement database. (DOT/FAA/CT-TN99/17).

Hendy, K. C. (1995). *Situation awareness and workload: birds of a feather?* : DTIC Document.

International Organization for Standardization. (1991). Ergonomic principles related to mental work-load. In *General terms and definitions* (Vol. ISO 10075:1991).

Jeannot, E. (2000). Situation awareness, synthesis of literature research. *EEC note*(16/00).

Jeannot, E., Kelly, C., & Thompson, D. (2003). *The Development of Situation Awareness Measures in ATM Systems*, Brussels, Belgium.

Kahneman, D. (1973). *Attention and effort*: Citeseer.
Karikawa, D., Aoyama, H., Takahashi, M., Furuta, K., Ishibashi, A., & Kitamura, M. (2014). Analysis of the performance characteristics of controllers' strategies in en route air traffic control tasks. *Cognition, technology & work, 16*(3), 389-403.
Kass, S. J., Cole, K. S., & Stanny, C. J. (2007). Effects of distraction and experience on situation awareness and simulated driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 10*(4), 321-329. doi:10.1016/j.trf.2006.12.002
Kerkau, F. (2005). *Biosignale der Pupille zur Steuerung intelligenter User-Interfaces: Untersuchung von Pupillenbewegungen zur Realisierung einer biopsychologischen Computerschnittstelle für die Mensch-Computer-Interaktion.* (PhD), Freie Universität Berlin, Berlin.
Kirwan, B., Bettignies-Thiebaux, B., & Scholte, J. (2014). *OPTICS: EXPERT WORKSHOP: HUMAN FACTORS.* Paper presented at the 1st Expert Workshop: Human Factors, Brussel.
Kontogiannis, T., & Malakis, S. (2013). Strategies in controlling, coordinating and adapting performance in air traffic control: modelling 'loss of control'events. *Cognition, technology & work, 15*(2), 153-169.
Kraemer, J., & Süß, H.-M. (2015). Real Time Validation of Online Situation Awareness Questionnaires in Simulated Approach Air Traffic Control. *Procedia Manufacturing*, 3152-3159.
Kuk, G., Arnold, M., & Ritter, F. E. (1999). Effects of light and heavy workload on air traffic tactical operations: a hazard rate model. *Ergonomics, 42*(9), 1133-1148.
Lange, M. (2014). *Information Search for Decision-Making in Tower Air Traffic Control: A Field Study.* (Master), Tu Chemnitz,
Lee, Y. H., Jeon, J.-D., & Choi, Y.-C. (2012). Air Traffic Controllers' Situation Awareness and Workload under Dynamic Air Traffic Situations. *Transportation Journal, 51*(3), 338-352. doi:10.5325/transportationj.51.3.0338
Loft, S., Bowden, V., Braithwaite, J., Morrell, D. B., Huf, S., & Durso, F. T. (2015). Situation awareness measures for simulated submarine track management. *Human Factors, 57*(2), 298-310.
Mensen, H. (2014). *Moderne Flugsicherung: Organisation, Verfahren, Technik[Modern Air Traffic Management: Organisation, Procedures, Methodes]*: Springer-Verlag.
Metzger, U., & Parasuraman, R. (2005). Automation in Future Air Traffic Management: Effects of Decision Aid Reliability on Controller Performance and Mental Workload. *Human Factors, 47*(1), 35-49. doi:doi:10.1518/0018720053653802
Papenfuss, A., & Friedrich, M. (2016). Head Up Only – A design concept to enable multiple remote tower operations. *35th Digital Avionics Systems Conference*.
Papenfuss, A., Friedrich, M., Möhlenbrink, C., Rudolph, M., Schier, S., Schmidt, M., & Fürstenau, N. (2010). Assessing Operational Validity of Remote Tower Control in High-fidelity Tower Simulation. *IFAC Proceedings Volumes, 43*(13), 117-122. doi:http://dx.doi.org/10.3182/20100831-4-FR-2021.00022
Papenmeier, F., & Huff, M. (2010). DynAOI: A tool for matching eye-movement data with dynamic areas of interest in animations and movies. *Behavior research methods, 41*, 179-187. doi:10.3758/BRM.42.1.179
Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making, 2*(2), 140-160.

Patten, C. J., Kircher, A., Ostlund, J., Nilsson, L., & Svenson, O. (2006). Driver experience and cognitive workload in different traffic environments. *Accid Anal Prev, 38*(5), 887-894. doi:10.1016/j.aap.2006.02.014

Redding, R. E. (1992). Analysis of Operational Errors and Workload in Air-Traffic-Control. *Proceedings of the Human Factors Society, 36th Annual Meeting, Vols 1 and 2*, 1321-1325.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring Situation Awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics, 39*(3), 490-500.

Shelton, C. L., Kinston, R., Molyneux, A. J., & Ambrose, L. J. (2012). Real-time situation awareness assessment in critical illness management: adapting the situation present assessment method to clinical simulation. *BMJ quality & safety*, bmjqs-2012-000932.

Sperandio, J.-C. (1971). Variation of Operator's Strategies and Regulating Effects on Workload. *Ergonomics, 14*(5), 571-577. doi:10.1080/00140137108931277

Sperandio, J.-C. (1978). The Regulation of Working Methods as a Function of Work-load among Air Traffic Controllers. *Ergonomics, 21*(3), 195-202. doi:10.1080/00140137808931713

Strater, L., Tinsley, M., Costello, A., Colombo, D., & Endsley, M. (2010). Situation awareness requirements for Optimized Profile Descent and integrated arrivals and departures. *NASA Ames Research Center*.

Taylor, R. M. (1990). *Situational Awareness Rating Technique(SART): The development of a tool for aircrew systems design.* Paper presented at the Situational Awareness: Limitations and Enhancement in the Aviation Environment, Neuilly-Sur-Seine, France.

Teutsch, J., & Postma-Kurlanc, A. (2014). *Enhanced Virtual Block Control for Milan Malpensa Airport in Low Visibility.* Paper presented at the Integrated Communications, Navigation and Surveillance Conference (ICNS) 2014.

Vidulich, M. A. (2000). *The relationship between mental workload and situation awareness.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Wickens, C. D. (2002). Situation awareness and workload in aviation. *Current directions in psychological science, 11*(4), 128-133.

Wickens, C. D., Mavor, A. S., & McGee, J. P. (1997). *Flight to the future: Human factors in air traffic control*. Washington, D.C.: National Academies Press.

Yang, C.-W., Yang, L.-C., Cheng, T.-C., Jou, Y.-T., & Chiou, S.-W. (2012). Assessing mental workload and situation awareness in the evaluation of computerized procedures in the main control room. *Nuclear Engineering and Design, 250*, 713-719. doi:10.1016/j.nucengdes.2012.05.038

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics, 58*(1), 1-17.

V Contributions to conferences

Friedrich, M., Pichelmann, S., Papenfuß, A., & Jakobi, J. (2017). The Evaluation of Remote Tower Visual Assistance System in Preparation of Two Design Concepts. In D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics: Performance, Emotion and Situation Awareness (1 ed., pp. 285). Vancouver: Springer International Publishing.

Friedrich, M., & Möhlenbrink, C. (2013). Which data provide the best insight? A field trial for validating a remote tower operation concept. Tenth USA/Europe Air Traffic Management Research and Development Seminar.

Sachs, F., Pruter, I., Walter, H., Friedrich, M., & Mielke, O. (2017). Simulator Study on Gyrocopter Pilot Performance in Different Takeoff Procedure Scenarios. Paper presented at the 66. Deutsche Luft- und Raumfahrtkongress, Munich.

Papenfuss, A., & Friedrich, M. (2016). Head Up Only – A design concept to enable multiple remote tower operations. 35th Digital Avionics Systems Conference.

VI Curriculum vitae and publications

Maik Friedrich

M.sc., PMP

Personal details

Date of birth	8th of November 1981
Nationality	German

Experience

Vocational

May 2008 - Present: **Researcher**, *German Aerospace Center*, Institute of Flight Guidance, Braunschweig, Germany.
- Task Leader of the SESAR Project 06.08.04 Single Remote TWR Ph1 V2
- Development of an eye tracking analysis software called EyeTA
- Development of work flows with Microsoft Workflow Foundation for handling critical situations
- Analysis of eye motion data with the application of dynamic areas of interest
- Execution and analysis of several experimental studies in the area of human factors

May 2007 - December 2007: **Internship**, *Pennylvania State University*, Department of Psychology, Penn State, USA.
Implementing cognitive models with Herbal and supporting the reseach of Frank Ritter

August 2003: **Internship**, *Daimler Chrysler*, Ulm, Germany.
Development of a workflow for handeling 3D data from Cat5 to Inventor

Education

Juni 2016: **Successful certification to Project Management Professional (PMP)**, *Project Management Institute*, United States.

May 2009 - Present: **PhD Student in Cognitive Modeling**, *Chemnitz University of Technology*, Germany, Department of Psychology.

March 2006 - May 2008: **Master in Applied Computer Science**, *Otto-Friedrich University Bamberg*, Germany, Department of Cognitive Systems.
Graduated with 1.7

June 2001 - Feburary 2006: **Diplom (FH) in Computer Science**, *University of Applied Sciences Schmalkalden*, Germany.
Graduated with 1.8

July 1998 - July 2000 **High School**, *Philipp-Melanchthon-Gymnasium in Schmalkalden*, Germany.

Master thesis

Title *Implementing diagrammatic reasoning strategies in a high level language: Extending and testing the existing model results by gathering additional data and creating additional strategies.*

Supervisors Ute Schmid, Frank Ritter

Computer skills

Languages C, C++, C#, Java, R

Platforms Windows, MacOS, Linux

Cognitiv Architectures Herbal, Soar

Tools Office, Latex, OpenOffice, Maya, GIMP, Photoshop

Languages

German native language

English fluent

Russian basic

Publications

M. Friedrich and F. E. Ritter. Implementing diagrammatic reasoning strategies in a high level language. unpublisched.

M. Friedrich and C. Möhlenbrink. How to evaluate remote tower metrics in connection to weather observations. an extension of the existing metrics. *Aviation Psychology and Applied Human Factors*, accepted, in press.

M. Friedrich, M. Biermann, P. Gontar, M. Biella, and K. Bengler. The maik influence of task load on situation awareness and control strategy at the atc tower environment. *Cognition, Technology Work*, 2018.

M. Friedrich, S. Pichelmann, A. Papenfuß, and J. Jakobi. The evaluation of remote tower visual assistance system in preparation of two design concepts. In *Engineering Psychology and Cognitive Ergonomics: Performance, Emotion and Situation Awareness*, page 285, Vancouver, 2017. Springer International Publishing.

M. Friedrich, N. Rußwinkel, and C. Möhlenbrink. A guideline for integrating dynamic areas of interests in existing set-up for capturing eye movement: Looking at moving aircraft. *Behavior Research Methods*, 2016.

M. Friedrich. Which metrics provide the insight needed? a selection of remote tower evaluation metrics to support a remote tower operation concept validation. In *Virtual and Remote Control Tower*, page 385, Berlin, Heidelberg, New York, 2016. Springer Publishers.

M. Friedrich, M. Schmidt, and N. Fürstenau. Autonome Kameras im Tower? Ein stufenweises Automatisierungskonzept zur Unterstützung der Lotsentätigkeit. In *49. Kongress der Deutschen Gesellschaft für Psychologie*, 2014.

M. Friedrich and C. Möhlenbrink. Which data provide the best insight? a field trial for validating a remote tower operation concept. In *Tenth USA/Europe Air Traffic Management Research and Development Seminar*, page 10, Chicago, 2013.

M. Friedrich. Weather in ATM and CDM. In *TAM Symposium*, 2013.

M. Friedrich. Application of dynamic AOI within AVIATOR II. In B. Carmen, editor, *profa Symposium on Psychological Requirements*, 2013.

M. Friedrich, B. Weber, S. Schätzle, H. Oberheid, C. Preusche, and B. Deml. Air traffic controller assistance systems for attention direction: Comparing visual, acoustical, and tactile feedback. In N. Merat, B. Yvonne, C. Oliver, J. Hamish, C. Weikert, and D. de Waard, editors, *Human Factors of Systems and Technology*, page 10, Leeds, 2011. Shaker Publishing.

M. Friedrich, A. Papenfuß, M. Jipp, L. Solf, and C. Möhlenbrink. Air Traffic Controller Assistance Systems for Attention Direction: Comparing Visual, Acoustical, and Tactile Feedback. In *9. Berliner Werkstatt Mensch-Maschine-Systeme Reflexionen und Visionen der Mensch-Maschine-Interaktion.*, page 10, Berlin, 2011.

M. Friedrich, B. Weber, and A. Papenfuß. Incorporating workflows for collaborative air traffic management. In *DGLR Fachausschusssitzung Anthropotechnik*, 2009.

M. Friedrich and F. E. Ritter. Reimplementing a diagrammatic reasoning model in herbal. In *International Conference on Cognitive Modelling*, 2009.

M. Friedrich, M. A. Cohen, and F. E. Ritter. A gentle introduction to xml within herbal. Tech. report, Applied Cognitive Science Lab, School of Information Sciences and Technology, Penn State. acs.ist.psu.edu/acs-lab/reports/cohenR04.pdf, 2007.

Publications as Co-Author

A. Friedrich and M. Friedrich. *Strukturierte Programmierung von Ablauf- und Zeitplansteuerungen. Problemanalyse und mathematische Modellierung.* Expert-Verlag, November 2006.

F. Sachs, I. Pruter, H. Walter, M. Friedrich, and O. Mielke. Simulator study on gyrocopter pilot performance in different takeoff procedure scenarios. In *66. Deutsche Luft- und Raumfahrtkongress*, 2017.

A. Papenfuss and M. Friedrich. Head up only – a design concept to enable multiple remote tower operations. *35th Digital Avionics Systems Conference*, 2016.

N. Fürstenau, M. Mittendorf, and M. Friedrich. Analysis of two-alternative decision errors in a videopanorama-based remote control tower working position for quantifying operator performance. In J. Brombach and H. Rausch, editors, *60. Frühjahrskongress d. Ges. f. Arbeitswissenschaft GfA*, pages 671–673. GfA-Press, 2014.

N. Fürstenau, M. Friedrich, M. Mittendorf, M. Schmidt, and M. Rudolph. Discriminability of flight man and risk of false decisions derived from dual choice decision errors in a videopanorama-based remote tower work position. 2013.

N. Rußwinkel and M. Friedrich. Zeitabhängige Entscheidungsfindung im Fluglotsenkontext. In *KogWis 2012*, 2012.

B. Weber, B. Deml, M. B. Friedrich, M. Schätzle, H. Oberheid, and C. Preusche. Vibrotaktiles Feedback zur Aufmerksamkeitslenkung bei komplexen Lotsentätigkeiten. In M. Grandt and A. Bauch, editors, *Innovative Interaktionstechnologien für Mensch-Maschine-Schnittstellen*, pages 71–85, Bonn, 2010. DGLR.

A. Papenfuß, C. Möhlenbrink, and M. Friedrich. Die SAGAT-Methode: Welche Informationen muss der Towerlotse erfassen? In *Tagung experimentell arbeitender Psychologen*, 2010.

H. Gross and M. B. Friedrich. A tool chain for gaze analysis using dynamic areas of interest. In D. de Waard, N. Gérard, L. Onnasch, R. Wiczorek, and D. Manzey, editors, *Human Centred Automation*, Maastricht, 2010. Shaker Publishing.

M. Schmidt, M. Rudolph, A. Papenfuß, M. Friedrich, W. Halle, S. Kaltenhäuser, C. Möhlenbrink, E. Schlüssler, and N. Fürstenau. Remote airport traffic control center with augmented vision video panorama. In I. AIAA, editor, *28th Digital Avionics Systems Conference (DASC 28)*. IEEE, USA, 2009.

C. Möhlenbrink, M. Friedrich, and A. Papenfuß. Remotecenter: Eine Mikrowelt zur Analyse der mentalen Repräsentation von zwei Flughäfen während einer Lotsentätigkeitsaufgabe. In *8. Berliner Werkstatt Mensch-Maschine-Systeme*, 2009.

N. Fürstenau, M. Schmidt, M. Rudolph, C. Möhlenbrink, M. Friedrich, A. Papenfuß, and S. Kaltenhäuser. Steps towards the virtual tower: Remote airport traffic control center (raice). In S. Nagaoka, editor, *ENRI International workshop on ATM/CNS*, pages 67–76, 2007.

N. Fürstenau, M. Schmidt, M. Rudolph, C. Möhlenbrink, M. Friedrich, A. Papenfuß, and S. Kaltenhäuser. Remote airport traffic control center. In FAA, editor, *ICNS Conference 2009*, 2007.

Eidesstattliche Erklärung

Die Promotion mit dem Titel

„Designing a workplace in the aviation domain: The transition to a remote air traffic control workplace by analysing the human-computer interaction.“

wurde von mir selbstständig verfasst.

Ich habe keine anderen als die angegebenen Hilfsmittel benutzt.

Maik Friedrich