



Analyse und Vergleich von Benutzerverhalten und Gehirnaktivität bei verschiedenen Versionen eines Memory-Spiels zum Test adaptiver Hilfestellungen

Bachelorarbeit am Cognitive Systems Lab
Prof. Dr.-Ing. Tanja Schultz
Fachbereich 3
Universität Bremen

von

Markus Ihrig

Betreuer:

Felix Putze, Mazen Salous

Gutachter:

Felix Putze
Johannes Schöning

Tag der Anmeldung: 7. Juni 2019
Tag der Abgabe: 1. Oktober 2019

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Bremen, den 1. Oktober 2019

Zusammenfassung

In dieser Arbeit wird eine App für das Spiel 'Memory' für die Android-Plattform um zwei Hindernisse kognitiver und visueller Art erweitert, um bei Probanden¹ eine Behinderung zu simulieren. Gleichzeitig werden zwei Assistenzen hinzugefügt, welche den jeweils hervorgerufenen Beeinträchtigungen entgegenwirken sollen. Die Auswirkungen der Hindernisse und Assistenzen werden experimentell durch Analyse von Fragebögen, Spieldaten sowie durch aufgenommene Hirnaktivität mittels EEG getestet. Es wird festgestellt, dass beide Hindernisse einen deutlichen negativen Einfluss auf die Spielerleistung und Spielerfahrung haben. Die visuelle Assistenz kann diesen Einfluss vollständig aufheben, die kognitive zum größten Teil nicht. Die Hirnaktivität wird jeweils im Delta- und im Alpha-Band untersucht und zeigt in ersterem signifikante Unterschiede in der Aktivität zwischen dem normalen Modus und einem Modus mit Hindernis. In Durchschnittsgraphen ist die Aktivität bei vorhandenen Hindernissen teilweise stärker ausgeprägt. Eine Klassifizierung der Modi anhand von Teilen der in der Arbeit aufgeführten Metriken erscheint möglich.

¹Ich verwende in dieser Arbeit das generische Maskulinum, um den Lesefluss nicht zu beeinflussen. Damit sind stets Personen jeglichen Geschlechts gemeint.

Inhaltsverzeichnis

1	Einleitung	1
2	Verwandte Arbeiten	3
3	Das Spiel	5
3.1	Bisherige Version	5
3.2	Kartenset	6
3.3	Hindernisse	7
3.4	Adaptionen	8
3.5	Spielmodi und Begrifflichkeit	9
4	Der Versuch	11
4.1	Versuchsaufbau	11
4.1.1	Versuchsablauf	12
4.1.2	Probanden	12
5	Auswertung der Verhaltensdaten	13
5.1	Vorgehensweise	13
5.2	Vergleich der Fragebögen	14
5.2.1	Einschätzung der drei Hauptmodi	14
5.2.2	Einschätzung der Assistenzen	15
5.3	Datenverarbeitung der Spieldaten	17
5.4	Vergleich der Spieldaten	17
5.4.1	Vergleich der Hauptmodi	17
5.4.2	Auswirkungen der Assistenzen	19
6	Auswertung der EEG-Daten	21
6.1	Biologische Grundlagen	21
6.1.1	Hirnareale	21
6.1.2	Gehirnaktivität	21
6.2	Hardware Setup	23
6.3	Mathematische Grundlagen	23
6.3.1	Diskrete Fouriertransformation	23
6.3.2	Hann-Fensterfunktion	25
6.4	Datenverarbeitung der EEG-Daten	25
6.5	Relevante Frequenzbereiche und Elektrodenpositionen aus der Literatur	26
6.6	Vergleich der EEG-Daten	28
6.7	Delta-Aktivität	28
6.7.1	Classic-Modus vs. Memory-Obstacle	28

6.7.2	Classic-Modus vs. Visual-Obstacle	30
6.7.3	Memory-Obstacle vs. Visual-Obstacle	30
6.8	Alpha-Aktivität	33
7	Ergebnisse	37
8	Fazit	39
8.1	Ausblick	40
A	Anhang	41
A.1	Versuchsdaten	41
A.2	Versuchsdokumente	41
A.3	Graphen bezüglich der Fragebögen	41
A.4	Zusätzliche Featuresuche	46
A.5	Skripte	47
	Literaturverzeichnis	49

Abbildungsverzeichnis

3.1	Normale Tierkarte	5
3.2	Tierkarte mit Visual-Obstacle	5
3.3	Normale Tierkarte	6
3.4	Visual-Obstacle mit Zoom	6
3.5	Alle Karten des Spektral-Sets	6
3.6	Alle Karten des Rot-Grün-Sets	7
3.7	Das Spektral-Set mit Visual-Obstacle	8
3.8	Das Rot-Grün-Set mit Visual-Obstacle	8
4.1	Spielmodi	11
5.1	Benötigte Zeit	18
5.2	Benötigte Züge	18
5.3	Fehler	18
5.4	Summe	19
6.1	Die 4 Lappen des Gehirns	22
6.2	Die empirische Einteilung von Hirnfrequenzen in Bänder	23
6.3	Elektrodenpositionen	24
6.4	Die Formen der Hamming- und Hann-Fensterfunktionen	25
6.5	Ausgeprägte Hirnregionen pro Band nach Kumar et al.	27
6.6	Delta-Aktivität Classic vs. Mem.-Obs.	28
6.7	Delta-Aktivität Beispiel	29
6.8	Delta t-Tests Classic vs. Mem.-Obs.	30
6.9	Delta t-Tests Classic vs. Vis.-Obs.	30
6.11	Delta t-Tests Mem.-Obs. vs. Vis.-Obs.	30
6.10	Delta-Aktivität Classic vs. Vis.-Obs.	31
6.12	Delta-Aktivität Mem.-Obs. vs. Vis.-Obs.	32
6.13	Alpha-Aktivität Classic vs. Mem.-Obs.	33
6.14	Alpha-Aktivität Classic vs. Vis.-Obs.	34
6.15	Alpha-Aktivität Mem.-Obs. vs. Vis.-Obs.	35
6.16	Ergebnisse der t-Tests im Alpha-Band	35
A.1	Beispielbild des Versuchsaufbaus	42
A.2	Die ersten Fragen des Fragebogens	43
A.3	weitere Fragen des Fragebogens	44
A.4	Die letzten Fragen des Fragebogens	45

Tabellenverzeichnis

5.1	Einschätzung der Spieler bezüglich ihrer Leistung	14
5.2	Einschätzung der Spieler bezüglich der Usability-Aspekte	15
5.3	Spielmodusspezifische Fragen	16
5.4	Die Leistungseinschätzung im Vergleich für den unassistierten und assistierten Memory-Obstacle-Modus	16
5.5	Die Leistungseinschätzung im Vergleich für den unassistierten und assistierten Visual-Obstacle-Modus	17
A.1	Elektroden-Scores im Delta-Band (1-4Hz)	46
A.2	Elektroden-Scores im Theta-Band (5-8Hz)	46
A.3	Elektroden-Scores im Alpha-Band (9-13Hz)	46
A.4	Elektroden-Scores im Beta-Band (14-30Hz)	46
A.5	Elektroden-Scores im Gamma-Band (31-40Hz)	46

1. Einleitung

Mit der voranschreitenden Digitalisierung steigt die Zahl der Computer und computerähnlichen Geräte in deutschen Haushalten immer weiter an [Bun18]. Infolge dieser steigenden Nutzung wird eine gute Mensch-Technik-Interaktion bei der Entwicklung von modernen Systemen immer wichtiger. Diese zeichnet sich neben Einfachheit, Intuitivität, Robustheit und Sicherheit auch durch den Aspekt der Barrierefreiheit aus. Durch die stetig steigende Lebenserwartung in Deutschland [vde19] verbringen mehr Menschen ihr Leben im Alter und leiden so potenziell länger an altersbedingter Sehschwäche und motorischen Einschränkungen. Zwar stellt diese Altersgruppe bisher keine Hauptzielgruppe von Computern dar, dies wird sich in Zukunft aber spätestens ändern, sobald die erste Generation, welche mit Rechnern aufgewachsen ist, dieses Alter erreichen.

Abgesehen von diesen altersbedingten Einschränkungen sollte eine Barrierefreiheit mit Blick in die Zukunft auch für individuellere Einschränkungen wie Taubheit, Farbenblindheit, Rechen- oder Konzentrationsschwächen erreicht werden. Für eine solche Anpassung können einerseits fundamentale Einschränkungen wie zum Beispiel Blindheit des Nutzers vor der eigentlichen Interaktion statisch vermittelt werden. Mit der Vielzahl an möglichen Einschränkungen ist aber besser, wenn sich das System dynamisch an den Nutzer anpasst, indem es diesen durch das Nutzerverhalten selbst klassifiziert.

Diese Problematik soll in dieser Arbeit anhand einer am CSL entwickelten Memory-App auf einem Android Tablet untersucht werden. Es wurde dieses Spiel gewählt, da es eine einfache Möglichkeit bietet, dem System einen Einblick in die visuellen und mnestischen Fähigkeiten des Nutzers zu geben. Zwei vielversprechende Metriken für die Klassifikation sind hierbei die objektiv erfassbare Leistung im Vergleich zum Durchschnitt, sowie die Hirnaktivität des Nutzers beim Spielen. Mit diesem Spiel wird daher eine Versuchreihe durchgeführt, bei welcher Probanden das Spiel spielen, Fragebögen diesbezüglich ausfüllen und gleichzeitig Spiellogs, sowie mittels EEG die Hirnaktivität der Probanden aufgezeichnet werden. Da es jedoch einerseits schwierig ist, genug Probanden mit oben genannten Einschränkungen in kurzer Zeit zu finden und andererseits die Bedingungen für eine wissenschaftlich haltbare Studie für alle Probanden gleich sein sollten, werden die Einschränkungen durch zwei im Spiel vorliegende Hindernisse simuliert. Diese Hindernisse sind visueller und

gedächtnisbasierter Natur und sollen die Spieler in diesen Bereichen so besonders fordern. Als Gegenmaßnahme, um die oben genannte Barrierefreiheit zu erreichen, wurden ebenfalls zwei Assistenzen implementiert, welche das entsprechende Hindernis neutralisieren sollen, ohne dabei einen signifikanten Einfluss auf die anderen Modi des Spiels zu haben. Ziel dieser Bachelorarbeit ist es, einerseits die Änderungen im Spielverhalten und der Spielerleistung zu untersuchen, um die Basis für ein System zu schaffen, die diese Unterschiede klassifiziert und Adaptionen dementsprechend anbietet. Andererseits soll die allgemeine Effektivität der Adaptionen in den jeweiligen Anwendungsbereichen untersucht werden.

Daraus folgen die zentralen Fragestellungen dieser Arbeit:

Frage 1: Wirken sich die Hindernisse im Memory-Spiel auf die Spielerdaten aus?

Hypothese 1: Die Hindernisse wirken sich negativ auf die Spielerleistung, die User Experience und die mentale Auslastung aus.

Frage 2: Wie wirken sich die Adaptionen auf die verschiedenen Spielmodi aus?

Hypothese 2: Die Adaptionen wirken sich im passenden Hindernismodus positiv auf die Spielerleistung und die User Experience aus. In den anderen beiden Modi führen sie zu keinem signifikanten Unterschied.

Frage 3: Kann man anhand der Verhaltens- und Hirnstromdaten eines Spielers auf den gespielten Modus schließen, um eine Adaption anzubieten?

Hypothese 3: Die Auswirkungen der Hindernisse sind in beiden Datenreihen groß genug, um zwischen Modus mit und ohne Hindernis zu unterscheiden. Darüber hinaus lassen sich beide Hindernismodi voneinander abgrenzen, sodass durch Verkettung ein Rückschluss möglich ist.

2. Verwandte Arbeiten

Diese Bachelorarbeit ist eine Fortführung zweier vorangegangener Abschlussarbeiten am CSL. Miranda [Mir] entwickelte die erste Version des Memory Spiels, fügte ein visuelles und ein gedächtnisbasiertes Hindernis sowie zwei entgegenwirkende Adaptionen hinzu und wertete die Spiel- und Fragebogendaten aus einer Experimentreihe mit diesem Spiel aus.

Klassen [Kla] führte die Arbeit fort indem sie die bei dem Experiment angefallenen EEG-Daten auswertete und eine Klassifikation der Hindernismodi mittels Support Vector Machines und Linear Discriminant Analysis untersuchte, wobei erstere gute Ergebnisse erzielten. Diese Arbeit basiert auf dem gleichen Spiel und versucht, nach Änderungen an Hindernissen und Adaptionen sowie einer neuen Experimentreihe, beide Bereiche zu untersuchen, wobei allerdings auf eine explizite Klassifikation durch Machine-Learning Verfahren verzichtet wird. Die genauen Änderungen am Spiel werden in Kapitel 3 beschrieben.

Ebenfalls, verwandt zu dieser Arbeit, untersuchten Putze et al. [PSS18] für ein ähnliches Memory Spiel mit und ohne gedächtnisbasierter Nebenaufgabe die Klassifikationsfähigkeit der Spielmodi durch Long Short Term Memory Networks, einer speziellen Form von Recurrent Neural Networks, gegenüber der einer Linear Discriminant Analysis. LSTMs sind speziell auf sequentielle Daten (hier: die Spielzüge der Spieler) ausgelegt und erreichten deutlich bessere Ergebnisse als der lineare Ansatz.

Zum Thema adaptive und inklusive User Interfaces gibt es diverse Forschungsansätze, wie diese gestaltet werden können:

Stephanidis et al. [SPS⁺98] entwickelten ein GUI-System welches modularisiert an Einschränkungen von Nutzern anpassbar ist (unter Anderem mittels Sprachin- und output, Braille-Display, Touchscreen und Joystick-Support) und dabei die Paradigmen *adaptable* und *adaptive* umsetzt. Es können also einerseits zum Start statische Informationen wie Sehbehinderungen oder fehlende Sprachkompetenzen kenntlich gemacht werden, aber auch zur Laufzeit verarbeitet das System das Verhalten des Nutzers wie etwa langes Zögern oder mangelnde Navigationsfähigkeit und bietet entsprechende Module dynamisch an.

Gajos et al. [GWW07] entwarfen eine UI mit skalierbaren Parametern um ebenfalls Menschen mit motorischen und visuellen Einschränkungen zu unterstützen. Das

Entwickeln einer guten UI wird hier als Optimierungsproblem der Bedienungszeit beschrieben, welches es algorithmisch zu lösen gilt. Dafür führten Nutzer mit verschiedenen Einschränkungen einen kurzen motorischen Test durch, bei welchem Klickpräzision, Schnelligkeit und ruhige Cursorführung untersucht wurde, um anhand dieser Features mithilfe eines Regressionsmodells die motorischen Fähigkeiten der Nutzer einzuschätzen. Die so spezialisierten GUIs erreichten für die Nutzer im Durchschnitt einen Zeitvorteil von 20% bei der Bedienung für die selbe Aufgabe.

Ebenso relevant für diese Arbeit gibt es verschiedene Ansätze, mentale Auslastung auf Basis von EEG-Daten zu klassifizieren. Diese werden später nach den EEG-Grundlagen in Kapitel 6.5 beleuchtet.

3. Das Spiel

Dieses Kapitel beschreibt die für das Experiment genutzte Memory-App und ihren Entwicklungsprozess.

3.1 Bisherige Version

Die ursprüngliche Version der App wurde von Rafael Miranda [Mir] in einer Bachelorarbeit am CSL geschrieben. Sie basiert auf dem Open-Source-Projekt von sromku [Kus19] und wurde in Kotlin für die Android-Plattform entwickelt. Das Spiel benutzte ein Miniatur-Tierkartenset und enthielt bereits ein visuelles und ein gedächtnisbasiertes Hindernis (im Folgenden werden diese Hindernisse als *Visual-Obstacle* und *Memory-Obstacle* bezeichnet) sowie jeweils eine passende Adaption. Das Visual-Obstacle stellte die Tierkarte in Graustufen dar und wendete darüber hinaus einen gaußschen Weichzeichner an (siehe Abb. 3.2), um die Ähnlichkeit zwischen den Karten signifikant zu erhöhen, da diese grundsätzlich in ähnlichen Posen und Proportionen dargestellt wurden.



Abbildung 3.1: Normale Tierkarte

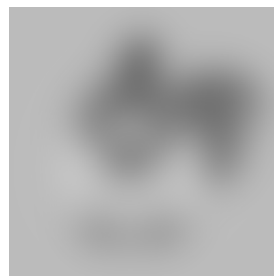


Abbildung 3.2: Tierkarte mit Visual-Obstacle

Die visuelle Adaption wurde in Form eines Zooms realisiert, der die Karte stark vergrößert in der Mitte des Bildschirms präsentierte. Dabei wurde die Stärke des Gauß-Filters reduziert, sodass die Konturen der Tiere wieder deutlich leichter zu unterscheiden waren (siehe Abb. 3.4).



Abbildung 3.3: Normale Tierkarte

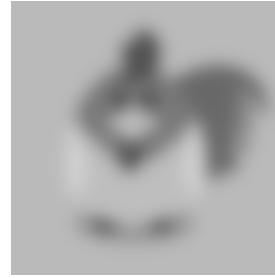


Abbildung 3.4: Visual-Obstacle mit Zoom

Das Memory-Obstacle bestand aus einer gänzlich hinzugefügten Nebenaufgabe für den Spieler: Bei jedem Kartenflip nennt ein Sprachsynthese-System eine Zahl zwischen 1 und 10. Der Spieler hat nun die Aufgabe, diese Zahlen jeweils aufzuaddieren und am Ende die Summe aller einzugeben. Als Adaption dazu werden beim Umdrehen von zwei unterschiedlichen Karten alle vorherig aufgedeckten Karten in der gleichen Reihenfolge noch einmal aufgezeigt.

3.2 Kartenset

Nachdem das Visual-Obstacle in der bisherigen Version eine allgemeine Sehschwäche erfolgreich modelliert hatte, wird in dieser Arbeit der Fokus auf eine Farbschwäche bzw. Farbenblindheit gewechselt. Etwa 8% aller Männer leiden an einer Form von Farbschwäche [Ins19a], was diese Beeinträchtigung häufiger als oft angenommen macht. Um das Visual-Obstacle dementsprechend anzupassen, wurde zu dem Tierkartenset ein Set aus gänzlich einfarbigen Karten hinzugefügt. Das Tierkartenset wurde nicht verworfen, in dieser Arbeit aber auch nicht weiter verwendet. Für das neue Set wurden sieben Spektralfarben gewählt, die möglichst unterscheidbar waren (siehe Abb. 3.5).

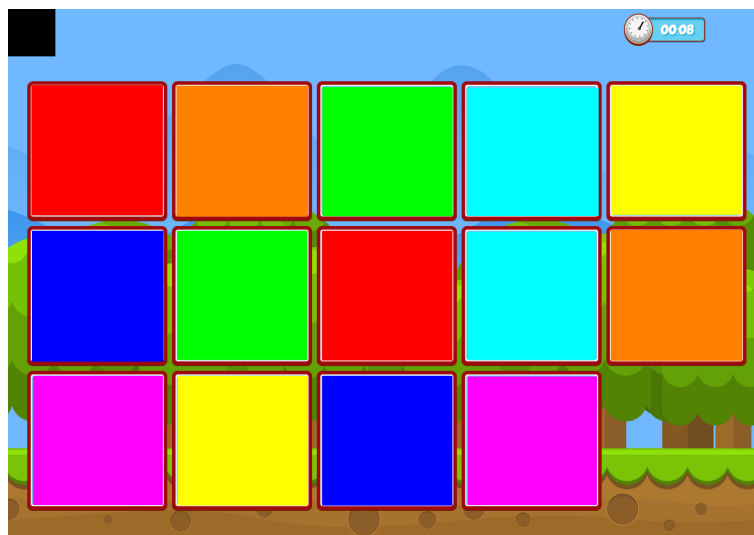


Abbildung 3.5: Alle Karten des Spektral-Sets

Dieses Kartenset wurde jedoch einerseits nach den ersten Testdurchläufen als zu einfach unterscheidbar empfunden, andererseits wollte ich das Visual-Obstacle nach

weiterer Recherche auf die häufigste Form der Farbschwäche, die Rot-Grün-Schwäche [Ins19a], reduzieren. Diese äußert sich unterschiedlich pro Person und schränkt die Rezeption von entweder Rot- oder Grüntönen ein. Um dies in einer universellen Form zu modellieren, habe ich ein weiteres farbiges Kartenset angelegt, diesmal jedoch mit jeweils drei unterschiedlich hellen Rot- und Grüntönen und einem Brauntönen (siehe Abb. 3.6). Im normalen Spielmodus sind diese Farben zunächst problemlos voneinander unterscheidbar.

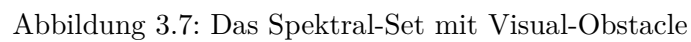


Abbildung 3.6: Alle Karten des Rot-Grün-Sets

3.3 Hindernisse

Das Memory-Obstacle mit der arithmetischen Nebenaufgabe ist unverändert geblieben. Nur das Visual-Obstacle wurde verändert. Ursprünglich, als noch das Spektralfarbenset genutzt wurde, bestand es darin, die Karten in Graustufen darzustellen (siehe Abb. 3.7). Im Gegensatz zu dem normale Farbenset wurde dies jedoch in den Testversuchen als sehr schwer empfunden und führte dazu, dass die Testkandidaten fast nur noch raten konnten, was sowohl der statistischen Auswertung als auch dem Spielspaß schadete.

Mit dem Rot-Grün-Kartenset wurde das Hindernis noch einmal überarbeitet, so dass es nun alle Farbtöne “Braunverschoben” darstellt. Dazu wurde ein Brauntönen mit 75% Deckkraft auf alle Farben aufaddiert (siehe Abb. 3.8). Dies soll bewirken, dass die Farbtöne im Farbraum näher zusammenrücken und die Spieler so, wie bei einer Rot-Grün-Schwäche, diese Farbtöne fast nur noch durch Helligkeitsunterschiede unterscheiden können.



Die Adaptionen beider Hindernisse wurden überarbeitet. Die ursprüngliche Memory-Assistance nahm durch das sequenzielle Aufzeigen aller Karten bei jedem Zug sehr viel Zeit in Anspruch und verlängerte die durchschnittliche Spieldauer damit in manchen Fällen auf das Doppelte [Mir]. Zusätzlich machte sie das Spiel, abgesehen von der Nebenaufgabe, mit jedem Zug trivialer, da die Spieler mehrfach alle Karten direkt in Folge sehen und sie sich so deutlich besser einprägen können.

Deshalb wurde die Memory-Assistance insofern geändert, dass sie nur noch den letzten Zug erneut aufzeigt. Dies soll dem Spieler nicht die gesamte Hauptaufgabe abnehmen, ihm aber durch das zweimalige Zeigen jedes Zuges eine faire Chance geben, sich die Karten trotz der Nebenaufgabe einzuprägen. Um die Zeitkosten dabei zu minimieren, werden die Karten außerdem nicht mehr sequenziell sondern gleichzeitig aufgedeckt und werden auch nach kurzer Zeit wieder umgedreht.

Die andere Adaption, für das Visual-Obstacle, wurde den neuen Umständen entspre-

chend gänzlich neu entworfen:

Jedem Kartenpaar wird ein Buchstabe zugeordnet, welcher beim Umdrehen durch Sprachsynthese genannt wird. Für die sieben Kartenpaare wurden die Buchstaben **A**, **C**, **J**, **Q**, **X**, **V** gewählt, da diese in der deutschen Aussprache durch die verschiedenen mitgesprochenen Vokale maximal unterscheidbar sind. Diese Adaption soll bei Spielen mit visuellem Hindernis eine neue Ebene hinzufügen, an welcher sich die Spieler statt der Kartenmotive orientieren können. Da die Unterscheidbarkeit der Motive in den anderen Modi nicht die größte Schwierigkeit ausmacht, sollte diese Adaption die anderen Modi auch nicht beeinflussen.

3.5 Spielmodi und Begrifflichkeit

Aus der obigen Beschreibung der Änderungen an dem Memory-Spiel ergeben sich die folgenden neun Spielmodi, welche in dieser Arbeit unter folgenden Namen weiter behandelt werden:

Classic: Ein normales Memory-Spiel mit dem Rot-Grün-Kartenset und der im Folgenden genannten *Hauptaufgabe*: Drehe alle passenden Kartenpaare um und erreiche dabei eine gute Spielleistung in Hinblick auf die benötigte Zeit, die Zuganzahl und die Fehleranzahl.

Memory-Obstacle: Ein Memory Spiel mit dem Rot-Grün-Kartenset mit zusätzlicher *Nebenaufgabe*: Bei jedem Kartenflip wird eine zufällige Zahl zwischen 1 und 10 genannt, welche der Spieler konsekutiv aufaddieren muss. Nach Finden des letzten Paares wird der Spieler nach einer Summe gefragt, welche im Gamelog neben der exakten Summe gespeichert wird.

Visual-Obstacle: Bei diesem Memory Spiel wird das Rot-Grün Kartenset durch das braunverschobene Kartenset ausgetauscht. Die Hauptaufgabe bleibt dabei unverändert.

Classic + Memory-Assistance: Ein normales Memory Spiel, bei welchem nach jedem unerfolgreichen Spielzug (also zwei umgedrehte, nicht zueinander passende Karten) der vorherige Spielzug wiederholt wird. Die Karten werden dabei gleichzeitig umgedreht und nach kurzer Zeit wieder umgedreht.

Classic + Visual-Assistance: Ein normales Memory Spiel, bei welchem jedem Kartenpaar einer von den sieben oben beschriebenen Buchstaben zugeordnet wird. Jedes Mal, wenn eine Karte umgedreht wird, wird der Buchstabe durch Sprachsynthese genannt.

Memory-Obstacle + Memory-Assistance: Ein Memory Spiel mit Nebenaufgabe und der Memory-Assistance, welche vergangene Spielzüge wiederholt.

Memory-Obstacle + Visual-Assistance: Ein Memory Spiel mit Nebenaufgabe und der Visual-Assistance, welche den Kartenpaaren Buchstaben zuweist.

Visual-Obstacle + Memory-Assistance: Ein Memory Spiel mit braunverschobenem Kartenset und der Memory-Assistance, welche vergangene Spielzüge wiederholt.

Visual-Obstacle + Visual-Assistance: Ein Memory Spiel mit braunverschobenem Kartenset und der Visual-Assistance, welche den Kartenpaaren Buchstaben zuweist.

In Graphen und Tabellen werden aus Platzgründen die Abkürzungen **Mem.**, **Vis.**, **Obs.**, und **Ass.**, sowie in einer **c,cm,cv,m,mm,mv,v,vm,vv** benutzt. In letzterem Fall steht der erste Buchstabe für den Spielmodus und der zweite für eine vorhandene Assistenz.

4. Der Versuch

In diesem Kapitel wird der Versuchsaufbau, -ablauf und die teilnehmenden Probanden beschrieben.

4.1 Versuchsaufbau

Ziel des Experiments war es, durch normales Spielen von Versuchsteilnehmern die folgenden Fragen zu beantworten:

- Haben die Hindernisse einen negativen Einfluss auf die Leistungen der Spieler?
- Gleichen die Assistenzen diese Einflüsse effektiv aus?
- Beeinflussen sie darüber hinaus die anderen Modi, für die sie nicht entwickelt wurden?
- Zeigen die 3 Hauptspielmodi *Classic*, *Memory-Obstacle* und *Visual-Obstacle* ein jeweils oder teilweise individuelles Spielverhalten auf, welches eine Klassifikation ermöglicht?
- Zeigen die 3 Hauptspielmodi Unterschiede in der Messgröße der Hirnaktivität auf, sodass dadurch Klassifikation möglich ist?

Spielmodus:	Classic	Memory-Obstacle	Visual-Obstacle
Ohne Assistenz	c	m	v
Memory-Assistance	cm	mm	vm
Visual-Assistance	cv	mv	vv

Abbildung 4.1: Alle Kombinationen aus Hindernissen und Assistenzen

Dazu wurden Probanden eingeladen, alle neun Kombinationen (siehe Abb. 4.1) aus Hindernissen und Adaptionen zu spielen. Die App speichert dabei automatisch nach jedem einzelnen Modus die Spieldaten, bestehend aus der *Sequenz* der vom Spieler

umgedrehten Karten, der *Spielzeit* und der *Endsumme* sowie der *spielererrechneten Summe* beim Memory-Obstacle. Die Sequenz der umgedrehten Karten bietet außerdem die Möglichkeit, die Anzahl der Fehler des Spielers später zu errechnen. Um die Hirnaktivität während des Spiels aufzuzeichnen, trugen die Probanden während des Versuchs außerdem eine EEG-Kappe. Die Erfassung und Verarbeitung der Hirnaktivitätsdaten wird in Kapitel 6 näher behandelt.

4.1.1 Versuchsaufbau

Zu Beginn des Versuchs unterschrieben alle Probanden eine Einverständniserklärung, sowohl für die Nutzung der Verhaltensdaten als auch für die EEG-Daten. Es wurde sichergestellt, dass keine Verständnisprobleme diesbezüglich vorlagen. Anschließend wurde die EEG-Kappe mit Elektrodengel befüllt. In diesem Zeitraum, der zwischen 30 und 45 Minuten andauerte, hatten die Probanden Zugang zur App und konnten sich mit allen verschiedenen Spielmodi vertraut machen. Damit sollte sichergestellt werden, dass alle Probanden den eigentlichen Versuch mit den gleichen Vorkenntnissen antreten. Zudem wurde erfragt, ob die Probanden an einer Art von Farbenblindheit leiden, was trotz der erwarteten Prävalenz [Ins19a] nicht der Fall war. Brillen mussten zwar für das Befüllen der Seitenelektroden kurz abgesetzt werden, konnten davor und danach aber über der EEG-Kappe problemlos getragen werden. Der Versuch wurde im elektromagnetisch abgeschirmten Biolab des CSL durchgeführt, um eventuelle Artefakte des EEG zu minimieren. Um die EEG-Daten mit den Verhaltensdaten zu synchronisieren, gibt es ein schwarzes Feld in der oberen linken Ecke des Spiels, welches beim Kartenumdrehen für 250ms weiß aufblinkt. Auf diesem Feld wurde ein Lichtsensor angebracht, welcher parallel zum EEG die Lichtimpulse aufzeichnet. Dafür musste der Raum während der Spielzeit abgedunkelt werden. Es wurde deshalb auch in der Vorbereitungszeit sichergestellt, dass die Probanden im Dunkeln problemlos und beschwerdefrei spielen konnten. Nach der Vorbereitungsphase wurden die neun oben gezeigten Modi in zufälliger Reihenfolge von den Probanden gespielt. Ein Spiel dauerte dabei je nach Modus zwischen 21 und 265 Sekunden. Nach jedem einzelnen Modus sollten die Probanden einen passenden Fragebogen dazu ausfüllen, welcher Einschätzungen über die eigene Leistung und die Schwierigkeit des Spielmodus nach dem Schema der Likert-Skala, in diesem Fall von 1-7, bietet [Lik32]. Da sich Probanden während der EEG-Aufnahme möglichst nicht bewegen sollten, durfte bei Bedarf auch eine kurze Pause zwischen den Modi eingelegt werden. Der erste Versuch wurde als Testversuch behandelt um den ungefähren Test-, Installations- und Erklärungsablauf kennen zu lernen und danach für alle Probanden gleich gestalten zu können. Er wurde nicht in die spätere Auswertung miteinbezogen.

4.1.2 Probanden

Das Experiment wurde mit 19 Probanden durchgeführt. Dafür wurden öffentliche Ausschreibungen innerhalb des Uni-Bereichs sowie Online-Anzeigen ausgestellt. Hierbei wurde keine Altersbeschränkung gesetzt, dennoch waren alle Probanden zwischen 18 und 27 Jahren. Von den Probanden waren 14 männlich und 5 weiblich.

5. Auswertung der Verhaltensdaten

In diesem Kapitel werden die Daten der Fragebögen und die der Gamelogs ausgewertet und verglichen. Dies soll zum einen den Einfluss der Hindernisse auf die Spielleistung darlegen und ermitteln, wie sehr die Adaptionen diesen Einfluss aufheben, verstärken, oder ob sie darüber hinaus die Leistung auf die Spielmodi beeinflussen.

5.1 Vorgehensweise

Die Ergebnisse der Fragen, welche wie in Kapitel 4.1.1 beschrieben mit einem Wert von 1-7 (Volle Ablehnung - Volle Zustimmung) beantwortet werden konnten, sowie die Werte der Gamelogs der Spieler wurden in jeweils einem Spreadsheet (im Anhang) zusammengetragen. Anschließend wurde für jeden Spielmodus der Mittelwert und die Standardabweichung aller Spieler berechnet. Diese Werte werden im Folgenden bei zwei entsprechend passenden Spielmodi miteinander verglichen. Für die Untersuchung der statistischen Signifikanz wird der gepaarte t-Test herangezogen [pyt19], welcher auf der studentschen t-Verteilung beruht [Stu08]. Dieser t-Test prüft für zwei Mittelwerte wie wahrscheinlich es ist, dass sie aus zwei unterschiedlichen Mengen stammen. Diese Wahrscheinlichkeit wird in Form der probabilistic-value (p) angegeben. Die Nullhypothese H_0 des t-Tests besagt, dass die Mittelwerte nicht aus unterschiedlichen Mengen stammen, die durchschnittliche Differenz aller Werte der beiden Gruppen zueinander also 0 beträgt. Für diese Arbeit wurde ein statistisch üblicher Konfidenzwert von p kleiner gleich 0.05 gewählt, für welchen die Nullhypothese verworfen wird und der Mittelwertunterschied als statistisch signifikant angenommen wird. Liegt das Wissen vor, dass die Werte einer der beiden Messgruppen im Durchschnitt höher oder niedriger sind, kann man auch einseitige (one-tailed) t-Tests durchführen, welche die doppelte p-value (hier 0.1) in nur einer statt beiden Richtungen nutzen. Damit kann man bei entsprechenden Messwerten mehr signifikante Zusammenhänge finden, wenn man die oben genannte Annahme garantieren kann. Zwar kann man bei diesem Versuch in Bezug auf die Adaptionen theoretisch davon ausgehen, dass sie die Leistungen des Spielers und damit auch seine Selbsteinschätzungen positiv beeinflussen, trotzdem verzichte ich, auch aufgrund der verhältnismäßig geringen

Frage	Classic	Memory-Obstacle	Visual-Obstacle
Das Spiel hat mich mental gefordert p-value (zu <i>Classic</i>)	3.8 (1.6) -	6.4 (0.7) <0.01	4.2 (1.4) 0.26
Ich konnte mir die Positionen der Karten gut merken p-value (zu <i>Classic</i>)	6.1 (0.8) -	3.4 (1.6) <0.01	3.9 (1.7) <0.01
Ich denke, ich habe das Spiel zügig durchgespielt p-value (zu <i>Classic</i>)	6.0 (1.0) -	3.1 (1.8) <0.01	4.7 (1.2) <0.01
Die Bilder der Karten waren leicht zu unterscheiden p-value (zu <i>Classic</i>)	6.4 (0.7) -	5.5 (1.5) 0.01	2.9 (1.5) <0.01
Ich denke, ich hätte besser spielen können p-value (zu <i>Classic</i>)	4.1 (2) -	5.6 (1.4) <0.01	4.7 (1.4) 0.27
Ich konnte mich gut an die Positionen der Karten erinnern p-value (zu <i>Classic</i>)	5.6 (1.4) -	2.8 (1.8) <0.01	4.4 (1.6) 0.02

Tabelle 5.1: Einschätzung der Spieler bezüglich ihrer Leistung. Dabei steht 1 für volle Ablehnung und 7 für volle Zustimmung. Die Werte in Klammern sind die dazugehörigen Standardabweichungen. Die p-Values ergeben sich aus t-Tests des Modusses mit dem Classic-Modus, wie in der Fragenspalte gekennzeichnet.

Stichprobe, auf den one-tailed t-Test.

5.2 Vergleich der Fragebögen

Im Folgenden werde ich zuerst die drei Hauptspielmodi *Classic*, *Memory-Obstacle* und *Visual-Obstacle* miteinander vergleichen und später auf die assistierten Modi eingehen:

5.2.1 Einschätzung der drei Hauptmodi

Betrachtet man die Leistungsfragen in Tabelle 5.1 lässt sich ein Trend erkennen: In allen Fragen, außer Frage 4 empfanden die Probanden den Classic-Modus als den einfachsten Modus, Visual-Obstacle als schwerer und Memory-Obstacle als den schwierigsten Modus. Bis auf zwei Fälle mit einem zu hohen p-Wert (Frage 1 und 5 bei Visual-Obstacle) sind diese Ergebnisse signifikant unterschiedlich. Bei der Frage “Das Spiel hat mich mental gefordert” erreicht der Spielmodus Memory-Obstacle mit relativ geringer Varianz nahezu höchstmögliche Beanspruchung. Frage 4, “Die Bilder der Karten waren leicht zu unterscheiden” bricht den sonstigen Trend verständlicherweise, da das Visual-Obstacle darauf ausgelegt ist, dies größtmöglich zu erschweren. Selbst bei dieser Frage wurde der Memory-Obstacle-Modus schwieriger als der Classic-Modus eingestuft, obwohl das Memory-Obstacle die Hauptaufgabe nicht verändert.

Frage	Classic	Memory-Obstacle	Visual-Obstacle
Das Design der App hat mir gut gefallen p-value (zu Classic)	6.5 (0.7) -	6.3 (0.9) 0.14	6.3 (0.8) 0.02
Mir hat das Spiel Spaß gemacht p-value (zu Classic)	6.6 (0.7) -	4.4 (1.5) <0.01	5.8 (1.2) <0.01
Die Darstellung der Karten hat mir dabei geholfen, die Positionen der Karten zu merken p-value (zu Classic)	5.1 (2.8) -	4.5 (1.5) 0.11	3.7 (1.6) <0.01

Tabelle 5.2: Die durchschnittliche Einschätzung der Spieler bezüglich der Usability-Aspekte mit Standardabweichungen in gleicher Darstellung wie in Tabelle 5.1

Dieser Trend der Schwierigkeit der Modi spiegelt sich auch im Spielspaß wider. In Tabelle 5.2 zeigt sich, dass die Spieler dem Classic-Modus fast einheitlich die beste Bewertung gegeben haben, dem Visual- und Memory-Obstacle-Modus dann jeweils zunehmend weniger. Frage 3 über die Darstellung der Karten lässt sich wahrscheinlich ähnlich wie Frage 4 in der vorherigen Tabelle interpretieren. Allerdings ist hier die Varianz beim Classic-Modus so groß, dass zumindest der Memory-Obstacle-Modus nicht signifikant davon unterscheidbar ist. Schlussendlich wurde das Design der App in allen Fällen sehr positiv aufgenommen und bietet wenig Deutungsmöglichkeiten.

5.2.2 Einschätzung der Assistenzen

In Tabelle 5.3 stehen die Einschätzungen für alle assistierten Modi und den unassistierten Memory-Obstacle-Modus. Aus Platzgründen wurden diese abgekürzt, dabei steht jeweils der erste Buchstabe für den Hauptmodus und der zweite für die Assistenz. Für den Memory-Obstacle-Modus mit Memory-Assistance fehlt der Eintrag bei Frage 2, da dieser von mir in den jeweiligen Fragebögen leider vergessen wurde. Es fällt jedoch auf, dass sich die Ergebnisse von Frage 1 und 2 ohnehin sehr stark ähneln. Ein t-Test zwischen beiden Fragen ergab für alle 5 möglichen Fälle ein $p > 0.05$.

Für den Classic-Modus wurden beide Assistenzen grundsätzlich als hilfreich eingestuft. Besonders die Visual-Assistance wurde mit einer Bewertung von 6.4/6.5 (Fragen 1 und 2) sehr positiv aufgenommen, auch wenn sie in diesem Modus eigentlich nicht nötig ist und diesen nicht beeinflussen sollte. Bei dem Memory-Obstacle-Modus wurden die Assistenzen unterschiedlich aufgenommen. Zwar wurde die extra dafür entworfene Memory-Assistance mit 5.1 gut bewertet, damit allerdings schlechter als fast alle anderen, sogar als die Assistenzen im Classic-Modus. Als einzig, aber deutlich schlechter wurde nur die Visual-Assistance bei diesem Modus empfunden, mit einer negativen Bewertung von 2.2/2.4 (Fragen 1 und 2). Dies war zu erwarten, da der gesprochene Buchstabe der Visual-Assistance bewusst direkt auf die gesprochene Zahl der Nebenaufgabe folgt und diese damit erheblich behindern kann. Bei dem Visual-Obstacle-Modus wurden die Assistenzen wieder ähnlich wie im Classic-Modus aufgenommen. Beide Assistenzen erhielten sehr gute Bewertungen, die hierfür entwickelte Visual-Assistance sogar fast maximale Punktzahl, welches die Intention positiv

Frage	cm	cv	m	mm	mv	vm	vv
Die Unterstützung empfand ich als hilfreich	5.7 (1.5)	6.4 (1.3)	-	5.1 (2)	2.2 (2)	5.9 (1.3)	6.7 (0.9)
p-value (zum Gegenmodus)	0.054	0.054	-	<0.01	<0.01	0.052	0.052
Durch die Unterstützung konnte ich mir die Karten besser merken	5.6 (1.5)	6.5 (1.3)	-	(-)	2.4 (2.1)	5.4 (1)	6.5 (1)
p-value (zum Gegenmodus)	0.01	0.01	-	-	-	<0.01	<0.01
p-value F1 zu F2	0.51	0.16	-	-	0.19	0.06	0.08
Ich konnte die genannten Zahlen gut zusammen addieren	-	-	3.5 (1.6)	3.5 (1.5)	3.1 (1.7)	-	-
p-value (zu m)	-	-	-	1	0.47	-	-
Ich konnte mir die Zwischensummen gut merken	-	-	3.4 (1.8)	3.1 (1.3)	2.7 (1.5)	-	-
p-value (zu m)	-	-	-	0.38	0.15	-	-
p-value F3 zu F4	-	-	0.84	0.13	0.38	-	-

Tabelle 5.3: Die durchschnittlichen Ergebnisse der spielmodusspezifischen Fragen mit Standardabweichungen. Die t-Tests wurden jeweils zu dem gekennzeichneten Modus durchgeführt. Der Gegenmodus ist der Modus mit dem selben Hindernis aber anderer Assistenz. M bezeichnet den unassitierten Memory-Obstacle-Modus.

bestätigt. Einerseits ist es verwunderlich, dass die Memory-Assistance hier so gut aufgenommen wurde, andererseits bringt sie auch in der jetzigen Form praktisch keine Nachteile mehr mit sich und wurde so im schlimmsten Fall nicht beachtet, jedoch auch nicht als störend empfunden.

In den Tabellen 5.4 und 5.5 werden jeweils nochmal die beiden Hindernismodi mit und ohne passender Assistenz verglichen. Hier zeigt sich noch deutlicher, dass die Spieler den Memory-Obstacle-Modus auch mit passender Assistenz nicht merklich einfacher fanden. Keine der Vergleiche weisen statistische Signifikanz auf, die mentale Forderung hat sogar die exakt gleiche Bewertung bekommen. Beim Visual-Obstacle-Modus hingegen haben die subjektive Spielgeschwindigkeit und Merkfähigkeit stark zugenommen, nur die mentale Forderung weist keine Signifikanz auf.

Zusammenfassend wurden beide Assistenzen in allen Fällen bis auf Memory-Obstacle

Frage	Ohne Assistenz	Mit Assistenz
Das Spiel hat mich mental gefordert	6.4 (0.7)	6.4 (0.7)
p-value (zum Gegenmodus)	-	1
Ich denke, ich habe das Spiel zügig durchgespielt	3.1 (1.8)	3.7 (1.7)
p-value (zum Gegenmodus)	-	0.27
Ich konnte mir die Positionen der Karten gut merken	3.4 (1.6)	3.5 (1.5)
p-value (zum Gegenmodus)	-	0.07

Tabelle 5.4: Die Leistungseinschätzung im Vergleich für den unassitierten und assistierten Memory-Obstacle-Modus

Frage	Ohne Assistenz	Mit Assistenz
Das Spiel hat mich mental gefordert	4.2 (1.4)	3.7 (1.6)
<i>p-value</i> (zum Gegenmodus)	-	0.23
Ich denke, ich habe das Spiel zügig durchgespielt	4.7 (1.2)	5.9 (1.2)
<i>p-value</i> (zum Gegenmodus)	-	<0.01
Ich konnte mir die Positionen der Karten gut merken	3.9 (1.7)	6 (1.3)
<i>p-value</i> (zum Gegenmodus)	-	<0.01

Tabelle 5.5: Die Leistungseinschätzung im Vergleich für den unassistenten und assistierten Visual-Obstacle-Modus

mit Visual-Assistance positiv aufgenommen, die Visual-Assistance dabei im Classic und im Visual-Obstacle-Modus am besten. Die für den Memory-Obstacle-Modus entworfene Assistenz hingegen fand wider Erwarten nur minimal positiven Anklang. Für all diese subjektiven Einschätzungen wird im Folgenden betrachtet, wie sehr sie mit den Leistungsunterschieden übereinstimmen.

5.3 Datenverarbeitung der Spieldaten

In den Gamelogs wurden für jedes Spiel einzeln die benötigte Zeit und die ersten 20 Züge des Spielers festgehalten. Die Züge bestehen dabei jeweils aus dem umgedrehten Kartentupel. Dadurch ist es anhand der Logs auch möglich, die Anzahl der “Fehler” zu rekonstruieren. Ein Fehler ist ein Zug, in dem zu einer Karte kein passendes Paar gefunden wurde, obwohl die dafür benötigte Karte schon einmal aufgedeckt wurde. Für den Memory-Obstacle-Modus wird außerdem noch festgehalten, welche Endsumme der Nutzer eingegeben hat und wie weit er damit von der korrekten Summe entfernt lag. All diese Werte wurden erneut in einem Spreadsheet eingetragen und für sie Mittelwerte und Standardabweichungen ausgerechnet.

5.4 Vergleich der Spieldaten

5.4.1 Vergleich der Hauptmodi

Die Graphen 5.1, 5.2 und 5.3 zeigen jeweils die Zeit, Zuganzahl und Fehleranzahl pro Spielmodus. Vergleicht man hier zunächst den Classic-Modus mit den beiden Hindernismodi, so lässt sich beobachten, dass Spieler für den Classic-Modus deutlich weniger Zeit ($p = <0.01$, 0.01) und etwas weniger Züge ($p = <0.01$, <0.01) benötigten, außerdem auch nur etwa halb so viele Fehler wie in den anderen Modi machten ($p = <0.01$, 0.01). In all diesen Metriken erzielt der Visual-Obstacle-Modus die zweitbesten Ergebnisse und der Memory-Obstacle-Modus die schlechtesten, wobei der Unterschied beider nur bei der benötigten Zeit statistisch signifikant ist ($p = <0.01$, 0.69, 0.55, in Reihenfolge Zeit, Züge, Fehler). Für die benötigten Züge ist außerdem fragwürdig, inwiefern das 20-Züge-Limit der Gamelogaufzeichnung die wahre Zuganzahl nach unten verfälscht, gerade für den Memory-Obstacle-Modus. Der in der Fragebogenauswertung beobachtete grundsätzliche Schwierigkeitstrend der drei Modi findet sich hier zumindest in Ansätzen ebenso.

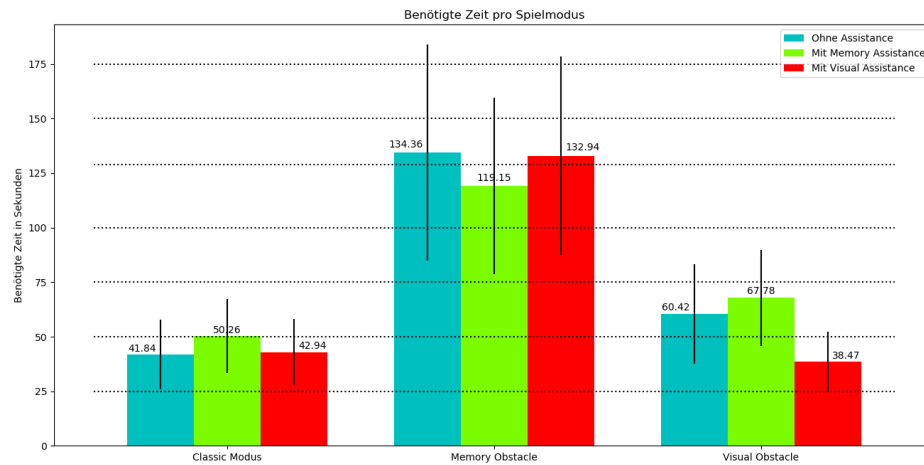


Abbildung 5.1: Benötigte Zeit für die verschiedenen Spielmodi

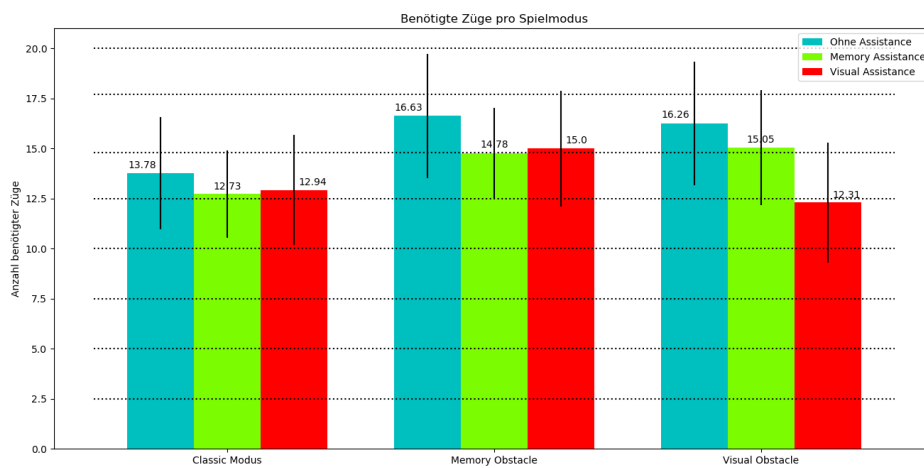


Abbildung 5.2: Benötigte Züge für die verschiedenen Spielmodi

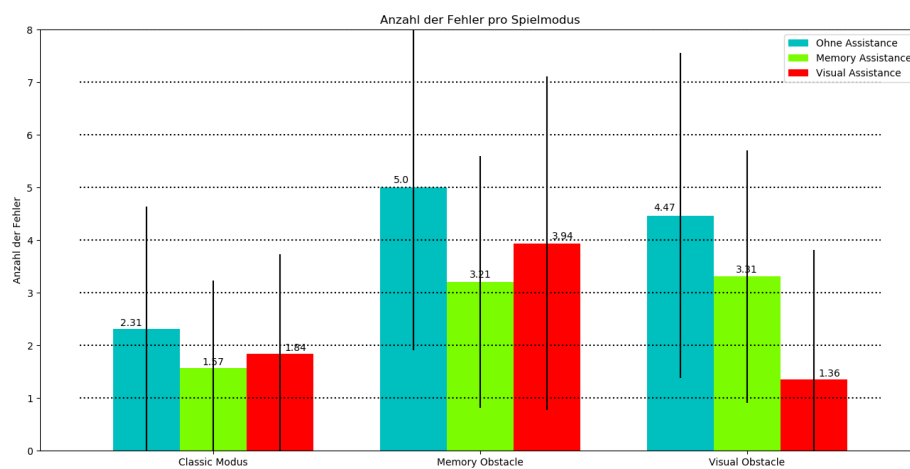


Abbildung 5.3: Anzahl Fehler nach Spielmodus

5.4.2 Auswirkungen der Assistenzen

Die Memory-Assistance hat sich bei dem Classic-Modus negativ auf die Zeit (+20%) und leicht positiv auf Züge und Fehler ausgewirkt, allerdings wird nur die Zeitzunahme vom t-Test gedeckt (in der Reihenfolge $p = 0.01, 0.12, 0.12$). Die Visual-Assistance hingegen hat im Classic-Modus kaum bis gar keinen Einfluss auf diese Metriken ($p = 0.84, 0.38, 0.51$). Dies widerspricht der subjektiven Einschätzung der Spieler, welche die Assistenz in diesem Spielmodus als sehr hilfreich (6.4 von 7) empfanden.

Bei dem Memory-Obstacle-Modus hat sich die Memory-Assistance auf keine der drei Metriken signifikant ausgewirkt ($p = 0.30, 0.07, 0.11$), auch wenn dies bei der durchschnittlich etwa 20% geringeren Fehleranzahl zunächst so erscheint. Dies deckt sich mit der Einschätzung der Spieler, die sich durch die Assistenz nicht weniger beansprucht fühlten. Es kann jedoch wieder argumentiert werden, dass die Zuganzahl im unassistierten Modus höher sein könnte als in der Statistik festgehalten. Da die Zuganzahl durch die Assistenz auch eigentlich nicht gesteigert werden kann, hätte man hier einen one-tailed t-Test anwenden können und damit einen signifikanten Zusammenhang erkannt. Der Abstand der berechneten zu der korrekten Summe wurde ebenfalls ausgewertet (siehe Graph 5.4). Im Mittel hat die passende Assistenz den Abstand um 30% vermindert und die unpassende sie um 40% erhöht. Dies war grundsätzlich zu erwarten, jedoch ist die Varianz zwischen den einzelnen Spielern viel zu groß, teilweise größer als der Abstand selbst, als dass man dies als signifikant festhalten kann ($p = 0.24$ für Ohne Assistenz vs. Memory-Assistance, $p = 0.32$ für Ohne Assistenz vs. Visual-Assistance, $p = 0.09$ für Memory-Assistance vs. Visual-Assistance).

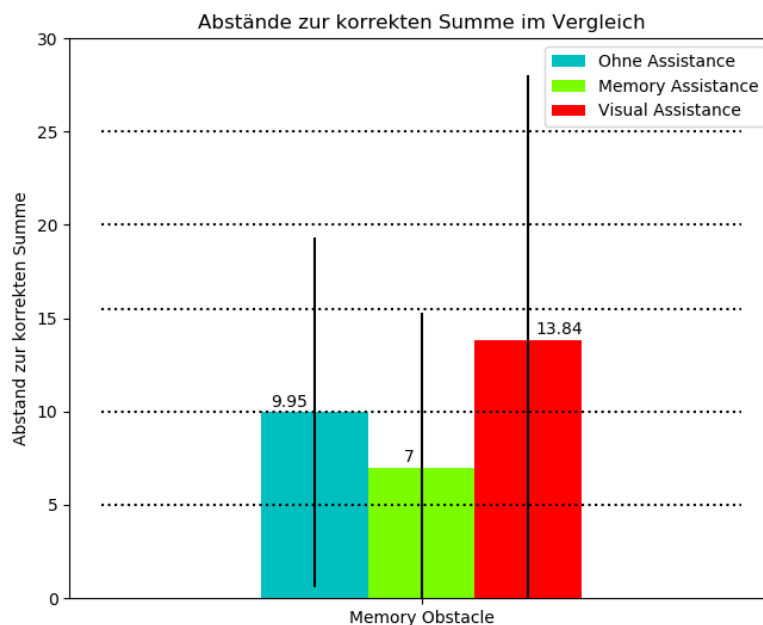


Abbildung 5.4: Errechnete Summe nach Spielmodus

Bei dem Visual-Obstacle-Modus hat sich die Memory-Assistance bei keiner der Metriken signifikant positiv oder negativ ausgewirkt ($p = 0.30, 0.23, 0.26$). Die

durchschnittliche Zeitzunahme lässt sogar vermuten, dass die Spieler die Assistenz einfach ignoriert haben und der Zeitunterschied nur durch die Animationszeit zu begründen ist. Die Visual-Assistance hingegen wirkte sich durchweg signifikant positiv aus und erreichte eine Zeitverbesserung von fast 40% ($p = <0.01$), 25% weniger Züge ($p = <0.01$) und 70% weniger Fehler ($p = <0.01$). Auch dies deckt sich teilweise mit den Ergebnissen der Fragebögen. Einerseits haben die Spieler die passende Assistenz als sehr hilfreich empfunden, was sich hier widerspiegelt, andererseits erhielt auch die Memory-Assistance gute Bewertungen, obwohl sie praktisch kaum einen Einfluss auf die Spielerleistung hatte. Bei Betrachtung aller Graphen fällt auf, dass der passend assistierte Visual-Obstacle-Modus in allen drei Metriken durchschnittlich bessere Resultate zeigte als der Classic-Modus und die Assistenz damit zu gut sein könnte. Zum einen erwiesen sich t-Tests zwischen diesen beiden Modi jedoch als negativ ($p = 0.5$, $p = 0.1$, $p = 0.25$), zum anderen zeigt sich im Classic-Modus selber durch die Visual-Assistance nicht eine solche Verbesserung.

Zusammenfassend lässt sich sagen, dass die Memory-Assistance im passenden Modus mit gering positiven, aber nicht signifikanten Auswirkungen daherkam, und im Classic-Modus in Hinblick auf die Zeit leicht störte. Damit hat sie ihren Zweck nicht hinreichend erfüllt. Die Visual-Assistance hingegen brachte im Visual-Obstacle-Modus massive Verbesserungen und hatte zusätzlich keine signifikanten Auswirkungen auf die anderen beiden Modi. Sie kann also als Erfolg angesehen werden. Beide Beobachtungen stimmen mit den Fragebögen grundsätzlich überein, allerdings wurden alle Assistenzen dort unverhältnismäßig positiv bewertet, auch wenn sie real kaum einen Einfluss hatten. Dies mag auch an der Formulierung liegen, da eine Assistenz, auch wenn sie keinen massiv positiven Einfluss ausübte, im Zweifelsfall immer eher positiv als störend empfunden wird.

6. Auswertung der EEG-Daten

In diesem Kapitel werden die während des Spiels aufgenommen Gehirnaktivitätsdaten ausgewertet. Hierbei soll die Intensität der Aktivität sowie die Lokalität untersucht werden. Im Folgenden werden dazu zunächst die relevanten Grundlagen zu Hirnarealen, Gehirnaktivitätserfassung und der damit verbundenen Datenverarbeitung erläutert.

6.1 Biologische Grundlagen

6.1.1 Hirnareale

Das Gehirn wird in seinen Grundaufgaben neben dem Stamm- und Kleinhirn in vier Bereiche, sogenannte Lappen, unterteilt (siehe Abb. 6.1), auf welchen später die Elektroden aufliegen. Diese vier Bereiche sowie ihre Funktionen werden von Lotte et al. [LBC99] wie folgt beschrieben:

Der *Frontallappen* beinhaltet unter Anderem motorische Lernfunktionen, Gedächtnis, Emotionen, Persönlichkeit, Sprache, Problembearbeitung und Planung. Für das Memory-Spiel könnten dabei natürlich das Gedächtnis, allerdings auch Problembearbeitung und Planung relevant sein.

Der *Parietallappen* beinhaltet die Verarbeitung von sensorischem Input wie Schmerzen, Temperatur und Druck, und ist ebenso für die Sprache zuständig. Dieser Lappen wird für das Memory-Spiel höchstwahrscheinlich nicht sehr relevant sein.

Im *Okzipitallappen* wird visueller Input verarbeitet und räumliches Denken durchgeführt. Für das Memory ist visuelle Verarbeitung grundsätzlich wichtig, allerdings ist es unwahrscheinlich, dass sich die Verarbeitung von Karten aus dem Rot-Grün-Set und dem braunverschobenen Set in ihrer Aktivität signifikant unterscheiden.

Der *Temporallappen* ist zuständig für die Verarbeitung von auditivem Input, Koordination und ebenso wie der Frontallappen Emotionen und Gedächtnis. Er könnte entsprechend ebenso relevant sein.

6.1.2 Gehirnaktivität

Gehirnaktivität kann auf verschiedene Arten gemessen werden. Drei Möglichkeiten stellen *Elektrokortikographie*, *funktionelle Nahinfrarotspektroskopie* und *Elektroen-*

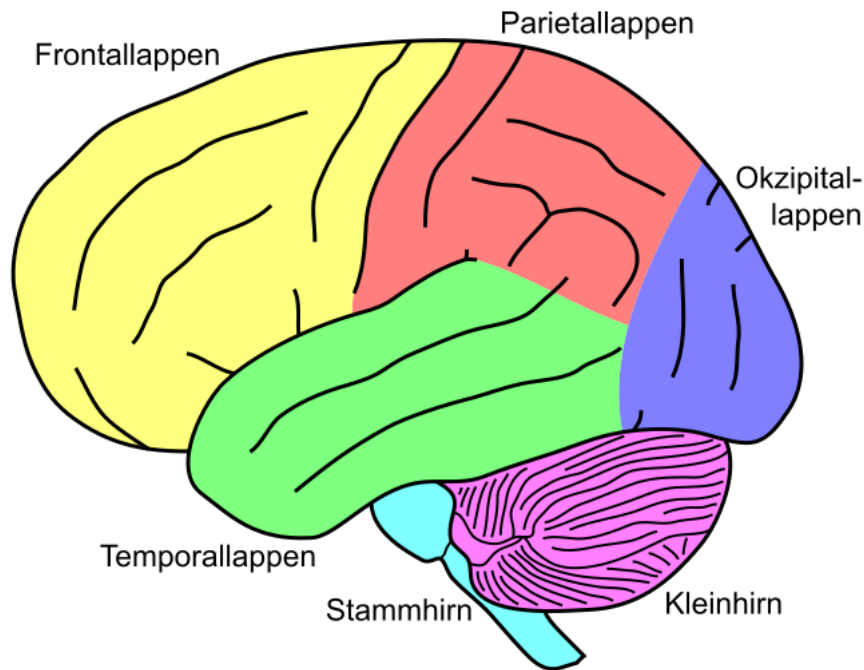


Abbildung 6.1: Die 4 Lappen des Gehirns [bra07]

zephalographie dar, welche in der Vorlesung Biosignale und Benutzerschnittstellen [Sch] folgendermaßen vorgestellt werden:

Elektrokortikographie ist ein invasives Verfahren, bei welchem ein Elektrodenarray direkt auf das offene Gehirn gelegt wird. Es ist sehr zeitpräzise und hat ein gutes Signal-zu-Rauschen-Verhältnis. Da hierfür allerdings die Schädeldecke geöffnet werden muss, ist es für die meisten Forschungszwecke unbenutzbar.

Bei der funktionellen Nahinfrarotspektroskopie (fNIRS) wird nicht-invasiv Infrarotlicht auf das Gehirn geleuchtet. Die meisten Teile des Gehirns sind dafür durchlässig, allerdings absorbiert das Sauerstoff transportierende Hämoglobin in den roten Blutkörperchen dieses Licht und liefert so ein Bild, welches aufzeigt, welche Hirnareale gerade gut durchblutet sind. Sauerstoffreiche Durchblutung von Hirnarealen ist aufgrund der neurovaskulären Kopplung ein Anzeichen für aktiviertes Nervengewebe. Dieses Verfahren ist jedoch zeitlich nicht so präzise wie die anderen Methoden, da der Blutkreislauf langsamer verläuft als die direkten elektrischen Prozesse im Gehirn. Die Elektroenzephalographie ist ebenfalls nicht invasiv. Bei ihr wird durch eine Elektrodenkappe das direkt von der kumulierten Neuronenaktivität hervorgerufene elektrische Feld gemessen [LBC99]. Da die Kappe nur auf dem Schädel, häufig auf Haaren aufliegt, ist das Signal-zu-Rauschen-Verhältnis vergleichsweise schlecht. Es besitzt außerdem eine geringe räumliche Auflösung (üblicherweise 16-128 Elektroden für die gesamte Schädeloberfläche), dafür jedoch eine gute zeitliche Auflösung von nur wenigen Millisekunden [LBC99]. Der größte Vorteil besteht in der Portabilität: Abgesehen von einem Gel, welches zugunsten der Leitfähigkeit an den Elektroden eingefüllt werden muss und einem Verstärker, an welchem passive Elektroden angeschlossen werden müssen, kann man eine EEG-Kappe sehr leicht auf- und absetzen und so Versuchsreihen ohne größeren Aufwand durchführen. Aus diesen Gründen wurde für diese Arbeit das EEG als Messverfahren gewählt.

Die Gehirnaktivität wird auf Frequenzbasis analysiert. Dabei werden zwischen fünf verschiedenen Frequenzbereichen unterschieden [Sch]. Diese werden in Tab. 6.2 dar-

Bezeichnung	Frequenzbereich	Mentaler Zustand
Delta	1-4Hz	Tiefschlaf, längere Aufmerksamkeitsaufgaben [KK16]
Theta	4-8Hz	Leichter Schlaf
Alpha	8-13Hz	Wachzustand, Augen geschlossen
Beta	13-30Hz	Augen offen, Aufmerksamer Zustand
Gamma	31-60Hz	Komplexe Aufgaben, Konzentration, hohe sensorische Signalverarbeitung [KK16]

Abbildung 6.2: Die empirische Einteilung von Hirnfrequenzen in Bänder

gestellt:

Zusätzlich zu den mentalen Zuständen in Tabelle 6.2 schreiben Kumar et al. [KK16], dass eine stärkere Ausprägung von Theta-, Alpha- und Beta- Wellen mit erhöhter mentaler Auslastung einhergeht.

6.2 Hardware Setup

Für die EEG-Datenaufnahme wurde die ActiCAP mit dem ActiCHamp-Verstärker und der PyCorder Software 1.0.9 genutzt ¹. Die 32 Elektroden auf der Kappe wurden nach dem 10-20 System [KLJ⁺99] angebracht. Abbildung 6.3 zeigt unterschiedliche Layouts für die ActiCAP, darunter das in diesem Versuch genutzte. Die in der Abbildung benutzten Bezeichner für die Elektrodenpositionen werden in der folgenden Analyse weiterverwendet. Elektrode Fz wirkte als Referenzelektrode und fällt somit aus der Messreihe heraus. Die Abtastrate der Datenaufnahme betrug 500 Hertz.

6.3 Mathematische Grundlagen

6.3.1 Diskrete Fouriertransformation

Die diskrete Fouriertransformation (DFT) ist eine mathematische Funktion, welche ein Eingangssignal von der Zeitrepräsentation in die Frequenzrepräsentation wandelt, im Falle des EEG also von $\mu V/s$ in $\mu V/Hz$. Sie basiert auf der Annahme, dass alle Signale als Überlagerung von vielen Sinusschwingungen dargestellt, und damit auch dekonstruiert werden können [Ins19b]. Da das Inputsignal häufig durch Sampling ermittelt wird, und damit diskret vorliegt, muss auch die diskrete Fouriertransformation angewandt werden, welche das Signal ebenfalls in diskrete Frequenzen (Frequenz-Bins) unterteilt. Bei der DFT wird vorausgesetzt, dass der Input genau eine Periode eines Signals ist, also Anfang und Ende stetig ineinander übergehen [Ins19b]. Dies ist allerdings bei vielen Signalen, wie auch den nachfolgenden EEG-Daten nicht der Fall. Ist der Übergang von Ende zu Anfang unstetig, kann die DFT abhängig von der Unstetigkeit Frequenzen aus dem Signal deuten, die in dem Signal eigentlich gar nicht vorkommen. Diese Schwäche wird **spectral leakage** genannt [Ins19b]. Diesem Effekt kann vorgebeugt werden, indem eine **Fensterfunktion** (siehe 6.3.2) genutzt wird. Die diskrete Fouriertransformation wird in python als rfft (real input fast fourier transform) in der numpy.fft library bereitgestellt. Dies ist eine recheneffiziente Variante für ein reelles Eingabesignal, bei welchem zudem die Symmetrie der DFT ausgenutzt wird und nur der positive Teil ausgegeben wird [Sci19].

¹Mehr Informationen dazu unter <http://www.iitg.ac.in/naveen.kashyap/pycorder.pdf>

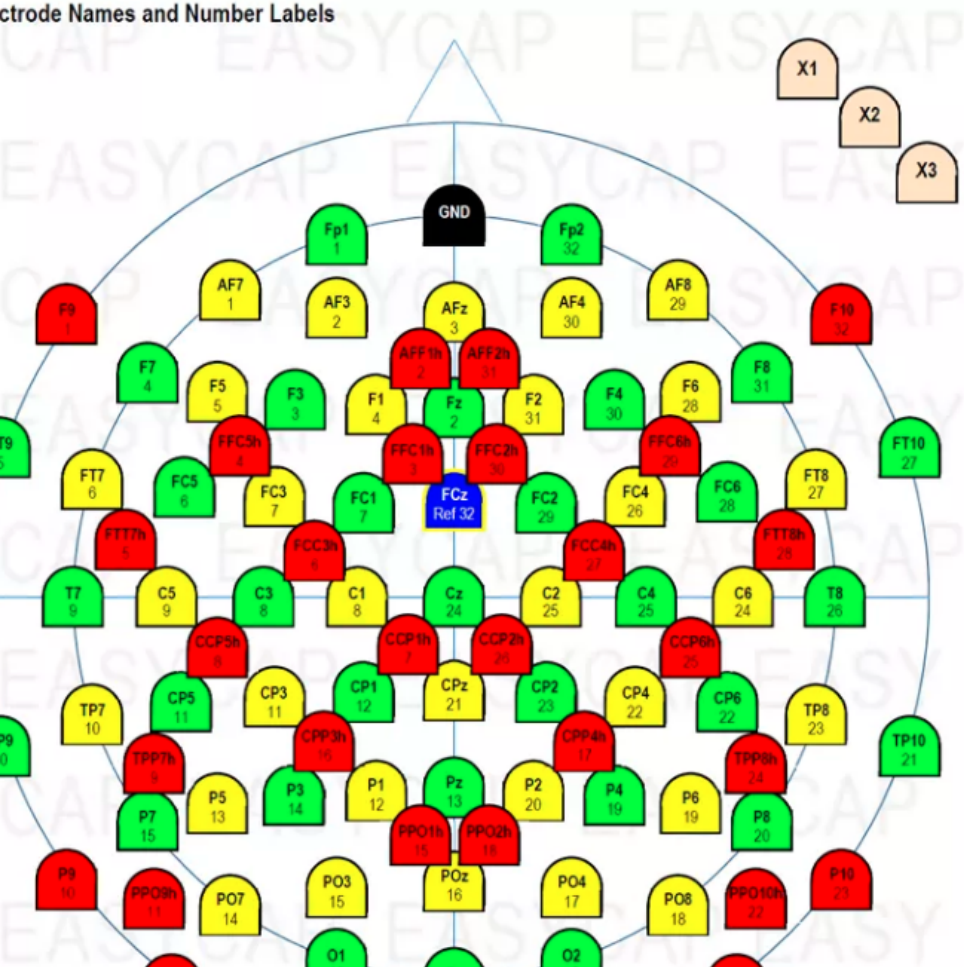


Abbildung 6.3: Die Positionen der Elektroden auf der EEG-Kappe. Es wurden nur die 32 grün eingefärbten Elektroden benutzt. Die Bezeichner F,C,T,P,O stehen dabei für Frontal, Central (Motor und Somatosensorischer Cortex), Temporal, Parietal und Okzipital. Z kennzeichnet die Mittellinie des Gehirns.

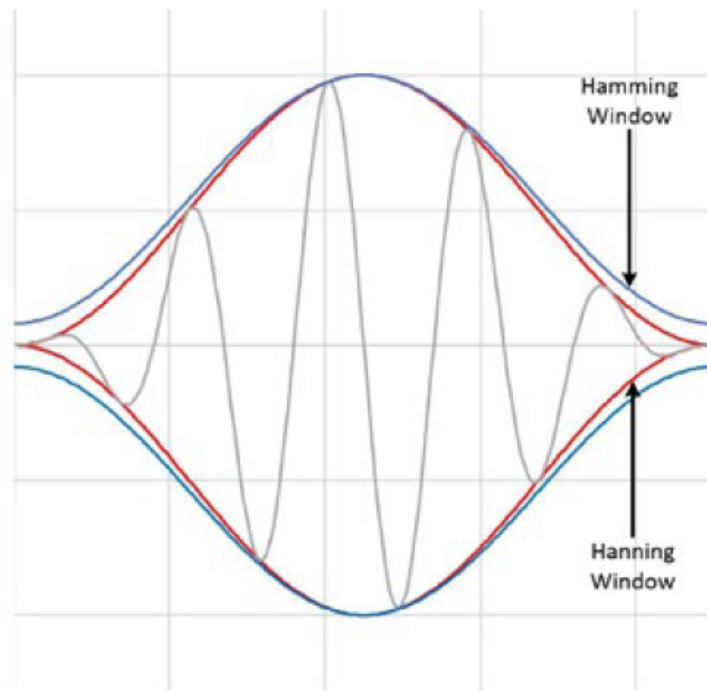


Abbildung 6.4: Die Formen der Hamming- und Hann-Fensterfunktionen. Der Unterschied besteht darin, dass die Hann-Funktion am Fensterrand den Wert 0 annimmt, die Hamming-Funktion sich dem nur annähert.
Quelle: [Ins19b]

6.3.2 Hann-Fensterfunktion

Eine Fensterfunktion ist eine Funktion auf einem Signalintervall (parametrisiert als Fenstergröße), welche schrittweise über ein Signal geschoben und angewandt wird und dabei die Amplitude mit Werten zwischen 0 und 1 multipliziert. In der Regel beträgt der Funktionsfaktor in der Mitte des Fensters 1 und nimmt in Richtung beider Enden gegen 0 ab [Ins19b]. Bei den gängigen Fensterfunktionen *Hamming* und *Hann* ist dies als Parabelform modelliert, wie in Abb. 6.4 gezeigt.

Dadurch, dass die Ränder des Fensters mit sehr kleinen Faktoren multipliziert werden, geht auch die Amplitude des Signals gegen 0, sodass sich Ende und Anfang des Fensters stetig annähern und das Signalintervall als periodisch angesehen werden kann. Somit wird der spectral leakage Effekt reduziert. Da bei der DFT sämtliche zeitkritischen Informationen verloren gehen, besteht ein weiterer Nutzen von Fensterung darin, den zeitlichen Zusammenhang wieder herzustellen, da zeitlich verschiedene Fenster auch unterschiedliche DFT-Ergebnisse liefern. In dieser Arbeit wird mit der Hann-Funktion gearbeitet. Sie zeichnet sich gerade im Vergleich zur Hamming-Funktion dadurch aus, dass die Werte an den Fensterrändern exakt 0 beträgt und die Stetigkeit somit garantiert ist.

6.4 Datenverarbeitung der EEG-Daten

Die EEG-Daten werden von der Actichamp Software als .eeg und .vmrk sowie einer dazugehörigen .vhdr-Header Datei mit Metadaten gespeichert. Diese wurden in Pythonskripten durch die mne-Bibliothek eingelesen. Da am CSL bereits vor dieser

Arbeit EEG-Daten vorheriger Memory-Spiele ausgewertet wurden, wurden für diese Datenverarbeitung Skripte aus dem memory-statistics Repository (in der digitalen Abgabe im gleichnamigen Ordner) genutzt. Im Folgenden wird erläutert, wie in den Skripten vorgegangen wird und warum.

Um die Analyse und den Vergleich der EEG-Daten zu vereinheitlichen, werden nur bestimmt lange Ausschnitte unmittelbar nach dem Umdrehen einer Karte beachtet. Um diese Zeitpunkte zu finden, wurde bei der Versuchsreihe ein Lichtsensor, ebenfalls am ActiCHamp angeschlossen, benutzt, welcher parallel zur EEG-Datenaufnahme einen Lichtimpuls bei jedem Kartenumdrehen gemessen hat. Anschließend wurde ein Schwellwert ermittelt, welcher einen solchen Lichtimpuls (*light onset*) in den Lichtsensordaten klassifiziert. Die Indizes der light onsets wurden dann wiederum auf die EEG-Daten übertragen und diese in einem Zeitfenster von einer Sekunde betrachtet und gespeichert. Zunächst wurden diese Hirnaktivitätsdaten normalisiert, also der Mittelwert der Daten von allen Datenpunkten abgezogen. Um aus diesen Daten anschließend die vorliegenden Frequenzen zu bestimmen, wird die Fouriertransformation benutzt. Dafür wird zuvor die Hann-Fensterfunktion auf das Signal angewandt (siehe Kap. 6.3.2). Abschließend wird das Signal logarithmiert, um die eventuelle Schiefe der Daten auszugleichen, also das Verhältnis einzelner stark ausgeprägter Datenpunkte gegenüber dem Rest zu verringern. Nach dieser Datenverarbeitung habe ich für die paarweisen Spielmodusvergleiche zwei weitere Skripte geschrieben, in welchen der verarbeitete Datensatz auf einzelne Elektrodenkanäle beschränkt wird und Bandpassfilter angewendet werden. Dabei erzeugt eines Mittelwerte aller Kartenflips und graphische Vergleiche sequenziell für alle Spiele einzeln, das andere die Mittelwerte und Graphen aller Spieler zusammen. Um die Klassifizierungsfähigkeit quantitativ zu ermitteln, wurden außerdem erneut gepaarte t-Tests mit $p \leq 0.05$ durchgeführt.

Anmerkung zur Datenverarbeitung: Von den 19 Messreihen wurde eine aufgrund zu starken Rauschens verworfen, darüber hinaus konnten zu einem Spieler bei dem Visual-Obstacle-Modus keine sinnvollen light-onsets ermittelt werden, wahrscheinlich weil sich der Lichtsensor während des Experiments unbemerkt gelöst oder verschoben hat. In den späteren t-Tests gibt es daher je nach Spielmodus nur 18 oder 17 Iterationen.

6.5 Relevante Frequenzbereiche und Elektrodenpositionen aus der Literatur

Zunächst wurde ein Bandpassfilter von 0-40Hz auf die Daten angewandt, um Artefakte wie Netzbrummen zu isolieren ([KK16]). Da der Featureraum mit 31 Elektroden mal 40 Frequenzen pro Kartenflip jedoch zu groß ist, um Signifikanztests auf den gesamten Daten durchzuführen, fasse ich im Folgenden Ansätze aus der Literatur zusammen, welche sich mit der geistigen Auslastung beschäftigt haben:

Kumar et al. [KK16] untersuchten zwei Stufen von mentaler Auslastung mit einer Multiplikationsaufgabe von jeweils zwei einstelligen und zwei dreistelligen Zahlen. Dabei erreichten sie an den Elektrodenpositionen AF3, AF4, T7 und T8 die besten Ergebnisse. Die Positionen AF3 und AF4 sind in meinem Layout nicht vorhanden, liegen aber sehr nahe bei den Positionen FP1 und FP2 (siehe 6.3). In ihrem Experiment erwies sich das Alpha-Band (8-13Hz) als signifikantestes Klassifikationsmerkmal,

Bezeichnung	Frequenzbereich	Areal
Delta	1-4Hz	Temporallappen
Theta	4-8Hz	Frontallappen
Alpha	8-13Hz	Temporal-, Okzipitallappen, bei komplexen Aufgaben auch Frontallappen.
Beta	13-30Hz	Temporal-, Okzipitallappen.
Gamma	31-60Hz	Zentrale Hirnregion (C-Elektroden),

Abbildung 6.5: Ausgeprägte Hirnregionen pro Band nach Kumar et al.

jedoch schrieben sie, dass Beta-Frequenzen ebenfalls ein häufiges Klassifikationsfeature für mentale Auslastung sind. In ihrer Tabellenübersicht zu Frequenzbändern, welche ich in meine Zusammenfassung (Abb. 6.2) miteinfließen lassen habe, erwähnen sie außerdem zu jedem Frequenzbereich jeweils Areale, in welchen diese besonders prominente Aktivität aufzeigen (siehe Abb. 6.5).

Klimesch [Kli99] beschreibt ebenfalls eine Korrelation zwischen kognitiver Auslastung und einer Kombination aus Alpha- und Theta-Aktivität. Des Weiteren erwähnt er, dass Alpha-Aktivität sehr individuell ausgeprägt ist und vom Alter sowie der Gedächtnisleistung beeinflusst wird.

Dolce et al. [DW74] führten ein Experiment mit lese- und arithmetischen Aufgaben durch und fanden hingegen bei letzteren einen Zusammenhang zu erhöhter Delta-Aktivität in Parietal- und Okzipitalregionen.

Dimitriadis et al. zeigen ebenfalls in zwei arithmetischen Versuchsreihen (Zahlen vergleichen und multiplizieren) einen Zusammenhang zwischen mentaler Auslastung und erhöhter Aktivität im Delta-Bereich [DLT⁺10].

Fernandez et al. [FHR⁺95] untersuchten die EEG-Aktivität von Probanden anhand von vier verschiedenen arithmetischen Aufgaben im Vergleich zur Ruheaktivität. Sie erhielten signifikante Ergebnisse im frontalen Delta- (1,5-3,5Hz) und Betabereich (12,5-19Hz) und beschrieben eine generelle Abnahme im Alphabereich bei den Aufgaben, bemerkten dabei aber wie Klimesch [Kli99] eine erhöhte Varianz in diesem Bereich.

Basierend auf den in Kapitel 6.1.1 vorgestellten verschiedenen Zugehörigkeitsbereichen der Hirnlappen, von welchen besonders der Frontal- und der Temporallappen relevant sind, sowie den Erkenntnissen von Kumar [KK16], Fernandez [FHR⁺95] und Dimitriadis [DLT⁺10], werden folgende Features im Anschluss näher behandelt:

- Delta-Aktivität in der Frontalregion (Fp1, Fp2) und Temporalregion (T7, T8) pro Spieler
- Delta-Aktivität in der Region gemittelt über alle Spieler
- Alpha-Aktivität in der Frontalregion (Fp1, Fp2) und Temporalregion (T7, T8) pro Spieler
- Alpha-Aktivität in den Regionen gemittelt über alle Spieler

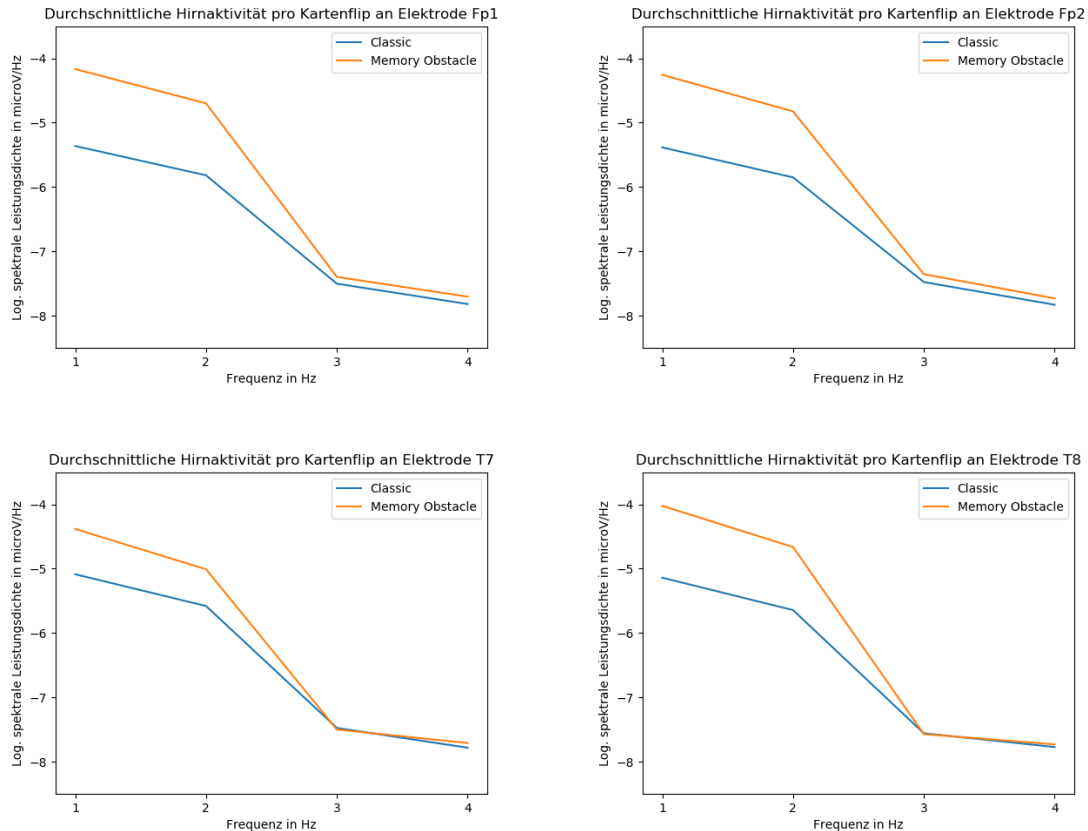


Abbildung 6.6: Die durchschnittliche Delta-Aktivität über alle Spieler gemittelt für Classic- vs. Memory-Obstacle-Modus

6.6 Vergleich der EEG-Daten

Die EEG-Daten werden analog zu den Leistungsdaten paarweise verglichen. Ziel soll es sein, zu untersuchen, ob je zwei Datensätze binär unterscheidbar sind. Dafür werden 3 Vergleiche durchgeführt: *Classic-Modus vs. Memory-Obstacle*, *Classic-Modus vs. Visual-Obstacle* und *Memory-Obstacle vs. Visual-Obstacle*.

6.7 Delta-Aktivität

6.7.1 Classic-Modus vs. Memory-Obstacle

Diese Vergleiche wurden an den Daten der Elektroden Fp1, Fp2, T7 und T8 im Frequenzbereich 1-4Hz durchgeführt.

Betrachtet man zunächst den Mittelwertsgraphen über alle Spieler (Abb. 6.6) so sieht man an allen 4 Elektrodenpositionen grundsätzlich höhere Aktivität im Frequenzbereich 1-3Hz bei dem Memory-Obstacle gegenüber dem Classic-Modus.

Mittelt man allerdings alle Daten, so fallen Spieler mit starken Aktivitätsunterschieden auch besonders ins Gewicht, wohingegen bei anderen eventuell gar keine eindeutige Diskriminanz vorhanden war. Dies zeigt Abb. 6.7 beispielhaft anhand der

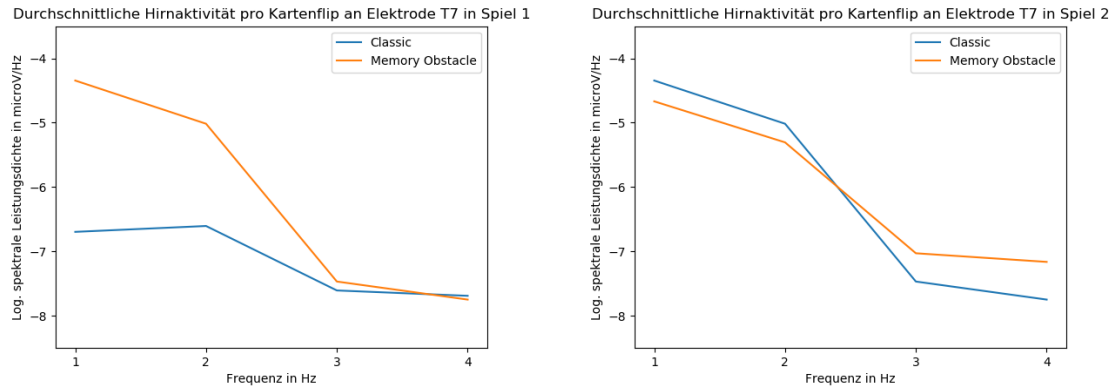


Abbildung 6.7: Die temporale Delta-Aktivität an Elektrode T7: Links deutlich unterscheidbar, rechts verwoben.

Daten von Spieler 1 und 2 an der Elektrode T7. Mittelwertsgraphen werden im Folgenden weiter zur allgemeinen Veranschaulichung gezeigt, um jedoch zu untersuchen, ob die Spielmodi anhand von Hirnaktivität klassifizierbar sind, muss dies aus oben genanntem Grund pro Spieler nachgewiesen werden.

Dafür wurde für jeden Spieler ein gepaarter t-Test seiner ersten 20 Kartenflips zwischen dem Classic-Modus und dem Memory-Obstacle-Modus durchgeführt. Dies wurde für alle vier Elektrodenpositionen individuell geprüft. Bei einer Kette von mehreren t-Tests (hier: 18 Spieler) kann allerdings der Fehler der *Alphafehler-Kumulierung* auftreten. Dieser entsteht dadurch, dass bei einer zunehmenden Anzahl an konsekutiven Tests sich die Chance multipliziert, trotz 95% Konfidenz eine Nullhypothese fälschlicherweise zu verwerfen oder zu akzeptieren. Deswegen müssen bei mehreren t-Tests die p-Werte angepasst werden. Naiv lässt sich dies mit der Bonferroni-Korrektur erreichen, welche den p-Schwellwert durch die Anzahl der Tests teilt [Abd07]. Dies ist ein strenges Kriterium und daher sinnvoll, wenn man False-Positives unbedingt vermeiden muss. Es erhöht dabei jedoch ebenso stark die Chance, dass eine Nullhypothese falsch akzeptiert wird, also ein False-Negative erzeugt wird ([Abd07], [Cho19]). Ein weicherer, verbreiteter Ansatz ist die Benjamini-Hochberg-Korrektur, welche die p-Werte aufsteigend sortiert und ebenfalls anhand der Anzahl der Tests anpasst, wobei allerdings größere Werte zunehmend schwächer angepasst werden [BH95]. Da ich in dieser Arbeit False-Negatives stärker gewichte als False-Positives, nutze ich die Benjamini-Hochberg-Korrektur. Eine Implementation dieser Korrektur wurde durch die *statsmodels*-Bibliothek in Python bereitgestellt [SP10]. Die Ergebnisse sind in Tab. 6.8 zusammengeführt. Die Tests fallen auch nach der Korrektur noch überwiegend positiv aus, besonders dabei an den Frontalelektroden. Über beide Elektroden gemittelt ergaben etwa 75% der t-Tests signifikante Ergebnisse. Anhand dessen halte ich diese beiden Spielmodi in diesem Frequenzbereich für klassifizierbar.

Elektrodenposition	Signifikanz bei n Spielern	nach Ben.-Hoch.-Korrektur
Fp1	15/18	15/18
Fp2	13/18	12/18
T7	13/18	12/18
T8	11/18	10/18

Abbildung 6.8: Ergebnisse der gepaarten t-Tests für Classic- vs. Memory-Obstacle im Delta-Band

6.7.2 Classic-Modus vs. Visual-Obstacle

Vergleicht man den Classic-Modus mit dem Visual-Obstacle-Modus, sind die Graphen der einzelnen Spieler etwas verwobener, zeigen jedoch im Mittel (siehe Abb. 6.10) insgesamt ähnliche Trends wie im vorherigen Vergleich auf. Die Ergebnisse der t-Tests finden sich in Tab. 6.9. Sie zeigen deutlich weniger Signifikanz als beim Memory-Obstacle Vergleich (etwa 40% an den Frontalelektroden), allerdings lag bei diesem Spielmodus auch keine Nebenaufgabe vor, die Schwierigkeit der Grundaufgabe hat sich lediglich erhöht. Eine Klassifikation könnte noch möglich sein.

Elektrodenposition	Signifikanz bei n Spielern	nach Ben.-Hoch.-Korrektur
Fp1	8/17	7/17
Fp2	7/17	7/17
T7	6/17	5/17
T8	6/17	5/17

Abbildung 6.9: Ergebnisse der gepaarten t-Tests für Classic- vs. Visual-Obstacle im Delta-Band

6.7.3 Memory-Obstacle vs. Visual-Obstacle

Die Modi Memory-Obstacle und Visual-Obstacle lassen sich wieder deutlicher voneinander abgrenzen, wie in Abb. 6.12 zu sehen ist. Die Ergebnisse der t-Tests sind in Tab. 6.11 zusammengefasst und fallen ebenfalls wieder signifikanter aus (etwa 64% an den Frontalelektroden). Das Memory-Obstacle weist demnach im Delta-Band die deutlichsten Unterschiede in der Hirnaktivität aus.

Elektrodenposition	Signifikanz bei n Spielern	nach Ben.-Hoch.-Korrektur
Fp1	10/17	10/17
Fp2	12/17	12/17
T7	8/17	7/17
T8	10/17	10/17

Abbildung 6.11: Ergebnisse der gepaarten t-Tests für Memory-Obstacle- vs. Visual-Obstacle im Delta Band

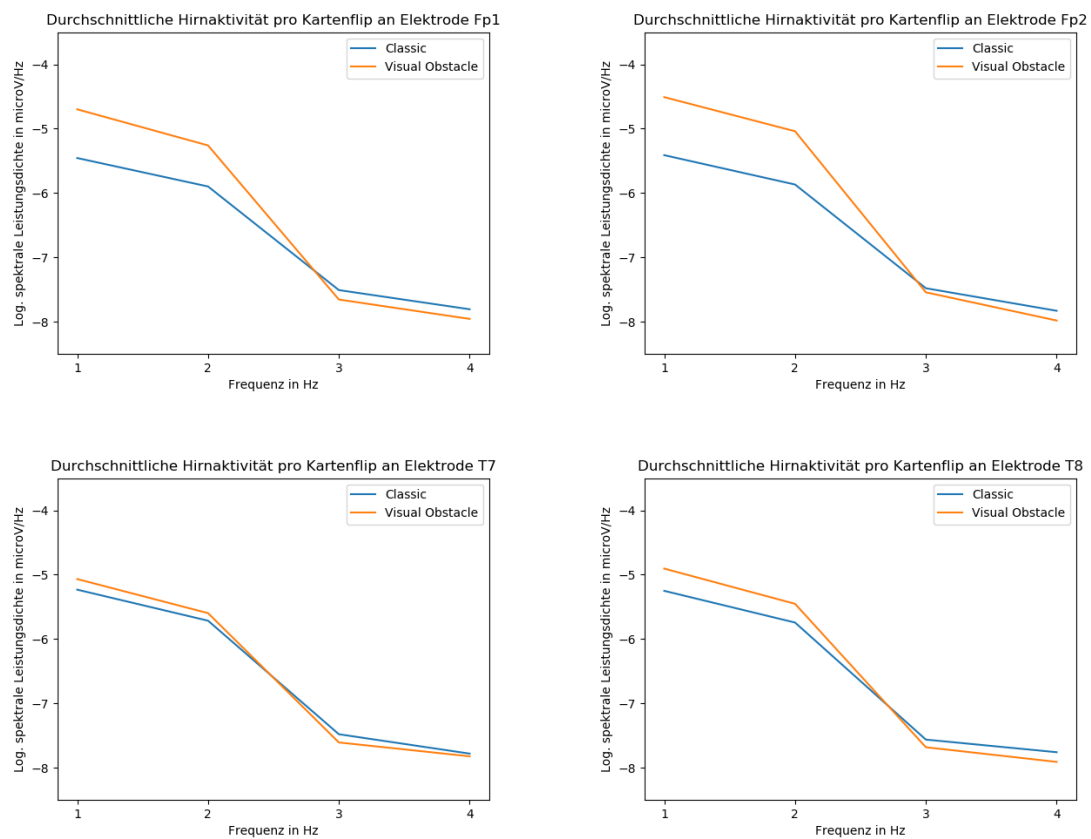


Abbildung 6.10: Die durchschnittliche Delta-Aktivität über alle Spieler gemittelt für Classic- vs. Visual-Obstacle-Modus

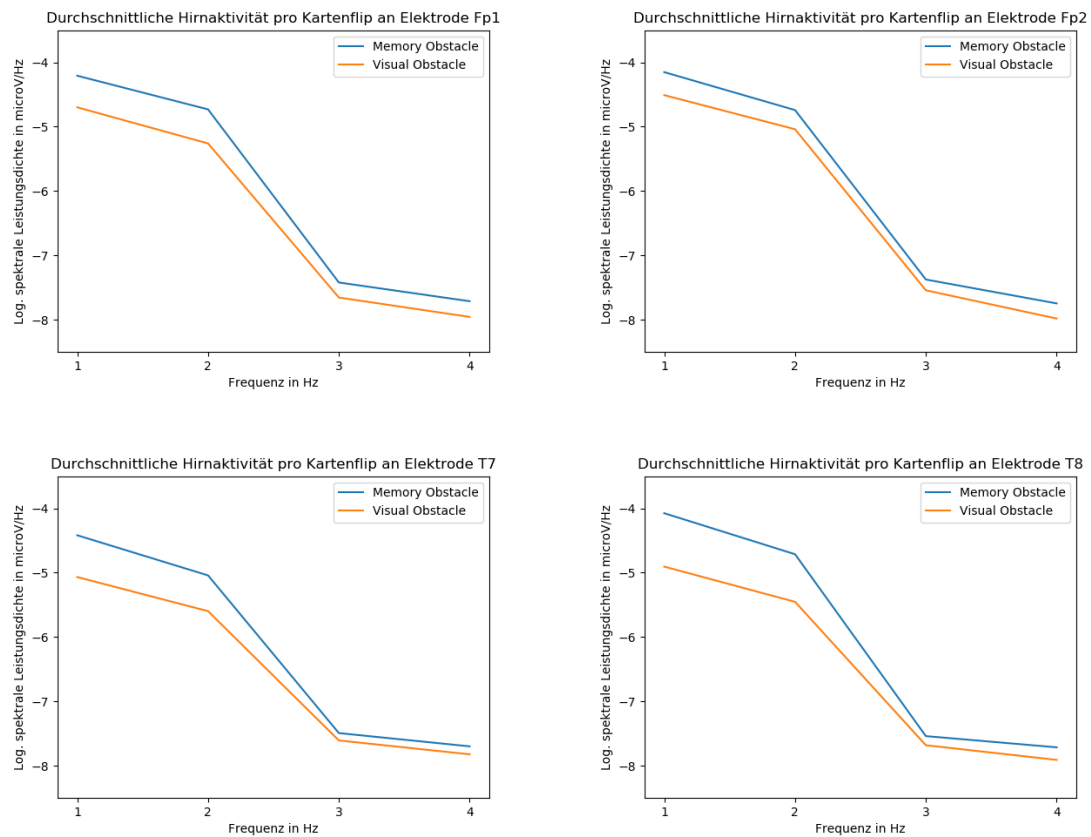


Abbildung 6.12: Die durchschnittliche Delta-Aktivität über alle Spieler gemittelt für Memory-Obstacle- vs. Visual-Obstacle-Modus

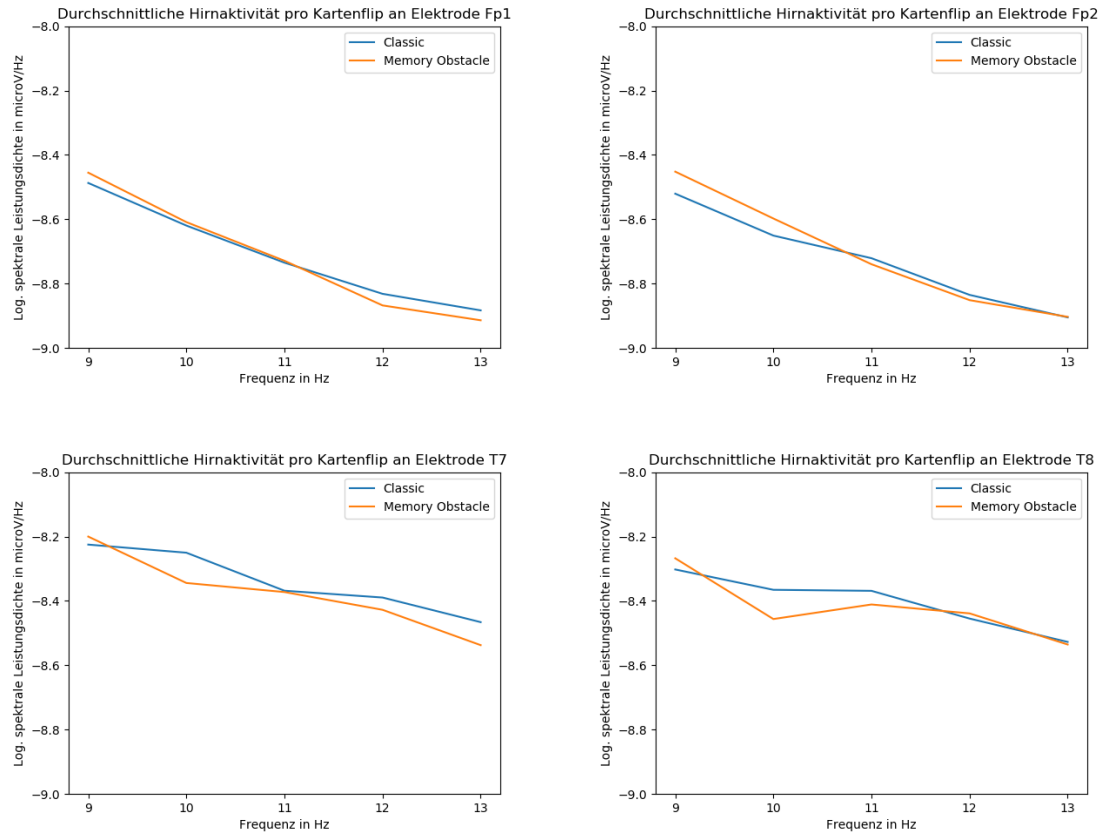


Abbildung 6.13: Durchschnittliche frontale (oben) und temporale (unten) Alpha-Aktivität über alle Spieler gemittelt für Classic- vs. Memory-Obstacle-Modus. Die Daten sind vollständig verwoben.

6.8 Alpha-Aktivität

Mit gleicher Vorgehensweise wurde anschließend die Alpha-Aktivität an den Elektrodenpositionen Fp1, Fp2, T7 und T8 geplottet und verglichen. Abb. 6.13 zeigt diesen Vergleich für den Classic- und Memory-Obstacle-Modus. Gemittelt über alle Spieler sind die Aktivitäten sehr ähnlich und verwoben. Die beiden anschließenden Average-Graphen für die Vergleiche Classic- und Visual-Obstacle-Modus (siehe Abb. 6.14), sowie Memory-Obstacle und Visual-Obstacle-Modus (siehe Abb. 6.15) legen zwar leicht unterschiedliche frontale Aktivität nahe, die temporalen Elektroden T7 und T8 zeigen jedoch erneut verwobene Aktivität. Die t-Tests zu allen Vergleichen wurden in der Tabelle 6.16 zusammengefasst. Im Vergleich zeigten die frontalen Elektroden zunächst vereinzelt signifikante Ergebnisse. Nach der Benjamini-Hochberg-Korrektur werden die Signifikanzkriterien jedoch fast gar nicht mehr erfüllt. Insgesamt liefert das Alpha-Band deutlich schlechtere Ergebnisse als das im vorherigen Kapitel untersuchte Delta-Band. Eine allgemeine Klassifikation anhand dieser Merkmale ist nicht möglich.

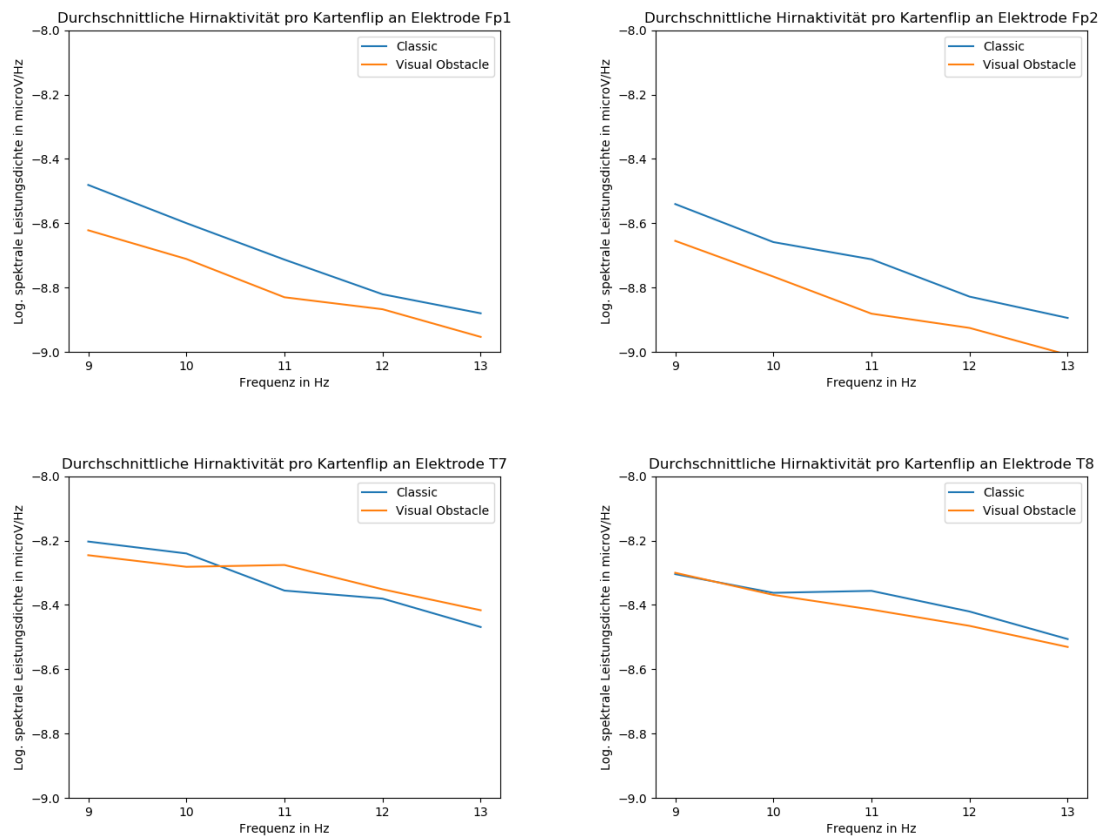


Abbildung 6.14: Durchschnittliche frontale (oben) und temporale (unten) Alpha-Aktivität über alle Spieler gemittelt für Classic- vs. Visual-Obstacle-Modus. Frontal sind die Daten leicht unterschiedlich, temporal jedoch verwoben.

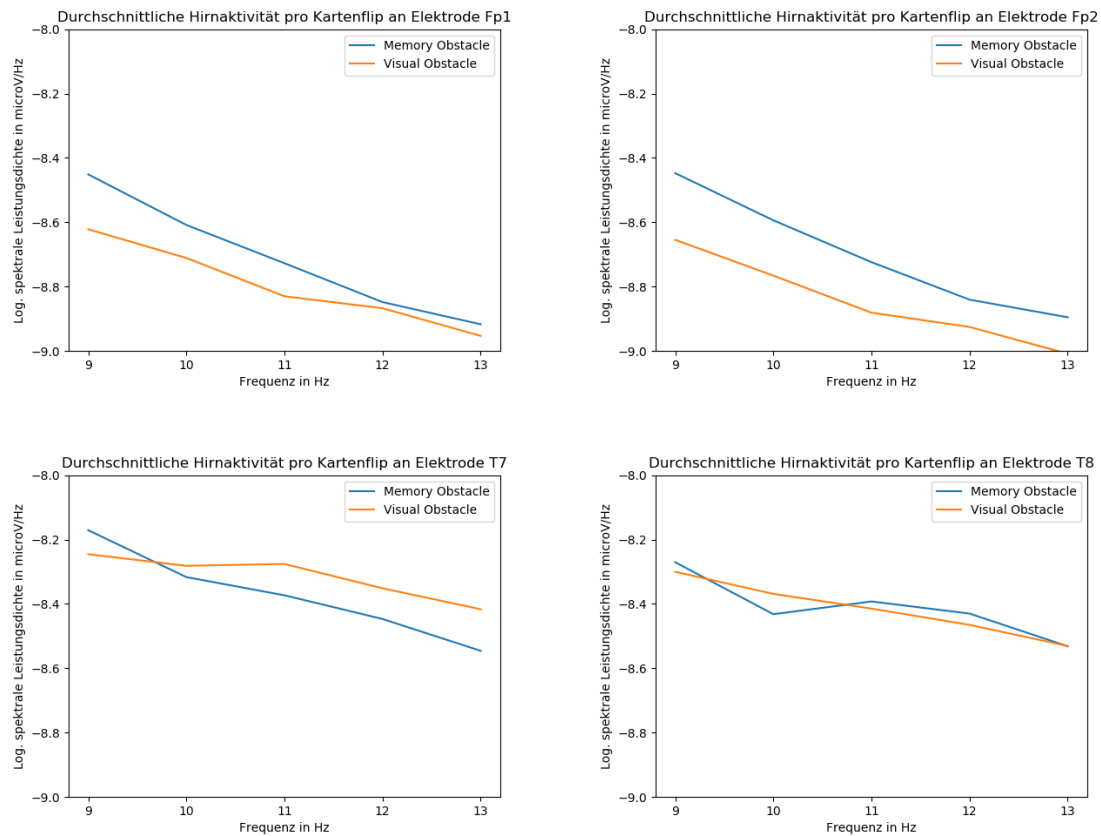


Abbildung 6.15: Durchschnittliche frontale (oben) und temporale (unten) Alpha-Aktivität über alle Spieler gemittelt für Memory-Obstacle- vs. Visual-Obstacle-Modus. Erneut zeigt nur die Frontalaktivität Unterschiede auf.

Spielmodus	Elektrode	Erfolgreiche t-Tests	nach Ben.-Hoch. Korrektur
Classic vs. Mem.-Obs.	Fp1	4/18	0/18
Classic vs. Mem.-Obs.	Fp2	1/18	0/18
Classic vs. Mem.-Obs.	T7	2/18	0/18
Classic vs. Mem.-Obs.	T8	1/18	0/18
Classic vs. Vis.-Obs.	Fp1	2/17	1/17
Classic vs. Vis.-Obs.	Fp2	4/17	0/17
Classic vs. Vis.-Obs.	T7	1/17	0/17
Classic vs. Vis.-Obs.	T8	1/17	0/17
Mem.-Obs. vs. Vis.-Obs.	Fp1	2/17	2/17
Mem.-Obs. vs. Vis.-Obs.	Fp2	5/17	4/17
Mem.-Obs. vs. Vis.-Obs.	T7	1/17	0/17
Mem.-Obs. vs. Vis.-Obs.	T8	0/17	0/17

Abbildung 6.16: Ergebnisse der t-Tests im Alpha-Band

7. Ergebnisse

Die zu Beginn dieser Arbeit aufgestellten Fragen und Hypothesen wurden in den letzten Kapiteln durch die statistischen Auswertungen der Fragebögen, der Nutzerdaten und der Hirnaktivität untersucht und einzeln beantwortet:

Frage 1:

Wirken sich die Hindernisse im Memory-Spiel auf die Spielerdaten aus?

Hypothese 1:

Die Hindernisse wirken sich negativ auf die Spielerleistung, die User Experience und die mentale Auslastung aus.

Ergebnis zu 1:

Beide Hindernisse wirkten sich negativ auf die Spielerleistung und die Spielerfahrung aus. Die Fragebogenanalyse gab einen klaren Trend, dass der klassische Modus am leichtesten wahrgenommen wurde, gefolgt vom Visual-Obstacle-Modus, und der Memory-Obstacle-Modus am schwierigsten. In den Gamelogs konnte dieser Trend teilweise bestätigt werden, immer jedoch waren die Leistungen mit Hindernis schlechter als ohne. Die Hirnaktivität unterlag Schwankungen je nach Person, was allgemeingültige Aussagen erschwerte, jedoch zeigten die Mittelwertsgraphen im Delta-Band, dass die Aktivität in beiden Hindernismodi über der des Classic-Modus lag. Im untersuchten Alpha-Band konnte sich die Hypothese allerdings nicht bestätigen.

Frage 2:

Wie wirken sich die Adaptionen auf die verschiedenen Spielmodi aus?

Hypothese 2:

Die Adaptionen wirken sich im passenden Hindernismodus positiv auf die Spielerleistung und die User Experience aus. In den anderen beiden Modi machen sie keinen signifikanten Unterschied.

Ergebnis zu Frage 2:

Die Memory-Assistance wurde von den Spielern im Fragebogen zwar positiv aufgenommen, erzielte aber weder in dem dafür vorgesehenen Modus, noch in den anderen beiden signifikante Veränderungen, bis auf eine Zeitzunahme im Classic-Modus, welche wahrscheinlich mit der Animation zu begründen ist. Diesbezüglich konnte die Hypothese nicht bestätigt werden.

Die Visual-Assistance hat hingegen wie vermutet funktioniert: Sie wirkte sich stark positiv auf die Leistungen im Visual-Obstacle-Modus aus und hatte darüber hinaus keinen signifikanten Einfluss auf die anderen beiden Modi. Gemäß der Fragebögen wurde sie im Memory-Obstacle-Modus als störend aufgenommen, was durch die konkurrierende Sprachausgabe zu erklären ist. Im Classic-Modus wurde sie jedoch wieder positiv bewertet, obwohl sie keinen statistischen Vorteil erbrachte. In den Fragebögen kann demnach auf Basis der Gamelogs ein gewisses Response-Bias (Vgl. [Mos08]) nicht ausgeschlossen werden.

Frage 3:

Kann man anhand der Verhaltens- und Hirnstromdaten eines Spielers Rückschlüsse auf den gespielten Modus schließen, um eine Adaption anzubieten?

Hypothese 3:

Die Auswirkungen der Hindernisse sind in beiden Datenreihen groß genug, um zwischen Modus mit und ohne Hindernis zu unterscheiden. Darüber hinaus lassen sich beide Hindernismodi voneinander abgrenzen, sodass durch Verkettung ein Rückschluss möglich ist.

Ergebnis zu Frage 3:

Die Metriken benötigte Zeit, benötigte Züge und Fehler ließen sich im Classic-Modus signifikant zu denen der beiden Hindernismodi abgrenzen. Es lässt sich also klassifizieren, ob ein Hindernis vorliegt. Danach müsste man einen weiteren binären Klassifikationstest zwischen beiden Hindernismodi anfügen, um die passende Adaption anzubieten. Zwischen beiden Hindernismodi konnte in dieser Auswertung allerdings nur anhand der Zeit signifikant unterschieden werden. Eine Klassifizierung auf Basis der Verhaltensdaten kann außerdem nur nach einem Durchgang, und nicht in Echtzeit erfolgen, es müssten also mehrere Durchgänge durchgeführt werden, bevor der Nutzer von einer Assistenz profitiert.

Die Unterschiede der mentalen Auslastung wurden auf zwei Literaturbeispielen jeweils im Alpha- und Delta-Band an vier ausgewählten Positionen durchgeführt. Die Aktivität des Alpha-Bands unterlag starken Schwankungen sowohl zwischen den Spielern als auch den Zügen Einzelner. In diesem Bereich konnte nicht zwischen den Modi unterschieden werden. Im Delta-Band hingegen ließ sich der Memory-Obstacle-Modus deutlich vom Classic-Modus, sowie in über der Hälfte der Fälle signifikant vom Visual-Obstacle-Modus abgrenzen. Damit könnte demnach eine sinnvolle Klassifikation erfolgen. Diese Tests wurden allerdings für alle Spieler einzeln durchgeführt und erfordern so für Echtzeitklassifikation bereits Messdaten von vorherigen Experimenten derselben Person und gleichen Umständen. Für eine einheitliche Klassifikation auf mittelwertsbasierten Daten waren die Schwankungen zwischen den Spielern zu groß. Die Individualität ist jedoch ein übliches Problem bei EEG-Klassifikation [LBC99].

8. Fazit

In dieser Arbeit wurde ein Memory-Spiel um ein farbiges Kartenset, ein visuelles Hindernis in Form einer Rot-Grün-Schwäche, eine passende Assistenz mit Sprachausgabe, und eine neue Assistenz für das gedächtnisbasierte Hindernis erweitert, welche den letzten Zug des Spielers erneut aufdeckt. Die Wirksamkeit dieser Assistenzen wurde anschließend experimentell untersucht. Dazu wurde eine Versuchsreihe mit 19 Probanden durchgeführt, wobei Spieldaten und Hirnaktivität mittels EEG gesammelt wurden und die Probanden darüber hinaus Fragebögen ausfüllten. Auf Basis dieser Daten wurden statistische Tests durchgeführt, um zu prüfen, ob die Hindernisse den Spieler merklich einschränken, die passenden Assistenzen diese Einschränkungen aushebeln, unpassende Assistenzen sich neutral auswirken und ob die Hindernisse einen Einfluss auf die Hirnaktivität der Probanden hat. Dabei wurde der EEG-Featureraum auf acht aus der Literatur genommenen Features beschränkt. In den Ergebnissen war zunächst zu sehen, dass die Hindernisse die Leistung der Spieler sinnvoll negativ beeinflussen und diese Unterschiede statistisch signifikant sind. Bezüglich der Assistenzen fiel auf, dass die subjektive Einschätzung der Probanden nicht immer mit den objektiven Ergebnissen übereinstimmte. Assistenzen wurden tendenziell gut bewertet, jedoch erfüllte nur die Assistenz für das Visual-Obstacle signifikante Verbesserungen der Leistung in diesem Spielmodus. Die Memory-Assistance war nicht gut genug, um das passende Memory-Obstacle auszuhebeln. Die Hirnaktivität wurde im Delta- und Alpha-Band an vier Elektrodenpositionen untersucht. Im Delta-Band konnte eine durchschnittliche Zunahme der Aktivität bei beiden Hindernissen erkennen, welche beim Memory-Obstacle stärker ausgeprägt als beim Visual-Obstacle war. Dies wurde anhand von spielerindividuellen t-Tests untersucht. Im Alpha-Band konnten aus den Ergebnissen keine sinnvollen Resultate erkannt werden. Die Daten waren verwoben und zeigten weder individuell noch im Mittel einen deutlichen Trend. Insgesamt ist zu beachten, dass für eine EEG-Studie der Experimentumfang eher zu klein war. Die großen Leistungsunterschiede der Probanden im Visual-Obstacle-Modus mit und ohne passender Assistenz zeigen die Wichtigkeit von adaptiven Systemen beim Design von Mensch-Maschine-Schnittstellen. Diese Arbeit bietet eine Basis, ein System zu schaffen, welches nach Ansammlung von benutzerindividuellen Daten eine Klassifikation vornehmen kann, in welchem Modus sich der Benutzer befindet und dementsprechend eine Assistenz automatisch einschalten kann.

8.1 Ausblick

In der Auswertung der Gamelogs könnte die Zeit pro Zug ein weiteres Klassifikationsfeature sein. Im Memory-Obstacle-Modus fällt diese sehr wahrscheinlich größer aus als in den Anderen, da die Spieler üblicherweise erst eine neue Karte umdrehen, wenn sie die neue Summe fertig berechnet haben. Dieses Feature ist allerdings nur sinnvoll messbar, wenn in den Gamelogs alle Züge, nicht nur die ersten 20, gespeichert werden.

In der EEG-Analyse dieser Arbeit wurden einzelne aus der Literatur genommene Features untersucht, der Featureraum war jedoch noch deutlich größer. Anhand der verhältnismäßig kleinen Experimentgröße von 17 bzw. 18 Personen ist es zwar nicht sinnvoll, gezielt nach den besten Features zu suchen und daraus neue wissenschaftliche Annahmen zu formulieren, es ist jedoch möglich, dass es bessere Features als die in dieser Arbeit untersuchten gibt. Zum einen kann man Bandgrenzen verschwimmen lassen und auf eine Mischung verschiedener einzelner Frequenzen setzen, zum anderen kann man für jeden Spielmodusvergleich, oder sogar für jede einzelne Versuchsperson individuelle Features und Elektrodenpositionen verwenden.

Davon abgesehen lässt sich die individuelle Klassifikationsfähigkeit sicher steigern, wenn die Spieler mehrere Durchgänge aller Spielmodi direkt hintereinander bzw. in ähnlichen Bedingungen ausführen.

Höhere Sicherheit bei der Klassifikation könnte man ebenfalls erreichen, indem man das Signal vor der Verarbeitung mehr von Artefakten trennt. Diese können z.B. durch eine Independent Component Analysis ermittelt werden und anschließend vom Signal abgezogen werden [DSM07].

In der EEG-Auswertung erwies sich die frontale Aktivität an Elektrode Fp1 und Fp2 durchweg als wirkungsvollstes Feature. Dies kann damit zusammenhängen, dass die Elektrode Fz, welche zu beiden benachbart liegt, als Referenzelektrode gewählt wurde, und die Ergebnisse so eventuell einen Bias vorweisen. Eine zentralere Referenzelektrode könnte die Ergebnisse in den Temporalregionen verbessern.

Insgesamt ließen sich jedoch trotzdem messbare Unterschiede in den Spielerleistungen und der Gehirnaktivität der Spieler finden. Für eine Klassifizierung auf Basis dieser Unterschiede bieten sich Verfahren wie Linear Discriminant Analysis oder Support Vector Machines an, welche für EEG-Daten in der früheren Version dieses Spiels bereits gute Ergebnisse lieferten (siehe [Kla]). Es wird jedoch empfohlen, den Messdatenbereich zunächst durch zusätzliche Versuchsreihen zu vergrößern.

A. Anhang

Im Folgenden werden weiterführende Dokumente und Grafiken eingefügt, die Teil dieser Arbeit waren. Des Weiteren wird noch eine Suche nach den Elektroden mit höchster Aktivitätsdifferenz pro Band für diesen Versuch durchgeführt.

A.1 Versuchsdaten

Die Ergebnisse der Fragebögen als Google-Spreadsheet:

<https://bit.ly/2mY9K3x>

Die Ergebnisse der Gamelogs als Google-Spreadsheet:

<https://bit.ly/2nbbarp>

A.2 Versuchsdokumente

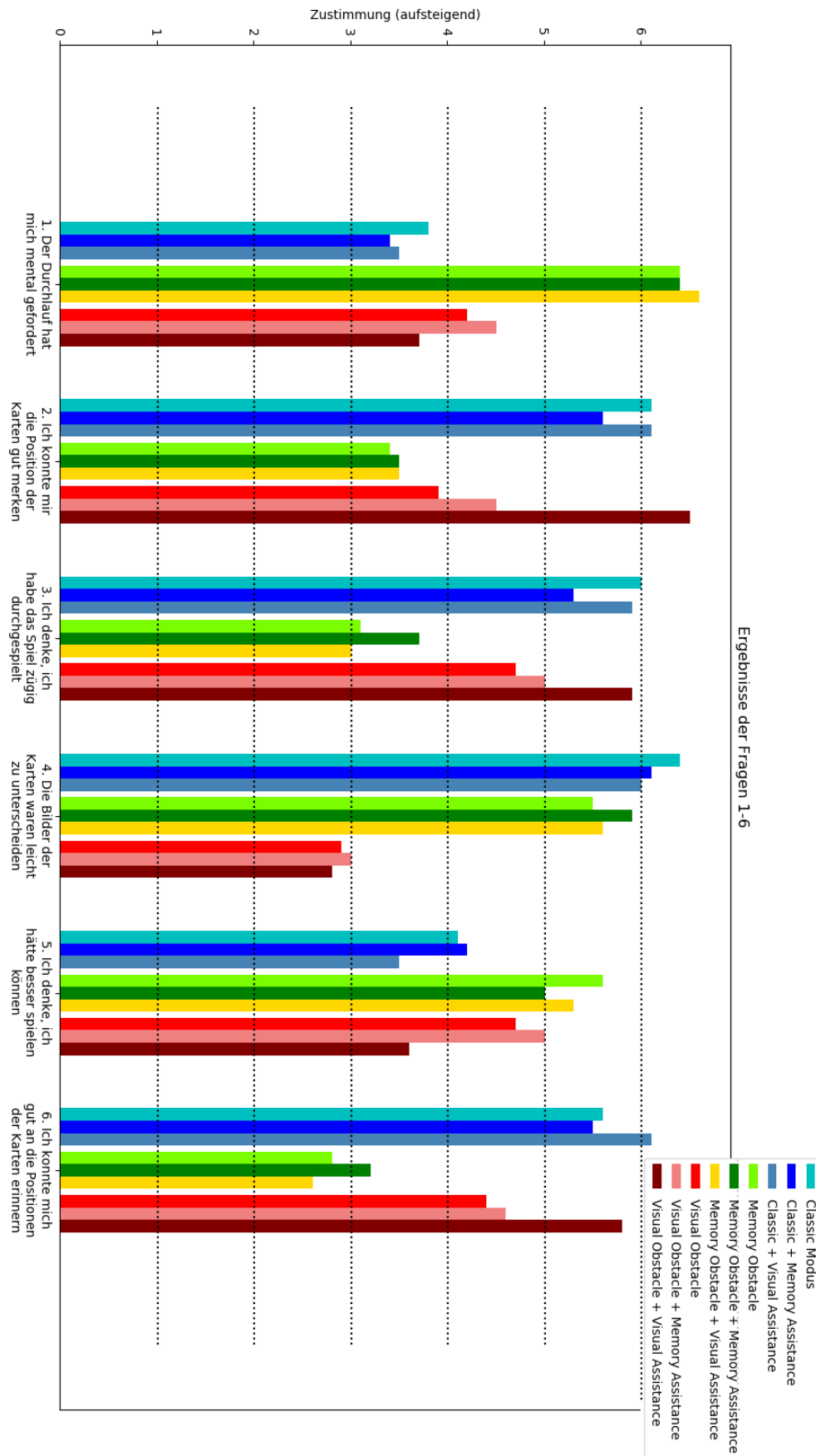
Die Versuchsdokumente umfassen das **EEG-Informationsblatt**, die **Versuchsbeschreibung**, die **Einverständniserklärung** und die **Fragebögen**. Sie alle sind in der digitalen Abgabe mitenthalten. Abb. A.1 zeigt außerdem einen exemplarischen Versuchsaufbau.

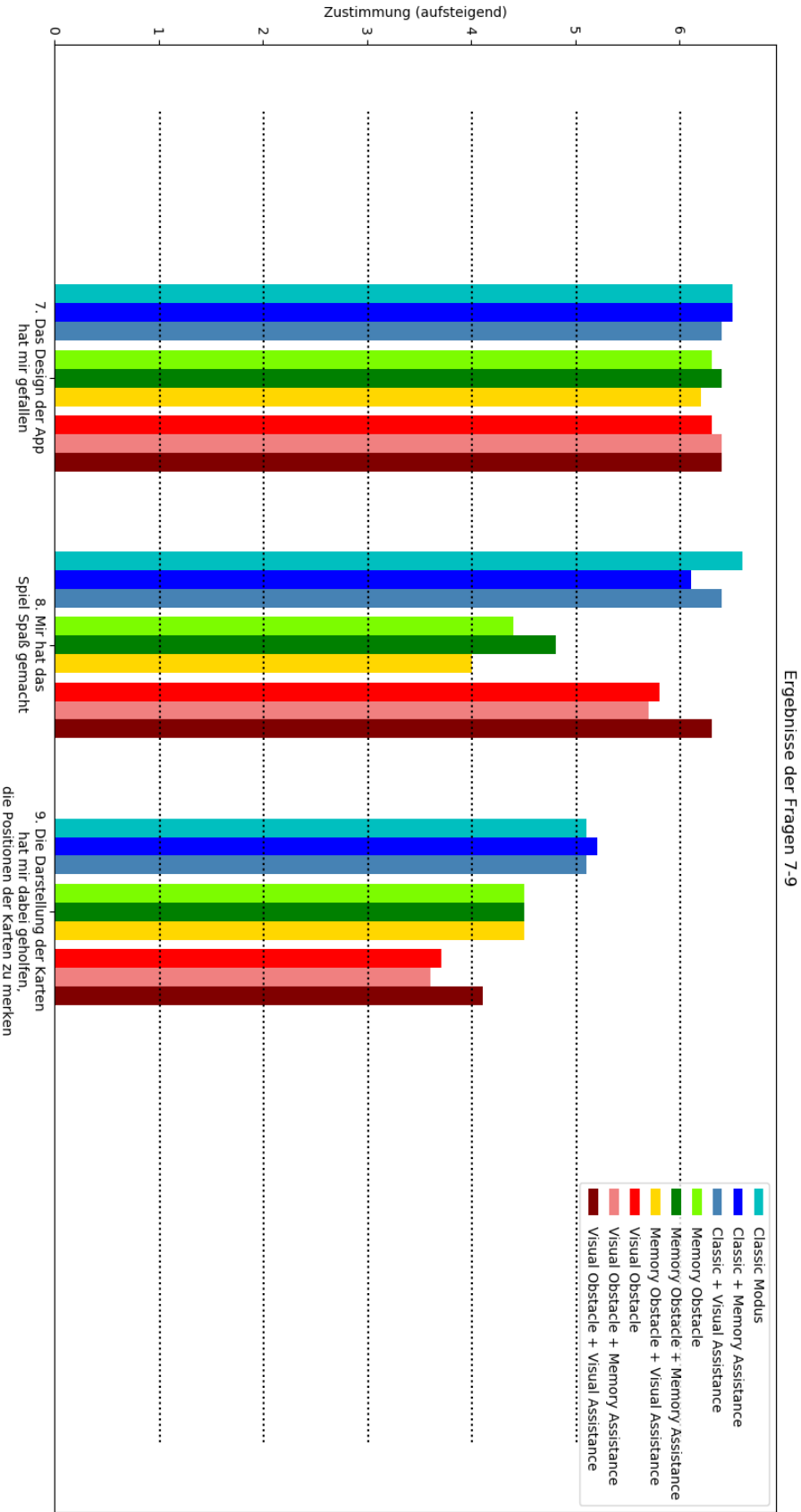
A.3 Graphen bezüglich der Fragebögen

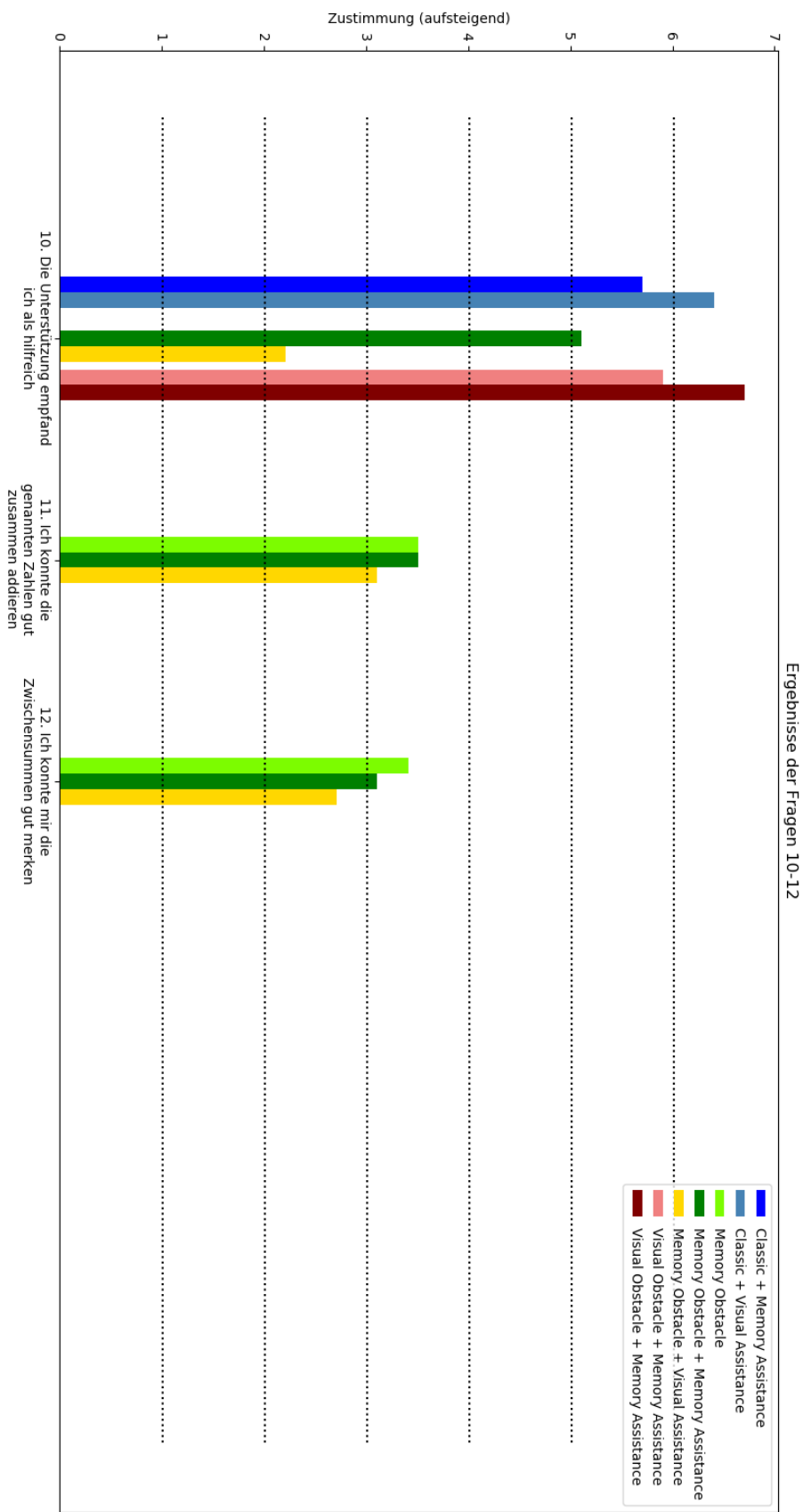
Im Folgenden sind die Ergebnisse der Fragebögen der besseren Lesbarkeit halber grafisch aufgearbeitet:



Abbildung A.1: Ein Beispielbild des Versuchsaufbaus. Anders als in diesem Bild war der Raum während der Spieldurchläufe für den Lichtsensor abgedunkelt.







A.4 Zusätzliche Featuresuche

Zusätzlich zu den untersuchten Features aus der Literatur wollte ich untersuchen, welche Elektroden empirisch nur auf Basis meiner Ergebnisse hohe Unterschiede aufweisen. Dafür habe ich für jedes Band die Daten aller Elektroden pro Spiel gemittelt und für jede Elektrode den vorzeichenbehafteten Abstand zwischen dem Classic und Memory Obstacle Modus errechnet, da dieser nach bisherigen Erkenntnissen den Unterschied der Hirnaktivität am deutlichsten zeigt. Jede Elektrode erhält einen Score, welcher sich pro Spiel bei den drei besten Elektroden um jeweils 3, 2 und 1 erhöht. Im Folgenden werden die Elektroden-scores nach allen Spielen für alle 5 Bänder gezeigt. Wichtig hierbei anzumerken ist, dass, gerade aufgrund der geringen Datenmenge daraus keine wissenschaftlichen Schlüsse gezogen werden sollten. Es dient lediglich als Anregung für weitere Arbeit auf diesen Daten und als Kontrolle, ob die näher untersuchten Elektroden in diesen Listen positiv herausstechen.

Elektrode	Fp1	F3	F7	FT9	FC5	FC1	C3	T7	TP9	CP5	CP1	Pz	P3	P7	O1	Oz
Score	3	2	0	2	3	8	3	6	6	0	3	3	4	3	4	1
Elektrode	O2	P4	P8	TP10	CP6	CP2	Cz	C4	T8	FT10	FC6	FC2	F4	F8	Fp2	
Score	1	4	0	1	1	5	4	11	1	5	4	5	4	2	9	

Tabelle A.1: Elektroden-Scores im Delta-Band (1-4Hz)

Elektrode	Fp1	F3	F7	FT9	FC5	FC1	C3	T7	TP9	CP5	CP1	Pz	P3	P7	O1	Oz
Score	16	5	8	1	2	4	0	3	0	0	0	2	0	0	0	0
Elektrode	O2	P4	P8	TP10	CP6	CP2	Cz	C4	T8	FT10	FC6	FC2	F4	F8	Fp2	
Score	1	2	3	4	1	1	7	9	1	2	4	5	0	1	26	

Tabelle A.2: Elektroden-Scores im Theta-Band (5-8Hz)

Elektrode	Fp1	F3	F7	FT9	FC5	FC1	C3	T7	TP9	CP5	CP1	Pz	P3	P7	O1	Oz
Score	12	3	5	5	2	5	3	6	7	0	0	0	0	0	7	2
Elektrode	O2	P4	P8	TP10	CP6	CP2	Cz	C4	T8	FT10	FC6	FC2	F4	F8	Fp2	
Score	3	0	0	1	1	0	3	6	3	1	1	13	3	8	8	

Tabelle A.3: Elektroden-Scores im Alpha-Band (9-13Hz)

Elektrode	Fp1	F3	F7	FT9	FC5	FC1	C3	T7	TP9	CP5	CP1	Pz	P3	P7	O1	Oz
Score	0	0	2	3	7	0	1	13	16	0	0	0	0	12	14	6
Elektrode	O2	P4	P8	TP10	CP6	CP2	Cz	C4	T8	FT10	FC6	FC2	F4	F8	Fp2	
Score	4	0	1	0	0	0	0	0	19	0	3	2	2	3	0	

Tabelle A.4: Elektroden-Scores im Beta-Band (14-30Hz)

Elektrode	Fp1	F3	F7	FT9	FC5	FC1	C3	T7	TP9	CP5	CP1	Pz	P3	P7	O1	Oz
Score	0	1	3	2	6	0	0	16	17	1	0	0	0	3	13	4
Elektrode	O2	P4	P8	TP10	CP6	CP2	Cz	C4	T8	FT10	FC6	FC2	F4	F8	Fp2	
Score	8	2	1	1	0	0	0	0	19	0	5	0	4	2	0	

Tabelle A.5: Elektroden-Scores im Gamma-Band (31-40Hz)

Die Arrays zeigen unterschiedliche Ergebnisse. Vermutlich gibt es, wie im Ausblick erwähnt, keine universale Elektrodenposition, welche auf allen Bändern gute Ergebnisse erzielt. Besonders im Theta-, aber auch im Alpha-Band stechen die Frontalelektroden wie nach Kumar [KK16] heraus. Im Beta- und Gamma-Band hingegen zeigten die Temporalelektroden T7, T8 und TP9 insgesamt die größten Unterschiede. Parietal- und Okipitalelektroden zeigen in den unteren drei Frequenzbändern wie vermutet wenig Unterscheidungspotenzial auf. Im Beta- und Gamma-Band erhalten P7 und O1 allerdings verhältnismäßig gute Scores. Aufgrund der starken Individualität der Daten sollten dieser Metrik aber auch nicht zu viel Gewicht zugesprochen werden.

A.5 Skripte

Die in dieser Arbeit benutzten Skripte sind in der digitalen Abgabe im Ordner *Scripts* zu finden. Darin finden sich einzeln die Skripte für die Gamelogs und Fragebögen:

- `game_time_turns_error_plot.py`
- `questions_results_updated.py`
- `ttest_gamelogs_questions.py`

Die ersten Beiden stellen die Ergebnisse grafisch dar, das Dritte berechnet die T-Tests.

Im Unterordner **eeg** befinden sich zwei Skripte:

- `eeg_fpx_bandx_gamemodex.py`
- `eeg_fpx_bandx_gamemodex_sequenced.py`

Diese stellen die EEG-Daten jeweils gemittelt und sequenziell für alle Spieler grafisch dar, berechnen T-Tests inklusive Benjamini-Hochberg-Korrektur und ermitteln die eben gezeigten Elektroden-Scores. In den EEG-Skripten müssen die Parameter (wie die beiden zu vergleichenden Spielmodi und das Band) in den globalen Variablen angepasst werden (in der Datei beschrieben). Ansonsten funktionieren alle Skripte ohne Zusatzargumente und können direkt ausgeführt werden (Die Funktionalität bzgl. z.B. der Datenpfade wurde nur auf Windows 10 getestet). Im Ordner *memory_statistics* befinden sich die aus dem CSL-Repository genutzten Skripte für das Einlesen und Verarbeiten von Daten, wie in Kapitel 6.4 beschrieben. Diese Skripte wurden nur auf die für meinen Versuch relevanten Bereiche gekürzt und um zwei Bandparameter für die Eingrenzung erweitert.

Literaturverzeichnis

- [Abd07] Hervé Abdi. Bonferroni and šidák corrections for multiple comparisons. *Encyclopedia of measurement and statistics*, 3:103–107, 2007.
- [BH95] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300, 1995.
- [bra07] Lateral view of a human brain, telencephalic lobes, cerebellum and brainstem colored, 2007. https://commons.wikimedia.org/wiki/File:Gehirn,_lateral_-_Lobi_%2B_Stammhirn_%2B_Cerebellum_deu.svg, Abruf am 27.08.2019.
- [Bun18] Statistisches Bundesamt. Anteil der privaten Haushalte in Deutschland mit Personal Computern von 2000 bis 2018., 2018. <https://de.statista.com/statistik/daten/studie/160925/umfrage/ausstattungsgrad-mit-personal-computer-in-deutschen-haushalten/>, Abruf am 04.09.2019.
- [Cho19] Z.R. Chow. False positives, FWER, and FDR explained, 2019. <https://lesslikely.com/statistics/multiplicity-explained/>, Abruf am 29.09.2019.
- [DLT⁺10] Stavros I Dimitriadis, Nikolaos A Laskaris, Vasso Tsirka, Michael Vourkas, and Sifis Micheloyannis. What does delta band tell us about cognitive processes: a mental calculation study. *Neuroscience letters*, 483(1):11–15, 2010.
- [DSM07] Arnaud Delorme, Terrence Sejnowski, and Scott Makeig. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage*, 34(4):1443–1449, 2007.
- [DW74] G Dolce and H Waldeier. Spectral and multivariate analysis of EEG changes during mental activity in man. *Electroencephalography and Clinical Neurophysiology*, 36:577–584, 1974.
- [FHR⁺95] Thalía Fernández, Thalía Harmony, Mario Rodríguez, Jorge Bernal, Juan Silva, Alfonso Reyes, and Erzsébet Marosi. EEG activation patterns during the performance of tasks involving different components of mental calculation. *Electroencephalography and clinical neurophysiology*, 94(3):175–182, 1995.
- [GWW07] Krzysztof Z Gajos, Jacob O Wobbrock, and Daniel S Weld. Automatically generating user interfaces adapted to users’ motor and vision capabilities.

- In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 231–240. ACM, 2007.
- [Ins19a] National Eye Institute. Color blindness, 2019. <https://nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness>, Abruf am 25.09.2019.
- [Ins19b] National Instruments. Understanding ffts and windowing, 2019. <http://download.ni.com/evaluation/pxi/Understanding%20FFTs%20and%20Windowing.pdf>, Abruf am 28.08.2019.
- [KK16] Naveen Kumar and Jyoti Kumar. Measurement of cognitive load in hci systems using eeg power spectrum: an experimental study. *Procedia Computer Science*, 84:70–78, 2016.
- [Kla] A. Klassen. Erkennung und unterscheidung von gedächtnis-basierten und visuellen interaktions-hindernissen mittels eeg. Bachelor Thesis, Universität Bremen, Germany, 2019.
- [Kli99] Wolfgang Klimesch. Eeg alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, 29(2-3):169–195, 1999.
- [KLJ⁺99] George H Klem, Hans Otto Lüders, HH Jasper, C Elger, et al. The twenty electrode system of the international federation. *Electroencephalogr Clin Neurophysiol*, 52(3):3–6, 1999.
- [Kus19] Roman Kushnarenko. memory-game, 2019. <https://github.com/sromku/memory-game>, Abruf am 17.09.2019.
- [LBC99] Fabien Lotte, Laurent Bougrain, and Maureen Clerc. Electroencephalography (eeg)-based brain-computer interfaces. *Wiley Encyclopedia of Electrical and Electronics Engineering*, pages 1–20, 1999.
- [Lik32] Rensis Likert. A technique for the measurement of attitudes. *Archives of psychology*, 1932.
- [Mir] R. Miranda. Recognising & dynamically adapting to visual and memory-based hci obstacles based on behavioural data. Bachelor Thesis, Universität Bremen, Germany, 2018.
- [Mos08] Simon Moss. Acquiescence bias. *Unter: http://www.psych-it.com.au/Psychlopedia/article.asp*, 2008.
- [PSS18] Felix Putze, Mazen Salous, and Tanja Schultz. Detecting memory-based interaction obstacles with a recurrent neural model of user behavior. In *23rd International Conference on Intelligent User Interfaces*, pages 205–209. ACM, 2018.
- [pyt19] Paired samples t-test, 2019. <https://pythonfordatascience.org/paired-samples-t-test-python/>, Letzter Aufruf am 18.09.2019.
- [Sch] T. Schultz. Vorlesung: Biosignale und benutzerschnittstellen. Universität Bremen, Germany, 2018.

- [Sci19] SciPy. rfft scipy documentation, 2019. <https://docs.scipy.org/doc/numpy/reference/generated/numpy.fft.rfft.html>, Abruf am 18.09.2019.
- [SP10] Skipper Seabold and Josef Perktold. Statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.
- [SPS⁺98] Constantine Stephanidis, Alex Paramythis, Michael Sfyrakis, A Stergiou, Napoleon Maou, A Leventis, George Paparoulis, and Charalampos Karagiannidis. Adaptable and adaptive user interfaces for disabled users in the avanti project. In *International Conference on Intelligence in Services and Networks*, pages 153–166. Springer, 1998.
- [Stu08] Student. The probable error of a mean. *Biometrika*, pages 1–25, 1908.
- [vde19] vdek. Entwicklung der lebenserwartung bei geburt in deutschland nach geschlecht in den jahren von 1950 bis 2060 (in jahren)., 2019. <https://de.statista.com/statistik/daten/studie/273406/umfrage/entwicklung-der-lebenserwartung-bei-geburt--in-deutschland-nach-geschlecht/>, Abruf am 04.09.2019.

