

HUMBOLDT-UNIVERSITÄT ZU BERLIN



**Verhaltensindikatoren von Operateuren  
zur Evaluation der Prozessüberwachungs- und -führungsgüte  
am Beispiel einer simulierten Leitwarte im ATEO Projekt**

Diplomarbeit zur Erlangung des akademischen Grades  
Diplom-Psychologin (Dipl.-Psych.)

Eingereicht am Institut für Psychologie  
Mathematisch-Naturwissenschaftliche Fakultät II  
Humboldt-Universität zu Berlin

Vorgelegt von: Melanie Janina Christine Stade  
geboren am 10.11.1986 in Berlin  
Matrikelnummer 514168

Gutachter: Prof. Dr. Hartmut Wandke  
Dr. Jens Nachtwei

Betreuende: Prof. Dr. Hartmut Wandke  
Dr. Jens Nachtwei  
Dipl.-Psych. Charlotte von Bernstorff

Berlin, 10.07.2012

Für M.

In einem dunklen Wort.

## INHALTSVERZEICHNIS

Zusammenfassung.....	i
Danksagung.....	vii
Eidesstattliche Erklärung.....	viii
Digitaler Anhang.....	ix
Forschungsartikel.....	x

## ZUSAMMENFASSUNG

Die vorliegende Arbeit beschäftigt sich mit der Frage, wie die Leistung von Operateuren bei der Prozessüberwachung und -führung analysiert werden kann. Ziel der Arbeit war es, Parameter zur Beschreibung und Bewertung des Verhaltens von Operateuren zu explorieren, die geeignet sind, Prozessüberwachungs- und -führungsgüte von Operateuren in einer simulierten Leitwarte abzubilden. Hierfür wurden Daten aus empirischen Untersuchungen mit einer simulierten Leitwarte aufbereitet und ausgewertet. Diese Daten waren bereits erhoben bzw. wurden von der Autorin der vorliegenden Arbeit in einer eigenen Studie mit der Leitwarte erhoben. Bevor näher auf das methodische Vorgehen bei der Identifikation und Analyse von Verhaltensindikatoren eingegangen wird, folgt zunächst eine Beschreibung des Laborsystems, welches die simulierte Leitwarte umfasst, sowie des Forschungsprojektes, welches mit diesem Laborsystem forscht.

Das in dieser Arbeit verwendete Laborsystem wurde im Projekt „ATEO“ („ArbeitsTeilung Entwickler-Operateur“) entwickelt für den empirischen Vergleich der Ressourcen von Entwicklern und Operateuren bei einer (antizipierten bzw. tatsächlichen) Prozessüberwachung und -führung. Teil dieses Laborsystems ist eine Mikrowelt mit zwei Probanden, die gemeinsam ein Fahrobjekt entlang einer virtuellen Fahrstrecke steuern (sog. Tracking). Hierbei sollen Abzweigungen in Gabelungen gewählt und Hindernissen ausgewichen werden. Bei den Studien der vorliegenden Arbeit wird das Tracking überwacht und geführt durch eine dritte Versuchsperson, die als Operateur agiert. Der Operateur ist über eine Leitwarte mit der Mikrowelt verbunden. Die Leitwarte beinhaltet einen Überwachungsbereich (z.B. aktuelle Position des Fahrobjektes auf der Fahrstrecke, Vorschau des Streckenverlaufs) und einen Eingabebereich (für z.B. Warnungen vor Hindernissen). Ziel des Operateurs ist es, maximale Geschwindigkeit bei maximaler Genauigkeit des Fahrobjektes zu erreichen und somit das Tracking zu optimieren.

Das Verhalten der Operateure bei der Prozessüberwachung und -führung wurde zum einen über Eingriffe der Operateure erfasst, d.h. es wurden Klicks der Operateure im Eingabebereich der Leitwarte registriert. Zum anderen wurde das Blickverhalten der Operateure auf der Leitwarte aufgenommen, inklusive einer Bildschirmaufzeichnung der Leitwarte. Dabei wurde eine Vielzahl an Verhaltensdaten durch die Versuchspersonen im beschriebenen Laborsystem generiert. Dies ist charakteristisch für Laborsysteme, die komplexe und dynamische Mensch-Maschine-Systeme simulieren. Herausfordernd ist hierbei, geeignete Ansätze für die Identifikation von Parametern zu finden, die das Verhalten und die Leistung der Operateure angemessen abbilden. Bei der Identifikation solcher Parameter müssen u.a. Auflösungsgrad, Aggregation und Kombination der aufgezeichneten Verhaltensdaten berücksichtigt werden.

Die vorliegende Arbeit umfasst drei Forschungsartikel, in denen unterschiedliche Ansätze für die Identifikation von Verhaltensindikatoren beschrieben und diskutiert werden. Die Ansätze und Ergebnisse der Forschungsartikel werden im Folgenden kurz vorgestellt. Abschließend wird erläutert, inwieweit die berichteten Ansätze geeignet sind, das Verhalten von Operateuren bei der Prozessüberwachung und -führung im Laborsystem des ATEO Projektes zu beschreiben und zu evaluieren. Außerdem werden Ansatzpunkte für zukünftige Forschungsarbeiten (im ATEO Projekt) aufgezeigt.

Im *Forschungsartikel Nr. 1* wird die Bewertung von Blickverhalten der Operateure mittels Eingriffsverhalten der Operateure untersucht und diskutiert. Bei der Auswahl geeigneter Parameter zur Beschreibung und Bewertung des Blickverhaltens mussten die Charakteristika der Leitwarte und der Mikrowelt berücksichtigt werden. Zunächst wurden Forschungsarbeiten diskutiert, die Blickverhalten von Operateuren in Laborsystemen erhoben und ausgewertet hatten, die vergleichbar sind mit dem ATEO Laborsystem. Schwerpunkt der Diskussion war dabei, inwieweit sich die Ansätze jener Forschungsarbeiten auf das ATEO

Laborsystem übertragen lassen. Auf der Grundlage dieser Diskussion wurden dann Parameter berechnet, die das Blickverhalten der Operateure auf der Leitwarte beschreiben. Der Fokus lag hierbei auf Blickverhalten der Operateure in der Vorschau des Streckenverlaufs. Es wird angenommen, dass diese Vorschau der Operateure über den zukünftigen Streckenverlauf einen Mehrwert für die steuernden Versuchspersonen in der Mikrowelt hat. Die Parameter des Blickverhaltens repräsentieren, wie oft und wie lange diese Streckenvorschau von den Operateuren betrachtet wurde. Dieser Ausschnitt aus dem Prozessüberwachungsverhalten der Operateure (es gibt weitere Bereiche in der Leitwarte, die von den Operateuren überwacht werden können) wurde in Zusammenhang mit Prozessführungsverhalten gebracht. Hierzu wurden Parameter zur Beschreibung des Eingriffsverhaltens gewählt, von denen angenommen wird, dass diese das Ausmaß des Systemverständnisses der Operateure repräsentieren. Die Ergebnisse zeigen, dass kein signifikanter Effekt zwischen den gewählten Parametern zur Beschreibung des Prozessüberwachungs- und -führungsverhaltens gefunden wurden. Zum einen kann dieser nicht ausgewiesene Effekt auf die für empirische Forschung im Bereich Mensch-Maschine-Systeme typische kleine Stichprobe zurückgeführt werden (es wurden Daten von 13 Operateuren analysiert). Außerdem ist der Ansatz der durchgeführten Studie dadurch limitiert, dass zum einen das Blickverhalten nur für ausgewählte Bereiche der Leitwarte analysiert wurde, zum anderen wurden exemplarisch zwei Verhaltensindikatoren für die Prozessführung ausgewählt. Darüber hinaus bleibt offen, ob und in welchem Ausmaß diese gewählten Verhaltensparameter in Zusammenhang mit der tatsächlichen Optimierung des Tracking stehen.

Demnach war es notwendig, den Zusammenhang zwischen Verhaltensindikatoren der Prozessführung und der Optimierung des Tracking im ATEO Laborsystem zu untersuchen. Der Fokus der zweiten Studie (*Forschungsartikel Nr. 2*) lag auf einer umfassenderen Analyse des *Eingriffsverhaltens* der Operateure. Hierfür wurde im ersten Schritt der Analyse die Fahrstrecke in normale (d.h. kurvige) und abnormale Streckenabschnitte (d.h. mit

auftretenden Hindernissen und Gabelungen) unterteilt. Für diese Unterteilung wurde berücksichtigt, dass der Operateur den Streckenverlauf weiter im Voraus sehen kann als die Steuerer des Fahrobjekts. Wie im *Forschungsartikel Nr. 1* wurde damit der angenommene Mehrwert der Streckenvorschau des Operateurs bei der Überwachung und Führung der Mikrowelt untersucht. Um für eine quantitative Auswertung des Eingriffsverhaltens der Operateure ausreichend Varianz in den Daten zu erzielen, wurden die Eingriffe der Operateure in normalen und abnormalen Streckenabschnitten unterschiedlich aggregiert. Bewertet wurde das Eingriffsverhalten der Operateure über die Teamleistung beim Tracking (Geschwindigkeit und Genauigkeit in normalen Streckenabschnitten, Genauigkeit in abnormalen Streckenabschnitten). Für das Eingriffsverhalten der Operateure in normalen Streckenabschnitten wurde zusätzlich der Zusammenhang mit Bewertungen durch die Steuerer des Fahrobjektes untersucht. Ein signifikanter Zusammenhang zwischen den berechneten Parametern der Prozessführung und den gewählten Parametern für Prozessführungsgüte konnte nur teilweise ausgewiesen werden. Erneut war die Stichprobe mit 40 Operateuren gering. Die Aggregation der Daten über normale bzw. abnormale Streckenabschnitte hinweg war demnach notwendig, um hinreichend Varianz in den Daten zu erhalten. Dennoch ist diese Form der Datenaggregation diskutierbar und eine mögliche Ursache für geringe bzw. nicht ausgewiesene Effekte. Selten auftretendes, möglicherweise aber (im Sinne der Optimierung des Tracking) besonders wirkungsvolles Blick- und Eingriffsverhalten einzelner Operateure muss bei dieser Art der quantitativen Datenanalyse untergehen.

Um auch den Mehrwert einer qualitativen Analyse bei der Beschreibung und Bewertung von Prozessüberwachungs- und -führungsverhalten aufzuzeigen, setzt *Forschungsartikel Nr. 3* an einem qualitativen Ansatz an. Beschrieben wird hier das Verhalten von Operateuren, das z.T. nur einmalig aufgetreten und den Versuchsleitenden bei der Versuchsdurchführung bzw. bei der Durchsicht der Bildschirmaufzeichnungen aufgefallen

ist. Hierzu gehören z.B. wiederholte Klicks von Operateuren auf (kurzzeitig) inaktive Bedienknöpfe im Eingriffsbereich der Leitwarte. Dies ermöglicht Operateuren, den Zeitpunkt rechtzeitig abzapfen, in dem der Bedienknopf wieder aktiv ist – und um dann ggf. rechtzeitig einen Eingriff vorzunehmen.

Die beschriebenen Ansätze zur quantitativen Analyse des Prozessüberwachungs- und Führungsverhalten der Operateure mittels Auswertung von Blick- und Klickverhalten sind als wertvoll zu beurteilen, um Verhalten von Operateuren im ATEO Laborsystem zu beschreiben und bewerten. Eine große Herausforderung stellt(e) dabei die Komplexität und Dynamik der Prozesse im verwendeten Laborsystem dar. Einerseits ist die Auswertung von Verhalten der Operateure mit einem hohen Auflösungsgrad sehr ressourcenaufwändig. Andererseits ist solch eine Analyse notwendig, um den (komplexen) Zusammenhang von Verhaltensursache und -wirkung in Laborsystemen, die Mensch-Maschine-Systeme simulieren, untersuchen zu können. Darüber hinaus ist eine kombinierte Auswertung von Blick- und Eingriffsverhalten unverzichtbar. Beobachtungen und Mitschriften von Versuchsleitenden während der Datenerhebung können helfen, Auffälligkeiten in den aufgezeichneten Daten und erste Erklärungsansätze hierfür zu finden. Auch eine nachträgliche Durchsicht von Aufzeichnungen (z.B. des Bildschirms der Leitwarte) kann das Herleiten von Erklärungen für (nicht) gefundene Effekte in der quantitativen Analyse unterstützen. Außerdem sollte die Analyse von aufgezeichnetem Verhalten der Operateure ergänzt werden durch umfassende Nachbefragungen der Operateure. Eine retrospektive Befragung von Operateuren, unterstützt durch das Abspielen von Videoaufzeichnungen (z.B. ihr Eingriffsverhalten in verschiedenen Streckenabschnitten) kann hilfreich sein, um Verhaltensstrategien der Operateure bei der Optimierung des Tracking zu identifizieren.



In der Diskussion der Ansätze und Ergebnisse der drei durchgeführten Studien wird deutlich, dass die Suche nach geeigneten Verhaltensindikatoren zur Beschreibung und Bewertung der Leistung von Operateuren im verwendeten Laborsystem noch nicht abgeschlossen ist. Über weitere, vertiefende Analysen im ATEO Laborsystem hinaus wäre es wünschenswert, wenn sich die in dieser Arbeit beschriebene Vorgehensweise zur Analyse von Verhaltensindikatoren der Operateure auf andere Laborsysteme mit komplexen und dynamischen Mikrowelten übertragen ließe. Weitere Forschungsarbeiten müssen ebenfalls zeigen, wie und in welchem Ausmaß sich das Laborsystem sowie die in der vorliegenden Arbeit gefundenen Ergebnisse für die Beantwortung von Fragen anderer Forschungsfelder anwenden lassen, z.B. wie Führungsverhalten in virtuellen Teams evaluiert werden kann.

Im Rahmen des ATEO Forschungsprojektes kann abschließend folgendes Fazit gezogen werden: Die detaillierte Auswertung des Verhaltens der Operateure über unterschiedliche Verhaltensindikatoren sowie das Bewerten des Verhaltens mittels verschiedener Parametern der Mikrowelt war ein erster, jedoch wichtiger und vorbereitender Schritt für weitere Untersuchungen im ATEO Forschungsprojekt.

## **DANKSAGUNG**

Danken möchte ich

Jens Nachtwei für eine herausragend gute Betreuung mit ansteckender Begeisterung und richtungsweisenden Anregungen

Hartmut Wandke für viele lehrreiche Veranstaltungen und den Aufbau einer breiten Wissensgrundlage für diese und folgende Arbeiten

Charlotte von Bernstorff für Ideen, Diskussionen, Motivierendes und wohlwollende Worte

Annekatriin Hoppe für Flexibilität und Förderung bei der Arbeit

Matthias Ziegler für frühzeitig entstandene Freude an der Statistik

Malte Sönksen für privaten und beruflichen Austausch jeder Zeit

Grit Scholz und Christine Kupper für zahlreiche Aufmunterungen zwischendurch

Michael Sengpiel, Jochen Heyden, Joachim Warning und Nicolas Niestroj für Anregungen im Kolloquium

Philipp Leppert und Sabrina Schulte für ständige Erreichbarkeit bei Fragen

Knut Polkehn und Jens Hüttner für das Andocken und Vertrauen

Jens Augner und Harald Paland für das vorgelebte Hinterfragen

Elisa und Laura für Pausenlacher

Dominik für Mailterror

Sascha für gelbe Fernsehabeude

Claudi für Vietnamesisches

Andreas für offene Worte

Oli für die reservierte Couch

Steffi für warme Zimtschnecken

Ronny für Vergangenes und Zukünftiges

meinen Eltern für Alles.

## **EIDESSTATTLICHE ERKLÄRUNG**

Hiermit erkläre ich, Melanie Janina Christine Stade, geboren am 10.11.1986, in Berlin, dass ich die vorliegende Diplomarbeit mit dem Titel

*Verhaltensindikatoren von Operateuren zur Evaluation der Prozessüberwachungs- und -führungsgüte am Beispiel einer simulierten Leitwarte im ATEO Projekt*

eigenständig und ausschließlich unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Inhalte und Gedanken sind als solche kenntlich gemacht. Die vorliegende Arbeit ist in dieser oder anderer Form zuvor nicht als Prüfungsarbeit zur Begutachtung vorgelegt worden.

---

Ort, Datum, Unterschrift

## DIGITALER ANHANG

Die beigefügten DVDs enthalten:

### Disk 1

- eine detaillierte Beschreibung des Aufbaus des ALS (inkl. Konfiguration) als Leitfaden für Versuchsleitende
- Logs des ALS (OA-III und OA-IV) sowie Aufzeichnungen der Versuchsleiterin
- Hilfstabellen für die Auswertung der Fragebögen und Messung von Situationsbewusstsein
- Aufzeichnungen zur Definition der Areas of Interest
- Exportierte Tabellen der Blickbewegungsmessung
- Hilfstabellen für die Berechnung von Messfenstern
- Hilfstabellen für die Datenaggregation
- Glossar zur Erläuterung des Logfileanalysetools
- die für die Analysen der vorliegenden Arbeit verwendeten Datensätze
- die vorliegende Arbeit als digitales Exemplar

### Disk 2

- Konfigurationsdateien für das Recording Design der Blickbewegungsstudie
- Logs der Blickbewegungsstudie, inkl. Aufzeichnung der Leitwarte (ohne Visualisierung der Blickbewegung)  
(OA-IV HO 1–7)

### Disk 3

- Logs der Blickbewegungsstudie, inkl. Aufzeichnung der Leitwarte (ohne Visualisierung der Blickbewegung)  
(OA-IV HO 8–13)

### Disk 4

- Aufzeichnung der Leitwarte inkl. Visualisierung der Blickbewegung  
(OA-IV HO 1–13)

**FORSCHUNGSARTIKEL****Forschungsartikel Nr. 1**

Stade, M. J. C., von Bernstorff, C., Nachtwei, J., & Wandke, H. (2012). *Forward-Looking: Measuring Attention Allocation for Evaluating Supervisory Control in a Complex and Dynamic Lab System*. Manuscript.

**Forschungsartikel Nr. 2**

Stade, M. J. C., von Bernstorff, C., Nachtwei, J., & Wandke, H. (2012). *Meeting a Well-Known Challenge: Measuring Operator Performance in a Complex and Dynamic Microworld*. Manuscript.

**Forschungsartikel Nr. 3**

Stade, M. J. C., Meyer, C., Niestroj, N., & Nachtwei, J. (2011). Ergonomische Gestaltung trotz oder aufgrund interdisziplinärer Teamarbeit? In M. Grandt, & S. Schmerwitz (Hrsg.), *Ergonomie im interdisziplinären Gestaltungsprozess (DGLR-Bericht 2011-1)*. Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt. 1–16.

## **FORSCHUNGSARTIKEL NR. 1**

Stade, M. J. C., von Bernstorff, C., Nachtwei, J., & Wandke, H. (2012). *Forward-Looking: Measuring Attention Allocation for Evaluating Supervisory Control in a Complex and Dynamic Lab System*. Manuscript.

**MANUSCRIPT**

**Forward-Looking: Measuring Attention Allocation for Evaluating Supervisory Control  
in a Complex and Dynamic Lab System**

Melanie J. C. Stade

Charlotte von Bernstorff

Jens Nachtwei

Hartmut Wandke

Humboldt-Universität zu Berlin, Germany

Word Count:

Abstract = 137; Total text (without references, figures, tables) = 3914.

Please address correspondence to Melanie J. C. Stade, Engineering Psychology / Cognitive Ergonomics, Department of Psychology, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. E-mail: melanie.stade@hu-berlin.de

### **Abstract**

The objective of the present study was to identify suitable eye movement parameters to measure and evaluate human operators' attention allocation behavior in our complex and dynamic lab system. This supervisory control lab system comprises a human operator and a microworld including the pursuit tracking performance of two other participants. Based on a discussion of approaches of research projects related to our lab system, several parameters were selected to investigate human operators' eye movements. These operators had to optimize the tracking performance of the other participants by the use of a specific interface. We examined the relationship between these eye movement parameters and two control action variables reflecting human operators' system understanding. The results indicate that adequate eye movement parameters could be identified in this first approach of evaluating human operators' attention allocation in our lab system.

*Keywords:* human-machine interaction, supervisory control, human operator, attention allocation, eye movement measurement, performance



Evaluating and optimizing human-machine interaction remains a challenge for both practitioners and researchers. In literature, it is generally accepted that focusing on either the human or the machine is not sufficient: Almost twenty years ago, Reason (1993) emphasized the pivotal importance of human-machine *interaction*. One approach of evaluating and optimizing human-machine interaction is to investigate the behavior of human operators who use a human-machine interface to operate a machine or process (e.g., operators of control rooms). In particular, researchers can investigate human operators' attention allocation during the interaction with such an interface. Such an investigation of attention allocation allows gaining insights into perceptual and cognitive processes effecting control actions. Among different measures of attention allocation, eye movement recording has been established as method of choice because "the eye tracker provides objective and quantitative evidence of the user's visual and (overt) attentional processes" (Duchowski, 2002, p. 455). In fact, eye movement parameters represent which region of a visual stimulus (e.g., interface display) is an area of interest. In human-machine interaction research, such eye movement parameters have proved their suitability in different research domains, for example in the evaluation of different versions of interfaces (e.g., Lin, Zhang, & Watson, 2003) or in the examination of expertise levels of human operators (e.g., Ottati, Hickox, & Richter, 1999).

Eye movement parameters hence indicate how much attention a human operator allocates on a piece of information in the interface. However, the link between attention allocation and its evaluation is to some extent indirect. The same attention allocation behavior can result in different task performance levels, and different attention allocation behavior in the same task performance level. Consequently, not the attention allocation behavior *itself* has to be evaluated, but the information's *effective use as a result of attention allocation* (i.e., information processing) (Lin, Zhang, & Koubek, 2004). Thus, it can be assumed that the relationship between human operators' attention allocation behavior and task performance level is mediated by human operators' control actions.

In the ATEO research project (“division of labor between developers and operators; ATEO as German acronym), a complex and dynamic lab system was developed to examine this mediated relationship. The ATEO lab system (ALS) includes a human operator and a microworld. The microworld consists of two participants (so-called microworld inhabitants) who perform a cooperative tracking task. This subsystem is supervised and controlled by a third participant acting as operator. This human operator is instructed to use inputs and outputs of a specific master display (ATEO master display; AMD) that is provided to the human operator for optimizing the tracking process. In previous studies the ATEO lab system has proved its suitability for investigating human-machine interaction (Nachtwei, submitted-a; von Bernstorff & Nachtwei, submitted).

The present study of the ATEO research project deals with the question which eye movement parameters are suitable to indicate human operators’ attention allocation on the master display. Based on our review and discussion of research projects which are related to our lab system, we selected eye movement metrics in order to investigate human operators’ attention allocation in our lab system. For this purpose, we examined the association between these eye movement parameters and control action parameters in a first analysis. Further studies investigating the relationship between human operators’ control actions and task performance in the ALS are currently in progress.

Thus, the present study focuses on the exploration of suitable eye movement parameters and should be regarded as a first approach to investigating the relationship between human operators’ attention allocation behavior (i.e., supervision) and control actions in a complex and dynamic lab system such as the ATEO lab system.

## **Method**

### **Participants**

Thirteen same sex triads (7 men triads and 6 women triads) participated in the experiment. The 13 operator participants aged from 20 to 34 years ( $M = 26.77$ ,  $SD = 4.19$ ),

the mean age of the 26 participants acting as microworld inhabitants (MWIs) was 25.04 years ( $SD = 2.97$ ), ranged from 20 to 32 years. Participants were recruited using an electronic recruitment system of the Department of Psychology at Humboldt University Berlin. More than 6200 students and non-students are registered in this system. All participants of the present study were German native speakers and right-handed. They had not studied psychology, had not received psychological / psychiatric treatment and had no red-green color blindness. Volunteers who had taken part in other studies with a similar approach (i.e., tracking paradigm) were excluded. Operator participants had at least German university entrance qualification, MWI participants had at least general certificate of secondary education. For their participation, operators were compensated with 20 euros (about US \$27) and each MWI with 15 euros (about US \$20). Afterwards, all participants received a debriefing with a short explanation of the rationale behind the study.

### **Simulation Tool**

The rationale behind the ATEO lab system includes a human operator who supervises and controls a complex and dynamic microworld. This microworld comprises the cooperative tracking of two real people and, therefore, is less deterministic for the human operator. As these two subjects augment the microworld by a social component, this microworld is called “Socially Augmented Microworld” (SAM). For a deeper description of SAM’s approach see Wandke and Nachtwei (2008). The MWIs have each 50 % control over the tracking object (i.e., vertical and horizontal joystick input). Thus, they have to cooperate in order to navigate the object safely and efficiently along a virtual track. This cooperative tracking also includes avoiding collisions with obstacles and choosing fork junctions. One MWI receives the instruction that the task requires rather high speed whereas the other MWI is instructed to focus on accurate tracking. Both MWIs believe in an identical (i.e., team) instruction and are not allowed to communicate with each other.

The task of the human operator is to achieve maximal tracking speed and accuracy of the tracking object. In order to successfully perform this task, the human operator can optimize the tracking process of the two MWIs by supervising several information sources and by controlling via inputs at the ATEO Master Display (AMD). For a description of the 3-year AMD design process see Nachtwei (submitted-a), and for an evaluation of the different interface extensions see von Bernstorff and Nachtwei (submitted).

Figure 1 shows the ATEO Lab System (ALS) including SAM (left) and the AMD (right).

*--- Insert Figure 1 about here. ---*

In the following, the AMD is described as the human operator's means for supervision and control. Primary, the AMD is subdivided into a screen area for supervision (left) and a screen area for control (right).

*--- Insert Figure 2 about here. ---*

**ATEO Master Display (AMD).** In the supervision area on the left, the system status displays the divergent tracking instructions of both MWIs (speed vs. accuracy) as well as the number of completed trials (e.g., 8 of 11 drives, see Figure 2). Below the system status, the efforts of both MWIs are visualized. This visualization bases on the MWIs' subjectively rated effort after each cooperative tracking trial and indicates the wear-out of the subsystem. In the bottom left-hand corner of the AMD, the human operator can observe MWIs' behavior through a real-time video stream. The human operator's track view includes the current view of the MWI on the track (black frame) and a track preview. This track preview comprises

information about the upcoming track (e.g., obstacle ahead). The joystick movements of the MWIs are illustrated at the left and right hand side of the lower part of the current track view.

In the control area (right side of the AMD) the human operator can use soft and hard inputs to optimize the tracking.

***Soft inputs.*** The soft inputs are differentiated into visual and auditory hints. Visual hints comprise warnings as well as advices referring to tracking speed and direction. These hints are shown on the MWIs' common screen. Hints regarding tracking speed and direction can also be given via split audio channels. Thus, the human operator can decide whether both MWIs or only one MWI (and which) perceive an auditory hint per earphones. Further auditory hints refer to behavior in curves, allocation of leadership and praise. Above the visualization of joystick movements, the human operator receives feedback on given hints.

***Hard inputs.*** By the use of hard inputs, the human operator can directly influence the MWIs' tracking behavior. In the upper-right corner of the AMD, the human operator can reduce the maximum speed (i.e., 100 %) via a slider in 1 % steps to a minimum value of 0 %. Below this area, the direction of the tracking object can be limited by blocking MWIs' tracking to the left or right. Additionally, the human operator can redistribute joystick input power between the MWIs. Clicking the reset button returns the default value of the corresponding hard input.

The upper part of the AMD is used for measuring human operators' situation awareness. See Nachtwei (submitted-b) for a detailed description and discussion.

### **Apparatus and Data Acquisition**

**ATEO lab system.** For both SAM and the AMD, the programming environment "SQUEAK" (Ingalls, Kaehler, Maloney, Wallace, & Kay, 1997) and the object-oriented programming language "Smalltalk80" were chosen. SAM and the AMD run on equal personal computers (CPU: Intel Core 2 Quad 2.40 GHz, RAM: 3 GB DDR3 1066 MHz, GPU: ATI Radeon HD 4850 512 MB, OS: MS Windows Vista Home Premium 32bit). The AMD is

designed for a 30-inch LCD computer display with a 2560x1600 resolution; the 19-inch LCD computer display of SAM has a resolution of 1024x768. The joystick movements of the MWIs, coordinates of the tracked object as well as human operator's mouse clicks were recorded into a log file (csv-format) with a resolution of 39 milliseconds.

**Eye tracker.** A remote eye tracker (Tobii X120 Eye Tracker, Tobii Technology AB) with a 120 Hz data rate was used. The eye tracker was located below the computer screen. Subjects sat in front of the master display with a distance of about 30 inches. Eye movement data were collected and analyzed using the i<sup>2</sup> Visualizer Implicit System 3.0.0 software from the eye square GmbH.

### **Procedure**

At the beginning of the experiment, each operator participant ran through an AMD training (about 85 minutes) with theoretical and practical modules. The two MWI participants were given a short introduction into their task and the steering means (i.e., joystick and display). After supervision of four solo and three following cooperative trials (around 30 minutes), the human operator acted in trial 8 to 11 as supervisory controller of SAM (operation phase about 30 minutes). The beginning of trial 6 was used for a short calibration phase (about 20 seconds). For this, the human operator gazed the movement of a dot on the master display. Subjective data of human operators and MWIs (e.g., rating of effort) were collected via paper questionnaires and a questionnaire tool running on laptops (see Figure 1). In sum, the experimental session lasted about 150 minutes for the human operator and about 90 minutes for the MWIs.

### **Selecting Eye Movement Parameters for the ALS: A Short Review of Related Research**

In this section, we describe and discuss research projects related to the ALS with regard to transferability of their methodical approaches. For this purpose, we focus on (1) simultaneous recording of monitoring and controlling behavior, (2) the so-called hand-eye measure as a computed combination of eye and hand movement data, and (3) requirements for

data integration and analysis. Based on this review, several eye movement parameters were selected and computed (with adaptations) for an examination in the ALS.

**Simultaneous supervision and control.** In their air traffic flow simulation tool, Grasshoff, Hasse, Eissfeld and Bruder (2010) investigated the effect of aviation operators' monitoring behavior in an automation phase on performance in a manual phase *afterwards*. For this, eye movement data were recorded in an automatic mode of a scenario and used to predict task performance in the following manual control mode of the same scenario. Consequently, data for both classes of behavior (i.e., monitoring and controlling) in the same scenario were collected *separately*.

In the ALS, we also examined attention allocation behavior and control actions in the same scenario, but the main trials 9 to 11 include (alternate and mutually dependent) supervision *and* control phases of the tracking process. Contrary to Grasshoff and colleagues (2010), we cannot *predict* human operators' control behavior from their attention allocation behavior in a well-defined and merely supervision phase. Thus, the analysis of simultaneously recorded eye movements and control actions is a challenge for the present study, especially with regard to interpretation of cause and effect of attention allocation and control actions.

**Several regions for supervision and control.** Lin, Zhang and Koubek (2004) examined eye movements, control actions and their relationship (via so-called hand-eye measures) of three human operators with different grades of performance. The human operators' task was to monitor and control a simulated plant system based on thermal-hydraulic processes. Data were analyzed for selected regions of the display, which includes information about system values and input devices to control these values. Thus, the slider that adjusted a system parameter was located right next to the displayed value of a system component.

Human operators in the ALS have many degrees of freedom to supervise an actual system value. For example, human operators can gaze at the visualization of the vertical

joystick movements or the set value of “limit speed” to supervise speed of the tracked object. Additionally, human operators have many degrees of freedom to change (e.g., increase) the speed. They can re-allocate the joystick control favoring MWI 1 (focus on speed), give hints “faster” or “cut short”, reset or increase the vehicle’s maximum speed. Compared to the simulated plant system, human operators cannot set a system value *directly* next to the tracking object. The descriptive results of Lin and colleagues (2004) are promising for future work, also in the ATEO project. But due to the human operators’ many degrees of freedom in supervision and control as well as the separated input area of the master display, the hand-eye measure is seen as not suitable for a first approach of evaluating attention allocation behavior in the ALS.

**Objective and quantitative data analysis of supervision and control.** Reviewing related research, several limitations regarding data analysis arise. Zülch and Stowasser (2003) examined the association of human operators’ eye movements and control strategies for task-solving in a simulated shop-floor control system. These control strategies were *subjectively* discriminated. The study of Lin and colleagues (2004) is limited by the *descriptive* data analysis of only three participants. Other studies are limited by the *short period* of recorded behavior. The presented scenarios lasted only one (Moray & Rotenberg, 1989) or five (Grasshoff et al., 2010) minutes. The examined trials 9 to 11 in the ATEO project lasted up to 25 minutes, each trial taking about 8 minutes. We argue that the statistical relationship between supervision, control and task performance level can be better examined using this longer period of time as well as objective metrics for human operators’ behavior.

### **Selected Eye Movement Parameters for the ALS**

Eye movements were recorded in order to measure human operators’ monitoring behavior. Several eye movement parameters have been used in research as indicators for perceptual and cognitive processes. Jacob and Karn (2003) have provided an overview of eye tracking metrics used in more than 20 studies of human-computer interaction and usability



research. As early as 1950, Fitts, Jones and Milton (as cited in Jacob & Karn, 2003) investigated the positive relationship between the importance of an interface region and the fixation frequency on this region. Moreover, longer fixations on a region are seen as indicator for a difficult extraction and interpretation of information. However, longer fixations not only reflect processing *difficulty* but also processing *depth* (e.g., planning, problem solving) (as cited in Eivazi & Bednarik, 2011). For the aim of the present study we also considered relative fixation *counts* and mean fixation *durations* on selected regions of the master display.

The analysis of human operators' fixations focused on the main element of the supervision area in the AMD: the track view, subdivided into track preview and MWIs' current track view. The fixation counts ratio "track preview / track view" represents the fixation counts in the track preview in relation to the fixation counts of the entire *track view*. A higher ratio indicates a higher percentage of fixations in the track preview. The absolute mean fixation duration was calculated for both the track preview and the MWIs' current view of the track. For a comparison of the mean fixation durations, the mean fixation duration of the track preview related to the mean fixation duration of the MWIs' current view was used (mean fixation duration ratio "track preview / current track view").

### **Selected Control Action Parameters for the ALS**

**Rate of timely warnings.** The human operator can inform MWIs about an obstacle or fork located on the upcoming track. The sum of warnings given in time (i.e., before MWIs themselves see the upcoming track abnormalities) was related to given warning hints. Repeated warning hints for the same obstacle respectively fork were ignored for timing indicator's calculation.

**Control actions regarding one MWI.** This indicator represents the proportion of control actions regarding either one of the MWIs. It comprises auditory hints given to MWI 1 or 2 and hard inputs to re-allocate the MWIs' amount of joystick input.

As mentioned above, results regarding the relationship between control actions and task performance level have not yet been submitted. Nevertheless, we argue that both selected control actions indicate system understanding including task understanding and AMD functionality. The timing of warnings might represent planning of supervisory and control behavior as well as human operator's knowledge that the preview adds value for the MWIs. Using control actions regarding either one of the MWIs might indicate that the human operator considers the MWIs' different instructions and performances. It is noted that the selected control action parameters might not only be evoked by attention allocation but also, for instance, by human operators' overload.

### **Statistical Analysis**

Data were aggregated for trial 9 to 11. Eye movement and control action parameters were inspected via scatterplots. We used Kendall's  $\tau$  as non-parametric correlation coefficient for small samples since this correlation coefficient is less vulnerable to outliers and is superior over Spearman's correlation coefficient  $r_s$  (Newson, 2002).

### **Results**

Correlations between human operators' attention allocation behavior and human operators' control actions and the corresponding effect sizes are represented in Table 1. All correlations were non-significant, and effect sizes were moderate. The moderate effect sizes we found here indicate that a bigger sample size would have been needed to reach significance. However, the sample size cannot be increased due to high time and financial costs as well as the equipment-consuming eye tracking assessment. Another reason for the moderate effect sizes which we found might be the calculation of the parameters (see discussion). In order to interpret results despite non-significance, we will hence focus on the effect sizes.

A higher "track preview over the entire track view" ratio was associated with a higher rate of timely warnings ( $\tau = .23$ ,  $p = .272$ ). Additionally, human operators with more eye

fixations in the track preview in relation to the entire track view used a higher percentage of control actions regarding only one of the MWIs ( $\tau = .23, p = .272$ ).

Considering the mean fixation duration, a higher mean fixation duration in the MWIs' current track view was related to a higher percentage of control actions addressing only one MWI ( $\tau = .31, p = .143$ ). The ratio "track preview over MWIs' current track view" was negatively correlated with the rate of timely warnings ( $\tau = -.31, p = .143$ ) and control actions regarding MWI 1 or 2 ( $\tau = -.26, p = .222$ ). That is to say, smaller mean fixation duration in the track preview related to the mean fixation duration in the current track view of the MWIs was associated with more promptly warnings and a higher percentage of control actions addressing MWI 1 or 2.

--- *Insert Table 1 about here.* ---

### **Discussion**

The present study aimed at exploring suitable eye tracking metrics in order to investigate the association between human operators' supervisory and control behavior in a complex and dynamic lab system such as our ATEO lab system. Thus, we discussed approaches of research projects which are comparable to our lab system. Based on this discussion, we selected and adapted eye movement parameters with the account of our lab system's characteristics.

This first analysis with these eye movement parameters is limited by the small sample size of our experimental study and is restricted to the chosen control action parameters. Nevertheless, the effect sizes are promising for future work, examining the relationship between eye movements and task performance level mediated by human operators' control actions.

The results show that a higher percentage of fixations in the track preview and smaller fixation durations in the track preview compared to those in the MWIs' current track view are associated with a better timing of warnings and more control actions regarding only one MWI. We assume that this attention allocation pattern describes short scanning of upcoming track characteristics indicating human operator's expertise. The positive relationship between longer mean fixation durations on MWIs' current track view and a higher percentage of control actions addressing only one MWI leads to the conclusion that human operators used this part of the track view for analyzing the current tracking process. Further research is needed to examine factors affecting the relationship between attention allocation and control actions. Both rate of timely warnings and control actions regarding only one MWI could be the result of cognitive processes (i.e., attention allocation) but also of states of the human operator (e.g., workload).

Human operators' behavior was examined during the whole trial and data were aggregated for three trials. However, Chapman and Underwood (1998) showed in their experiments with driving scenes, varying in dangerousness and complexity, that dangerous events affected long fixation durations whereas visual complex scenes evoked short fixation durations. In addition, drivers with different skill levels varied in their vertical scanning pattern. Consequently, future work in the ATEO project should not only examine supervisory control behavior during different *track sections* but also eye movements on different *regions* of both track preview and MWIs' current track view.

Eye movement parameters indicating the dynamics of human operator's attention allocation should also be considered. Number and frequency of transitions between track preview and MWIs' current track view as well as the order of fixated elements can be assumed to represent these dynamics. Yet, not only the interpretation but also the analysis of eye movement parameters in dynamic scenes is a challenge. Nonetheless, first approaches to solving this problem exist (e.g., Bruder, Leuchter, & Urbas, 2003; Reimer & Sodhi, 2006).

From the explorative approach of the presented study, we conclude that the selected eye movement and control action parameters are suitable parameters to investigate the relationship between human operators' supervisory and control behavior. Further studies based on our approach should extend the explored parameters by those which take the dynamic and complexity of human operators' behavior and different track sections more into account. Finally, we believe that future work considering the discussed findings will take another step towards an examination of the mediated relationship between human operators' attention allocation behavior and task performance level in our complex and dynamic lab system. Beyond our simulation tool, this anticipated relationship offers interesting application opportunities for personnel development and selection (e.g., ability tests) in the human-machine interaction domain.

### **Acknowledgements**

We thank the eye square GmbH to enable the presented study. We would also like to thank the eye square team of the headquarter Berlin for their support to set up the eye tracker apparatus. Special thanks go to Sabrina Schulte and Philipp Leppert.

### **References**

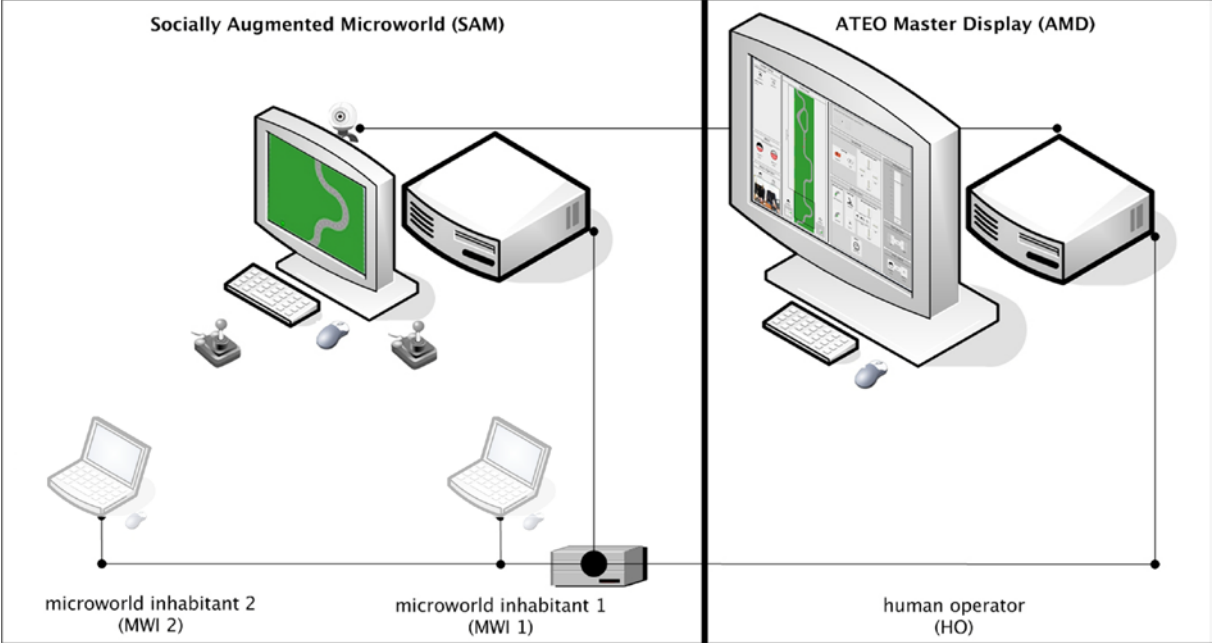
- Bruder, C., Leuchter, S., & Urbas, L. (2003). Konzept für automatische Auswertungen von Blickbewegungen bei dynamischen Informationen [Concept for automatic evaluation of eye movements on dynamic information]. *MMI-Interaktiv*, 6, 25–34.
- Chapman, P.R. & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, 27(8), 951–964. doi: 10.1068/p270951
- Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4), 455–470. doi: 10.3758/BF03195475
- Eivazi, S., & Bednarik, R. (2011). Predicting problem-solving behavior and performance levels from visual attention data. *Proceedings of the 2<sup>nd</sup> Workshop on Eye Gaze in*

- Intelligent Human Machine Interaction (Palo Alto, CA, USA, February 13)* (pp. 9–16). Retrieved from [http://www.ci.seikei.ac.jp/nakano/GAZEWS\\_IUI2011/proceedings/eye.pdf](http://www.ci.seikei.ac.jp/nakano/GAZEWS_IUI2011/proceedings/eye.pdf)
- Grasshoff, D., Hasse, C., Eissfeldt, H., & Bruder, C. (2010). Identifying operators monitoring appropriately. In A. J. Spink, F. Grieco, O. E. Krips, L. W. S. Loijens, L. P. J. J. Noldus, & P. H. Zimmermann (Eds.), *Proceedings of Measuring Behavior 2010, 7<sup>th</sup> International Conference on Methods and Techniques in Behavioral Research (Eindhoven, the Netherlands, August 24-27, 2010)* (pp. 334–337). Wageningen, the Netherlands: Noldus Information Technology.
- Ingalls, D., Kaehler, T., Maloney, J., Wallace, S., & Kay, A. (1997). Back to the future: The story of Squeak, a practical Smalltalk written in itself. *Proceedings of the 12<sup>th</sup> ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications* (pp. 318–326). Atlanta, GA: Association for Computer Machinery.
- Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573–605). Amsterdam, the Netherlands: Elsevier.
- Lin, Y., Zhang, W. J., & Watson, L. G. (2003). Using eye movement parameters for evaluating human-machine interface frameworks under normal control operation and fault detection situation. *International Journal of Human-Computer Studies* 59(6), 837–873. doi: 10.1016/S1071-5819(03)00122-8
- Lin, Y., Zhang, W. J., & Koubek, R. J. (2004). Effective attention allocation behavior and its measurement: a preliminary study. *Interacting with Computers*, 16(6), 1195–1210. doi: 10.1016/j.intcom.2004.086
- Moray, N., & Rotenberg, I. (1989). Fault management in process control: Eye movements and action, *Ergonomics*, 32(11), 1319–1342. doi: 10.1080/00140138908966910

- Nachtwei, J. (submitted-a). *A multi-level design approach for a supervisory control master display in human factors experiments*. Manuscript submitted for publication.
- Nachtwei, J. (submitted-b). *SAGAT versus SPAM? It depends! An aid for situation awareness method selection (aSAmS)*. Manuscript submitted for publication.
- Newson, R. (2002). Parameters behind “nonparametric” statistics: Kendall’s tau, Somers’ D and median differences. *The Stata Journal*, 2(1), 45–64.
- Ottati, W. L., Hickox, J. C., & Richter, J. (1999). Eye scan patterns of experienced and novice pilots during visual flight rules (VFR) navigation. *Proceedings of the Human Factors and Ergonomics Society, 43<sup>rd</sup> annual meeting* (pp. 66–70). Santa Monica, CA: Human Factors & Ergonomics Society.
- Reason, J. (1993). The identification of latent organizational failures in complex systems. In J. A. Wise, V. D. Hopkin, & P. Stager (Eds.), *Verification and validation of complex systems: Human Factors Issues* (pp. 223–237). Berlin, Germany: Springer.
- Reimer, B., & Sodhi, M. (2006). Detecting eye movements in dynamic environments. *Behavior Research Methods*, 38(4), 667–682.
- Von Bernstorff, C., & Nachtwei, J. (submitted). *Between keyhole and clutter effect – A multi-level evaluation of interface extensions using the ATEO master display (AMD)*. Manuscript submitted for publication.
- Wandke, H., & Nachtwei, J. (2008). The different human factor in automation: The developer behind vs. the operator in action. In D. de Waard, F. O. Flemisch, B. Lorenz, H. Oberheid, & K. A. Brookhuis (Eds.), *Human factors for assistance and automation* (pp. 493–502). Maastricht, the Netherlands: Shaker Publishing.
- Zülch, G., & Stowasser, S. (2003). Eye tracking for evaluating industrial human-computer interfaces. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind’s eye: Cognitive and applied aspects of eye movement research* (pp. 531–553). Amsterdam, the Netherlands: Elsevier.

Figures

Figure 1. Experimental Setup of the ATEO Lab System (ALS) With the Socially Augmented Microworld (SAM) and the ATEO Master Display (AMD)





**Figure 2.** Current Version of the ATEO Master Display (AMD) for Supervision and Control of SAM (Translated From German)



**Tables****Table 1.** Correlations Between Human Operators' Attention Allocation and Control Actions

Control behavior	Attention allocation behavior			
	Relative fixation count	Mean fixation duration		
	Track preview / track view	Track preview	Current track view	Track preview / current track view
Rate of timely warnings	.23	.08	.15	-.31
Control actions regarding one MWI	.23	.18	.31	-.26

*Note.*  $N = 13$ . Kendall's  $\tau$  as correlation coefficient. All coefficients were non-significant (two-tailed).

## **FORSCHUNGSARTIKEL NR. 2**

Stade, M. J. C., von Bernstorff, C., Nachtwei, J., & Wandke, H. (2012). *Meeting a Well-Known Challenge: Measuring Operator Performance in a Complex and Dynamic Microworld*. Manuscript.

**MANUSCRIPT**

**Meeting a Well-Known Challenge: Measuring Operator Performance in a Complex and  
Dynamic Microworld**

Melanie J. C. Stade

Charlotte von Bernstorff

Jens Nachtwei

Hartmut Wandke

Humboldt-Universität zu Berlin, Germany

Word Count:

Abstract = 159; Total text (without references, figures, tables) = 5651.

Please address correspondence to Melanie J. C. Stade, Engineering Psychology / Cognitive Ergonomics, Department of Psychology, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. E-mail: [melanie.stade@hu-berlin.de](mailto:melanie.stade@hu-berlin.de)

### **Abstract**

In human-machine interaction research, analyzing operator performance in microworlds is connected with several methodological difficulties. In particular, dealing with both a high amount and a high variability of data that are generated by complex and dynamic microworlds is a challenge. In our lab system, such a complex and dynamic microworld is used to test strengths and weaknesses of human operators. These human operators supervise and control the microworld which comprises the pursuit tracking of two participants. This paper presents an approach for investigating the logs generated by our complex and dynamic microworld. Our objective was to examine different parameters in order to identify differences between operators in supervision and control. We analyzed human operators' inputs during different sections of the track and aggregated data differently. Concerning the assessment of operators' interventions, different outcomes (i.e., states of the microworld) are discussed. The results show that the proposed approach seems suitable for measuring operator performance in our complex and dynamic microworld.

*Keywords:* human-machine interaction, microworlds, performance measurement, supervisory control, human operator

Microworlds have been successfully established as object of research in human-machine interaction science due to their favorable characteristics. On the one hand, these small-scale computer-based simulations are complex enough to be representative of real human-machine systems as required for human factors research (Meister, 2001). On the other hand, microworlds allow for the control of selected variables. Thus, microworld research is striking a balance between the controllability in experiments and the complexity of field studies (Brehmer & Dörner, 1993).

Although microworlds offer the opportunity to effectively deal with several human-machine research questions, one major problem still persists: How to *analyze* the generated data? Studying operators' behavior in such microworlds raises a number of methodological problems (Howie & Vicente, 1998). In general, microworlds generate many data, making an *identification* of operators' differences in supervision and control challenging. Moreover, a *comparison* of operators' performance is difficult because operators generally have many degrees of freedom to perform a given task. Another aggravating factor for performance measurement is that many microworld systems are characterized by slow dynamics, which often results in *rare observable actions* by operators.

Some of these methodological difficulties mentioned by Howie and Vicente (1998) also arose in our experimental setting. The main objective of the present study was to identify performance differences between operators by using different measurement parameters in our complex and dynamic lab system. This lab system was developed in order to test strengths and weaknesses of operators which supervise and control a complex and dynamic human-machine system. The so-called ATEO Lab System (ALS; ATEO is the German acronym of our research project "Division of labor between developers and operators") includes a particular microworld. Two participants act as microworld inhabitants (MWIs) and perform a cooperative tracking task that is supervised and controlled by the operator (a third participant). In order to optimize MWIs' tracking performance, the operator uses output

elements (e.g., visualization of the system state) and input elements (e.g., interventions regarding speed), which are presented on operator's master display.

It should be noted that the cooperation of the two MWIs augments the complexity of the microworld by a social interaction component (for a deeper description of the approach see Wandke and Nachtwei, 2008). This social interaction makes the tracking process less deterministic for the operator. However, the social interaction is not only a challenge for the operator. It also presents a challenge for researchers which aim for analysis and evaluation of operators' performance regarding the methodical difficulties mentioned by Howie and Vicente (1998). First, operators in our lab system have many degrees of freedom to optimize the tracking process (e.g., several inputs to reduce speed). Thus, we would expect a high variability in operators' actions. This high variability in data of small samples (which are characteristic for studies with human-machine systems) makes it difficult compare operators' performance. Contrary to Howie and Vicente (1998), we expected that operators in our lab system acted *more* frequently. One reason is that the dynamic of our microworld is relatively high. Moreover, the dynamic of the microworld is partly unforeseeable for the operator due to the social component of the microworld. Accordingly, we believe that these characteristics of our lab system will cause more frequent support actions by operators. These support actions as well as the system states are logged with a very high resolution over a long period. Thus, a meaningful aggregation of these data is challenging, but essential.

In this paper, we present an approach for dealing with these many and highly variable data in our lab system. First, we divided the track in different sections and subsections. Second, we calculated different parameters which describe operators' interventions in these track sections and subsections. Finally, we evaluated operators' interventions by objective measures in terms of team performance. Team performance includes efficiency (i.e., total time of tracking) and safety (i.e., number of errors during tracking). Moreover, we selected

subjective measures which consider the MWIs as the social component of our particular microworld.

## **Method**

### **Participants**

In the experimental study, 40 same-gender teams participated (one team consisted of one operator and two microworld inhabitants). The 40 operator participants aged from 18 to 34 years ( $M = 26.45$ ,  $SD = 4.09$ ), the 80 microworld inhabitant participants' (MWIs') age ranged from 20 to 34 years ( $M = 25.45$ ,  $SD = 3.72$ ). Participants were recruited via an electronic recruitment system (developed and applied at the Department of Psychology at Humboldt University Berlin), comprising more than registered 6200 students and non-students. The 120 selected participants for the current study were German native speakers and right-handed. They had not studied psychology, had no red-green color blindness and had not taken part in comparable studies (i.e., based on a tracking paradigm). Volunteers with at least German university entrance qualification were selected as operator participants, MWI participants had at least general certificate of secondary education. For their participation, operators were paid 20 euros (about US \$27) and each MWI was compensated with 15 euro (about US \$20). Afterwards, all participants received a short debriefing about the study's rationale.

### **Simulation Task**

In the ATEO lab system (ALS), the team (i.e., one operator and two MWIs) is split for work in separate rooms (see Figure 1).

*--- Insert Figure 1 about here. ---*

In one room, the two MWI team members perform a cooperative tracking task (left part of Figure 1). For this, they cooperatively steer a tracking object with one joystick each,



while each of the MWIs has 50 % steering input power. The MWIs are instructed to navigate the tracking object efficiently and safely along a virtual track, which is displayed on a shared monitor in front of them (19-inch LCD computer display with a 1024x768 resolution). The tracking task also includes mastering abnormal track sections (i.e., avoiding collisions with obstacles and choosing directions at fork junctions). It should be noted that MWI 1 is instructed to focus on *fast* tracking, whereas MWI 2 is instructed to focus on *accurate* tracking. These converse instructions should provide more complexity and dynamic of MWIs' tracking. However, both team members believe in an identical instruction. In addition, MWIs were told not to communicate with each other.

The tracking team members are supervised and controlled by the operator (see right part of Figure 1). The task of the operator is to support MWIs' tracking performance. For this purpose, the operator is provided with a special master display. This so-called ATEO Master Display (AMD) was developed over 3 years (see Nachtwei, submitted-a for the design process) and different interface versions were extensively tested (see von Bernstorff and Nachtwei, submitted). The master display was designed for and runs on a 30-inch LCD computer display (resolution of 2560x1600).

**ATEO master display (AMD).** The AMD connects the operator with the other two team members and is used for supervision and control of MWIs' tracking performance. Thus, the AMD comprises a screen area for supervision (see outputs on the left, Figure 2) and a screen area for control (see inputs on the right).

--- *Insert Figure 2 about here.* ---

In the system status on the left, MWIs' divergent tracking instructions (focus on speed vs. accuracy) and the number of completed trials are displayed. The visualization of MWIs' effort below is based on MWIs' ratings after each cooperative tracking trial. The area in the

bottom left-hand corner of the AMD shows the behavior of the MWIs' in a real-time video stream. The central track view of the operator comprises MWIs' current view on the track (black frame) and a preview of the track. The track preview includes information about the upcoming track (e.g., obstacles ahead). The joystick movements of each MWI are illustrated next to the bottom part of the track view.

In order to successfully optimize the tracking process, the operator can use inputs to support the team members.

**Soft inputs.** By the use of visual and auditory hints, the operator can indirectly control the tracking behavior of the MWIs. Visual hints regarding warnings (i.e., approaching obstacle or fork) as well as advices referring to tracking speed or direction are displayed on MWIs' shared screen. The operator can also give hints related to tracking speed or direction via split audio channels, so as to control only one of the two MWIs. Further auditory hints which address either both MWIs or one MWI refer to behavior in curves ("more accurate", "cut short"), allocation of leadership ("assume" or "leave leadership") and praise. The operator receives immediate feedback on given hints which are illustrated above the visualization of MWIs' joystick movements.

**Hard inputs.** The area on the far right contains hard inputs for a direct intervention in the tracking process. For this, the operator can (1) gradually reduce the maximum speed of the tracking object, (2) block the tracking to the left or right, or (3) re-allocate the joystick input power between the two MWIs to strengthen the input power of either one of the MWIs. A click on the reset button causes the default value of the corresponding hard input.

It should be noted that no direct verbal communication between operator and MWIs exists because auditory hints are pre-recorded mp3-files. Moreover, the communication is restricted by predetermined input options. In addition, MWIs were informed about an automatic system which supervises and assists them and hence were not aware about the support by a human.

The area in the upper part of the AMD is used for the assessment of operators' situation awareness in the ALS (see Nachtwei, submitted-b).

SAM and the AMD run on equal personal computers (CPU: Intel Core 2 Quad 2.40 GHz, RAM: 3 GB DDR3 1066 MHz, GPU: ATI Radeon HD 4850 512 MB, OS: MS Windows Vista Home Premium 32bit). For both SAM and the AMD, the programming environment "SQUEAK" (Ingalls, Kaehler, Maloney, Wallace, & Kay, 1997) and the object-oriented programming language "Smalltalk80" were used.

### **Division of the Track for Measurements**

Each track of the identical trials 9 to 11 has a length of 119.610 pixels and contains different track sections which vary in complexity. A track section refers to a defined pixel interval on the track. We assume that these different track sections require different (combinations of) operator's interventions, which in turn affect team performance (i.e., speed and accuracy) in different ways. Accordingly, for data analysis we wanted to consider operator's support in different track sections. Thus, we divided the track in two types of sections: *normal* track sections and *abnormal* track sections. Normal track sections comprise different narrow curves and few very short straight subsections (i.e., routine track sections). Abnormal track sections include obstacles as well as combined obstacles and forks which occur on the track (i.e., crisis track sections). Normal and abnormal track sections were defined in consideration of operators' and MWIs' visibility of upcoming track sections. The calculation is described in the following.

**Normal track sections.** A normal track section starts at the beginning of each trial or after an obstacle or fork (latter is shown in Figure 3-1). Normal track sections last until the end of the track or until upcoming obstacles or forks become visible to the operator in the track preview of the AMD (see Figure 2). We chose this position as endpoint of the normal track sections because we would argue that, from the moment obstacles or forks become visible to the operator, his or her interventions will refer to this upcoming track section. Thus,

we would expect such interventions to be a support for abnormal track sections, rather than for normal track sections.

Overall, ten normal track sections per trial were defined. However, six track sections were too short for analyze and thus excluded. In fact, the average length of the excluded track sections is 854 pixels and hence comparable to length of MWIs' track view. These track sections are therefore too short for an investigation of operators' support during normal track sections over a longer period. In contrast, the four sections that were included in our analyses comprise 13.093 pixels on average, and their sum makes up 44 % of the total track.

**Abnormal track sections.** As suggested above, abnormal track sections start when an obstacle or fork become visible to the operator, and last until the (last) obstacle or fork of the given abnormal track section. There were three classes of abnormal track sections with increasing complexity. The first class included a single moving obstacle or a combination of two static obstacles. The second class was an obstacle-fork combination which comprised a dynamic obstacle respectively two static obstacles and a subsequent fork. Finally, the third class of abnormal track sections was a fork-obstacle-fork combination.

We did not analyze operators' support in the first class of abnormal track sections (i.e., obstacle section) because test directors' observations showed that various teams perform similar while avoiding these obstacles, which thus produces insufficient variance in operators' supporting behavior and MWIs' tracking performance. Likewise, the third class of abnormal track sections (i.e., fork-obstacle-fork section) was excluded from data analysis. This combination of forks and obstacles is too complex and challenging for data aggregation, analysis and interpretation, especially if we want to consider which part of the fork-obstacle-fork combination is visible to whom (i.e., operator and MWIs afterwards). Thus, we selected the second class of abnormal track sections (i.e., the obstacle-fork combination; see Figure 3-4) for an analysis of operators' support in abnormal track sections. We assume that running through obstacle-fork sections represents a moderate task difficulty for the entire team and,

therefore, will cause sufficient variance in both operators' support and MWIs' tracking. In addition, in the selected abnormal track sections we can (partly) take into account if the upcoming obstacle-fork section is visible to the MWIs or not.

The first calculated window ( $w1$ ) of the obstacle-fork section starts when the (first) obstacle appears in the track preview of the master display (see Figure 3-2) and ends just before the (first) obstacle becomes visible to the MWIs (see Figure 3-3). Thus, this window takes operators' preparative for their teams into account. We expect that the preview of the track and operators' preliminaries add value for the MWIs. If the (first) obstacle occurs on MWIs' shared display, the second window starts ( $w2$ ) (see Figure 3-4). The following window ( $w3$ ) starts after the (last) obstacle and ends just before the object is tracked to the left or right branch. Finally, window four ( $w4$ ) includes the entire fork. This window ends when the tracking object leaves the fork (see Figure 3-1; concurrent starting point of the subsequent normal track section). Overall, each track contained three obstacle-fork sections which represent 14 % of the total track.

Predictors and criteria were calculated with regard to normal and abnormal track sections as described below.

### **Measures of Criteria**

We used team performance in terms of speed and accuracy as criteria since these indicators consider operators' maximizing and balancing of speed and accuracy. In addition, effort of both MWI team members and MWIs' evaluated quality of assistance served as criteria. The calculation of the criteria is described in the following.

**Team performance.** Speed and accuracy were measured as indicators for team performance. The speed variable represents the time of tracking in milliseconds. The accuracy variable is based on errors of tracking which were measured as the distance between the tracking object's position and the center of the track (total of pixels). For normal track sections, we ran analyses with both team performance indicators. In the case of abnormal

track sections, operators and MWIs were informed that they should focus on safety. Consequently, accuracy was chosen as safety indicator while (1) passing the obstacle(s) ( $w2$ ), and (2) running through the subsequent fork ( $w4$ ). Note that these two windows were analyzed separately. The time and error variables were aggregated across trials 9 to 11, inverted and z-standardized.

**Effort of MWIs.** Directly after a cooperative trial, both MWIs rated the effort they invested in the last tracking step. To do so, the MWIs used a visual analogue scale displayed on their laptops (see Figure 1). The one-dimensional scale is based on the Rating Scale Mental Effort (RSME) by Zijstra and van Doorn (1985) and was adapted for the ATEO project ( $RSE_A$ ). Thus, not only MWIs' *mental* effort but also components with regard to physical effort (e.g., joystick movement) and emotional stress (e.g., joint responsibility for team performance) are assessed subjectively. MWIs' set values were fed back to the AMD as percentaged visualization and value in the output area (see Figure 2). Comparable to a technical process component, the operator should keep the MWIs' "wear out" and its (possibly) negative effect on the tracking performance low. The measured effort was examined separately for the two MWIs and was averaged across trials 9 to 11.

**Quality of assistance.** We also assessed the subjectively evaluated quality of assistance as a second outcome variable which comprises psychological components. After the three trials, both MWIs assessed separately their perceived support in tracking. This evaluation of operator's support was measured with the German grading system (ranged from "1" = "excellent" to "6" = "unsatisfactory").

Although MWIs' effort was measured after each trial and quality of assistance was assessed after all trials, we chose these measurements as criteria in normal track sections. We assume that MWIs' ratings refer to a global assessment and depend on operators' support over a longer distance. As noted above, the analyzed normal track sections represent almost half of

the total track. Thus, we believe that operators' control actions across all normal track sections are suitable indicators to predict MWIs' effort and assessed quality of assistance.

### **Measures of Predictors**

The AMD offers many degrees of freedom for the operator to achieve the task. Thus, operators can use various (combined) intervention options in order to support the MWIs. As required, a high variability in operators' supporting behavior was observed. These differences include also outstanding behavior of single operators. However, we focus on *commonly* used intervention options as predictors since we selected a quantitative approach and therefore need enough variance in data for a statistical evaluation of operators' support. In general, the analyzed intervention options were used at least once by most operators, but operators differed in how frequently they used the different intervention options. Consequently, the calculated indicators represent *which* inputs operators used as well as how *often* and *when* operators supported. The calculation of these indicators is described in the following. Note that interventions of operators during normal track sections and abnormal track sections were aggregated in different ways.

**Normal track sections.** Hints regarding a higher tracking speed (visual or auditory hints "faster") and hints with regard to a higher tracking accuracy ("more accurate" in curves) were selected as predictors. We chose the commonly used interventions regarding the direct re-distribution of joystick input power between the MWIs as a third predictor. For data aggregation, control actions of the same type were summed up across all track sections and trials (e.g., 21 times hint "faster" given). This sum was then related to the sum of all interventions (e.g., 43 inputs). It stands to reason to believe that the computed variables thus represent which kind of intervention an operator emphasized (e.g., 49 % of interventions were hints "faster"). The computed variables might also indicate operator's support to maximize speed and accuracy. The described three predictors were analyzed to determine the effect of *specific* interventions on team performance and subsystem's outcomes. We selected the *total*

*number* of interventions as additional predictor because we suppose that the intervention frequency itself might influence the tracking performance and subsystem's outcomes.

**Abnormal track sections.** The first predictor of accuracy during the obstacle-subsection (*w2*) respectively fork-subsection (*w4*) was the *rate of timely warnings*. This variable comprises in how many abnormal track sections warnings were given in time (i.e., before the upcoming obstacle or fork is visible to the MWIs). This proportion was calculated separately for obstacles and forks.

The second predictor for accuracy while passing the obstacle- or fork-subsection contains hints with regard to tracking speed (i.e., hints "faster" or "slower"). This variable represents speed hints if the (first) obstacle is visible to operators until the tracking object reaches the (first) obstacle (*w1* and *w2*). One should note that we did not differentiate between hints "faster" and "slower". It may be assumed that a request of speed decrease is advantageous for the purpose of tracking accurately around static obstacles. In contrast, both decrease and increase of tracking speed might be an adequate strategy for the purpose of avoiding dynamic obstacles which cross the track from left to right (i.e., previously pass or let pass).

The third predictor comprises hints regarding the tracking direction (i.e., hints "left" or "right"). Direction hints which were set shortly before the tracking object reaches the first obstacle (*w2*) might support MWIs' obstacle avoidance maneuvers. Thus, direction hints served as predictor variable for accuracy during the obstacle-subsection. Direction hints after the (last) obstacle and before the following fork (*w3*) might support MWIs' fork selection. These hints constituted the predictor variable for accuracy in the fork-subsection (*w4*).

In general, operators intervened rather rarely and did not use various intervention options in the very short abnormal track sections. Thus, we were interested in the proportion of abnormal track sections a specific input was used: We related set intervention options to the sum of obstacle-fork combinations. For instance, one operator set direction hints before



obstacles in four of nine obstacle-fork combinations. Again, our objective was to reach enough variance for data analysis. Note that repeated inputs of the same type (e.g., obstacle warnings) in the same track sections were ignored for the calculation of control action parameters.

### **Procedure**

Each operator participant ran through an AMD training comprising theoretical and practical modules at the beginning of the experiment (around 85 minutes). MWI participants received a short briefing about their tracking task. After a supervision phase (around 30 minutes), the operator acted as supervisory controller of the MWI team in trials 8 to 11 (around 30 minutes). Subjective assessments of all three participants were collected via paper and digital questionnaires (see laptops in Figure 1) before, between, and after the experiment. All in all, the experimental session took about 150 minutes for the operator and about 90 minutes for the MWIs. The position of the tracked object on the track as well as each input of the team members were recorded into log files (csv-format) with a resolution of 39 milliseconds. Team members' inputs comprise operators' clicks on the AMD and the MWIs' joystick movements.

## **Results**

### **Analysis of Data**

We used a special tool developed in the ATEO project to aggregate the high-resolution logs (Hampel, 2009). These aggregated data were analyzed with IBM PASW 18. We ran eight multiple linear regressions<sup>1</sup> to determine the influence of track section related interventions on track section related team performance, MWIs' effort and evaluated quality of assistance. Following an explorative approach, predictors were entered simultaneously. Operators' interventions during normal track sections were entered as predictor variables in six

---

<sup>1</sup> For a discussion of multiple linear regression analysis as method of choice and the value of assumption checking see Stade, Meyer, Niestroj and Nachtwei (2011).

regressions models. Concerning accuracy in abnormal track sections, the rate of timely warnings as well as the proportion of hints “speed” and “direction” were included as predictors in two regression models (i.e., during obstacle- respectively fork-subsections). In fact, the repeated use of the same predictors generally requires a Bonferroni correction to counteract significance level inflation. However, we did not adjust the significance level due to our explorative approach and the small sample size. Moreover, we will interpret marginal not significant relations between predictors and criterions. The tabulated results display both standardized regression weights and part correlation coefficients (two-tailed; unadjusted  $\alpha$ ). Part correlation coefficients represent the unique relationship between each predictor and criterion and can be directly interpreted as effect sizes (Aiken & West, 1991).

### **Normal Track Sections**

Table 1 to 3 present the regressions of operators’ interventions in normal track sections on team performance, MWIs’ effort and evaluated quality of assistance. There was no partial relation between these criteria and operators’ intervention frequency, as shown by the corresponding weak part correlations (ranged from  $r_{part} = |.02|$  to  $|.15|$ ).

**Speed and accuracy.** None of the intervention indicators were significant predictors of speed or accuracy in normal track sections. The part correlations ranged from  $r_{part} = |.05|$  to  $|.24|$  (see Table 1).

--- *Insert Table 1 about here.* ---

**Effort of MWIs.** Table 2 shows that the part correlation of hint rates “faster” with effort of MWIs 1 (focus on fast tracking) was low ( $r_{part} = .15$ ,  $p = .353$ ), whereas this predictor showed a moderate effect on effort of MWIs 2 (focus on accuracy) ( $r_{part} = -.29$ ,  $p = .052$ ). For both MWIs, we found that their effort was significantly associated with the

percentage of hints “more accurate” displaying moderate effect sizes (effort of MWIs 1:  $r_{part} = .38$ ,  $p = .019$ ; effort of MWIs 2:  $r_{part} = .31$ ,  $p = .038$ ). The proportion of inputs “limit control” was not a significant predictor of MWIs’ effort.

--- Insert Table 1 about here. ---

**Assessed quality of assistance.** As Table 3 shows, there was no significant relation between operators’ control actions and the assistance evaluated by MWIs 1 (part correlations ranged from  $r_{part} = .00$  to  $.08$ ). Also for MWIs 2 (second block in Table 3), we found no significant effect of operators’ interventions on assessed quality of assistance.

--- Insert Table 3 about here. ---

### **Abnormal Track Sections**

Table 4 shows the regressions of operators’ control actions on tracking accuracy in obstacle-subsections. The part correlations of rates of timely obstacle warnings ( $r_{part} = .15$ ,  $p = .305$ ) and direction hints ( $r_{part} = -.25$ ,  $p = .089$ ) were not significant. In contrast, the hint rate regarding speed (i.e., hints “faster” or “slower”) significantly predicted accuracy ( $r_{part} = .38$ ,  $p = .012$ ) in obstacle-subsections.

--- Insert Table 4 about here. ---

Concerning tracking accuracy in the fork-subsections, we found no significant effect of operators’ interventions showing low part correlations (ranged from  $r_{part} = |.02|$  to  $|.13|$ ).

--- Insert Table 5 about here. ---

## DISCUSSION

The present study investigated intervention parameters in order to get more insights into operator's support in a virtual team task. For this purpose, we analyzed the relationship between track section specific interventions of operators and track section related outcomes. The described process of data aggregation and analysis was a first approach to deal with the amount of data generated by our particular microworld.

### Normal Track Sections

Surprisingly, operators' intervention frequency was not associated with the measured outcomes. At least an association with MWIs' effort would have been possible.

Another unexpected aspect was that no relationships were observed between the proportions of hints regarding greater tracking speed as well as more accurate behavior in curves and MWIs' tracking. We believe that these hints include operator's assessment of MWIs' past and actual tracking. These hints could also refer to a (visible) track section in the near future. In addition, we suppose that hints like "faster" or "more accurate" – from MWIs' point of view – imply a *request* to change or adapt tracking behavior, but not *how* to realize this request. Latter aspect might be strengthened by (1) the found lack of an association between hints "faster" respectively hints "more accurate" and MWIs' tracking performance, and (2) the found positive association between hints "more accurate" and MWIs effort. Surprisingly, higher proportions of hints "faster" were related to lower effort of MWIs 2 (focus on accuracy). A possible explanation for the lower effort of MWIs 2 could be that hints "faster" prioritized one sub-goal. The prioritization of one sub-goal might reduce MWIs 2 intra-personal goal conflict and thereby affect their performance (Locke, Smith, Erez, Chah, & Schaffer, 1994).

In contrast, an empowerment of one of the MWI teammates offers the operator the opportunity to directly confirm the tracking style of one MWI. By using the hard input "limit

control”, the operator can also balance speed and accuracy. It should be noted that we analyzed how *often* operators used this hard input. We did not analyze to *whom* the operator allocated more joystick control because an averaging across track sections and trials seems not very meaningful. On the other hand, a separate and in-depth analysis of each normal track section considering which MWI achieved more input power seems very time-consuming and challenging. Consequently, our results with regard to hard inputs re-allocating joystick control are limited in interpretation. Nevertheless, if we assume that MWIs 2 (focus on accurate tracking) received more joystick control, it would be stressful for MWIs 1 to reach their task goal “focusing on speed” (expressed in higher effort and lower assessed quality of assistance), whereas MWIs 2 would be confirmed in their tracking style (expressed in lower effort and higher assessed quality of assistance). Future work is needed to examine these relationships.

### **Abnormal Track Sections**

Results related to abnormal track sections are also unexpected. The finding that the timely warning rate did not predict accuracy in both obstacle- and fork-subsections is important and needs to be explained. One possible cause is the calculation of the timing variable. We took into account *if* the MWIs received obstacle (respectively fork) warnings before the imminent abnormal track sections became visible to the MWIs. In contrast, Abe and Richardson (2004) investigated the effect of several *time-lags* between alarm warnings and an upcoming collision on drivers’ behavior within the scope of the configuration of Forward Collision Warning Systems. Referring to our lab system, future work is needed to explore the effect of such different time-lags on team performance. For instance, early warnings might offer the MWIs the opportunity to prepare avoidance maneuvers. Beyond the examination of varied time-lags, it should be investigated if *warnings themselves* improve MWIs’ tracking in the subsequent abnormal track sections as well as for longer periods of tracking (Ben-Yaacov, Maltz, & Shinar (2002).

A second possible explanation for the missing effect of the timely warning rate with accuracy is that operators' assistance beyond or instead of warnings could have a greater stake in team performance than warnings themselves. However, none of the predictors were associated with accuracy in the obstacle-subsections respectively fork-subsections. In particular, operators' direction hints immediate before forks did not predict tracking accuracy while passing through forks. Thus, future work should examine how MWIs commonly decide on a fork since the quality of this selection might strongly influence tracking accuracy in forks.

### **Limitations and Extensions of Investigated Parameters**

One methodical limitation of the present study is the assumption of a linear relationship between predictors and criteria. Using a linear term to express an *existing* curvilinear relationship will result in underestimated regression parameters (Cohen, Cohen, West, & Aiken, 2003). Another methodical limitation is that "speed" and "accuracy" were separately investigated as task performance indicators although operators were instructed to *simultaneously* optimize efficiency and safety. Thus, we need further indicators which represent the extension to which operators differently weights speed and accuracy.

Beyond that, our findings should be interpreted with caution. First, we have restrictions regarding data recording and data aggregation. Predictors and team performance outcomes were averaged across track sections and trials. Thus, one should consider that we didn't investigate the *immediate* effect of interventions on MWIs' proximate tracking performance. Moreover, MWIs' effort and evaluated quality of assistance were not measured in normal vs. abnormal track sections but after each trial respectively after all trials. Second, causal relations cannot be proved by correlational data (Cohen et al., 2003). For instance, a possible reason for a relation between MWIs' effort and operators' interventions could be that this (fed back) rating is not the consequence of, but the cause for, operators' inputs. Latter causal relation could be investigated by an examination of operators' eye movements. Stade,

von Bernstorff, Nachtwei and Wandke (2012) reported a first approach in our lab system to measure and analyze operators' attention allocation behavior via various eye tracking parameters. A detailed analysis of operators' eye movements on the AMD might be helpful to understand which information operators use to prepare and set their control actions (e.g., observation of the real-time video or effort visualization). In addition, detailed post-experience interviews of operators (for example enriched by the playback of a recorded trial) would be informative to gain closer insights into operators' supporting behavior.

In order to better understand the link between operators' interventions and MWIs' effort respectively MWIs' evaluated quality of assistance, we suggest to explore under which circumstances (1) MWIs have need for supervision as well as (2) the supervisor (= operator) affects subordinates' (= MWIs') behavior (de Vries, Roe, & Taillieu, 1998). For this purpose, the ATEO Lab System could be extended by additional components. For instance, the MWIs could beckon the operator to support and give immediate feedback during the ongoing tracking. Associated future works should also investigate whether MWIs' rating of perceived support is affected by team performance (Ayman & Chemers, 1983).

## **Conclusion**

As shown, the current configuration of the ATEO Lab System and the actual test procedure can be applied as well as adapted for the investigation of various research questions, in particular for the assessment of supervisory and leadership behavior. The division of the track in several (sub)sections as well as a consideration of different visibility ranges of the team members can be seen as an adequate approach in our lab system for measuring operator performance. Beyond that, the measurements proposed in this paper are useful for investigating differences in operators' interventions. From the results of our present study we also conclude that our approach for data aggregation and data analysis was a first but important step towards measuring operator performance in our lab system. The use of further parameters (e.g., concerning sequences of interventions) as well as combined data

sources (e.g., logs of clicks and eye movements) will enhance this approach. Moreover, such future research will enable us to better predict and explain performance of operators in any complex and dynamic microworld.

### References

- Abe, G., & Richardson, J. (2004). The effect of alarm timing on driver behavior: An investigation of differences in driver trust and response to alarms according to alarm timing. *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(4-5), 307-322. doi: <http://dx.doi.org/10.1016/j.trf.2004.09.008>
- Aiken, L.S., & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: Sage.
- Ayman, R., & Chemers, M. M. (1983). Relationship of supervisory behavior ratings to work group effectiveness and subordinate satisfaction among Iranian managers. *Journal of Applied Psychology*, 68(2), 338-341. doi: 10.1037/0021-9010.68.2.338
- Ben-Yaacov, A., Maltz, M., & Shinar, D. (2002). Effects of an in-vehicle collision avoidance warning system on short- and long-term driving performance. *Human Factors*, 44(2), 335-342.
- Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, 9(2-3), 171-184.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- De Vries, R. E., Roe, R. A., & Taillieu, T. C. B. (1998). Need for supervision: Its impact on leadership effectiveness. *The Journal of Applied Behavioral Science*, 34(4), 486-501.
- Hampel, T. (2009). *Programmierung eines Auswertungsprogramms für eine komplexe, dynamische Versuchsumgebung [Programming of a log analyzer in a complex,*

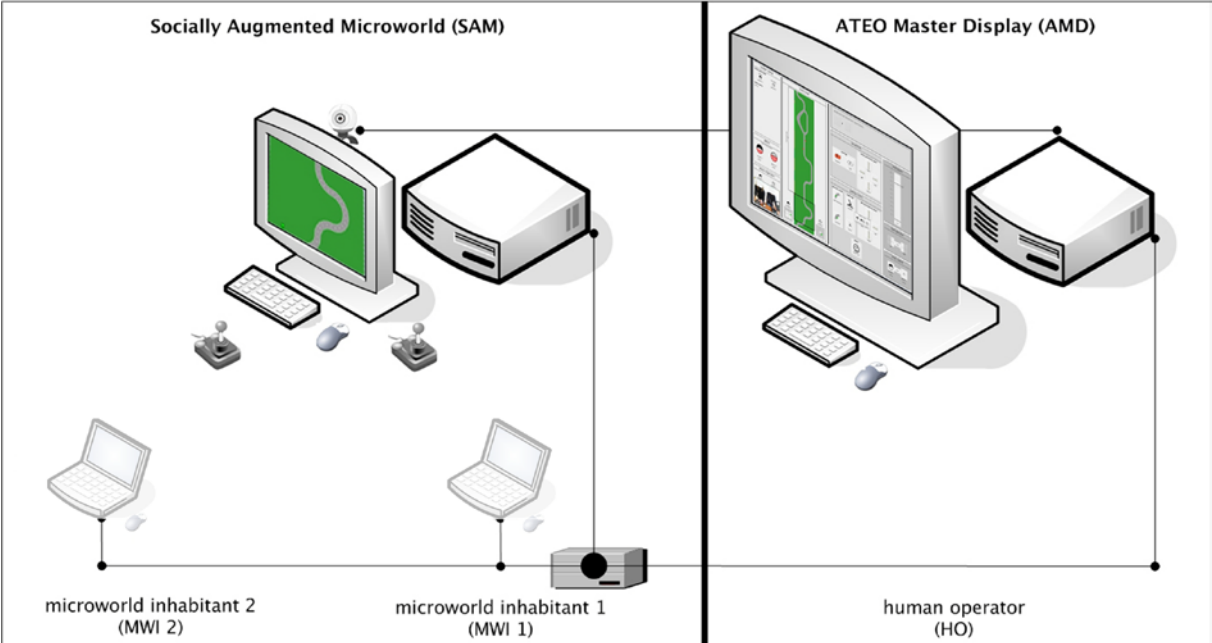


- dynamic lab system*]. Unpublished diploma thesis, Humboldt-Universität zu Berlin, Berlin, Germany.
- Howie, D. E., & Vicente, K. J. (1998). Measures of operator performance in complex, dynamic microworlds: Advancing the state of the art. *Ergonomics*, 41(4), 485–500.
- Ingalls, D., Kaehler, T, Maloney, J., Wallace, S., & Kay, A. (1997). Back to the future: The story of Squeak, a practical Smalltalk written in itself. *Proceedings of the 12<sup>th</sup> ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications* (pp. 318–326). Atlanta, GA: Association for Computer Machinery.
- Locke, E. A., Smith, K. G., Erez, M., Chah, D.-O., & Schaffer, A. (1994). The effects of intra-individual goal conflict on performance. *Journal of Management*, 20(1), 67–91.
- Meister, D. (2001). Basic premises and principles of human factors measurement. *Theoretical Issues in Ergonomics Science*, 2(1), 1–22.
- Nachtwei, J. (submitted-a). *A multi-level design approach for a supervisory control master display in human factors experiments*. Manuscript submitted for publication.
- Nachtwei, J. (submitted-b). *SAGAT versus SPAM? It depends! An aid for situation awareness method selection (aSAmS)*. Manuscript submitted for publication.
- Stade, M., Meyer, C., Niestroj, N. & Nachtwei, J. (2011). (Not) Everybody's darling: Value and prospects of multiple linear regression analysis and assumption checking. In B. Krause, R. Beyer, & G. Kaul (Eds.), *Empirische Evaluationsmethoden Band 15* (pp. 17–34). Berlin, Germany: ZeE Verlag.
- Stade, M. J. C., von Bernstorff, C., Nachtwei, J., & Wandke, H. (2012) *Forward-looking: Measuring attention allocation for evaluating supervisory control in a complex and dynamic lab system*. Manuscript in preparation.
- Von Bernstorff, C., & Nachtwei, J. (submitted). *Between keyhole and clutter effect – A multi-level evaluation of interface extensions using the ATEO master display (AMD)*. Manuscript submitted for publication.

- Wandke, H., & Nachtwei, J. (2008). The different human factor in automation: the developer behind vs. the operator in action. In D. de Waard, F. O. Flemisch, B. Lorenz, H. Oberheid, & K. A. Brookhuis (Eds.), *Human factors for assistance and automation* (pp. 493–502). Maastricht, the Netherlands: Shaker Publishing.
- Zijlstra, F. R . H., & van Doorn, L. (1985). *The construction of a scale to measure perceived effort*. Delft, the Netherlands: Department of Philosophy and Social Sciences, Delft University of Technology.

Figures

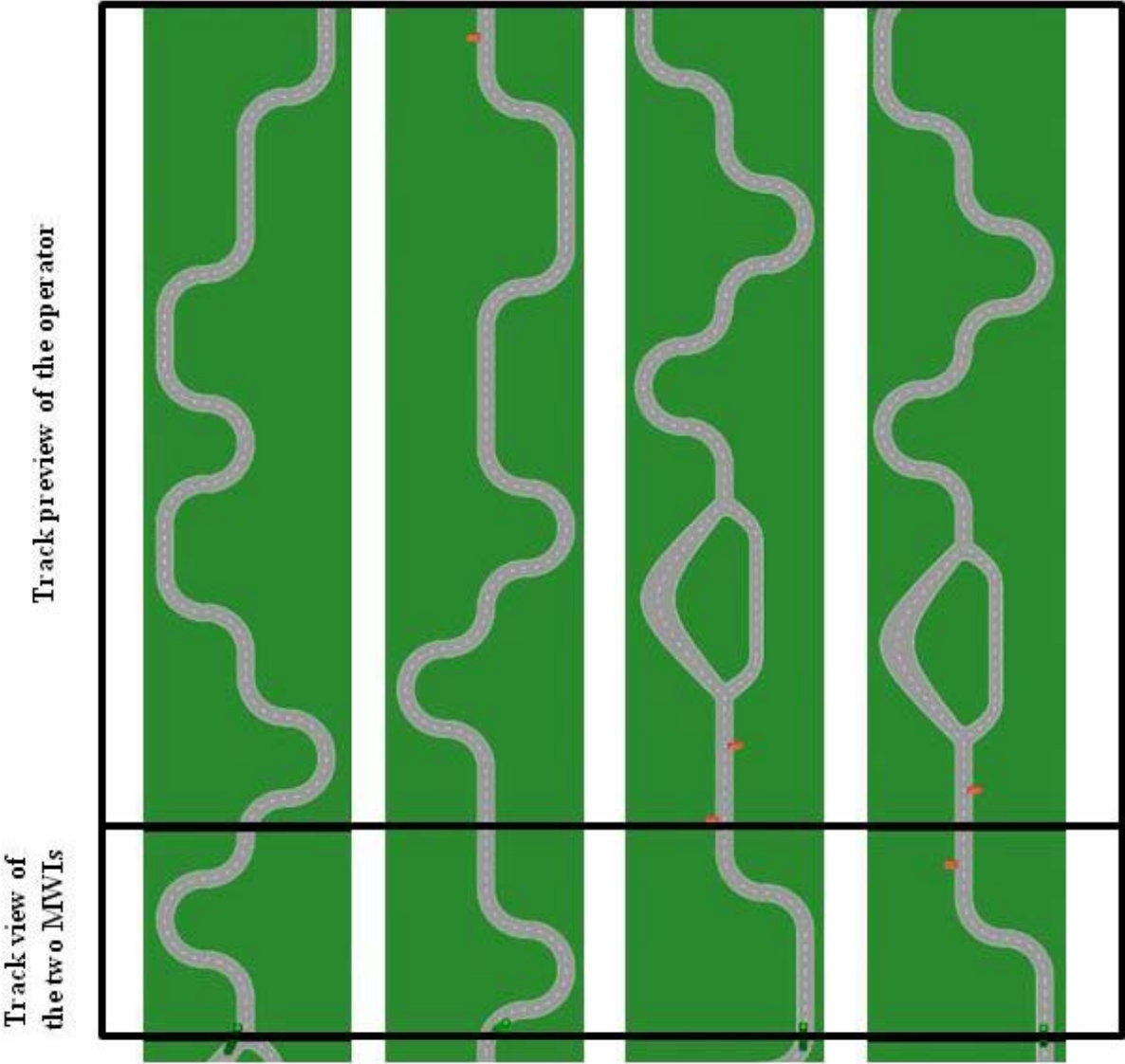
Figure 1. Experimental Setup of the ATEO Lab System (ALS) Containing the Socially Augmented Microworld (SAM) and the ATEO Master Display (AMD)



**Figure 2.** Actual Version of the ATEO Master Display (AMD) for Supervision and Control of SAM (Translated From German)



**Figure 3.** The calculation of the different track sections based on operators' and MWIs' visibility of the upcoming track section. From left to right: (1) Starting point of a normal track section, (2) the first part of an abnormal track section is visible to the operator, (3) immediately before the first part of an obstacle-fork combination is visible to the two MWIs and (4) the first part of an abnormal track section is visible to the MWIs.



### Tables

**Table 1.** Regressions of Operators' Interventions on Team Performance (Normal Track Sections).

Operators' inputs	Speed			Accuracy		
	$\beta$	$r_{part}$	$p$	$\beta$	$r_{part}$	$p$
Proportion of hints "faster" (%)	.15	.14	.395	-.06	-.06	.739
Proportion of hints "more accurate" (%)	-.06	-.05	.739	-.07	-.06	.712
Proportion of "limit control" (%)	-.26	-.24	.144	.06	.05	.768
Intervention frequency	-.13	-.12	.475	-.06	-.05	.753

*Note.*  $N = 40$ . Speed:  $R^2 = .12$  ( $p = .356$ ); accuracy:  $R^2 = .02$  ( $p = .965$ ); reported values of  $R^2$  are unadjusted.

**Table 2.** Regressions of Operators' Interventions on MWIs' Effort (Normal Track Sections).

Operators' inputs	Effort of MWIs 1 (focus on speed)			Effort of MWIs 2 (focus on accuracy)		
	$\beta$	$r_{part}$	$p$	$\beta$	$r_{part}$	$p$
Proportion of hints "faster" (%)	.16	.15	.353	-.32	-.29	.052
Proportion of hints "more accurate" (%)	.43	.38	.019	.35	.31	.038
Proportion of "limit control" (%)	.20	.18	.255	-.18	-.16	.276
Intervention frequency	.11	.10	.538	-.02	-.02	.890

*Note.*  $N = 40$ . Effort of MWI 1:  $R^2 = .16$  ( $p = .177$ ); effort of MWI 2:  $R^2 = .28$  ( $p = .020$ ); reported values of  $R^2$  are unadjusted.

**Table 3.** Regressions of Operators' Interventions on MWIs' Evaluated Quality of Assistance (Normal Track Sections).

Operators' inputs	Quality of assistance (rated by MWIs 1; focus on speed)			Quality of assistance (rated by MWIs 2; focus on accuracy)		
	$\beta$	$r_{part}$	$p$	$\beta$	$r_{part}$	$p$
Proportion of hints "faster" (%)	.01	.01	.971	-.21	-.19	.228
Proportion of hints "more accurate" (%)	.00	.00	.984	-.17	-.15	.348
Proportion of "limit control" (%)	.01	.01	.961	-.32	-.29	.070
Intervention frequency	.09	.08	.639	.17	.15	.344

*Note.*  $N = 40$ . Quality of assistance rated by MWIs 1:  $R^2 = .01$ ,  $p = .991$ ; quality of assistance rated by MWIs 2:  $R^2 = .17$  ( $p = .160$ ); reported values of  $R^2$  are unadjusted.

**Table 4.** Regressions of Operators' Interventions on Team Performance (Abnormal Track sections; Obstacle-Fork-Combination, Sub-Section Obstacle)

Operators' inputs	Accuracy		
	$\beta$	$r_{part}$	$p$
Rate of timely warnings (%)	.16	.15	.305
Rate of hints regarding speed ( $w1+w2$ ) (%)	.40	.38	.012
Rate of hints regarding direction ( $w2$ ) (%)	-.26	-.25	.089

*Note.*  $N = 40$ .  $R^2 = .24$  ( $p = .019$ ); reported values of  $R^2$  are unadjusted.

**Table 5.** Regressions of Operators' Interventions on Team Performance (Abnormal Track sections; Obstacle-Fork-Combination, Sub-Section Fork)

Operators' inputs	Accuracy		
	$\beta$	$r_{part}$	$p$
Rate of timely warnings (%)	.02	.02	.923
Rate of hints regarding speed ( $w1+w2$ ) (%)	-.12	-.12	.464
Rate of hints regarding direction ( $w3$ ) (%)	-.14	-.13	.445

*Note.*  $N = 40$ .  $R^2 = .04$  ( $p = .730$ ); reported values of  $R^2$  are unadjusted.



### **FORSCHUNGSARTIKEL NR. 3**

Stade, M. J. C., Meyer, C., Niestroj, N., & Nachtwei, J. (2011). Ergonomische Gestaltung trotz oder aufgrund interdisziplinärer Teamarbeit? In M. Grandt, & S. Schmerwitz (Hrsg.), *Ergonomie im interdisziplinären Gestaltungsprozess (DGLR-Bericht 2011-1)*. Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt. 1–16.

## **Ergonomische Gestaltung und Evaluation trotz oder aufgrund interdisziplinärer Teamarbeit**

M.J.C. Stade, C. Meyer, N. Niestroj & J. Nachtwei

### **Zusammenfassung**

Im Forschungsprojekt ATEO (Arbeitsteilung Entwickler-Operateur) wurde ein Laborsystem entwickelt, mit dem ein fairer Vergleich von Stärken und Schwächen von Operateuren und Entwicklern ermöglicht werden soll. Konzipiert und gestaltet wurde dieses System in einem interdisziplinären Team von Diplomanden und Doktoranden der Psychologie und Informatik. Gestaltungsziel im Rahmen eines Laborsettings war ein sogenanntes Masterdisplay zur Supervision und Kontrolle eines komplexen, dynamischen Systems. Auf Basis von Analysen der Ressourcen von Operateuren (modellgeleitet und in Interviews mit realen Operateuren erfragt) sowie anhand von aufgabenorientierten Anforderungen an Anzeige- und Bedieneinheiten wurden erste Prototypen des Masterdisplays entworfen. Diese wurden iterativ bewertet, überarbeitet und ausgebaut. Eine Vielzahl an Indikatoren für die Bewertung des Masterdisplays, aber auch für die Interaktion des Operateurs mit dem Masterdisplay sowie seiner verwendeten Strategien zur Aufgabenbewältigung stehen zur Verfügung. Die Herausforderung besteht in der Auswahl und Kombination der Daten, gewonnen über Logfiles, Blickbewegungsaufzeichnungen, Fragebögen, Beobachtungen und Interviews. Exemplarisch werden Möglichkeiten und Schwierigkeiten für die Interpretation und Ableitung von weiteren Ausbaustufen des Masterdisplays aufgezeigt. Dieser Beitrag fungiert als Werkstattbericht mit starkem Bezug zu empirischen Arbeiten im Projekt ATEO. Aus Priorisierungs- und Platzersparnisgründen wird auf die Diskussion des Forschungsstandes bewusst verzichtet, da diese in den im Projekt entstandenen und hier zitierten Artikeln bereits vorgenommen wurde.

### **1 Interdisziplinäre Entwicklung eines Master Displays**

#### **1.1 Rahmen und Bedarf: Das Projekt ATEO**

Ziel des Projektes ATEO (Arbeitsteilung Entwickler-Operateur) ist ein empirischer Vergleich der Ressourcen von Operateuren und Entwicklern im Experiment. Um den Vergleich von Stärken und Schwächen von Operateuren und Entwicklern fair zu gestalten (Nachtwei, submitted-a), wird ein System benötigt, das sowohl durch die Operateure als auch durch eine Automatik – konzeptioniert und umgesetzt durch Entwickler – supervidiert und kontrolliert wird und dies-

bezüglich vergleichbar anspruchsvoll ist (Wandke & Nachtwei, 2008). Teil dieses Systems ist eine Mikrowelt, in der zwei sogenannten Mikroweltbewohner (MWB) in einer Trackingaufgabe gemeinsam ein Fahrobject eine virtuelle Fahrstrecke entlang steuern. Diese soziale Komponente, welche die Mikrowelt anreichert („Socially Augmented Microworld“, abgekürzt SAM), soll verhindern, dass die Mikrowelt zu deterministisch und zu wenig komplex ist – ein Problem, das bei Wandke und Nachtwei (2008) beschrieben wird und die Vorhersehbarkeit durch den Operateur als auch durch den Entwickler betrifft. Operateure müssen Systemverhalten ebenso vorhersagen wie Entwickler, lediglich die zeitliche Auflösung ist eine andere (von Sekunden und Minuten bei Operateuren bis Monate und Jahre bei Entwicklern).

Die Rolle der Mikroweltbewohner nehmen zwei Versuchspersonen ein. Ihre Aufgabe ist es, das Fahrobject gemeinsam möglichst schnell und sicher an das Ziel der Fahrstrecke zu bringen. Hierfür haben beide MWB jeweils einen Joystick zur Verfügung mit einer Steuerkontrolle, die sich auf Richtung (Links-Rechts-Auslenkung) und Geschwindigkeit (Vorn-Hinten-Auslenkung) bezieht und anfänglich für jeden MWB 50% beträgt (vgl. Bild 2, Bereich 16 für eine exemplarisch davon abweichende Aufteilung im Zuge des Experiments). Obgleich das Tracking kooperativ und koordiniert erfolgen soll, dürfen die MWB weder verbal noch nonverbal miteinander kommunizieren. Aufgrund dieser Einschränkung erfahren sie auch nicht von der jeweils unterschiedlichen Fahrinstruktion: MWB 1 ist instruiert, genau, aber vor allem schnell zu fahren, MWB 2 erhält die gegenläufige Instruktion. Beiden MWB wird versichert, dass sie eine identische Instruktion für ihr Team erhalten haben. Zudem werden beide MWB instruiert, sich bei Gabelungen für eine Abzweigung zu entscheiden und Hindernissen auszuweichen.

Eine dritte Versuchsperson in der Rolle des Operateurs überwacht und kontrolliert über das sogenannte ATEO Masterdisplay (AMD, vgl. Bild 1) die Systemleistung, das heißt das Absolvieren der Trackingaufgabe durch die MWB. Wie bereits geschildert stellen die MWB nur ein Mittel zum Zweck dar, das zu überwachende und kontrollierende System mit einer Eigendynamik zu versehen, die nicht zu 100% antizipierbar ist. Ziel des Operateurs ist die Optimierung des Systemprozesses, welche beinhaltet, maximale Geschwindigkeit bei maximaler Genauigkeit des Fahrobjectes zu erreichen. Aufgrund der sozialen Komponente durch die kooperative Aufgabenausführung durch die MWB ist der Systemprozess nicht nur für den Operateur schwer vorhersehbar. Auch ein Entwickler muss bei der Entwicklung einer Automatik die Interaktion von zwei realen Personen bzw. bei der Entwicklung einer unterstützenden Automatik für den Operateur drei reale Personen berücksichtigen.

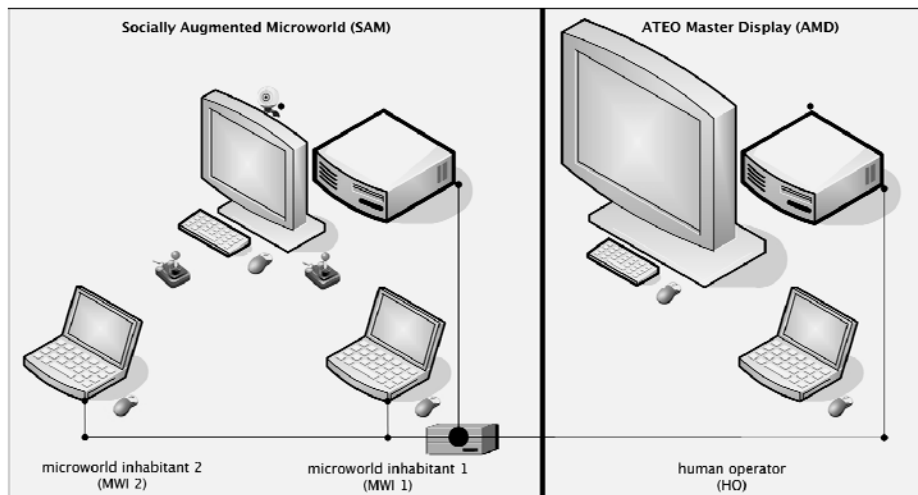


Bild 1: ATEO Labor System (ALS), bestehend aus der Socially Augmented Microword (SAM) und dem ATEO Master Display (AMD)

## 1.2 Das AMD in der aktuellen Version

Das AMD besteht in der aktuellen Version aus einem Überwachungs- und einem Eingabebereich, die klar voneinander durch einen prominenten Balken in der Mitte des Interfaces getrennt sind (vgl. Bild 2). Auf der linken Seite befinden sich Anzeigen zum Prozessverlauf, auf der rechten Seite Eingriffsmöglichkeiten für den Operateur.

Zum Informationsangebot (Bereiche 1 bis 6) für den Operateur gehören:

- (1) Systemstatus: Zuordnung der Fahrinstruktion (Fokus auf Geschwindigkeit versus Genauigkeit) zum jeweiligen MWB sowie eine Anzeige über die aktuelle Fahrtenzahl (1 bis 11)
- (2) Anstrengung: Nach jeder Fahrt bewerten die MWB jeweils an ihrem Laptop (vgl. Bild 1), wie anstrengend sie die gerade absolvierte Fahrt empfanden. Das Ergebnis wird kurz vor Beginn der nächsten Fahrt auf den AMD übertragen. Hierfür werden sowohl die Angaben der MWB in Prozenten quantifiziert als auch durch Ab- und Zunahme der roten Fläche in den kreisrunden Elementen veranschaulicht, die jeweils einen Kopf eines MWB symbolisieren sollen.
- (3) Videobild der Personen: In einem Echtzeitvideo der MWB kann der Operateur deren Körperhaltung, Gesichtsausdruck, Joystickausrückung usw. beobachten. Im Versuch sind die Gesichter der MWB zu erkennen, die im Bild 2 aus Datenschutzgründen unkenntlich gemacht wurden.
- (4) Streckenansicht: Der Operateur sieht den Bereich der Fahrstrecke, den auch die MWB auf ihrem Monitor sehen. Darüber hinaus hat der Operateur eine Streckenvorschau zur Verfügung, welche die vierfache Länge der MWB-Ansicht

umfasst. Dem Operateur wird damit ermöglicht, schon vor den MWB Hindernisse und, wie in diesem Beispiel, nachfolgende Gabelungen zu identifizieren und diese den MWB mit entsprechenden Hinweisen anzukündigen.

(5) Joystick Input: Das jeweilige Ausmaß der von den MWB vorgenommenen Joystickausrückung (Kreis in der Mitte) ebenso durch den Operateur ermöglichte Ausprägung der Joystickausrückung (weißes Kästchen) sind rechts und links neben der aktuellen Streckenansicht platziert.

(6) Hinweisfeedback: Oberhalb des Bereichs für den Joystickinput der MWB wird derjenige Hinweis angezeigt, der vom Operateur für einen oder beide MWB gegeben wurde.

Zum Funktionsangebot gehören drei Bereiche: Messung von Situationsbewusstsein (oben), weiche Eingriffe (mittig) und harte Eingriffe (rechts).

(7) Messung Situationsbewusstsein: Für eine genaue Beschreibung der Methode im ALS siehe Nachtwei (submitted-b).

Visuelle Hinweise (Bereiche 8 und 9) vom Operateur werden am gemeinsamen Monitor der MWB angezeigt (vgl. Bild 1).

(8) Warnungen: Der Operateur kann den MWB Hindernisse und Gabelungen ankündigen.

(9) Fahrtrichtung und Geschwindigkeit: Über diese Hinweise kann der Operateur die MWB auffordern, schneller bzw. langsamer zu fahren sowie nach links oder rechts zu steuern.

Auditive Hinweise (Bereiche 10 bis 13) werden über gesplittete Audiokanäle der Kopfhörer an die MWB übermittelt. Die auditiven Hinweise werden über das Abspielen eines vorher aufgenommenen MP3s realisiert. Aufgrund der gesplitteten Audiokanäle kann der Operateur bestimmen, ob ein Hinweis nur an einen oder an beide MWB übertragen wird. Hierfür muss der Operateur in einem kleinen Fenster, das den ausgewählten Hinweis überlappt, die Allokation bestimmen.

(10) Kurvenverhalten: In Kurven kann die Aufforderung erfolgen, diese genauer zu fahren oder abzukürzen.

(11) Führung: Hiermit kann der Operateur die Anweisung erteilen, die Führung im gemeinsamen Steuern des Trackingobjektes zu übernehmen oder zu überlassen.

(12) Fahrtrichtung und Geschwindigkeit: Wie mittels der visuellen Hinweise können hier Empfehlungen zur Geschwindigkeit und zur Richtung gegeben werden.

(13) Lob: Als positives Feedback kann der Operateur ein „Weiter so!“ abspielen lassen.

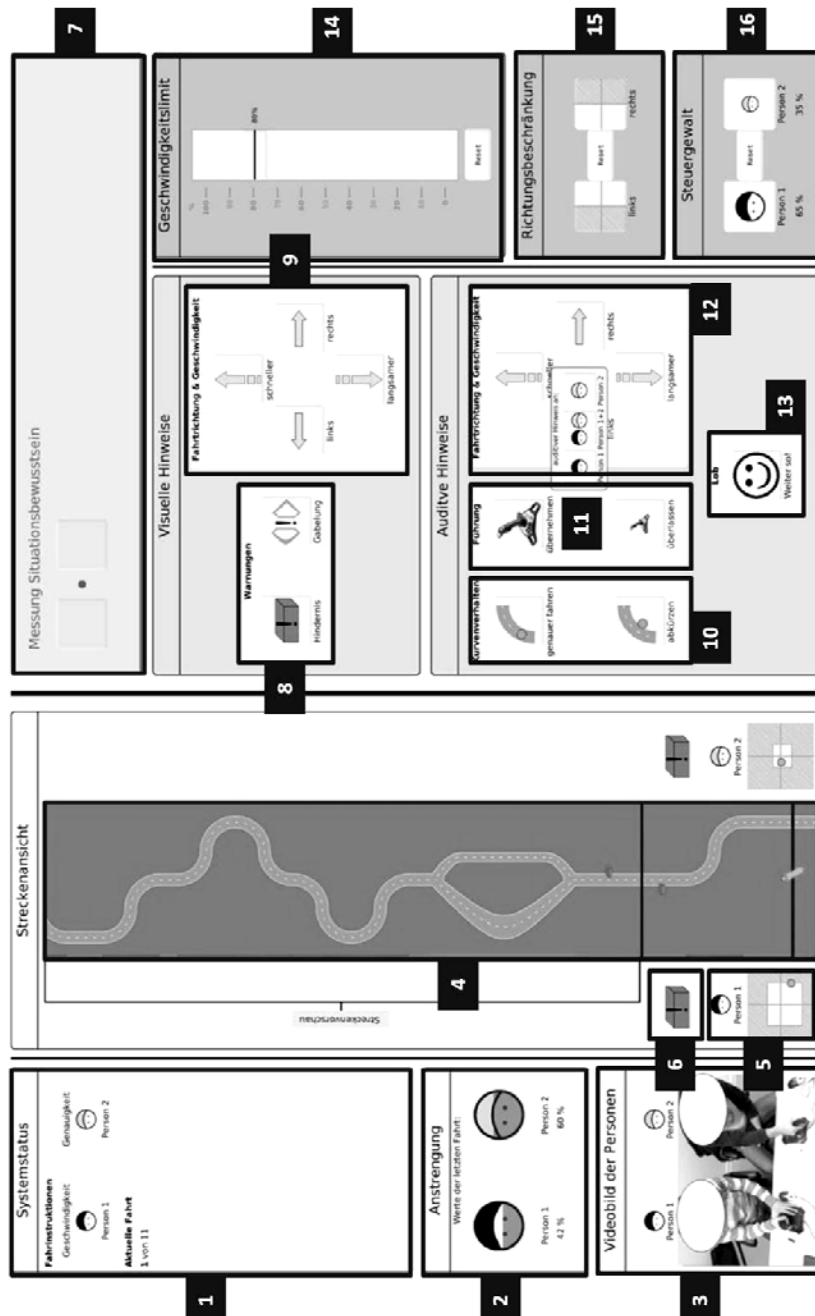


Bild 2: ATEO Master Display (Version v2.8.4); für eine hochaufgelöste und farbige Darstellung des AMDs siehe <http://jensnachtwei.blogspot.com/>

Mit harten Eingriffen (Bereich 14 bis 16) greift der Operateur direkt in den Prozess ein und kann so Manipulationen an der Steuerung des Fahrobjektes vornehmen. Diese Eingriffe können durch Drücken des Reset-Buttons in den Ausgangszustand versetzt werden.

(14) Geschwindigkeitslimit: Die voreingestellte maximale Geschwindigkeit des Fahrobjektes von 100% kann durch Ziehen des Schiebereglers bis zu 0% verringert werden. Die Skala mit 10% Intervallen an der linken Seite des Schiebereglers dient als Orientierung. An der rechten Seite des Schiebereglers wird jeweils der aktuelle Wert mit einer Auflösung von einem Prozent dargestellt.

(15) Richtungsbeschränkung: Die Steuerung des Fahrobjektes kann nach links oder rechts blockiert werden.

(16) Steuergewalt: Die anfängliche Gleichverteilung der Steuerkontrolle über die Joysticks kann in fünf Prozentschritten auf die MWB umverteilt werden bis zu einer Verteilung von 0% zu 100%.

### 1.3 Versuchsablauf

Nach einem kurzen Überblick über den Versuchsablauf und seine Aufgabe durchläuft der Operateur eine Trainingssequenz, in der er sich sowohl theoretisch als auch praktisch mit dem AMD vertraut macht. Während der ersten vier Fahrten überwacht er das Fahrverhalten der MWB ohne Eingriffe zu tätigen, wobei Fahrt 1 und 2 ausschließlich von MWB 1 und Fahrt 3 und 4 ausschließlich von MWB 2 absolviert werden. Nach dieser individuellen Trainingssequenz fahren die MWB ab Fahrt 5 gemeinsam. Ab Fahrt 8 kann der Operateur auch in den Prozess eingreifen, jedoch werden nur Fahrt 9, 10 und 11 für Analysen verwendet, da Fahrt 8 als Übungsfahrt dient. Im Verlauf des Experiments beantworten die drei Versuchspersonen Fragebögen (digital und in Papierform). Zur Nachbefragung des Operateurs gehört u.a. seine Einschätzung der Sinnhaftigkeit und Relevanz spezifischer Bereiche des AMD für die Aufgabenbearbeitung (vgl. Abschnitt 3). Darauf aufbauend wird er um Verbesserungsvorschläge bezüglich Inhalt, Design und Anordnung von Funktionen und Informationen gebeten (vgl. Abschnitt 4).

Die Interaktion des Operateurs mit dem AMD (Art, Dauer und Zeitpunkt von Eingriffen und Blindclicks) sowie der Input der MWB (Bewegungen der Joysticks, Wahl der Abzweigung bei Gabelungen, Kollision mit Hindernissen) wird mit einer Auflösung von 39 Millisekunden in einem Logfile (CSV-Format) gespeichert.

## 2 Der Weg zum AMD: Entwicklungsprozess und Ausbaustufen

Im folgenden Abschnitt werden die Entwicklungsschritte des AMD beschrieben. Hierbei handelt es sich nur um eine überblicksartige Darstellung, für eine genaue Betrachtung der Entwicklungsschritte sowie eine methodische Reflexion siehe Nachtwei (submitted-a)

Im ersten Entwicklungsschritt wurden nach Literatursichtung Modelle aufgestellt, welche die Fähigkeiten des Operators vor und während der Interaktion mit einem technischen System beschreiben (Nachtwei, 2010). Zudem wurden Faktoren zusammengetragen, welche diese Interaktion beeinflussen (Gérard et al., 2011). Für das konkrete Gestalten von Mensch-Maschine-Systemen wurden außerdem Leitfäden herangezogen und ihre konkrete Umsetzung am AMD diskutiert.

Basierend auf diesem theoretischen Hintergrund wurden im zweiten Entwicklungsschritt realen Operateure aus unterschiedlichen Bereichen interviewt. Hierbei wurde vor allem das Ausmaß an Informationen erfragt, das dem Operateur zur Verfügung stehen soll, um seine Aufgabe zu bewältigen.

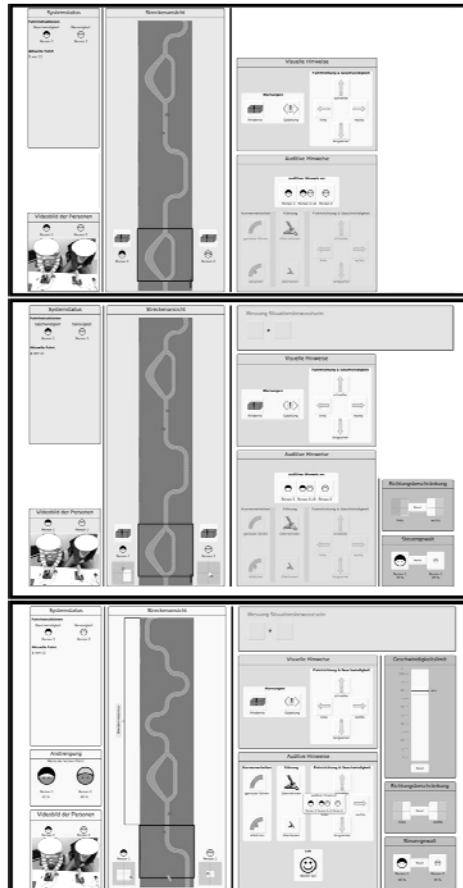


Bild 3: Ausbaustufen des AMD

An diesen Entwicklungsschritt anschließend wendete die interdisziplinäre Arbeitsgruppe, bestehend aus Psychologen und Informatikern, eine hierarchische Aufgabenanalyse an. Bei Hildebrandt et al. (2010) ist das genaue Vorgehen



beschrieben. Vorteil für dieses ressourcenschonende Vorgehen ist das Verwenden einer gemeinsamen Sprache innerhalb eines Designteams.

Basierend auf den beschriebenen Entwicklungsschritten wurde ein erster arbeitsfähiger Prototyp entwickelt, der von Experten bewertet wurde. Die Evaluation erfolgte aufgabenbasiert, Grundlage hierfür waren die in der Hierarchischen Aufgabenanalyse herausgearbeiteten Ziele und Aufgaben des Operateurs. Dieser Prototyp wurde angeregt durch die Expertenevaluation überarbeitet und erneut evaluiert.

Mehrere Iterationen (in Summe 24) folgten. Die Ergebnisse der Usability Tests, bestehend aus der Auswertung von Logfiles, Beobachtungen, Fragebögen und Interviews sowie auf dem Forschungsstand basierende Überlegungen zu Belastungsgrenzen von Operateuren beim Umgang mit komplexen dynamischen Systemen führten zu einem schrittweisen Ausbau des AMD (vgl. Bild 3).

### **3 Interdisziplinäre Evaluation des AMD**

#### **3.1 Auswahl von Indikatoren für die Bewertung des AMD und Strategien der Operateure**

Für die Bewertung des AMD sowie der Leistung des Operateurs stehen eine Vielzahl von Parametern zur Verfügung: Generische (z.B. Anzahl von Fehlern) und systemspezifische Leistungsparameter (z.B. Blindclicks bei Hinweisen, die zwei Einstellungen erfordern), mentale Zustände des Operateurs (z.B. mentale Anstrengung) oder die subjektive Einschätzung des Masterdisplays durch den Operateur (z.B. Zufriedenheit). In einer Untersuchungsreihe wurden zusätzlich Blickbewegungsdaten erhoben sowie basierend auf der Idee des Teach Backs von einigen Operateuren weitere qualitative Daten erhoben. Darüber hinaus sind Notizen der Versuchsleitenden vorhanden, in denen Anmerkungen sowie vom Versuchsleitenden als auffällig interpretierte Verhaltensweisen der Versuchspersonen dokumentiert sind. An dieser Stelle soll betont werden, wie bereichernd letztere Quelle der Datengewinnung war: Durch die Anwesenheit der Versuchsleitenden konnten wertvolle Implikationen für eine tiefere Analyse der Parameter gewonnen sowie mögliche Erklärungen für Verhaltensweisen der Operateure gesammelt werden.

Neben der großen Auswahl an Parametern für eine Evaluation besteht eine weitere Herausforderung für die Interpretation darin, dass Indikatoren für die Güte des AMD und die Variabilität im Verhalten der Operateure nicht unabhängig voneinander betrachtet werden können.

Für die Aufgabe des Operateurs, den Prozess zu überwachen und zu kontrollieren, gibt es keine vorgegebene Ideallösung. Durch das relativ breite Informations- und Funktionsangebot (vgl. Abschnitt 3.3) ergibt sich für den Operateur eine große Bandbreite an Verhaltensstrategien. Dies beinhaltet auch, dass er bestimmte Informationen und Funktionen nicht zwangsweise nutzen muss und trotzdem gute Leistung erzielen kann. In die Bewertung des gesamten AMD sowie seiner

Teilbereiche durch den Operateur fließen sowohl Bewertungen der Handhabbarkeit als auch der Sinnhaftigkeit für seine Aufgabenbewältigung ein. So brachte nach Meinung der Mehrzahl der Operateure die Funktion der Richtungsbeschränkung (vgl. Bild 2, Bereich 15) keinen Nutzen oder die Benutzung dieser Funktion wurde als unklar beschrieben. Auch zum Teil widersprüchliche Ergebnisse bei der Gegenüberstellung von Parametern wurden gefunden: Operateure, die keine Einstellung am Geschwindigkeitslimit des Fahrobjektes (vgl. Bild 2, Bereich 14) vornahmen, beurteilten die Sinnhaftigkeit dieser Eingriffsmöglichkeit für das Bewältigen ihrer Aufgabe dennoch mit bis zu 100 Prozentpunkten, das heißt als äußerst sinnvoll.

### 3.2 Visualisierung von Indikatoren für die Bewertung von AMD und Strategien der Operateure

Eine Herausforderung besteht in der Auswahl und Kombination von Indikatoren für die Beurteilung der Gestaltung des AMD sowie der von den Operateuren verwendeten Strategien. Darüber hinaus stellt sich die Frage, wie eine Gegenüberstellung dieser Indikatoren erfolgen kann, um eine Interpretation zu ermöglichen. Eine tabellarische Auflistung und Gegenüberstellung der Indikatoren ist für das Laborsetting z.T. zu komplex und erschwert, Auffälligkeiten zu entdecken.

In der Arbeitsgruppe wurde daher ein Werkzeug entwickelt, mit dem – angelehnt an die Funktionsweise von Analysesoftware für Blickbewegungsdaten – Areas of Interest über das AMD gezeichnet werden können. Diesen Blöcken werden nun Werte wie z.B. eingeschätzte Sinnhaftigkeit zur Aufgabenbewältigung zugeordnet. Der Forscher kann Wertebereiche für diese Blöcke definieren und festlegen, welche Farben diesen Wertebereichen zugeordnet werden sollen. Die so generierten Screens können als JPEG-Dateien gespeichert werden. Für den Ausbau des Werkzeuges ist geplant, dass Werte nicht per Hand vom Forscher in das Programm eingegeben werden müssen, sondern das Werkzeug auf eine Excel-Tabelle referenziert.

Muster und Unterschiede im Verhalten von Operateuren (Strategien in der Aufgabenbewältigung und Bewertung der Bereiche des AMD) konnten so leicht identifiziert werden. Beispielsweise gibt es Operateure, welche die Streckenvorschau als weniger sinnvoll für ihre Aufgabenbearbeitung einschätzten, sowohl absolut als auch in Relation zur aktuellen Streckenansicht (vgl. Bild 4; für eine Interpretation vgl. Abschnitt 3.3.3).

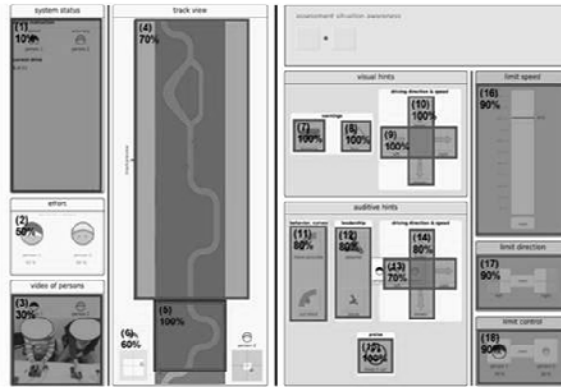


Bild 4: Visualisierung der eingeschätzten Sinnhaftigkeit durch den Operateur

### 3.3 Forschung über und mit dem AMD – Trade-off zwischen maximaler Usability und diagnostischer Aussagekraft

#### 3.3.1 Blindclicks – schlechte Usability oder gute Strategie des Operateurs?

Im AMD sind Blindclicks als solche Klicks definiert, die neben Buttons oder auf temporär inaktiven Buttons ausgeführt werden. Die Größe der Buttons als Ursache für verfehlte Klicks kann ausgeschlossen werden, da die Buttons sehr groß angelegt wurden. Die Operateure hatten im Training die Möglichkeit, sich mit der Maus und deren Handhabung im AMD vertraut zu machen (vgl. Abschnitt 1.3). Weder von den Versuchsleitenden beobachtet, noch von den Operateuren angemerkt waren grundsätzliche Probleme beim Umgang mit der Maus. Bei der letzten Ausbaustufe des AMD zeigte sich aber, dass Operateure in den letzten drei (für die Analyse relevanten) Fahrten des Versuchs bis zu 12 Blindclicks produzierten. Ist dies doch ein Indikator für eine schlechte Gestaltung des AMDs? Solch eine Schlussfolgerung auf Basis der Analyse der Logfiles darf nur unter Vorbehalt erfolgen.

Beim Verfolgen der Mausbewegung im Versuch durch die Versuchsleitenden bzw. bei der Auswertung der Screenrecordings wurde deutlich, dass produzierte Blindclicks nur teilweise unbeabsichtigt vorgenommen wurden. Nach dem Auslösen eines visuellen Hinweises müssen 1.500 ms vergehen, bis ein weiterer visueller Hinweis gegeben werden kann. Zwei aufeinanderfolgende visuelle Hinweise wurden von den Operateuren z.B. in Situationen mit Gabelungen genutzt: Operateure warnten vor der Gabelung und gaben anschließend eine Richtungsanweisung. Die Systemverzögerung wurde von manchen Operateuren umgangen, indem die Richtungsanweisung auditiv an die MWB gegeben wurde. Eine andere Strategie wählte ein Operateur, der den zweiten visuellen Button trotz bestehender Inaktivität hintereinander mehrmals anklickte, um so rechtzeitig den Zeitpunkt abzapfen, wann der Button wieder aktiv wurde.

Trotz Trainingssequenz, in der die Funktionsweise der Interaktionselemente im AMD sehr ausführlich beschrieben wurde (vgl. Abschnitt 1.3), bediente ein Operateur den Schieberegler zur Einstellung der maximalen Geschwindigkeit am Fahrobjekt nicht korrekt. Statt wie im Training erläutert den Schieberegler mit der linken Maustaste anzuklicken und bei gedrückter Maustaste diesen nach oben oder unten zu ziehen, klickte der Operateur mehrmals unterhalb des Schiebereglers. Um den Versuchsablauf nicht zu beeinflussen, intervenierte die Versuchsleitende nicht; der Operateur erkannte sein fehlerhaftes Bedienverhalten selbst nach kurzer Zeit und bediente den Schieberegler im weiteren Versuch korrekt. Eine Intention des Operateurs könnte gewesen sein, mit seinem Vorgehen den Wert der Geschwindigkeit des Fahrobjektes schneller zu ändern, indem er den Schieberegler größere Intervallsprünge vollziehen lässt. Dieses Verhalten wurde bei anderen Versuchspersonen im Umgang mit dem Fragebogentool an den Laptops beobachtet (vgl. Bild 1), obgleich die Instruktion lautete, den Schieberegler zum Einstellen der intendierten Werte zu benutzen.

### 3.3.2 Zwei Klicks für einen auditiven Hinweis – mehr Aufwand nötig oder mehr Strategie möglich?

In der aktuellen Version des AMD muss der Operateur – wie oben beschrieben – bei auditiven Hinweisen zunächst bestimmen, welchen Hinweis er überhaupt geben möchte. Daraufhin aktiviert er den auditiven Hinweis durch die Auswahl, welcher der MWB den Hinweis erhalten soll. Einige Operateure scheinen diese Allokationsauswahl zu vergessen. In welchen Situationen das Ausbleiben des notwendigen zweiten Klicks erfolgte, wie und ob dies bemerkt wurde, wird aktuell anhand von Daten der Blick- und Mausbewegung ausgewertet (Stade, in Vorbereitung-b). Da der Eingriffsbereich bei geöffnetem Allokationsfenster weiterhin aktiv bleibt, können z.B. andere auditive Hinweise gewählt und dann die Allokation bestimmt werden. Eine andere Lösung, um das Allokationsfenster zu deaktivieren, bestünde im Abwarten einer Situation auf der Fahrstrecke, für welche der bereits ausgewählte Hinweis ebenfalls sinnvoll erscheint.

Eine große Zeitspanne zwischen dem ersten und zweiten Klick bei der Auswahl von auditiven Hinweisen ist aber nicht zwangsläufig ein Indikator für eine vergessene Auswahl der Allokation. So wurden Operateure beobachtet, deren Strategie darin bestand, erst den Hinweis auszuwählen, den Prozessverlauf weiter zu verfolgen und dann gezielt den Hinweis abzuschicken. Am Beispiel des Loben-Buttons soll verdeutlicht werden, welche Idee hinter dieser Strategie der Operateure stecken könnte. Möchte der Operateur das Fahrverhalten der MWB positiv bewerten, so kann er das Bewältigen der Situation abwarten und dann gezielt denjenigen MWB loben, der aus Sicht des Operateurs besonders zur Bewältigung der Situation beigetragen hat. Aufgrund weiterer geplanter Eingriffe unmittelbar nach dieser Situation kann dieses Vorgehen ressourcensparend für den Operateur sein. Gezieltes Loben kann aber nicht nur zur zeitnahen Spezifikation des Adressaten genutzt werden, sondern sich auch auf den gerade durchfahrenen Streckenabschnitt beziehen. Durch die beschriebene Strategie der

verzögerte Allokationsauswahl kann der Operateur ein positives Feedback an den/die MWB senden, ohne dabei zu riskieren, dass das Feedback verzögert zur Situation oder zum Verhalten des/der MWB registriert und ggf. falsch interpretiert wird. Für eine Diskussion der Verwendung von Lob und mögliche dahinter liegende Strategien des Operateurs verweisen wir ebenfalls auf Stade (in Vorbereitung-b).

Offensichtlich bereitet die zweistufige Auswahl von auditiven Hinweisen einigen Operateuren Schwierigkeit, anderen dient diese Form der Aktivierung des Hinweises möglicherweise als Strategie zur Verbesserung ihrer Aufgabebearbeitung. Zeigt sich dieser Unterschied auch in der Bewertung und Anwendung im Vergleich von visuellen und auditiven Hinweisen?

Nachfolgende Analysen beziehen sich auf eine Stichprobe von 40 Operateuren (40% Männer, 60% Frauen, Alter zwischen 18 und 34 Jahren,  $M=26.45$ ,  $SD=4.09$ ), die nicht von Ausreißern bereinigt wurde. Bei der Bewertung der Sinnhaftigkeit der AMD-Bereiche hinsichtlich der Erreichung ihres Ziels (Erzeugung maximaler Genauigkeit bei gleichzeitig maximaler Geschwindigkeit des Fahrobjektes) schätzten Operateure auf einer Prozentskala von 0 bis 100 visuelle Hinweise höher ein ( $Mdn=54.79$ ) als auditive Hinweise ( $Mdn=24.85$ ),  $U=189$ ,  $p=.000$ ,  $r=.92$ . Auch wurden visuelle Hinweise häufiger genutzt ( $Mdn=52.58$ ) als auditive Hinweise ( $Mdn=28.43$ ),  $U=317$ ,  $p=.000$ ,  $r=-.74$ . Ein direkter Vergleich ist jedoch nur bedingt möglich, da lediglich Richtungs- und Geschwindigkeitshinweise sowohl visuell als auch auditiv gegeben werden können (vgl. Bild 2). Vergleicht man Richtungs- und Geschwindigkeitshinweise, so gibt es für Geschwindigkeit keine signifikanten Unterschiede im Rating zwischen visuellen ( $Mdn=42.13$ ) und auditiven ( $Mdn=38.88$ ) Hinweisen,  $U=735$ ,  $p=.526$ ,  $r=-0.10$  sowie keine signifikanten Unterschiede in der Anzahl der Klicks zwischen visuellen ( $Mdn=38.92$ ) und auditiven ( $Mdn=42.08$ ) Hinweisen,  $U=737$ ,  $p=.542$ ,  $r=-0.10$ . Hingegen wurden visuelle Richtungsvorgaben signifikant höher bewertet ( $Mdn=46.70$ ) als auditive ( $Mdn=34.30$ ),  $U=552$ ,  $p=.016$ ,  $r=-.38$ . Außerdem wurden Aufforderungen, schneller oder langsamer zu fahren, signifikant häufiger genutzt über visuelle Hinweise ( $Mdn=49.03$ ) als über das auditive Pendant ( $Mdn=31.98$ ),  $U=459$ ,  $p=.001$ ,  $r=-.53$ .

Eine mögliche Erklärung für die höhere Bewertung und Nutzung visueller Richtungshinweise ist, dass Richtungshinweise häufig in Kombination mit Warnungen gegeben werden. Bei einem visuellen Richtungshinweis ist der Aufwand für den Operateur geringer, da dieser sich in räumlicher Nähe zum Warnungs-Button für ein Hindernis oder eine Gabelung befindet. Die zurückzulegende Distanz mit der Maus vom Warnungs-Button zum Richtungshinweis ist also für visuelle Richtungshinweise geringer als für auditive. Außerdem entfällt bei der Verwendung von visuellen Hinweisen der oben beschriebene zweistufige Auswahlprozess bei auditiven Hinweisen, was insbesondere in Gefahrensituationen mit Hindernissen und Gabelungen zeitkritisch ist. Gleichzeitig muss

der Operateur aber die systembedingte Inaktivität von 1.500 ms zwischen der Auswahl von zwei visuellen Hinweisen berücksichtigen (vgl. Abschnitt 3.3.1).

Betrachtet man hingegen nur ausgewählte Abschnitte der Strecke, dann hebt sich die bevorzugte Verwendung von visuellen Richtungs- und Geschwindigkeits-hinweisen auf. So zeigt sich in normalen Situationen, das heißt Situationen ohne Hindernisse und Gabelungen, kein Unterschied in der Verwendung von visuellen und auditiven Hinweisen, weder bei Geschwindigkeitsvorgaben ( $Mdn=38.21$  für visuelle Hinweise,  $Mdn=42.79$  für auditive Hinweise,  $U=709$ ,  $p=.373$ ,  $r=-.14$ ) noch bei Richtungsangaben ( $Mdn=42.20$  für visuelle Hinweise,  $Mdn=38.80$  für auditive Hinweise,  $U=732$ ,  $p=.41$ ,  $r=-.13$ ).

Erklärt werden kann dieser Befund dadurch, dass normale Streckenabschnitte dem Operateur die Gelegenheit bieten, das Fahrverhalten der MWB an sein übergeordnetes Ziel, maximale Genauigkeit und Geschwindigkeit zu erreichen, anzupassen, indem er gezielt Hinweise nur an einen der beiden MWB adressiert. Ob diese Differenzierung in der Allokation zu einer gesteigerten Leistung des Systems führt und welche Eingriffe durch Operateure auf normalen Streckenabschnitten sinnvoll sind, wird bei Stade (in Vorbereitung-b) ausführlich diskutiert. Hierbei wird insbesondere der Effekt von Richtungshinweisen in normalen Situationen auf die Leistung der MWB untersucht.

### 3.3.3 Einseitige Nutzung des Informations- und Funktionsangebots des AMD – eine angepasste Strategie des Operateurs?

*Verschiedenartigkeit der Eingriffe:* Große Varianz gibt es in der Bewertung und Anwendung der harten Eingriffe, also solchen am Fahrobjekt durch den Operateur (vgl. Bild 2, Bereiche 14 bis 16). Einige Operateure nutzten dieses Funktionsangebot gar nicht oder sehr selten, andere Operateure hingegen häufig ( $N=40$ ,  $M=26.1$ ,  $SD=56.3$ ,  $Min=0$ ,  $Max=326$ ; Anzahl der Klicks). Erste Analysen der Blickbewegungsdaten zeigen ebenfalls, dass bestimmte Bereiche des AMD, insbesondere harte Eingriffe, von einigen Operateuren ignoriert wurden. Aufschluss hierüber sollen anstehende Analysen geben, die eingeschätzte Sinnhaftigkeit, Verwendung harter Eingriffe sowie Blickbewegungsparameter gegenüberstellen. Von den Versuchsleitenden beobachtet sowie von den Operateuren in der Nachbefragung rückgemeldet ist das Vergessen der eingestellten Werte bei harten Eingriffe, z.B. das Aufheben von Richtungsbeschränkungen (vgl. Bild 2, Bereich 15).

Aktuelle Analysen sollen zeigen, ob nicht nur die Anzahl von Eingriffen der Operateure, unterteilt in verschiedenen Situationen, sondern auch die Variation des Operateurs in der Art der Eingriffe Indikatoren für die Leistung der MWB als Subsystem sind.

Trotz intensiven Trainings der Operateure mit theoretischen und praktischen Anteilen wurde von vielen Operateuren die Supervision und Kontrolle der MWB (vgl. Abschnitt 1.3) im Erstversuch als herausfordernd beschrieben. Im Vergleich über die Fahrten bezogen auf die Anzahl und Verschiedenheit der Eingriffe soll in

anstehenden Analysen gezeigt werden, ob die Operateure ihre Strategien zur Lösung ihrer Aufgabe anpassen. Um die eigene Anstrengung und die der MWB möglichst gering zu halten, könnte eine Strategie des Operateurs darin bestehen, sich auf wenige Eingriffe, ggf. sogar des gleichen Typs (vgl. Abschnitt 3.3.2) zu beschränken.

*Nutzung der Streckenvorschau:* Ein Mehrwert für die MWB besteht in der Vorausschau des Streckenverlaufs durch den Operateur. Hierdurch erhält der Operateur die Möglichkeit, die MWB rechtzeitig vor schwierigen Situationen zu warnen und sie zu instruieren. In ersten qualitativen Analysen der Blickbewegungsdaten konnten drei Typen von Operateuren identifiziert werden: 1) Operateure, die hauptsächlich den für die MWB sichtbaren Streckenverlauf verfolgen, 2) Operateure, welche die Streckenvorschau nur kurz oberhalb der aktuellen Streckenansicht betrachten, 3) Operateure, welche die gesamte Streckenvorschau bis zum oberen Rand des Displays nutzen. Ferner unterscheiden sich die Operateure darin, wie sinnvoll sie die Streckenansicht für die Bewältigung ihrer Aufgabe einschätzen. Ob es einen Zusammenhang zwischen der Intensität in der Nutzung der Streckenvorschau und der Leistung der Operateure gibt, wird noch untersucht. Eine Verbindung zu den Analogien von Nachtwei (2010) wird gezogen, die die Stärken und Schwächen von Operateuren im Umgang mit komplexen dynamischen Systemen verdeutlichen.

Fokussiert ein Operateur die aktuelle Ansicht der Strecke, so muss diese verengte Sicht per se keine schlechte Strategie des Operateurs sein. Zum einen besteht die Möglichkeit, dass Operateure z.B. rot eingefärbte Hindernisse peripher wahrnehmen können. Zum anderen wiederholt sich der Streckenverlauf, sodass ein Lerneffekt nicht ausgeschlossen werden kann, nachdem der Operateur den Streckenverlauf verinnerlicht hat. Eine Fokussierung auf die aktuelle Streckenansicht und damit auf das Fahrverhalten der MWB kann ebenfalls Indikator dafür sein, dass der Operateur den Prozess genau verfolgt, um adäquate Handlungsanweisungen für die MWB zu geben (vgl. Abschnitt 3.3.2), was eine gute Strategie wäre.

#### **4 Fazit für einen (möglichen) Ausbau des AMD**

In der Nachbefragung äußerten Operateure Verbesserungsvorschläge für das Informations- und Funktionsangebot sowie die Gestaltung des AMDs. Es wurde z.B. vorgeschlagen, das zweistufige Auswählen bei auditiven Hinweisen durch mehrere Buttons zu ersetzen, die für jeden auditiven Hinweis zwischen der Allokation MWB 1, MWB 2 und MWB 1+2 unterscheiden. Ob dies zu einer Vereinfachung der Auswahl von auditiven Hinweisen führt, ist fraglich unter der Annahme, dass mehr Buttons auch mehr Suchaufwand erfordern und aufgrund der Größe des AMD zu kleineren Buttons führen würde.

Ebenfalls als Verbesserung von den Operateuren angeregt wurde, einen kombinierten Button aus Warnung vor Gabelungen und Richtungsvorgaben zu entwerfen. Aber auch hier ist zu diskutieren, ob eine Ein-Button-Lösung zu

präferieren ist. Zwar würde die Mausbewegung zwischen dem Button Warnung vor Gabelungen und Richtungsvorgabe sowie ein Klick wegfallen, jedoch gibt es Operateure, deren Strategie in Gabelungssituationen darin besteht, nur vor der Gabelung zu warnen. Dies kann zwei Ursachen haben: a) Die Operateure möchten eine Richtungsvorgabe geben, verpassen dies aber oder b) sie haben erkannt, dass die MWB z.B. stets die linke Abzweigung wählen und sehen daher keine Notwendigkeit für ein Eingreifen. Die Analyse von Kombination, Wirkung und Ursache von Eingriffen von Operateuren kann also aufschlussreich sein für die Identifikation von guten und schlechten Operateuren und würde durch eine Ein-Button-Lösung eingeschränkt werden.

Eine Vielzahl weiterer Informationen und Funktionen für die Operateure über die von ihnen genannten Vorschläge hinaus sind denkbar. Jedoch muss berücksichtigt werden, dass, wie an den beiden genannten Vorschlägen verdeutlicht wurde, hierbei womöglich (andere) Gestaltungsmängel auftreten können und Freiheitsgrade für den Operateur eingeschränkt werden. Diese eingangs beschriebene Vielzahl an Möglichkeiten, die Aufgabe der Supervision und Kontrolle des Trackingprozesses zu bewältigen, ist aber Grundlage für eine Variabilität in der Leistung der Operateure und damit für die Diagnostik von guten und schlechten Operateuren essentiell.

## **5 Fazit über die interdisziplinäre Arbeit als Herausforderung und Bereicherung**

Der in Abschnitt 2 beschriebene Entwicklungsprozess des Masterdisplays war geprägt durch interdisziplinäre Arbeiten. Herausforderung hierfür war das Sprechen einer gemeinsamen Sprache für das Formulieren von Anforderungen an das AMD sowie die konkrete Umsetzung in das Informations- und Funktionsangebot. Als hilfreich und als Basis für den gesamten Entwicklungsprozess dienend war die bei Hildebrandt et al. (2010) beschriebene Hierarchische Aufgabenanalyse. Eine weitere Stärke des Projektes bestand und besteht im regen und regelmäßigen Austausch der Beteiligten beider Disziplinen. Das gegenseitige Vermitteln von Wissen über Möglichkeiten und Einschränkungen auf Seiten der Technik unter Berücksichtigung der zur Verfügung stehenden Ressourcen als auch über notwendige Voraussetzungen für eine optimale Interaktion des Operateurs mit dem AMD aus Sicht der Psychologie war und ist produktiv für die Entwicklung und Evaluation. Obgleich nicht alle Ideen zum Funktions- und Informationsangebot sowie dem Design in den ersten Versionen des AMDs umgesetzt werden konnten, wurden sie keineswegs verworfen, sondern gesammelt und bieten damit eine kreative Grundlage für mögliche weitere Ausbaustufen und damit verbundene Forschungsfragen, sowohl für die Analyse des AMDs als auch das Verhalten der Operateure.



## Literatur

- Gérard, N., Huber, S., Nachtwei, J., Satriadarma, B. & Schubert, U. (2011). A Framework for Designers to Support Prospective Design of Human Computer Interaction. *International Journal on Human-Computer Interaction*, 2(7), 17–38.
- Hildebrandt, M., Kain, S., Kesselring, K., Nachtwei, J., Niestroj, N. & Schwarz, H. (2010). Die Hierarchische Aufgabenanalyse im Babel interdisziplinärer Softwareentwicklungsprojekte. *Zeitschrift für Arbeitswissenschaft*, 4, 374–380.
- Nachtwei, J. (2010). The many faces of human operators in process control – A framework of analogies. *Theoretical Issues in Ergonomics Science*, iFirst, 1-21. Retrieved November 25, 2010 from <http://dx.doi.org/10.1080/14639221003728609>.
- Nachtwei, J. (submitted-a). *A multi-level design approach for a supervisory control master display in human factors experiments*. Manuscript submitted for publication.
- Nachtwei, J. (submitted-b). *SAGAT vs. SPAM? It depends! An Aid for Situation Awareness Method Selection*. Manuscript submitted for publication.
- Stade, M.J.C. (in Vorbereitung-a). *Behind and through the eyes of human operators in process control: Chances and constraints of eye movement analysis*. Manuskript in Vorbereitung.
- Stade, M.J.C. (in Vorbereitung-b). *Skilled human operators in process control: Strategies they found and strategies to find them*. Manuskript in Vorbereitung.
- Wandke, H. & Nachtwei, J. (2008). The different human factor in automation: the developer behind vs. the operator in action. In D. de Waard, F.O. Flemisch, B. Lorenz, H. Oberheid & K.A. Brookhuis (Eds.), *Human factors for assistance and automation* (pp. 493–502). Maastricht, the Netherlands: Shaker Publishing.

## Autoren

cand. Psych. M.J.C. Stade	Humboldt-Universität zu Berlin Ingenieurpsychologie/Kognitive Ergonomie
Dipl.-Psych. C. Meyer	Technische Universität Berlin
Dipl.-Inf. N. Niestroj	GRK prometei
Dipl.-Psych. J. Nachtwei	Humboldt-Universität zu Berlin Sozial- und Organisationspsychologie

Kontakt:

[melanie.stade@student.hu-berlin.de](mailto:melanie.stade@student.hu-berlin.de)