

Modellierung und Vorhersage von mentaler Arbeitsbeanspruchung in einem Fluglotsenaufgabenexperiment

Martina Truschzinski¹

Abstract: Diese Veröffentlichung präsentiert die Ergebnisse einer dynamische Modellierung von Workload, welche auf der Basis von Pupillendaten eines psychologischen Experimentes entwickelt wurde. Es zeigt, wie kognitive Verarbeitungsprozesse mit Hilfe von mathematischen Funktionen innerhalb eines dynamischen Systems abgebildet und validiert werden können.

Keywords: Workload; Dynamisches System; Fluglotsenaufgabe; Pupillometrie

1 Motivation

Derzeitige Statistiken zeigen, dass der Flugverkehr und die zu transportierenden Fluggästepersonenanzahl steigt und auch in den nächsten Jahren steigen wird [Bo17]. Um einen reibungslosen und sicheren Flugverkehr zu garantieren, sind Fluglotsen, welche den gesamten Flugraum tagtäglich überwachen, und deren mentale Verfassung von enormer Bedeutung. Aus diesem Grund beschäftigt sich das BMBF geförderte interdisziplinäre Projekt "StayCentered: Methodenbasis eines Assistenzsystems für Centerlotsen" (MACeLot) in Zusammenarbeit mit der Deutschen Flugsicherung mit der Untersuchung von grundlegenden Methoden und Verfahren, die für die Realisierung eines Assistenzsystems für Fluglotsen verwendet werden können.

Kognitive Aufgaben in einer Fluglotsentätigkeit, sind sehr komplexe und kognitiv belastende Aufgaben, weil sie sehr viele mentale Ressourcen benötigen [Mo95]. Dabei beeinflussen vor allem Faktoren wie Flugverkehrsaufkommen oder Frequenzüberlastungen die mentale Arbeitsbeanspruchung der Fluglotsen [Mo95]. Des Weiteren gilt im Allgemeinen, dass vor allem Faktoren, wie Schwierigkeitsgrad der Aufgabe und Zeitdruck der beim Lösen der Aufgabe entsteht, die mentale Arbeitsbeanspruchung maßgeblich verändern kann [GCM12]. Demzufolge kann eine Erhöhung des Schwierigkeitsgrades einer zu lösenden Aufgabe, zu einer Erhöhung der erforderlichen mentalen Ressourcen führen und daraus auch eine erhöhte mentale Arbeitsbeanspruchung entstehen. Eine solche entstandene mentale Arbeitsbeanspruchung kann aufgrund dessen, dass mentale Kapazitäten zur Informationsverarbeitung

¹ Technische Universität Chemnitz, Professur Prozessautomatisierung, Reichenhainerstr. 70, 09126 Chemnitz, Deutschland mtru@hrz.tu-chemnitz.de

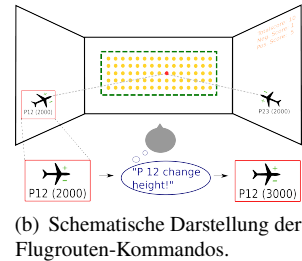
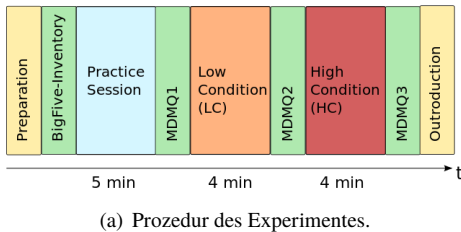


Abb. 1: Experimentalaufbau und Aufgabenstellung der Versuchspersonen.

beim Menschen begrenzt sind, anhand von verbrauchten oder noch vorhandenen mentalen Kapazitäten berechnet werden [GD86, Wi02].

Um die vorhandene mentale Beanspruchung während einer Aufgabe zu messen, wurde innerhalb des Projektes das Verfahren von Beatty und Lucero-Wagner verwendet. Diese entdeckten im Jahr 2000, dass nicht-reflexive phasische Pupillenbewegungen als Indikator für Gehirnprozesse gesehen werden können, da sie sich nachweislich bei dynamischen und beanspruchungsintensiven Aufgaben korrespondierend zum Schwierigkeitsgrad veränderten. Diese aufgabenbezogene Veränderungen der Pupille (TERPs), wurden in den folgenden Jahren von vielen Autoren zur Untersuchung kognitiver Funktionen verwendet [Ha10, Pa03].

2 Experiment

Mit Hilfe eines psychologischen Experimentes sollte anhand der aufgezeichneten Pupillendaten ein Workload-Modell erstellt werden, welche die mentale Arbeitsbeanspruchung in Fluglotsenaufgaben mathematische beschreiben und vorhersagen soll. Dafür wurde innerhalb des Projektes eine Simulation erstellt, welche es Novizen ermöglichte Flugzeuge auf einem Radar-ähnlichen Bildschirm zu steuern. An dem Experiment nahmen 25 Studenten an dem Experiment teil. Diese wurden auf dem Campusvorplatz der Universität Pompeu Fabra gesucht und nahmen freiwillig an der Untersuchung teil, ($M_{Age} = 28.12$; $SD = 5.67$, 64% männlich). Die Mehrheit der Befragten hatte keine Erfahrungen mit Fluglotsentätigkeiten, welche auch Computerspiele enthielten.

Experiment Design Das Experiment umfasste zwei verschiedenen Aufgabenschwierigkeiten, welches in intra-individuellen Prüfungsdesign über alle Versuchspersonen, ausgewertet wurde. Dabei war jedes Experiment in drei Phasen geteilt: der Übungsphase, in welcher die Versuchspersonen lernten die Flugzeuge auf dem Bildschirmen zu steuern, der ersten Phase, welche eine leichte Version der Fluglotsenaufgabe beinhaltete und der zweite Phase, welche die schwierige Version der Fluglotsenaufgabe beinhaltete (siehe Abb. 1(a)).

Um die von Beatty und Lucero-Wagner [BLW00] berichteten Vergrößerung der Pupille

aufzuzeichnen, wurde während des gesamten Experimentes die Pupillenerweiterung mit einem mobilen Eye-Tracker von *Pupil Labs* mit einer Frequenz von 60 Hz aufgezeichnet.

Prozedur Das Experiment wurde in einem Virtual-Reality-Raum durchgeführt (XIM). Nachdem jede Versuchsperson die Einverständniserklärung zur Teilnahme an dem Experiment ausgefüllt hatte, wurde sie in den Virtual-Reality-Raum geführt. In einer kurzen Vorbereitungsphase wurde der Versuchsperson die Eye-Tracker-Brille aufgesetzt, der Persönlichkeitsfragebogen ausgefüllt und eine Erklärung zum Experiment gegeben. Danach startete eine Kalibrierungsphase, die sicher stellte, dass die Beleuchtungssituation während des gesamten Experimentes gleichbleibend war. Anschließend startete die Übungsphase, in welcher die Versuchsperson lernte die Flugzeuge auf den Bildschirm mit seiner Stimme zu steuern. Die Flugzeuge wurden – wie reale Fluglotsen – mit vordefinierten Kommandos gesteuert. Es gab dabei zwei verschiedene Möglichkeiten die Flugroute zu ändern zum einen konnte die Höhe der Flugzeuge geändert: entweder eine Erhöhung von 2000 auf 3000 oder eine Verringerung der Flughöhe von 3000 auf 2000. Zum anderen konnte die Flugrichtung um 45° gegen oder mit dem Uhrzeigersinn gewechselt werden. Die Versuchspersonen waren angehalten ihre Kommandos klar und deutlich in englische Sprache zu tätigen (siehe 1(b)). Der Experimentleiter/in gab anschließend mittels Tastatur – als simulierter Pilot – die gegebene Flugroutenänderung in das System ein und die Flugzeuge änderten ihren Flugweg. Die Übungssession endete, wenn die Versuchsperson 10 korrekte Flugroutenänderungen durchgeführt hatte.

Im Anschluss folgte das eigentliche Experiment, in welchem in einer statischen Reihenfolge eine leichte und schwere Aufgabe zu lösen waren. Der Schwierigkeitsgrad der Aufgabe wurde durch die Frequenz gesteuert, mit welcher die Flugzeuge auf dem Radar erschienen. Dabei erschienen die Flugzeuge paarweise, wobei in der leichten Aufgabe alle 4–8 Sekunden und in der schwierigen Aufgabe alle 1–5 Sekunden ein neues Flugzeugpaar am Rand des Bildschirms auftauchte und in den für die Versuchspersonen verantwortlichen Flugbereich steuerte (grün umrandetes Rechteck in Abb. 1(b)). Das schnellere Erscheinen der Flugzeuge erzeugte in der schwierigen Aufgabe einen erhöhten Zeitdruck, der zum einen durch die höhere Anzahl wahrgenommener und mental verarbeiteter Flugzeuge entstand und zum anderen durch weitaus komplexere Flugsituation die einen erhöhten Bedarf an Flugroutenänderungen nach sich gezogen hat. Die Versuchspersonen hatten die Aufgabe, während des gesamten Experimentes, Kollisionen der Flugzeuge zu vermeiden indem sie die Flugrouten mit den in der Übungsaufgabe gelernten Kommandos veränderten. Eine solche fluglotsenähnliche Aufgabe ist insofern sehr anspruchsvoll, da die Darstellung der Höhe nur in einer 2,5D-Darstellung erfolgt. Diese Darstellung entsteht, weil die dritte Dimension, die Höhe, nur durch die Nummer in der Klammer unter den Flugzeugen (siehe Abb. 1(b)) dargestellt wurde. Aus diesem Grund war es möglich, dass zwei Flugzeuge visuell betrachtet kollidieren - das heißt auf dem selben Punkt auf dem Bildschirm erscheinen, aber aufgrund der unterschiedlichen Höhe lediglich übereinander hinweg flogen. Es musste also jeder Zeit eine mentale Repräsentation über die Höhen im Kopf behalten werden, damit eine reibungslose Flugzeugkoordination stattfinden konnte.

3 Workloadmodell

Die während des oben beschriebenen Experimentes aufgezeichneten Pupillendaten wurden anschließend für die Entwicklung eines mathematischen Modells verwendet. Hierzu mussten die Daten zunächst vor-verarbeitet werden. Hierzu wurden verschiedene Matlab-Funktionen, wie ein Medianfilter, benutzt. Dieser konnte aufgrund von fehlerhaften Daten nur auf 21 von 25 Datensätzen angewendet werden. Um das Modell aufzubauen und es anschließend zu validieren, wurden diese Datensätze in einen Modellierungsdatensatz - 16 Versuchspersonen - und einen Validierungsdatensatz - 5 Versuchspersonen - aufgeteilt. Diese Datensätze bildeten die Grundlage für die folgende Modellierung und Kreuzvalidierung des Modells. Die Hypothese, dass der unterschiedliche Schwierigkeitsgrad der beiden Experimentphasen zu einem unterschiedlichen Workload führt, welcher mit der Pupillengröße gemessen werden kann, konnte statistisch nachgewiesen werden. Ein durchgeführter paarweiser t-Test zeigt, dass die mittlere Pupillenerweiterung in der schweren Aufgabe des Experimentes signifikant größer war als in der leichten Aufgabe ($t(219) = 3.73, p < .001, d = 0.83$). Da dieser Unterschied auch ein Resultat der statischen Reihenfolge sein kann, wurde mit einem paarweisen t-Test auf Unterschiede der mittleren Pupillenerweiterung zu Beginn der Aufgaben getestet. Dafür wurde ein 1s-Fenster, in welchen die Versuchspersonen noch keine Flugzeuge sehen konnten, ausgewählt. Es zeigte sich hierbei kein signifikanter Unterschied in der Pupillengröße zu Beginn der leichten und schweren Aufgabe ($t(21) = -1.92, p > .05$). Dies zeigt, dass der gefundene Unterschied des Workloads in den zwei Aufgaben, wie vermutet, ein Resultat des unterschiedlichen Schwierigkeitsgrades der Aufgaben ist. Die dynamische Modellierung baut auf diesen statistisch gefundenen Effekt auf. Dabei wurde auf der Grundlage der aufgezeichneten Pupillen-Daten mithilfe der Matlab System-Identification Toolbox eine mathematische Beschreibung des Zusammenhangs zwischen Aufgabe und gemessenen Workload/ Pupillengröße gesucht.

Modellbeschreibung In der Signaltheorie werden dynamische Systeme unter anderem als Multiple-Input-Single-Output-Modell (MISO) beschrieben. Im Falle der in dem Experiment aufgezeichneten Daten bedeutete dies, dass die Pupillenerweiterung als Messparameter für den Workload der Output bzw. die Ausgangsgröße war. Diese Ausgangsgröße wurde durch verschiedene Aktionen oder Ereignisse der Aufgabe verändert. Aufgrund dessen wurden Ereignisse, wie Erscheinen von Flugzeugen auf dem Radar, Eingaben von Flugroutenänderungen, Auftreten von Kollisionen, als Eingaben innerhalb des Modells formuliert (für eine detaillierte Darstellung des Modells siehe [TW]).

Anhand der einzelnen Ereignisse wurden zunächst temporäre Veränderungen in der Pupillenerweiterung untersucht und diese als Grundlage für die mathematische Beschreibung der Pupillenerweiterung in der gesamten Aufgabe verwendet. Das Ergebnis des Modells ist in Abbildung dargestellt. Die rote Linie repräsentiert die Berechnung des Workloads vom Modell und die schwarze Linie die tatsächlich aufgezeichnete Pupillenerweiterung. Höhere Werte der Pupillengröße traten meistens innerhalb der gesprochenen Kommandos auf, da

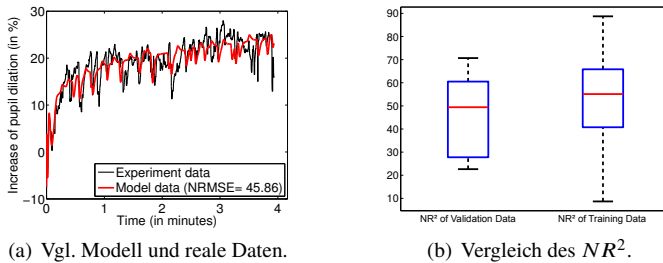


Abb. 2: Ergebnisse der Kreuzvalidierung des Modells.

hier wie vermutet die meisten mentalen Ressourcen zur Auflösung möglicher Kollisionssituationen benötigt wurden. In solchen Situationen musste die Versuchsperson die Flugroute der Flugzeuge vorausberechnen und entscheiden welches Flugzeuge wohin geleitet werden sollte. Verringerungen im Workload wurden zumeist nach Kollisionen verzeichnet. Dies war insofern überraschend, da eine Annahme war, dass nach einer Kollision die Versuchsperson über mögliche Ursachen die zu dieser Situation geführt hatte nachdenkt. Dies hätte zu einem Anstieg der mentalen Arbeitslast führen müssen. Doch im Gegensatz dazu, hatten die Versuchspersonen wenig oder gar keine Zeit darüber nachzudenken und es blieb nur der Effekt, dass nach einer Kollision weniger Flugzeuge auf dem Bildschirm waren und somit weniger Flugzeuge im Gedächtnis behalten werden mussten, welches wiederum zu einer Freigabe von mentalen Ressourcen führt und der zu verzeichnende Workload sinkt.

Kreuzvalidierung Um die Güte der mit Matlab gefundenen Beschreibung des Workloadmodells zu ermitteln wurde zunächst für jede Versuchsperson des Modellierungsdatensatzes eine Modellvorhersage getroffen und zwischen Modell und realen Daten der normalisierte Determinationskoeffizienten (NR^2) berechnet. Dieser gibt an wie viel Varianz in den Pupillendaten durch das vorliegende dynamische Workloadmodell erklärt werden kann. Im Mittel liegt der NR^2 beim Modelldatensatz bei 49.42% und beim Validierungsdatensatz, welcher für das Modell komplett neue Datensätze enthält, liegt der NR^2 bei 55.10 %. Im Vergleich zu technischen Systemen, ist der so berechnete Prognosefehler des Modells relativ hoch, allerdings ist zu beachten dass das Modell nur anhand der verzeichneten Ereignisse eine Berechnung des Workloads tätigen kann. Alle anderen Denkprozesse, die kein Kommando oder Kollision zur Folge hatten, kann das Modell nicht abbilden, da es keinen Zugriff darauf hat. Aus diesem Grund stellt ein Modell dieser Güte einen guten Fortschritt für die Untersuchung von Workloadprozessen dar. Allerdings zeigt es auch, dass bisher noch unerklärte Prozesse die Pupillengröße bzw. den Workload verändern, welche im weiteren Forschungsarbeiten untersucht werden sollten.

4 Fazit

Das in dieser Veröffentlichung vorgestellte dynamische Modell für die Modellierung und Vorhersage von Workload wurde anhand einer Simulation von Fluglotsentätigkeiten mit Novizen erstellt und getestet. Sowohl die Modellierung als auch die vorangegangene statistische Analyse bestätigte die Hypothese, dass eine Erhöhung des Schwierigkeitsgrades einen erhöhten Workload zur Folge hat. Mit Hilfe von aufgezeichneten Ereignissen, wie Kollisionen und Flugroutenkommandos, konnte ein mathematisches Modell erstellt werden, welches diese Erhöhung des Workloads innerhalb der Aufgabe vorhersagen und beschreiben kann. Dabei zeigte die Kreuzvalidierung, dass 49,4 – 55 % der Varianz aufgeklärt wurde. In der Betrachtung der Entwicklung des Workloads innerhalb der einzelnen Versuchspersonen zeigte sich, dass der Zusammenhang zwischen Workload und Schwierigkeit über die Zeit allerdings nicht linear sondern logarithmisch ansteigt. Diese Erkenntnis ist insofern für weitere statistische Analysen hilfreich, als dass das gewählte Analyseverfahren angepasst bzw. die entsprechenden Daten vorher entsprechend vor-verarbeitet werden können. Des Weiteren bestätigt die Analyse des Modells die von Beatty und Lucero-Wagoner getroffene Aussage, dass visuelle Stimuli TERPs hervorrufen. Allerdings bedeuten diese im Detail, nicht immer eine Erhöhung der mentalen Arbeitsbeanspruchung, wie in Beatty und Lucero-Wagoner (2000) dargestellt, vielmehr kann ein visueller Stimuli auch eine Verminderung der mentalen Arbeitsbeanspruchung hervorrufen, wie in dem Beispiel der Kollisionen gezeigt.

Literaturverzeichnis

- [BLW00] Beatty, J.; Lucero-Wagoner, B.: The pupillary system. In (Cacioppo, John T.; Tassinari, Louis G.; Berntson, Gary G., Hrsg.): *Handbook of Psychophysiology*. Cambridge University Press, Cambridge, 2nd. Auflage, 2000.
- [Bo17] Estimated annual growth rates for passenger and cargo air traffic from 2016 to 2035, by region. In Statista - The Statistics Portal.
- [GCM12] Galy, E.; Cariou, M.; Mélan, C.: What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology*, 83:269–275, 2012.
- [GD86] Gopher, D.; Donchin, E.: Workload - An examination of the concept. In (Boff, K. R.; Kaufmann, L.; Thomas, J. P., Hrsg.): *Handbook of Perception and Human Performance*. Vol. II. Cognitive Processes and Performance. Wiley & Sons, New York, 1986.
- [Ha10] Haapalainen, E.; Kim, S.; Forlizzi, J. F.; Dey, A. K.: *Psycho-Physiological Measures for Assessing Cognitive Load*. Copenhagen, 2010.
- [Mo95] Mogford, R. H.; Guttman, J. A.; Morrow, S. L.; Kopardekar, P.: The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature. DTIC Document, 1995.
- [Pa03] Paas, F.; Tuovinen, J. E.; Tabbers, H.; Van Gerven, P. W. M.: Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, 38(1), 2003.
- [TW] Truschzinski, M.; Wirzberger, M.: A dynamic process model for predicting workload in an air traffic controller task. (in press).
- [Wi02] Wickens, C. D.: Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3:159–177, 2002.