

The influence of age on the ability to use Brain-Computer-Interfaces

Use-case: TV remote control

Masters Thesis

To obtain the academic degree M.Sc.

Julius Neudecker

2025850



University of applied sciences Hamburg
Faculty of Design, Media und Information
Department of Media Engineering

First examiner: Prof. Dr.Roland Greule

Second examiner: Dipl. Inf. Rüdiger Höfert

preliminary version from October 8, 2021

Contents

Abstract

In recent years, Brain Computer Interfaces - BCI in short - evolved to a level of maturity which allows for these devices to be produced cheaply and thus being available to consumers. This study uses a device from a manufacturer called *Nextmind* to examine whether age has an effect on the ability to use such a device. A study was carried out with 30 participants from different age groups. They were confronted with a task to use a Graphical User Interface to select elements by looking at them. The goal was to select this corresponding element as fast as possible to reach good results. The study showed that in fact age does have an impact on the overall performance by proving a statistical significance in regards to the Information Transfer Rate.

1 Introduction

In recent years significant progress has been made on the development of interfaces which relies on direct interaction with the brain itself. The latest popular example is Elon Musks *Neuralink* with their monkey learning to play the game *Pong* only by using its brain (?). Apart from a solid scientific methodology, this study also presented a good media coverage including a showcase video which went viral. However there are more examples of a working interface, which will be discussed in section ??, since this vast area of research is an intersection between several areas of research: medical engineering, neuroscience, computer science and HCI¹. These interfaces are generally called *Brain-Computer-Interface* or *BCI* in short. ? has a very precise definition of the scope:

Brain-Computer Interface (BCI) is a system that measures central nervous system (CNS) activity and converts it into artificial output that replaces, restores, enhances, supplements, or improves the natural CNS output and thereby changes the ongoing interactions between the CNS and its external or internal environment. BCI is direct communication pathway between an enhanced or wired brain and an external device.

As of Q2 2021 there are already devices available for consumers to buy, which fall into this category. This opens up possibilities for a widespread application of these kind of interfaces. Nevertheless, new ways of interacting with computers require some degree of research to define useful and user-friendly ways to interact with such technology. This study aims to provide insight into one aspect of this process.

After a thorough discussion about the state of research in this field, the research hypothesis will be defined based on considerations about future use cases. Subsequently a user survey will be designed, carried out and conclusively evaluated to put the results into context.

¹Human Computer Interaction

1.1 Management Summary

In the *2018 Gartner Hype Cycle* report (?), which is shown in figure ??, BCIs are denoted as to be on the brink of inflated expectations:

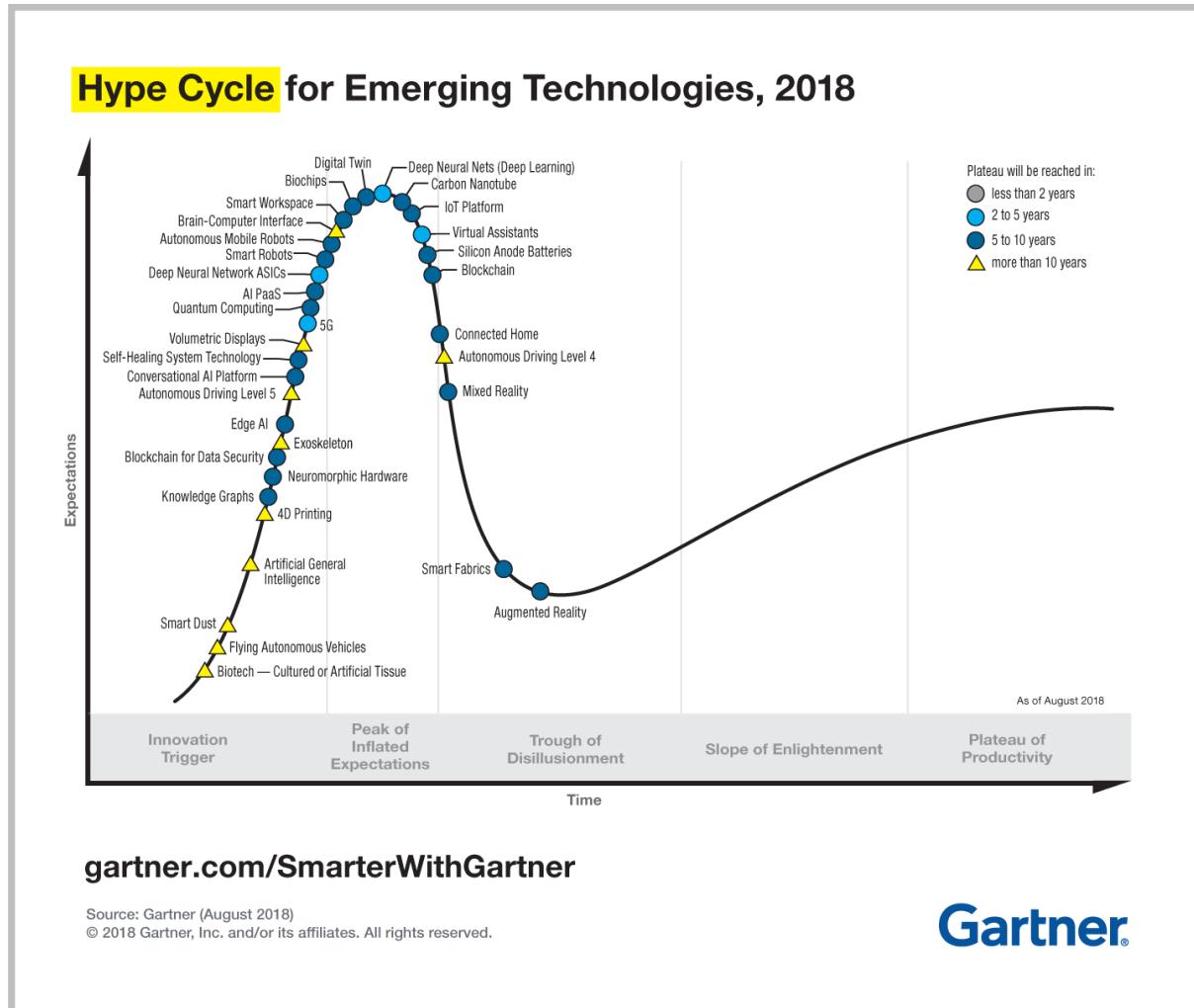


Figure 1.1: Gartner report of emerging technologies 2018

It is important to note though that as of 2018, it'll still take more than 10 years to reach a plateau of productivity. Although there is no mention about this technology in subsequent reports in the following year, two market revenue forecasts from 2015 until 2022 and 2018 until 2022 show a similar pattern in figure ??.

Essentially the market revenue expectation has been very inflated from 2015 on so that it was corrected downwards in 2018. But although the absolute growth was projected to only a small fraction, the relative growth potential stayed about the same of doubling within the next seven years. This is very indicative for the technology being overhyped, as Gartner explains: (?)

A wave of “buzz” builds and the expectations for this innovation rise above the current reality of its capabilities. In some cases, an investment bubble forms, as happened with the web and social media

Nevertheless, what this technology sets apart from other featured technologies is the fact

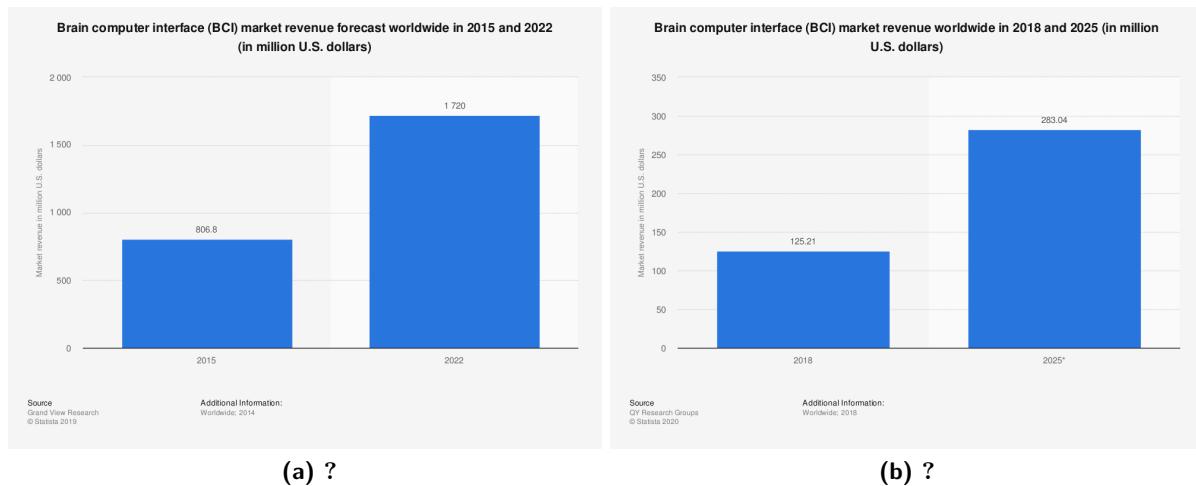


Figure 1.2: Statista revenue forecast as of 2015 and 2018

that is has been around for a few decades and has been continuously researched upon. A strong indicator is the amount of organizations and conferences held about this entire discipline, as can be seen in section ???. The fact that is has only been on the radar of early adopters and tech-enthusiasts in conjunction with market revenue projections is a strong indicator that this technology has reached a level of maturity which makes a widespread application outside of laboratories somewhat feasible.

The latest *2020 Gartner Hype Cycle* report shows already the enhanced version of bidirectional BCIs (titled "*2-Way Brain Machine Interface*") on the slope of innovation:

All in all, there are strong indicators that the technology gained traction over the last few years and could be considered a worthwhile investment if approached with care.

1.2 Brain-Computer-Interfaces

In this section a general overview of the working principle of these interfaces will be provided. Since this study is aimed at computer science and HCI², the neuroscience and medical domain will be only covered very briefly.

First studies began by ?, who investigated the possibility to use EEG³ waves, which were first recorded by ?, as a way to create a direct interaction between a machine and a human brain.

There are three types of BCIs: invasive, partially invasive and non invasive. This depicts the degree of intrusion into the skull and brain tissue. *Invasive* BCIs are electrodes, which are implanted directly into or onto the grey matter of the brain. This can cause long term issues like scars and also degraded signal strength according to ?. Partially invasive BCI however are although located within the skull not in direct contact with the grey matter. Non-Invasive BCI are only placed on the head without intrusion of any tissue. Due to the direct contact, invasive BCI provide the best resolution of the measured signals. Non-invasive BCI in comparison suffer from signal degradation and deformation of the cranial bone tissue. Therefore partially invasive

²Human Computer Interaction

³Electroencephalogram

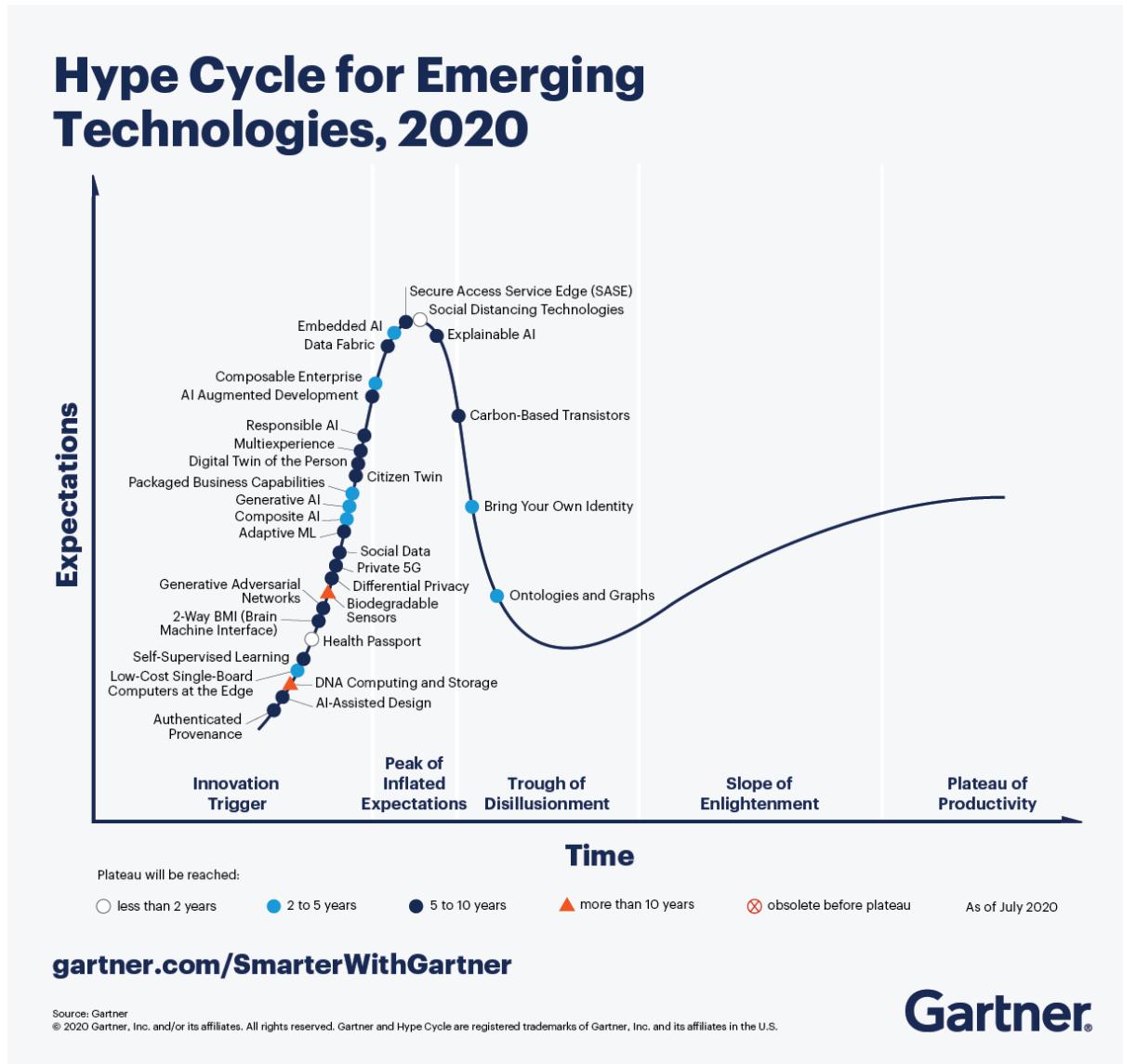


Figure 1.3: Gartner report of emerging technologies 2020

BCI are a compromise between good signal strength and the risk of medical conditions. Another potential advantage of non-invasive BCIs is that these Interfaces could be easier mass-produced and become affordable to consumers. Also they don't require specialized medical knowledge and equipment to operate.

The way these interfaces work is based on the same principle: A human brain emits electrical signals, which can be picked up. According to ?, they can be described as follows:

"Embedded in this sustained "spontaneous" or "ongoing" electrical activity, short, distinctive (0.5-2 sec) waveforms can be found that are evoked, for instance, when a brief sensory message (stimulus) such as a brief illumination of the visual field or a tap on the forearm is received by the subject."

Based on the origin within the brain, these can be correlated to certain stimuli, mental and emotional states (?) and according to ? been used to drive *an external effector or affecting internal body parts and functions*. The external effector is the use case which is being examined

in this study.

Without a BCI, interaction with a computer requires some physical interaction with devices such as keyboards, mouses or gestures on a touch screen. There are mainly two different reasons, why these devices are a constraint to speed and efficiency of HCI. The first reason is a limitation on interaction speed: Although there is no definitive consensus about the speed of thinking, alone being able to type along the spoken word is unattainable for non-professional typists. A professional typist has to be able to type at 180 - 220 WPM⁴ according to ?. ? made a survey with 168.000 volunteers, where the fastest typists weren't even able to come close to this mark with 120 WPM. Therefore it is safe to assume that typing in the same speed as thinking is impossible except for rare individuals who devoted a significant time practicing. Secondly: in applications such as games, where reaction time and accuracy is the fundamental element for success or failure, an interaction based on motoric interaction with a physical pointing device has some significant drawbacks like limited accuracy, if the whole chain of wrist movement in conjunction with a mouse is under scrutiny.

If a BCI was to replace these types interaction, these constraints could potentially be alleviated and interaction based on physical interaction rendered obsolete.

1.3 Working principle

In order to understand the reasoning behind this study, the working principle of a BCI has to be established on a general level. Section ?? already established a high-level overview of different BCI technologies and how these are triggered by certain external stimuli. In this study only visual stimuli will be used and therefore explained in detail. The vendor, where the BCI in this study is from - *NextMind* - does not disclose any details of the inner workings itself, it is safe to assume that the underlying technique used is the so called *Steady State Visually Evoked Potential* - SSVEP in short. ? provides detailed insight into the topic from a neuroscientific point of view. The general principle however is that any visual stimuli cause a certain pattern of waves within the visual cortex of the brain. These patterns are being evaluated by different means to assess if a certain pattern is being seen *and* in focus of the person.

According to *NextMind* this is being done by feeding the sensor data through a trained neural network, which is able to return a confidence score for each target. The objects, which are being seen by the person, have been labeled *neurotags* (?) from *NextMind*. These neurotags can provide two different readouts: If it is triggered (i.e. *seen*) and the confidence, which depicts the level of *focus* of the user on the neurotag (?).

The physical layout of the sensor is shown at figure ???. It has 18 electrodes, which are arranged in pairs to cover the area, where the visual cortex is located at the back of the cranium. These electrodes have a certain degree of freedom to their movement, which is beneficial since this mitigates the risk that small movements from the sensor alter the sensor position too much. It is battery driven and communicates via the Bluetooth LowEnergy protocol. A headstrap keeps it in place.

⁴Words Per Minute

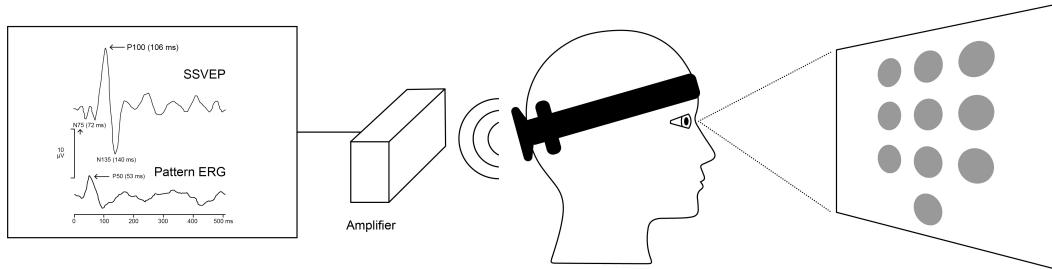


Figure 1.4: How VEP works in principle

1.4 Related work

As previously mentioned in section ??, research on BCIs is partitioned between four different domains: *medical engineering, neuroscience, computer science and HCI*. Apart from commercial entities such as *Microsoft* or *IBM* and scientific journals, the majority of the research community is clustered in three organizations:

- ICBCI (*International Conference of Brain Computer Interfaces*), which is a department of the WASET (*World Academy of Science Engineering and Technology*)
- EMBS (*Engineering in Medicine and Biology Society*), which is a department of the IEEE
- BCI Society, which is an entity of its own

There are also research efforts in the east-asian region, according to corresponding tech-sites such as ? and ? but due to a language barrier, these sources cannot be considered.

To narrow the scope, where this research paper is located at, the considerations from section ?? are taken into account. As already established in section ??, the sensor used in this study uses SSVEP and is non-invasive in nature. Therefore the general scope of this research is located in the realm of *non-invasive BCI based on SSVEP used for HCI*.

? and ? investigated the effects of different twinkle frequencies and duty cycles on the efficiency on precision of SSVEP BCI. They found that a certain combination of these parameters in fact could improve the ITR⁵. ? used a similar approach and found the ideal combination in conjunction with Korean characters. ? used a consumer ready BCI by *EMOTIV* to create a

⁵Information Transfer Rate



Figure 1.5: Physical layout of the sensor

Matrix-Speller in the Bengali-Language to allow people who have lost the ability to communicate to express themselves again. ? also used a SSVEP BCI to implement a BCI-speller and scrutinized the tradeoff between responsiveness and accuracy. ? designed an interface which is operated by a SSVEP BCI to control a robot arm, which could administer food to disabled people. ? developed a SSVEP BCI which overcomes the necessity for training the sensor to the user who wears it. The prototype reached a similar precision as *trained* interfaces. ? investigated and selected certain visual stimuli which work best with certain use cases. ? made a study, where participants controlled a UAV by using a SSVEP. ? used simulated impairments to examine if usage of a SSVEP is still possible with medical conditions which affects speech and ocular impairments.

Although not strictly within the SSVEP domain, the study by ? showed very promising results by not using visual stimuli but mechanical ones, where he had teenagers playing a racing videogame with the aid of mechanical stimuli.

There is a massive ongoing research effort to make the life of people who are suffering under ALS⁶ better and improve their ability to communicate normally, by using SSVEP, a hybrid between an SSVEP and P300⁷ or purely P300 based BCI. A significant number of relevant studies has been published in the BCI Society Journal: ?, ?, ?, ?, ?, ?, ?. All these studies aimed to provide a better understanding and performance of using BCI on people with medical conditions, which cause serious physical impairments.

⁶Amyotrophic lateral sclerosis

⁷An Event Related Potential (ERP) BCI

1.5 Establishing a use-case

Establishing a use-case in the context of this study is a pivotal point hence this section will be split up into several sections:

1. General considerations
2. Deduction of the core points of this study
3. Evaluating the constraints of a participatory study
4. Deduction of use-case
5. Conclusion into hypothesis and use-case

1.5.1 General considerations

The first step in conceiving any potential way of using such a device is to evaluate the way any user could interact with a BCI with a computer. According to (?:4.13) the way users interact with a device require an agent of control i.e. a hand, what is being sensed by the device (position, motion or pressure) and the number of dimensions being sensed (1, 2, 3). This results in a different input taxonomy for any given device. However, a BCI does not have either of these parameters, since the interaction does not require physical interaction. Hence a classification by means of using a taxonomy cannot be achieved. Where the interactions of BCIs can be compared to those classified by taxonomies is by the way they function they apply in relation to a user interface.

The API⁸ endpoints of the NextMind sensor offers two different modes of interaction. These are explained in the SDK⁹ of the sensor in detail: ?. They are depicted as *tracking results* with a *hit* property and a *confidence* metric. Where *hit* is a two state interaction: the neurotag is being seen by the user and subsequently recognized by the sensor and its backend or it is not. The *confidence* property depicts the attention which the user is paying to the *neurotag*. This is a continuous decimal value between 0 and 1. The fact that these types of interaction are based on neural activity raises the question if a pure mapping of continuous and discrete input modalities to established interfaces would be beneficial to the user experience. Under the reasonable assumption that without any training the metric *focus* can only be deliberately controlled on a very coarse level, the necessary sensitiveness required for modern GUIs¹⁰ can not be achieved with this particular sensor. The remaining two state property, which can be utilized to select or deselect certain objects also only allows for limited interaction. However, these neurotags can be placed in arbitrary places. Although a *toggle*-like behavior is not mentioned explicitly, it might be possible to de-select any activated neurotag when the *focus* property falls under a certain value.

⁸Application Programming Interface

⁹Software Development Kit

¹⁰Graphical User Interface

1.5.2 Corepoints of this study

Based on the previous reasoning, the following questions can be raised in regard to the feasibility of any interface which could potentially be conceived with this technology:

1. How fast is the perceived and measured reaction time of these neurotags?
2. What is the minimum size the neurotags have to have in order to be recognizable?
3. Is the interface usable for brains of all ages or do gerontological effects have an effect on usability?
4. Do certain medical conditions (i.e. attentiveness disorder) have an impact on the usability?
5. How fast can a user switch between neurotags?
6. Is a BCI controlled GUI intuitive to use?
7. Does a personal affinity to technology have an influence on the perceived difficulty of interaction?

These questions can be clustered into two groups: *neurological* and *interaction*. But all these considerations open up a vast space of potential cases, hence the priority is to examine whether these interfaces are generally usable by the majority of users and if these interfaces are intuitive to use. Out of this list only the points 1, 2, 3, 5, 6 and 7 can be applied to a general audience. Considered the possible interactions with the *Nextmind* sensor, a study which focusses on the neurological domain makes more sense in the context. This leaves the points 3 and 5. Although these are two different topics, the setup of the experiment itself will show that in fact number 5 is actually the driving metric to answer this question, which will be discussed in section ??.

1.5.3 Constraints of a participatory study

Although this study won't use the *AttrakDiff* survey, which was designed by ?, it is still worth considering how a potential use-case might perform in the context of an *AttrakDiff*. The reasoning is that a survey, which is entirely constrained to serve the purpose of producing results in favour of the study, might be harder to grasp for the participants in the experiment. The reason being that an interaction purely for the sake of interacting with something does not provide an incentive for the user to do so. As a consequence, the results which are produced by the participants, might be skewed due to a lacking frame of reference.

Figure ?? shows two axes which depict the hedonistic quality of an interaction, which is a metric of pleasure and the pragmatic quality which depicts a metric of *ease of use* or *technical quality*. Even without any deeper knowledge it is safe to assume that it is preferential for the interaction to be located in the upper right corner, since this makes it "*desireable*" instead of "*unnecessary*". Given that, it is certain that any interaction needs a way to facilitate pleasure in the user.

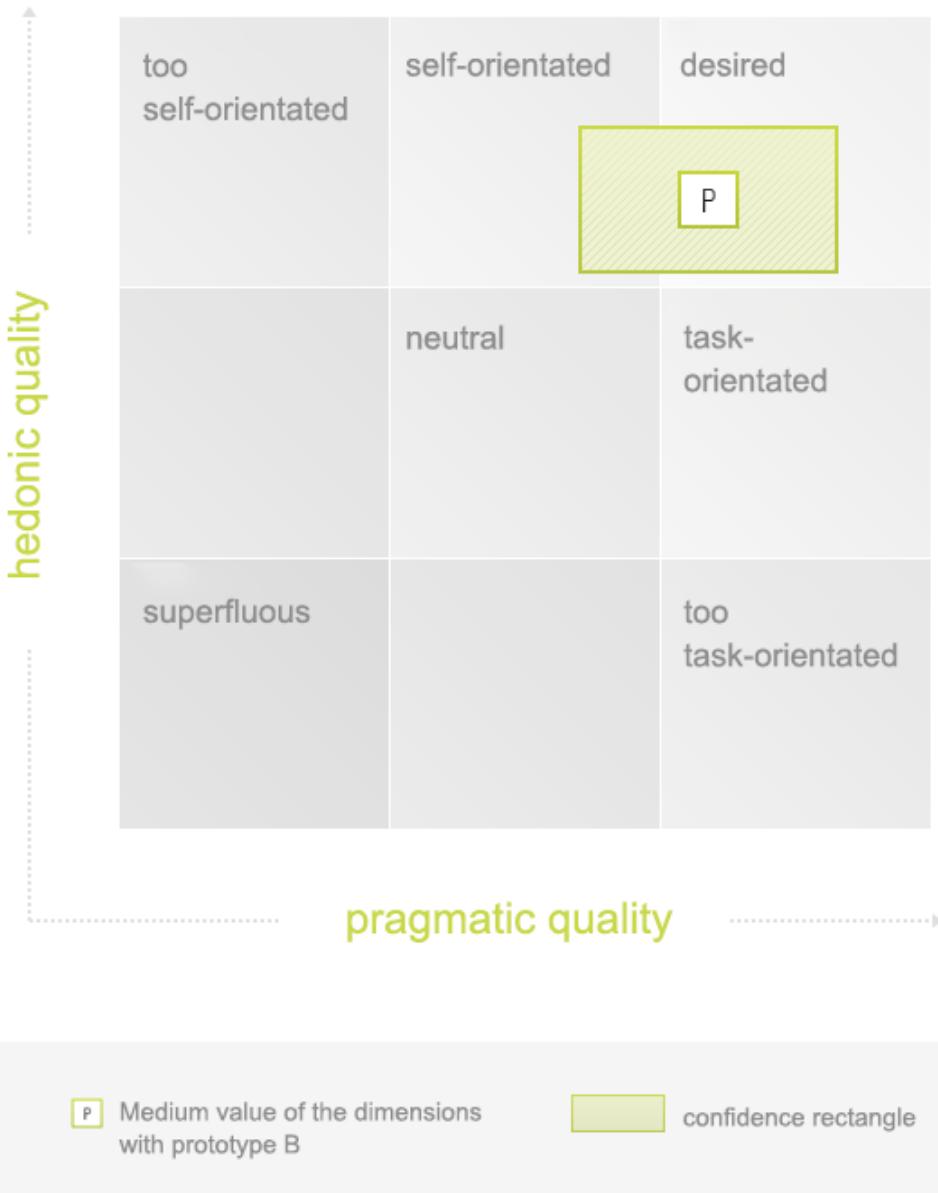


Figure 1.6: AttrakDiff results for a given interaction model, Source: ?

1.5.4 Deduction of use-case

Concluding the two parameters *frame of reference* and *facilitation of pleasure* into a coherent picture, it can be inferred that the experiment has to be set up in a way that gives the participants a familiar use case which facilitates positive feedback.

As already established, this study will examine gerontological effects in the context of an interaction. Therefore the age of participants will very likely vary to a wide degree, what necessitates a use case which is common to all age groups alike. Because people above a certain age did not grow up with computers, an experiment which relies heavily on the usage of computer as main point of interaction is very likely not a good choice due to the difference in proficiency. One example where both these criteria are met is a remote control for a television. Although nowadays there are computers involved, the interaction hasn't really changed in the last decades on a general level.

Nevertheless the working principle of the BCI is a visible flashing pattern, as established in section ???. Therefore the representation of some kind of GUI on a display is still necessary. Since VR goggles provide better isolation from external visual stimuli, the representation within a VR application was chosen.

1.5.5 Concluding hypothesis and use-case

Section ?? established that age is the first parameter to examine in this context, therefore the fundamental research hypothesis can be defined as follows:

"Age does not have a detrimental effect on the ability to use a non-invasive BCI based on VEP technology."

In Section ?? the fundamental reasoning behind the use case, which provides the necessary framework for this study was established:

A GUI which represents a tv remote control will be presented to the participants within a VR environment.

1.6 Beyond the scope

The case in this study is fairly limited in the realm of BCI. Section ?? already briefly mentioned that the general research on BCI has several fascinating topics to explore. This section aims to broaden the horizon on the realm of BCI technology.

Picture synthesis In 2019 ? published a paper which outlined the reconstruction of images seen by individuals by means of a BCI. They used a 128-channel EEG cap to record brainactivity in a 1...35Hz bandwidth while being shown videos of different subjects. The signals were treated with a PCA¹¹ to create a 20-dimensional feature vector which is feeded into a classification model, which yields the results:

? wrote a good summarized article about the paper alongside results and a video which features a demonstration. On the other hand ? used a similar approach to enhance grasping of robotic arms by means of deciding on the correct technique to grasp based on image reconstruction from a BCI.

Controlling vessels, aircraft and spacecraft Since BCI like the P300 have been used as neuroprosthetics for physically impaired people, there are efforts made to study a broader application of BCI to control machines or vessels. (?) published a study, where invasive BCIs for direct control of machines were examined. They came to the conclusion that this is very much possible but requires several advancements to gain really high quality data for fine-motoric movements. This also requires that electrodes have to be implanted deep inside the brain in order to gain high resolution results.

(?) published an article on the SAE News homepage, describing a report published by the DARPA where an individual was able to control three warplanes in a simulator. The novelty

¹¹Principal Component Analysis

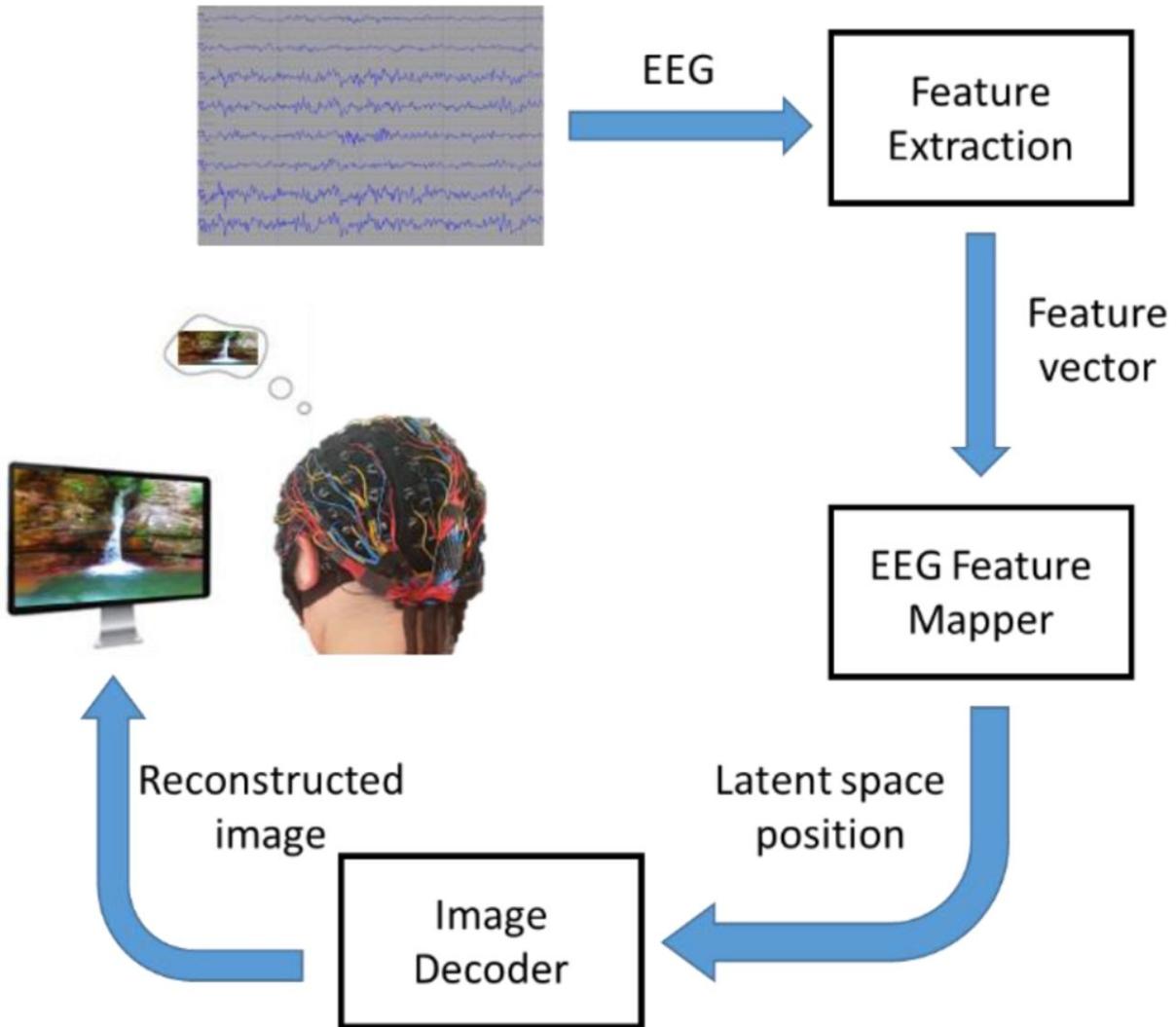


Figure 1.7: General scheme of neurofeedback model, Source: ?

in this research was also the feedback given to the individual, thereby posing a bidirectional BCI. As of Q2 2021, this article is no longer available on the DARPA homepage but several other news organizations wrote about this: (?), (?) and (?). Since the DARPA is well known for being decades ahead with their research, it is very well feasible outlook to have BCIs in the future, which allows for control of not only aircraft but also ships and spacecraft.

Improving human learning rate and augmenting system performance In an article published by the Journal of neuroscience methods by (?) the efforts by the DARPA¹², which is a US federal agency to advance technology for defensive efforts, were outlined. These grouped into prosthetics and rehabilitation as well as efforts to improve human training and performance. The latter has fascinating research on Improving the learning rate of the human brain:

The Narrative Networks (N2) program is developing new techniques to quantify the effect of narratives on human cognition and behavior, including initial development of a closed-loop BCI system that adapts a narrative in response to a listener's EEG signals. Such a system would have numerous applications to training and human performance domains. (?)

¹²Defense Advanced Research Projects Agency

Further applications were the integration of a *Human-in-the-loop*, scenario, where intelligence analysis and threat warning. A 10-fold improvement in analysis throughput was achieved in the former case. In the latter case of threat detection the incorporation of a BCI-enabled human increased the probability of threat detection to 91% compared to 53% detection rate solely relying on computer vision.

Transferring thoughts over the internet A paper published by (?) outlined the possibility of using nanotechnology to insert sensors to every single neuron and synapse in the human brain to have a real time stream from the within the human brain with a resolution down to a single neuron. This would allow for every human having a digital twin of himself in the cloud, having access to the whole repository of human knowledge without delay and even share experiences, thoughts and memories with other people.

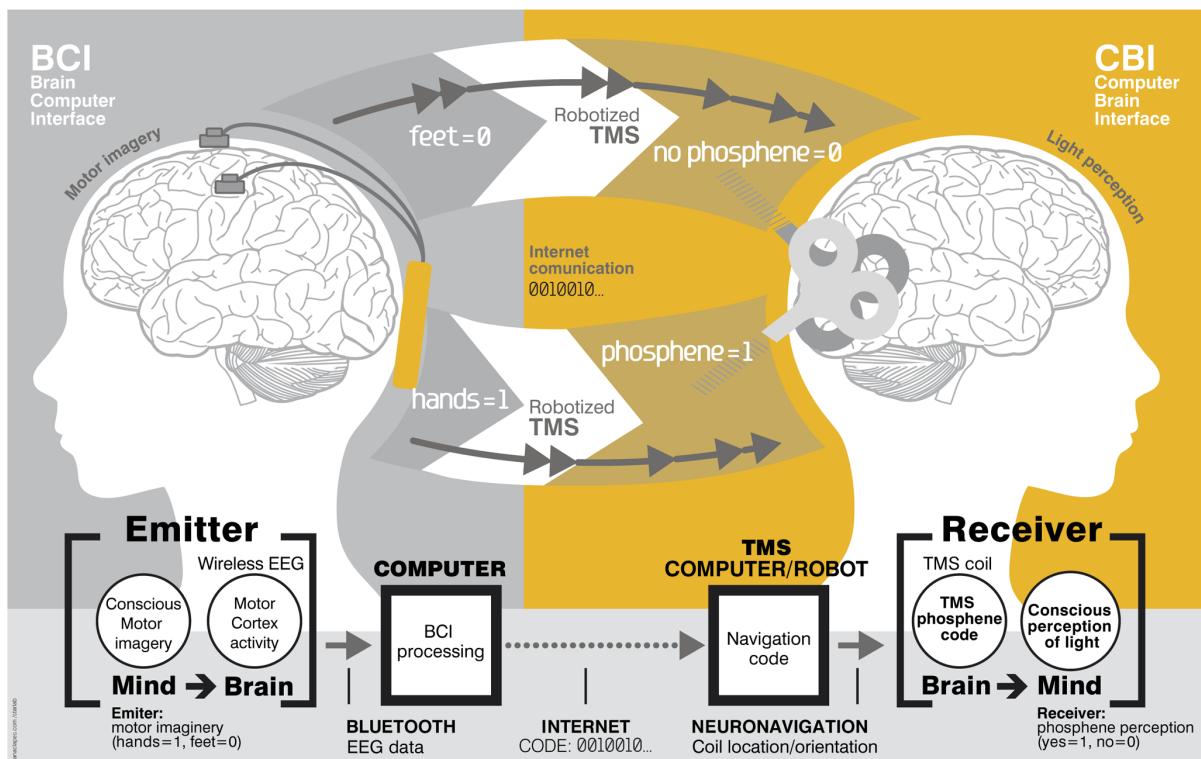


Figure 1.8: B2B communication signal chain, Source: ?

A little less futuristic is a study published by ABC News: (?) and outlined in figure ???. In this case for the first time in human history, conscious communication from one brain to the other without the involvement of sensory or motoric stimuli was achieved. This is called a B2B or Brain-to-Brain communication. The according paper was written by (?). In this case two humans exchanged the greetings *Hola* and *Ciao* over the internet from India to France.

Data privacy With the latest innovation in the field of BCI and BMI a new concern arises along with them: privacy. (?) published a paper, where she outlined the current status and outlook on data privacy in regard to brain data, where the jurisdictions of the EU, United States and Canada were compared with their different approaches:

In relation to the governance of personal data in the private sector, the United States adopts a self-regulatory mechanism, while the European Union assumes a rights-based approach, with Canada's position being intermediate to the two extremes., ?

She concludes that *As BMIs enter the marketplace, legal and ethical questions pertaining to brain data privacy are certain to arise(?: 43)* and makes policy recommendations to be implemented in order to guarantee privacy. These recommendations outline a *privacy by design*, improvement of transparency and general sharpening of existing policies to better suit the requirements of brain data applications.

1.7 EEG acquisition devices

Concluding section ?? and ??, it is reasonable to talk about some technical implications of BCIs and their limitations. In technical terms, these are EEG acquisition devices from the medical domain and only gather EEG brainwaves from certain regions of the brain. The placement of the electrodes on the skull determines which domain of brain activity will be recorded.

There is a great variety of these different EEG acquisition devices on the market. Ranging from consumer-grade cheaply and readily available sensors which cost a few hundred Euros up to medical grade devices. According to ? the difference in price and quality are mainly determined by the factors amplifiers, electrode type and count and transmission technique, which could be wired or wireless.

Since these sensors measure electrical potentials, the actual resolution of the electrodes and amplifiers is the single most influential parameter governing quality. ? used the SNR¹³ of different sensors to compare them to one another, because quality in electrical measurements is best compared using the SNR value of each measurement:

Figure ?? shows the picture that *cheaper* sensors tend to have a worse SNR than more expensive ones with one exception which performed better. However, the cheaper sensors were perceived more comfortable and their set-up time was significantly less compared to fully fledged medical devices, which made them the better choice for experimenting and rapid prototyping. Also the expensive sensors required special treatment with gels and saline solutions, which wasn't necessary on the dry-electrode setup of the cheaper sensors. The conclusion of ? was that either of those two categories have their advantages in their own domain and there is no *best choice*.

¹³Signal to Noise Ratio

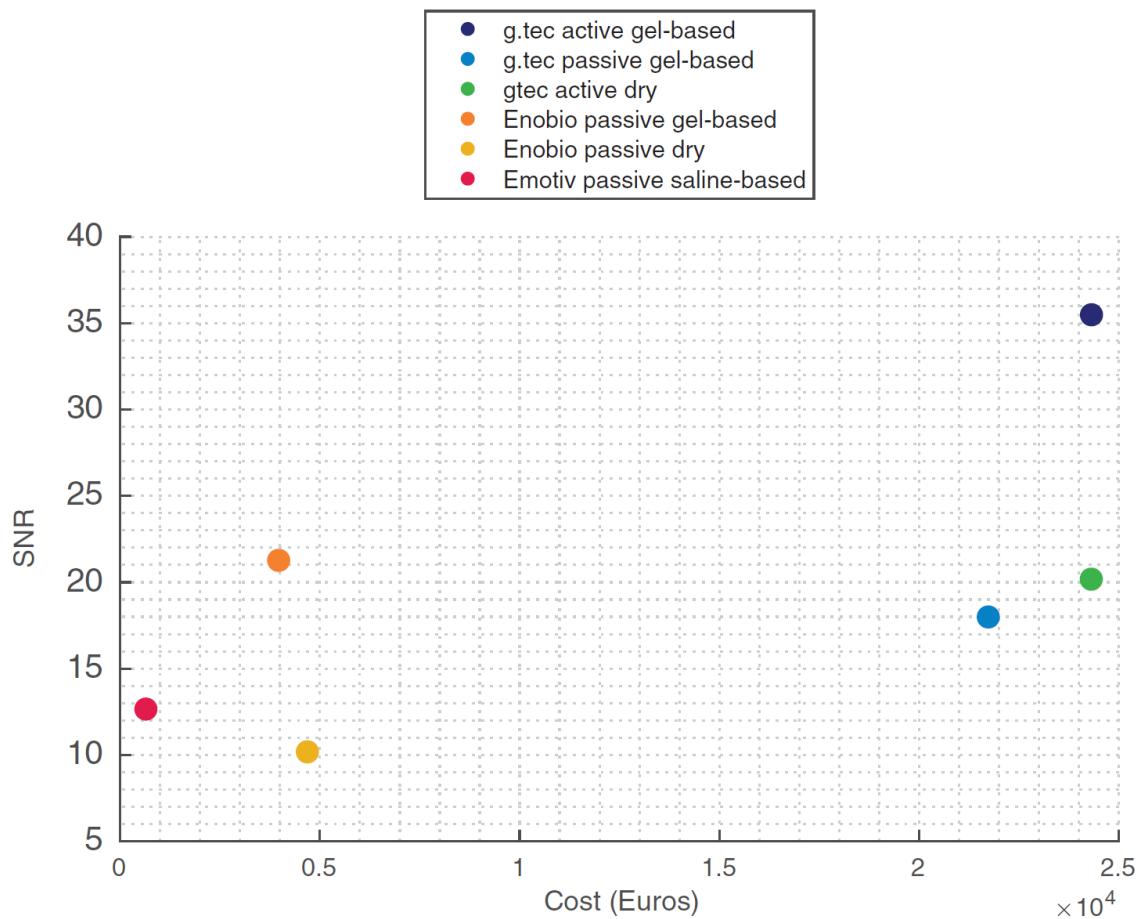


Figure 1.9: SNR plotted against price, Source: ?

2 Study Framework

Based on section ?? the study, which investigates whether the hypothesis holds true or not has to be defined in a theoretical way. This is being done by answering four separate questions in this context:

1. The general scope of this study
2. Considerations for the questionnaire
3. Statistical significance
4. Time vs Age correlation

2.1 Scope of the study

The first aspect which defines the general layout of a study is determined by its scope. Apart from domain specific studies like for example clinical trials, there are two general types of studies: *qualitative* and *quantitative* ones. Where the former one depicts a study where the number of conducted experiments are in the range of one or lower two digit numbers like 20-30, which is common according to (?:302). Quantitative studies often have four to five digit numbers of testcases according to (?:305). The latter ones are far more resource intensive however than their qualitative counterpart. A subset of qualitative studies are explorative studies. In these studies there have not been any previous research efforts so they're often considered preliminary to evaluate the scope and nature of the problem. Since all of these studies have a different objective target, the requirements regarding the samples are different in order to mitigate the risks of a previously unknown factor to skew the results. This might result in rendering a complete quantitative study irrelevant due to previously undetected errors. For this reason *quantitative* studies have to make careful deliberation about the characteristics of the participants in order to avoid a systematical error. This is often done by using *Qualitative* studies which are taking a different approach in regards to sampling the population. Since the number of participants is fairly low, they are often hand selected to pick certain characteristics deliberately. For the scope of this study, these deliberations are discussed in section ??.

Regarding the scope of this study, two main questions have to be answered, to evaluate whether a quantitative, qualitative or even exploratory study is the best approach:

- Is there a previous study, which already established key-findings to build upon?
- How many resources are available to conduct the study?

2 Study Framework

To answer the first question regarding previous studies to build upon, none of the studies discussed in section ?? had the age as a core point. Nor were there any conclusions drawn from a study, which found that age had a significant impact on the results. Age did always play a major role, because BCI are often used as neuroprosthetics, whose usage is being found more often among the older population. More so the *NextMind* BCI has not been subject of a scientific study at the time this study has been written.

In regards to the secound question, is has to be taken into account that the amount of people involved in this study is very limited and there was no additional funding granted to help with procuring more resources. All preparations and the experiment are being done by a single person, which severly limits effects of scale to sample a larger population.

When these two facts are taken into account, the clear picture of a qualitative study can be drawn. More so, it is reasonable to design this study in an explorative way such that it ca be determined, if there is evidence which supports a quantitative study in order to determine exact resons why the hypothesis might or might not hold true.

2.2 Questionnaire Considerations

The primary goal of the survey is to find empirical evidence that age does not affect the ability to use a BCI. In order to exclude certain parametres, which might cause an unwanted effect, potential disturbance parametres have to be identified and discussed:

- Age
- Gender
- Quality of the sensor readings
- User wears glasses
- Cognitive impairments
- Medication or drugs known to affect EEG readings

Furthermore participants of the study should disclose if they had previous or ongoing conditions of seizures or epilepsy, especially in case of light or flashing related cases since the visual cues in the experiment might trigger an attack.

Apart from the demographic parametres of age and gender, the other factors have to be considered to prevent potential malformed data. Firstly, a condition which causes a detrimental effect on the ability to see and identify patterns such as deteriorated eyesight might have a dampening effect on the ability of the visual cortex to create the required brain waves, if the person can't see clearly. However, since the setup of the experiment allows for glasses to be worn, this will not be considered a concern. Medical conditions or drug intake might also have a huge impact on the results. Diseases such as dementia or alzheimers or drug intake such als alcohol or caffeine have to be considered in the individual performance of each subject. ? showed that caffeine intake has an effect on brain wave activity, therefore this has to be taken into account. Lastly the readings of the sensor will very likely be different for each experiment. Since the

sensor provides quality readings, these will be considered in the data evaluation. Nevertheless, there will be a questionnaire provided which will ask the user after experiment if he experienced any of these effects to have a possible explanation for potential outliers.

This study also assumes that gender has no effect on the study, because the brain of men and women is at least structurally identical. Therefore age is the only parameter which is the variable in this study.

2.3 Statistical significance

The solid proof that the hypothesis holds true or not can not be made based on looking at diagrams and educated guesses. Hence the need for design of experiment to provide a numerical framework which defines thresholds and quantities to make results reproducible. This section is based on the theoretical framework which is described in (?: 87ff) under the considerations in the previous section. There are two main parameters to consider, when it comes to the DoE¹:

- Where is the threshold which determines whether an effect is significant or not?
- How many samples are necessary to reliably prove that the hypothesis holds true?

These values are often calculated by means of statistics. The first one is usually depicted as α and the second one as β :

Significant Level Alpha - α : Alpha is the governing parameter, which decides how big the risk in a given sample size is that the hypothesis is wrong and therefore falsified. Since the hypothesis can be either true or false, the probability that it is wrong decreases with the number of samples taken in a binomial fashion (?: 103). That means in this case that the likelihood p of a participant to be above or below the threshold, which will be discussed in section ??, is not determined by age. According to (?: 110) common values for alpha tend to be chosen such that the probability of a falsified hypothesis is at 1%, 5% and 10%. Where a smaller value means a more strict threshold. To account for the explorative character of this study the value of 5% is chosen. Therefore the susceptibility for unaccounted side-effects is reduced and the total number of participants is still not unfeasibly large. This value cannot be calculated - it is often chosen based on experience or strict requirements. This parameter is often also called a *Type I Error* or *False Positive*.

Parameter beta - β : This parameter depicts how effective a potential effect is in a given sample size. The likelihood of a significant effect not being recognized decreases with an increasing number of samples, since it becomes increasingly unlikely that a systematic effect affects the samples always in a way which makes it undiscoverable. Table ?? in section ?? shows the beta values for this study. Usually the β parameter is not used but instead the *Power* of an experiment is calculated as $1 - \beta$. This parameter is often called *Type II Error* or *False Negative*.

¹Design of Experiment

2.4 Time vs Age correlation

After establishing the necessary framework for the survey and how the results have to be interpreted in order to draw a meaningful conclusion, the underlying metric to examine and compare the performance of the individual subjects in the experiment has to be defined.

The only meaningful metric, which can be evaluated with the programming library is the trial time between two cues. This is the time between the visual cue and the successful detection of a target. Figure ?? shows the connection between cue time t_1 and detection time t_2 . The cue time is assumed to be constant for each participant because the interface is presented on a laptop screen (see section ??), which does not require head movement and there is no reason to assume that eye movement of a participant slows down significantly over a period of a few minutes on a healthy individual. Other metrics provided from the sensor, including quality telemetry and confidence values for each target proved to be unreliable at best and not relevant at worst.

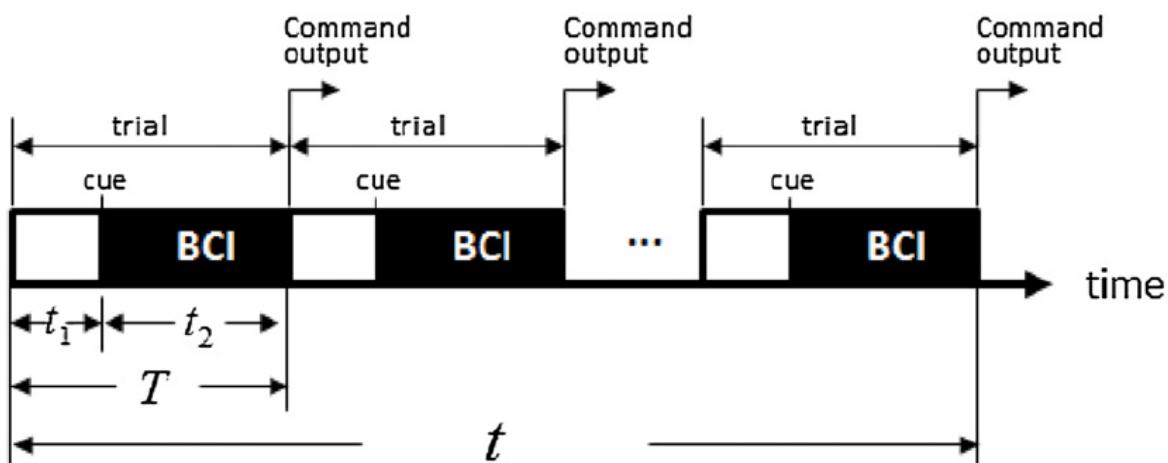


Figure 2.1: Timing of symbol recognition and activation

By collecting several timing data points for each participant, the goal is to draw a regression graph between the age of the participants and the mean timing for each visual cue. By calculating the statistical significance of the correlation, which was discussed in section ??, the hypothesis can be proven or rejected.

3 Survey Contents

Chapter ?? established the theoretical framework of the survey. This chapter aims to define the experiment which will be used to gather the data. To do so, the following considerations have to be made:

1. What tools will be used?
2. How is the interface going to look like?
3. What data will be collected?
4. How are the age groups structured?
5. How is the data collected?
6. What are the questions in the questionnaire?

3.1 What tools will be used

Laptop with experimental application A Laptop will be used to conduct the study and collect the necessary timings as shown in figure ???. This application will be used to show the user the visual stimuli in the experimental *remote control interface*. This application was purposefully built with the *NextMind SDK* in Unity. The timings will be collected as timestamps, which are in turn created by the application itself upon the corresponding events with a resolution down to a millisecond. The code is made available at Github(?).

Nextmind BCI Shown in figure ?? is the Nextmind BCI, which will be used to gather the data from the VEP. The vendor provides a SDK which integrates really well in Unity and in consequence with the Quest2. During the experiments, this sensor will be located at the back of the participant's head, where the visual cortex is located at (figure ??.).

The Sensor will be strapped to the VR headset and thus these two devices will be easy to put on or off.

Pen and paper questionnaire A questionnaire will be used to collect anonymized demographic data from the participant. This is necessary for the evaluation of the results. Among demographic data and self assessments the questionnaire will also ask for opinions. See section ?? for details and the full survey at section ??.

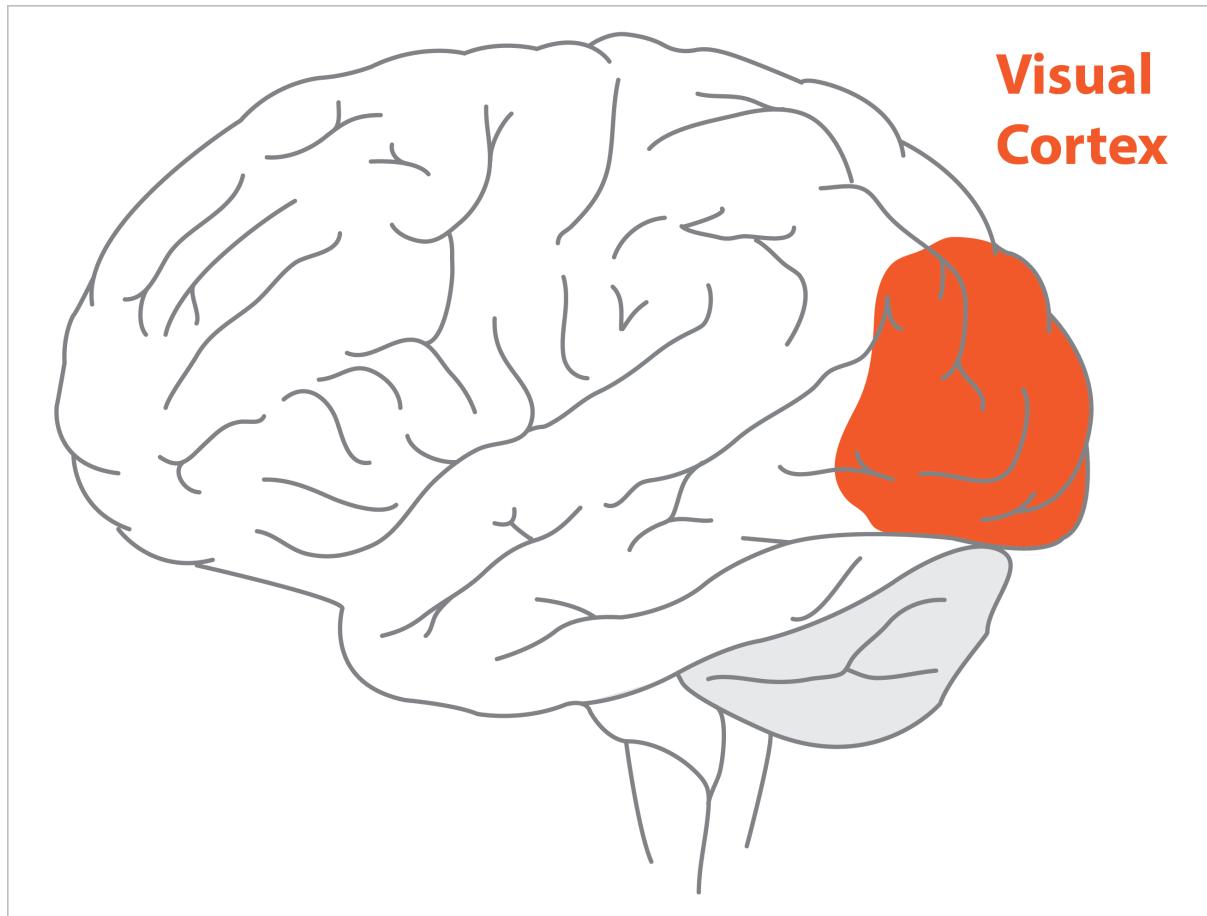


Figure 3.1: Position of the visual cortex. Left is front. Source: ?

3.2 Remote control interface

Figure ?? shows the interface, which the user will see in the experiment. It resembles the number pad on a conventional TV remote control in structure. The circles with the numbers will be the neurotags, which the participant has to look at. To measure the activation time, one neurotag will be visually cued - represented by the highlighted 6. As soon as this neurotag has been activated, a new visual cue will be given. The sequence for each activation will be the same for every participant.

A total of 50 different cues will be given. The order of the activation is randomly so there is no way the participant could know beforehand which neurotag will be cued next. Since the t_1 cannot be measured without significant effort and the absolute time is not important, it is assumed that the t_1 time for each individual is about the same for every age. Under this assumption only the t_2 will vary, which is the desired goal

3.3 What data will be collected

The data collection will be structured in three blocks: Two questionnaires before and after the experiment which the participant had to answer and data collection during the experiment. The first questionnaire consisted of general questions regarding demographics and medical conditions:

3 Survey Contents

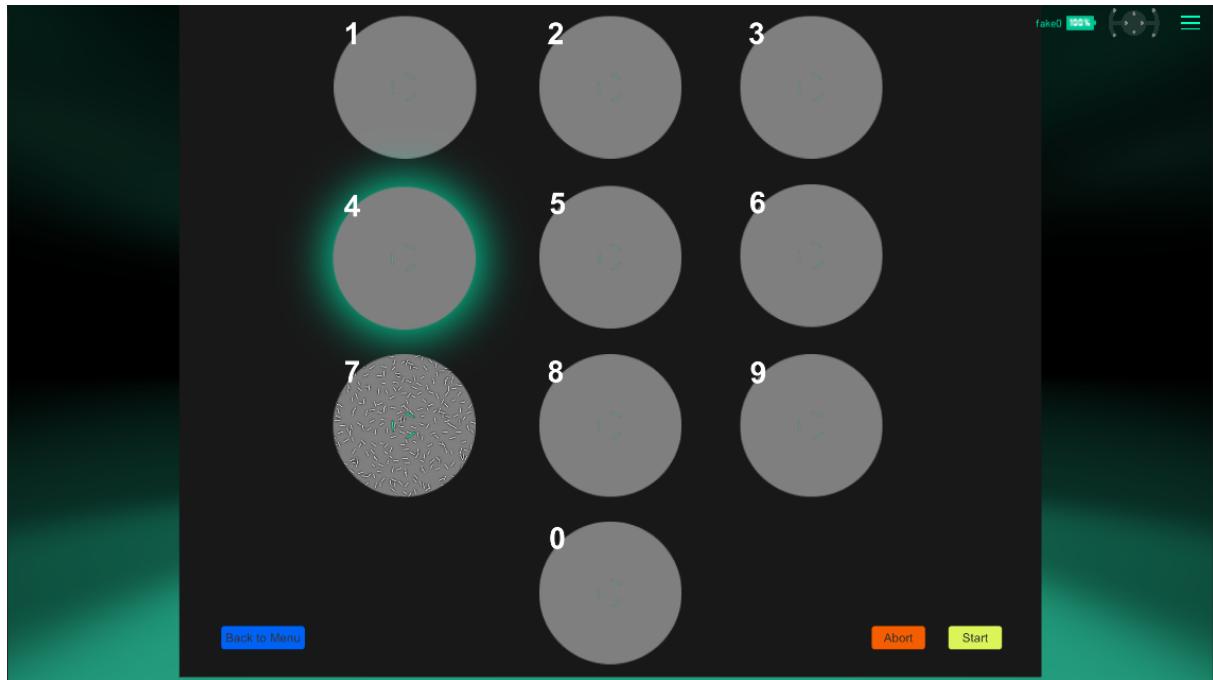


Figure 3.2: Interface of the remote control experiment. Ques are highlighted.

- Demographics: Age and Sex
- If the participant wears glasses
- If the user felt motion sickness during the experiment
- Cognitive impairments
- Medication known to affect EEG readings
- Drug intake within the last 12hrs

The collected data during the experiment was collected on the VR headset (**application**) and stored in a file, which was available after the finished experiment. This data contained the following readings:

- Speed of each neurotag activation from the point where a cue was given. This is the start of t_1 until the end of t_2 , where the sensor is confident that a certain neurotag was looked at.
- Confidence score of each activated neurotag upon activation until another neurotag is activated.
- Quality readings from each single electrode to monitor the tracking quality of the sensor, in order to explain outliers in the collected data.

After the experiment the participant answered the second block of questions, which consisted of questions about the general experience of using the BCI:

- If the user felt motion sickness during the experiment (rated 1...5)
- comfort of wearing the sensor (rated 1...5)
- perceived time to setup the sensor (rated 1...5)
- ease of use (rated 1...5)

3.4 Experimental Protocol

The data is collected in a single session lasting between 10-20mins for each participant. Each session starts with the participant having to complete the first questionnaire ?? and the form of consent ?? for anonymized data processing. At this stage potential hazards, problems and dispositions as discussed in section ?? are identified. Afterwards each participant is introduced into the experiment. The working principle of a SSVEP is explained and questions regarding data privacy are discussed. Subsequently the GUI of the experiment is shown to the participant with a short introduction about the nature of the task.



Figure 3.3: Experimental setup

Then the participant was equipped with the BCI Sensor and placed in front of the screen as shown in figure ???. The distance to the screen was varying since each participant needed to adjust according to eyesight. The first step of the experiment was the calibration of the sensor

to the participant. In case of non-optimal readings, the position of the sensor was adjusted accordingly until all electrodes showed at least *very good* contact (figure ??).

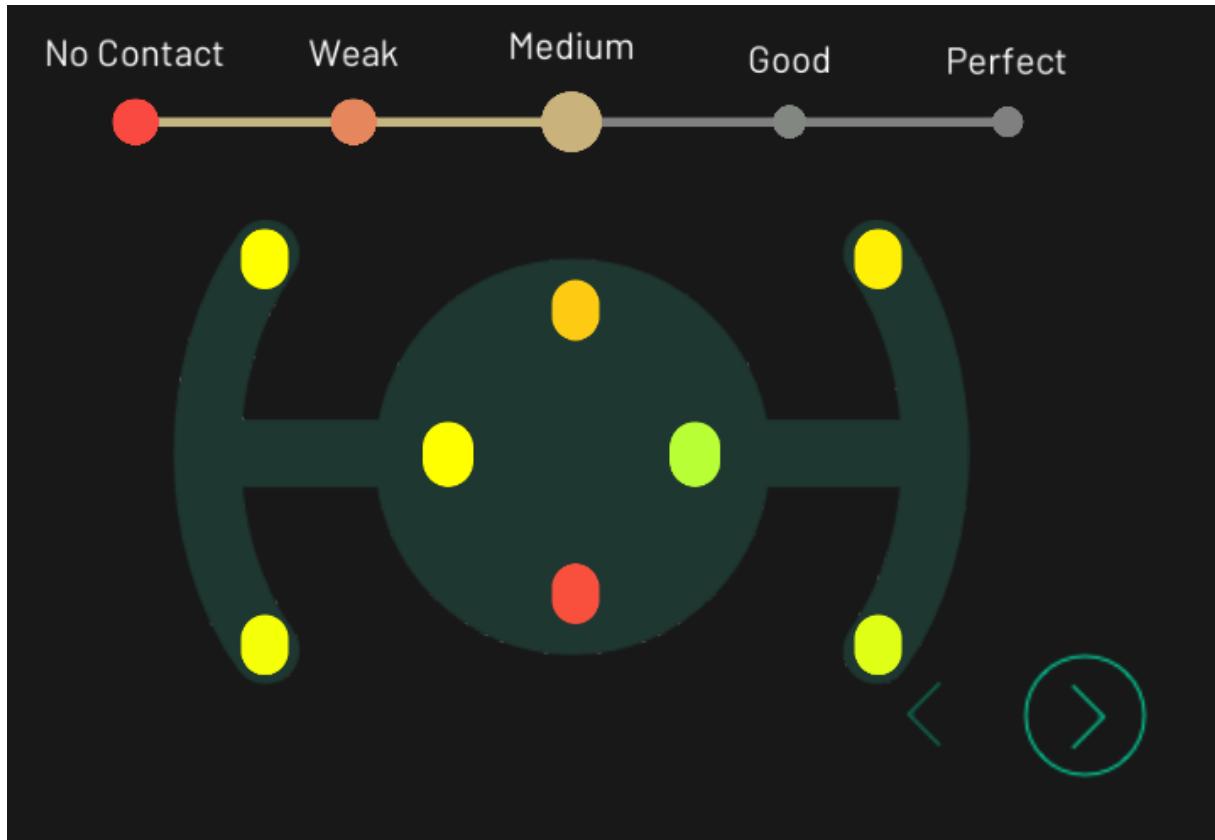


Figure 3.4: Quality readings of the sensor

Then the experiment was started. The participant had to perform 50 neurotag activations. Each one consisted of a visual cue (see figure ??) on which the participant had to focus on. Upon a successful activation of the cued neurotag, the readings were saved to the logfile (see ??) and the next cue was given in order to keep the $t1$ timing as short as possible.

After completion of the experiment, each participant had to fill the second questionnaire ??, which gathered information about the experiment such as wearing comfort, perceived difficulty and questions about the opinion towards the technology in general.

4 Survey results

This chapter gives a brief insight into the collected data by means of visually examining them. The main focus point are the detection times for different clusterings of the population by the criteria provided by the questionnaires

1. Raw detection times
2. Clustered by gender
3. Clustered by glasses or no glasses
4. Clustered by age group
5. Clustered by different target

Also the other results of the questionnaires are discussed: Wellbeing, Dizziness, Focus, Comfort, Time and perceived Difficulty. Also the personal opinion and whether or not the participants found this technology useful or would use it themselves.

4.1 Raw results

The study gathered 1370 different target hits datapoints from 28 participants. Figure ?? shows a histogram of the achieved detection times. The y-axis is scaled logarithmically because approximately 870 out of all detections were made in under three seconds. If scaled linearly, this would have heavily skewed the overall plot. Also the occurrence of detection times longer than 15-20 seconds are extremely rare in comparison.

This is also shown very clearly in figure ??, where the time window shows only the results between two and ten seconds, which also constitutes for the majority of data points (over 95%). Some outliers where the detection took over 120s, were removed from the sample. The clustering of the detection right at the 2.3s mark, which is indicated by the red line with almost no other data points below that threshold might indicate, that this threshold is the minimum amount of time required to confidently identify a target. The logfiles show that a confidence score of >0.3 qualifies for a successful identification. See Appendix ?? for details. This score is not displayed for targets, which are identified under a certain amount of time, since there aren't enough samples supposedly.

The fact that most of the detection times center around that value might also be an identification that 2.3s is the minimal achievable time by the BCI in conjunction with the SDK on that particular testsetup. In order to have a sufficient certainty without too much noise, it is reasonable to assume that the confidence value has to be above the 90% threshold for a defined timeframe in order to qualify for a positive detection.

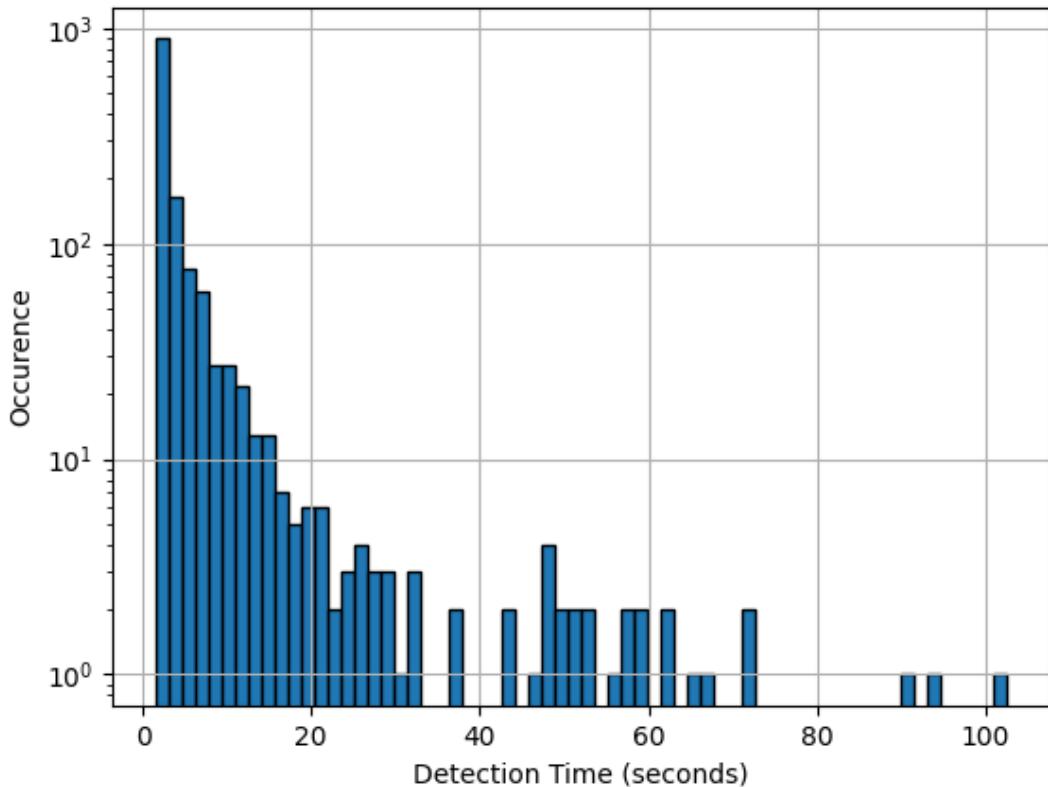


Figure 4.1: Histogram of achieved detection times

If put into perspective, figure ?? shows the same distribution but with mean times for every participant as histogram. What this data shows is that the few outliers in the achieved detection times are likely caused by a few individuals and the the majority of the participants had a very consistent timing below 10s and more so close to the minimal achievable timing of about 2.3 seconds.

This gets even clearer, when the mean times for each individual are shown in context to each other with the age dimension added. Conclusively what these results clearly show is that the majority of the participants who were younger than 40 showed a very consistent timing at or close to the minimum time.

4.2 Telemetry

The *NextMind* Engine provides telemetry readings on request. Figure ?? shows a boxplot of all quality readings prior to all successful conducted experiments. Each boxplot stands for one distinct electrode at the sensor. The scale is *Worst to Best* in the typical *Red to Green* convention with the indicated steps in between. The SDK provides a readout with a score from 0...100.

Apart from some outliers, there are no significant systematic errors, which might cause a significant error in the acquired data. There is a very small tendency in the outermost electrodes to have slightly less quality over the course of the whole experiment what can be attributed

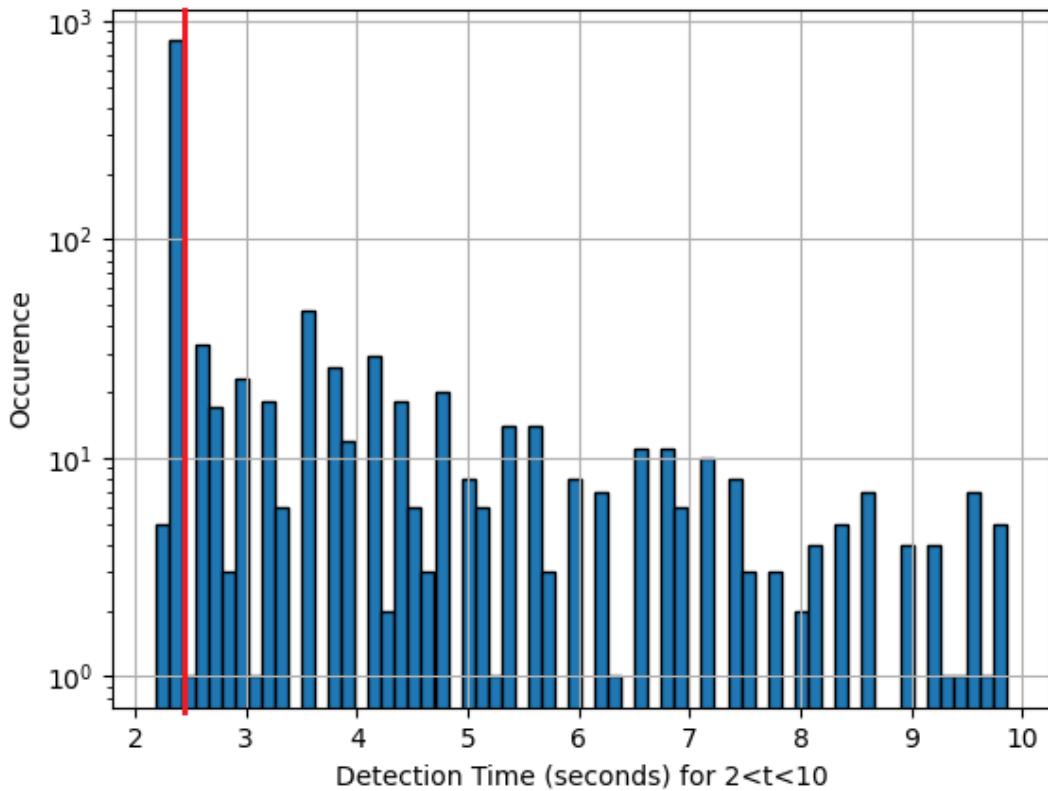


Figure 4.2: Narrowed window of achieved detection times Occurrence [log s], Move Red Line

to being the outer and bottom electrodes which might not have ideal contact on an individual basis if the skull shape is more elliptical than usual.

4.3 Demographic differences

To rule out potential side effects, the first questionnaire required answers for gender, whether the participant wears glasses, neurological conditions, alcohol or drugs and colourblindness. However the latter four questions were answered negative entirely. The only exception was a neurological condition which resulted in the experiment not being conducted altogether in that particular case. This leaves the gender and if the participant wears glasses as the only major demographic parametres.

When the genders are compared side by side in figure ??, there is a tendency of women having a more consistent timing. The average value of both is approximately on the same level but the male results have a larger upper quartile and whisker therefore being less consistent in timing. As in figure ??, the window was narrowed down to 10s max, to cut large outliers. However since the age distribution across genders is about the same, this won't have an effect on the outcome in regards to the hypothesis. Nevertheless a quantitative study with a large number of participants could prove if there actually is a significant effect on detection times.

Figure ?? also shows the result that participants without glasses have a higher consistency in

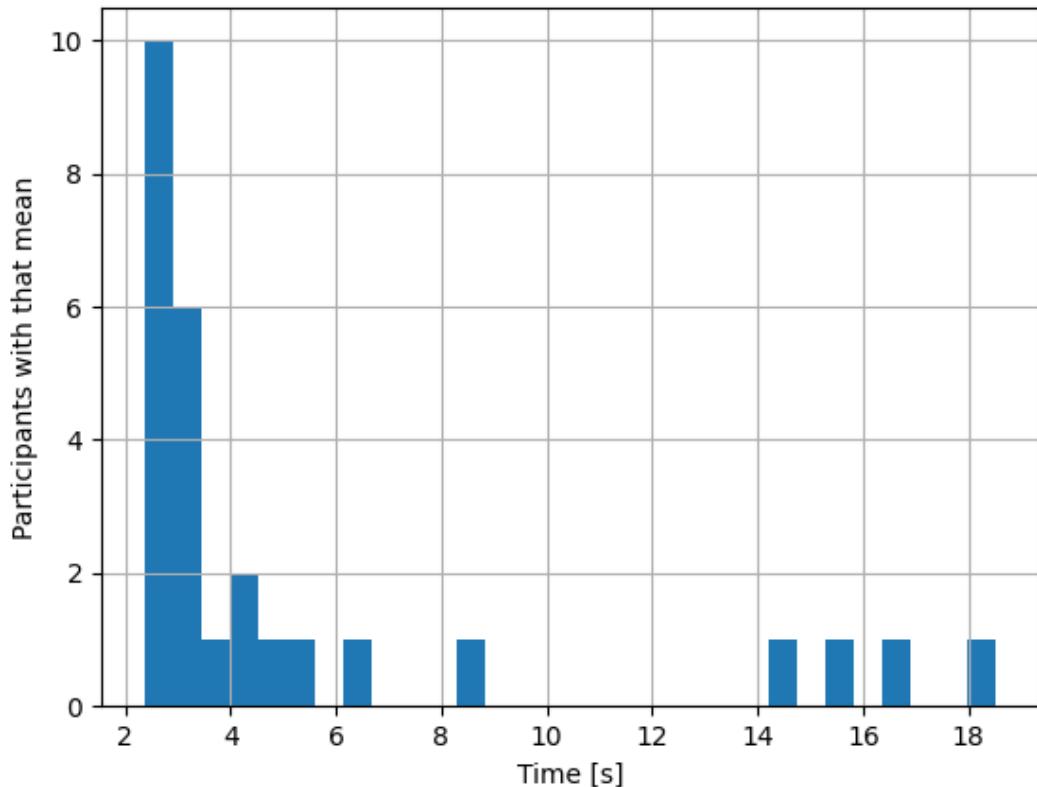


Figure 4.3: Achieved mean times per participant histogram

their target times. Although the average being at roughly the same level, the upper quartile and whisker is far more spread with the participants who are wearing glasses. Since the participants which were wearing glasses tend to be about 10 years older in average than their non glass wearing counterparts, this effect seems to be tied to the age itself and not the fact the participants actually cannot see as good.

In order to have a better understanding of the difference between younger and older participants, the whole population of the study has been binned into three different age groups:

- 0-30
- 31-60
- >61

The results are show in figure ???. The most prominent feature is the age group above 61 years old. Their average time to detection of the target is between 5 and 7.5 seconds. However almost indistinguishable from the 0-30 age group is the 31-60 age group but on a close look the graph actually shows a slightly slower detection time here as well. So the first preliminary finding is that there in fact is a tendency that with increasing age the detection of a target takes longer.

These results will be later set into context with other results from the survey to show a clear picture of the findings of this study.

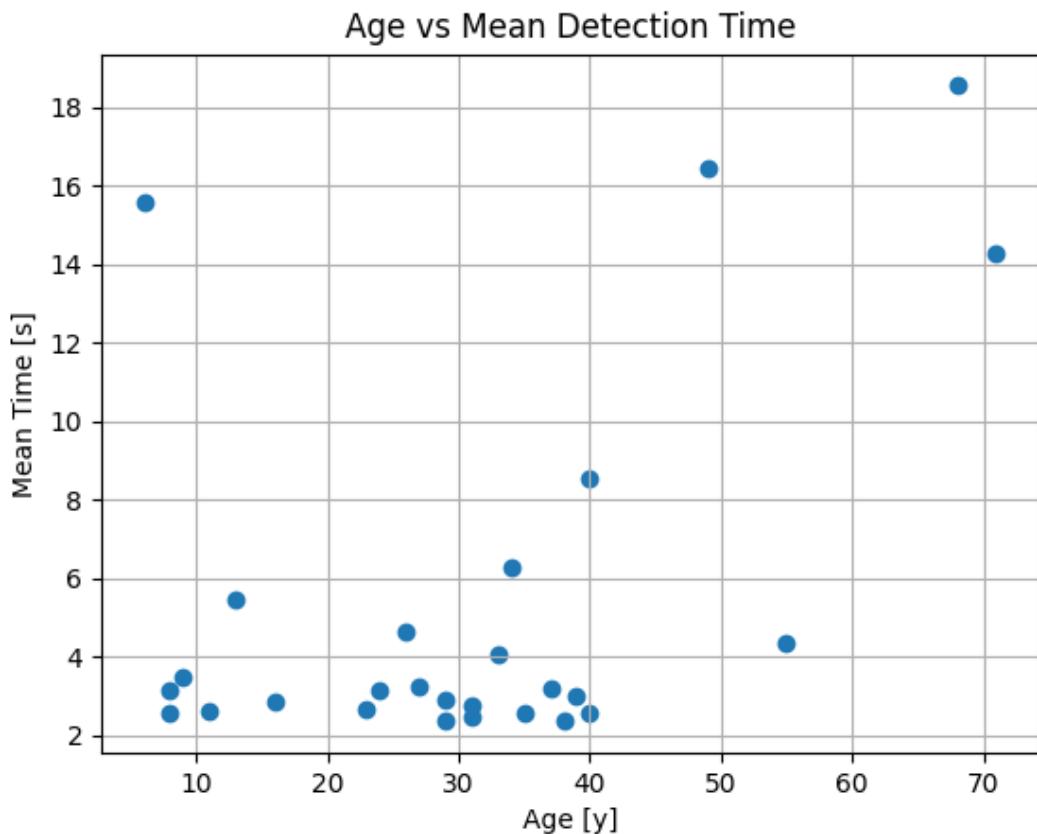


Figure 4.4: Mean detection times plotted against age

4.4 Detection time differences for each target

One effect observed during the experiment was, that some targets were generally faster detected than others. Figure ?? shows this phenomenon:

This effect comes presumably from the layout of the experiment itself, where the participant sat in front of a laptop and the interface was placed in a different angle. Therefore the targets, which were arranged in rows and columns (see figure ??), were being seen in different angles. Since the test series of cues was the same for each experiment and close to evenly distributed in their occurrences, there is no systematic error in the experiment itself.

4.5 Conduction of the Survey

Two participants had not undergone the experiment itself due to an insufficient calibration score. One case was due to missing glasses for a clear view of the screen and another case had concentration problems. Both cases were in the above 60 age group. This is especially noteworthy since in section ?? the assumption was made that reduced eyesight does not have an influence, since glasses can be worn. This is a strong indicator that without proper visual resolution of the neurotags, the sensor does indeed not work. One participant also from the above 60 age group had to end the experiment prematurely due to concentration difficulties and eye strain.

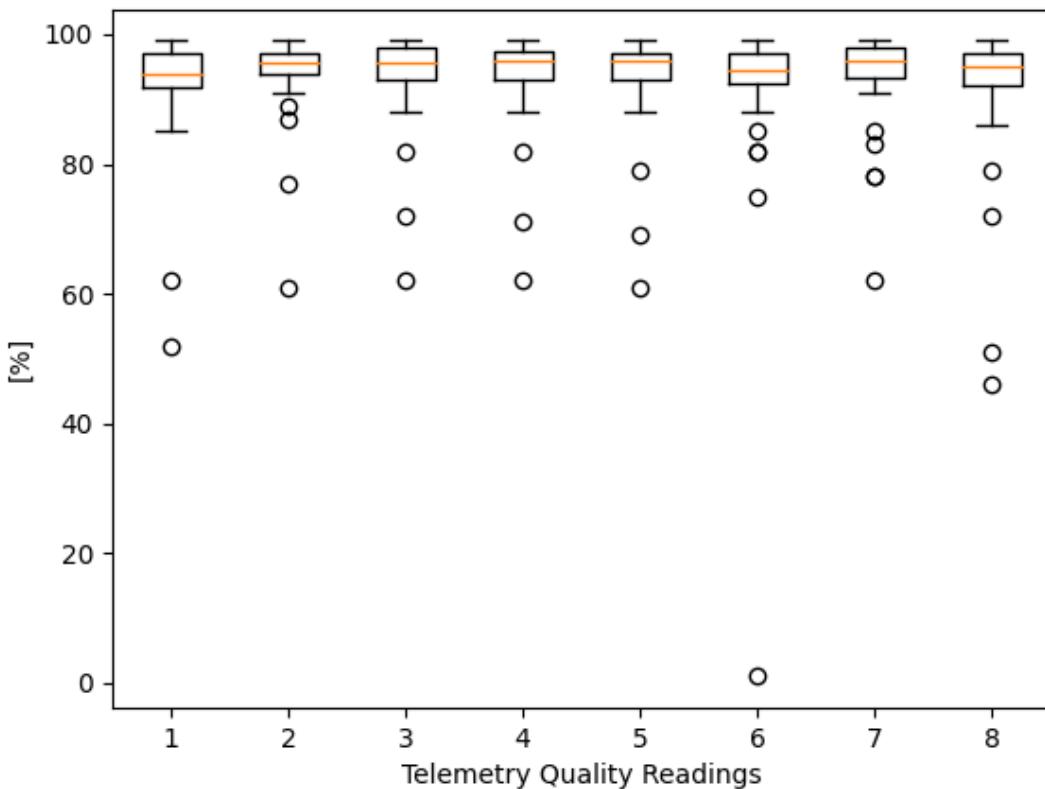


Figure 4.5: Quality readings of the sensor prior to the experiments

Eye strain was reported from four participants. However there was no majority among any age group in this case which means that this does not indicate a correlation. Two participants wore contact lenses. One participant reported that focussing on the experiment made her forget about blinking her eyes, causing dry eyes.

Two participants reported that they felt the SSVEP physically inside their brain with a pulsating or tickling sensation. However, this was not reported to be painful.

In many cases the calibration wasn't successful at first try. Usually a sufficient score was reached after 3-5 attempts. Also for some participants the optimal sensor position was in a different place on the back of the skull than recommended by *NextMind*. In these cases the results were perfect instantly after a relocation of the sensor to a different position.

One participant inadvertently altered the position of the sensor which did not have a significant effect on the detection time. Also with this same participant the calibration score was comparatively low. Although not explicitly tested in this survey, this might indicate that the *NextMind* sensor has a resilient design to a shifting position. This would allow for applications which includes physical movement.

One individual could not participate as she had epilepsy, what might have gotten triggered by the experiment.

Figure ?? shows the distribution of the age of all the participants in the study. There is no age group which is particularly under- or overrepresented. The range of all participants was

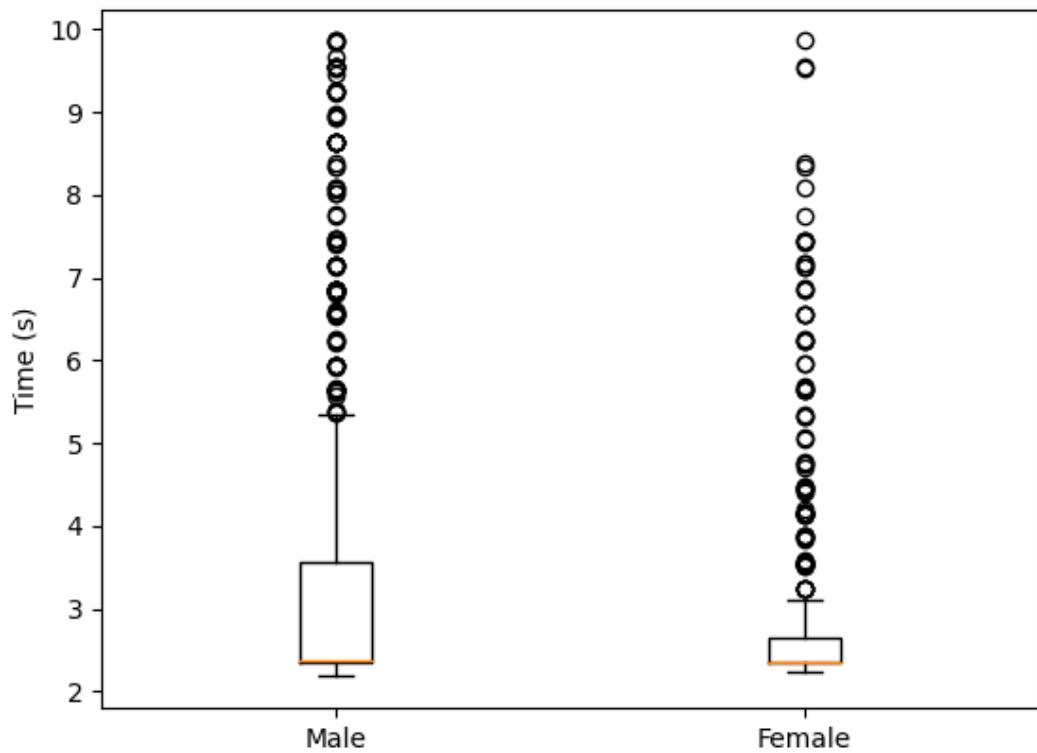


Figure 4.6: Detection times between Male and Female participants

between 6-81 years old, resulting in a big cross section. It is noteworthy however, that not all of the participants represented in this chart also generated results for the regression curve in figure ???. The reason being that predominantly in the above 60 age group, some participants had problems with the calibration due to reduced eyesight or concentration problems.

4.6 Questionnaire Results

While the first survey had the primary goal to evaluate the Demographic data for the experiment and whether the subject was even able to conduct the experiment from a medical standpoint, the second survey ?? asked questions about the parametres of the experiment itself.

These were:

- Wellbeing: How the participant felt that way, which might impact the ability to concentrate
- Dizziness: If the participant feld dizzy after the experiement
- Focus: How easy it was for the participant to focus on the task
- Comfort: The comfort of wearing the headset.
- Time: How the participant perceived the procedure of calibrating the sensor

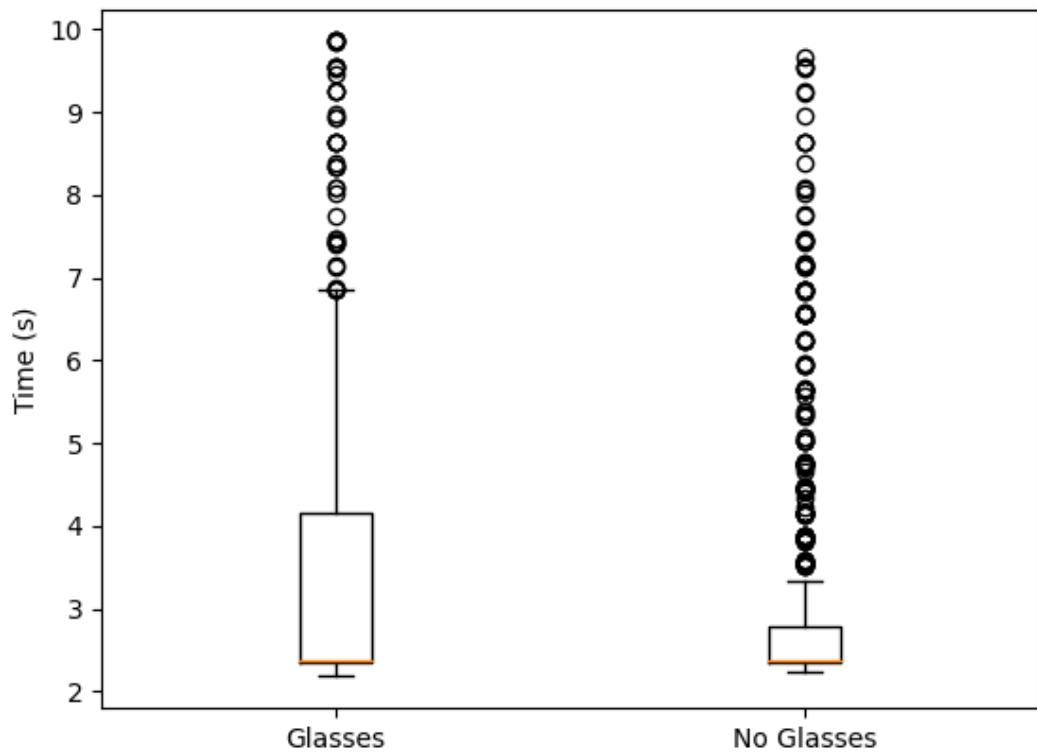


Figure 4.7: Detection times between participants with and without glasses

- Difficulty: If the task of targeting Neurotags felt difficult.

Figure ?? shows the results from the survey on a 1...5 scale. The **Wellbeing** on the day of the experiment of most participants was generally good. Only one participant had back pain, which however didn't impair the ability to concentrate. Almost no participant felt **dizziness** after the experiment. **Focus**, **Difficulty** and **Time** didn't show any noteworthy results. Interestingly there was no real tendency to the **Comfort** of wearing this sensor. The opinions about the device are very evenly distributed, which is interesting because some participants apparently found it very comfortable to wear whereas others found it really uncomfortable.

In terms of opinion about the technology, which is not part of this study per se but interesting to gauge a general opinion nevertheless. Only 3 out of 25 participants are sceptical towards this technology. Three persons abstained from giving an opinion and two persons didn't fill the second questionnaire. 15 out of 30 participants would use the technology themselves with the remaining 15 answers consisting of two abstentions, two not filled questionnaires and the remaining 11 being negative answers. However, 24 participants think that this technology provides a benefit in the right domain. Again two persons didn't fill the questionnaire and four refrained from answering the question.

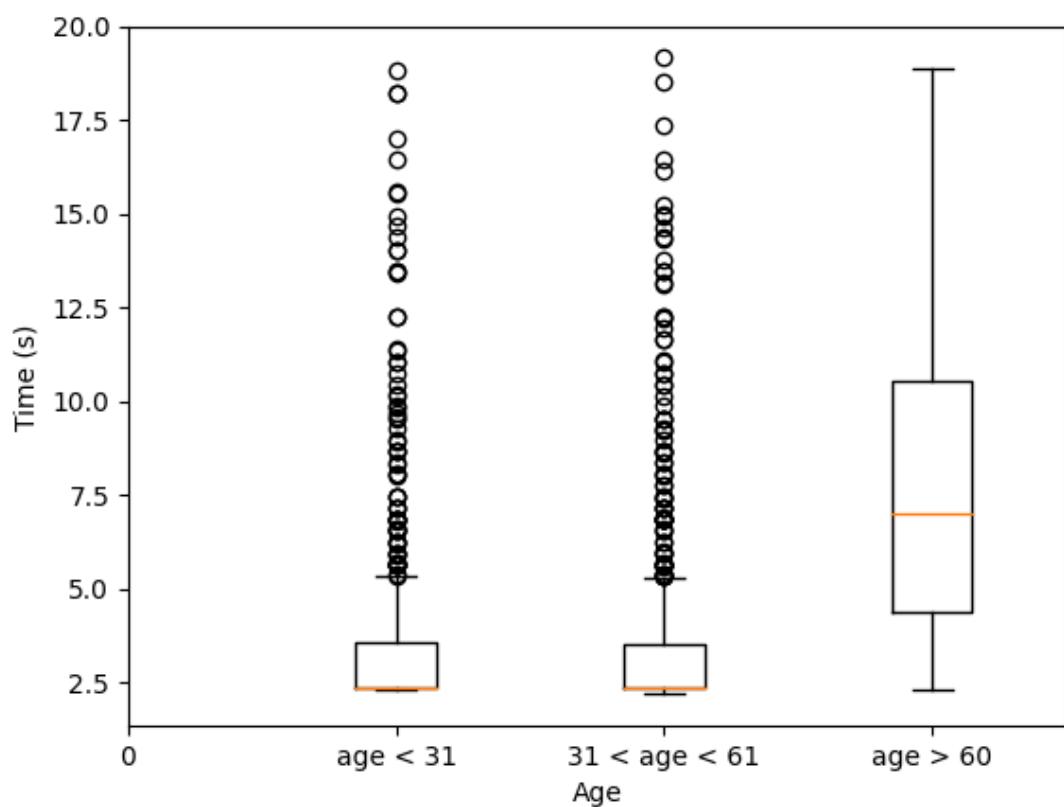


Figure 4.8: Detection times for three age groups

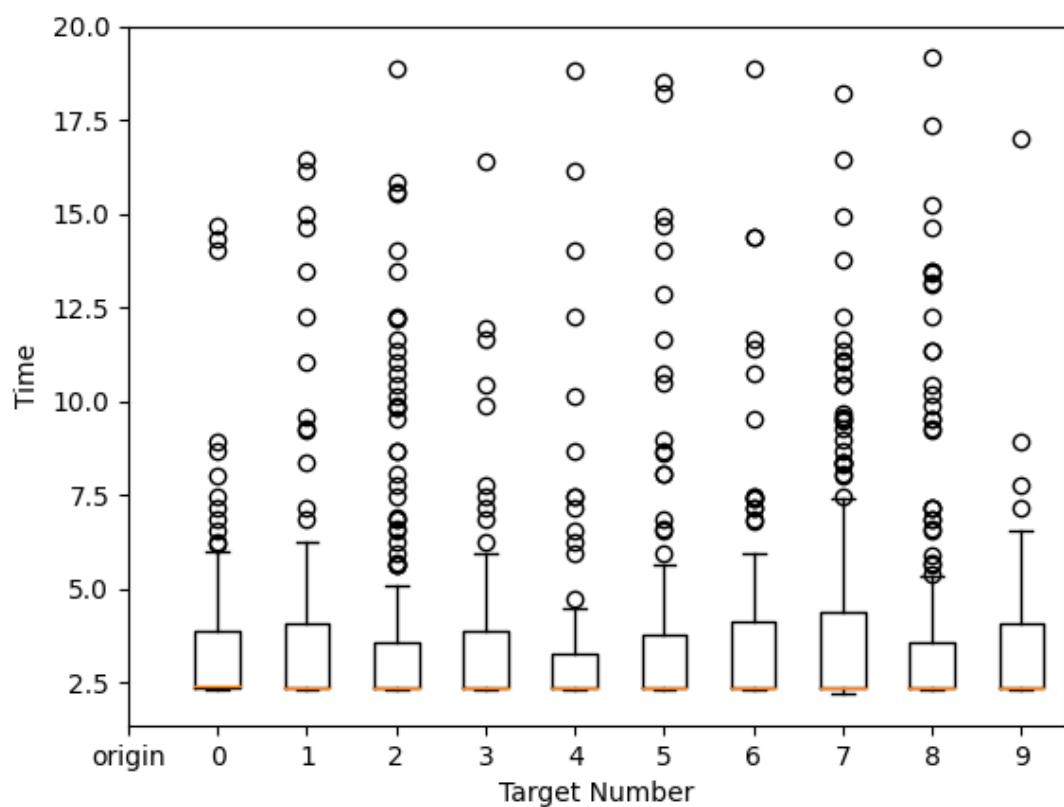


Figure 4.9: Detection times for the different target

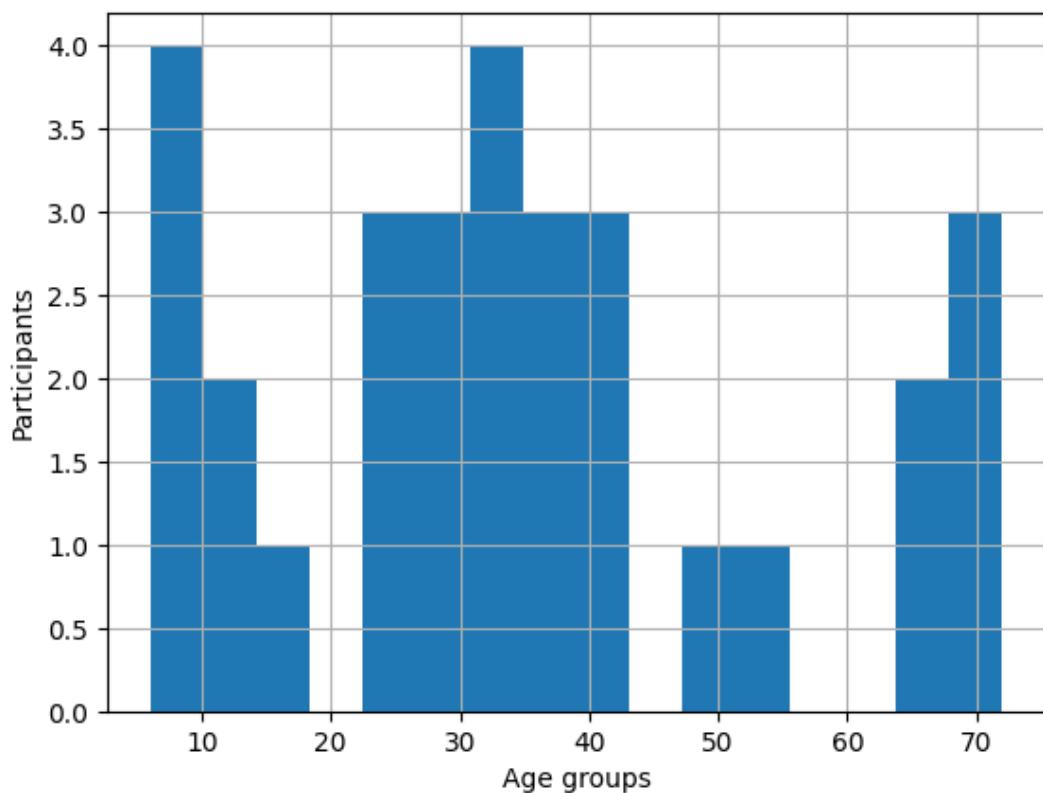


Figure 4.10: Age distribution of the participants

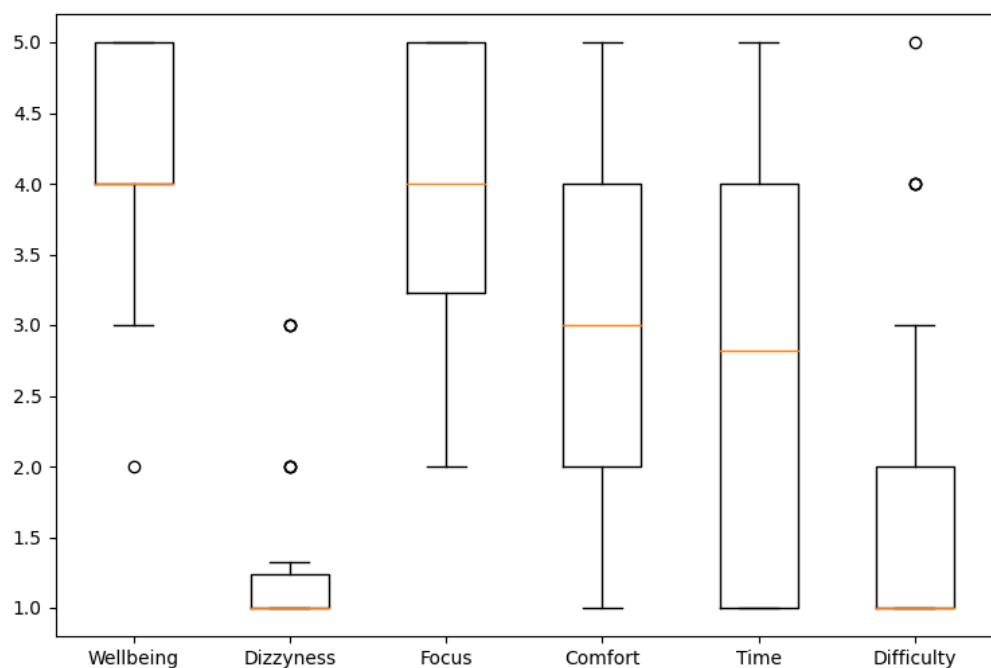


Figure 4.11: Results of post experiment user survey

5 Findings

5.1 Age as a factor in BCI usage

Section ?? discussed, how a regression graph will be used in order to identify a trend between age and expected performance in regards to BCI usage. Figure ?? shows the same values as figure ?? but with added regression graph and the defined 95% confidence interval:

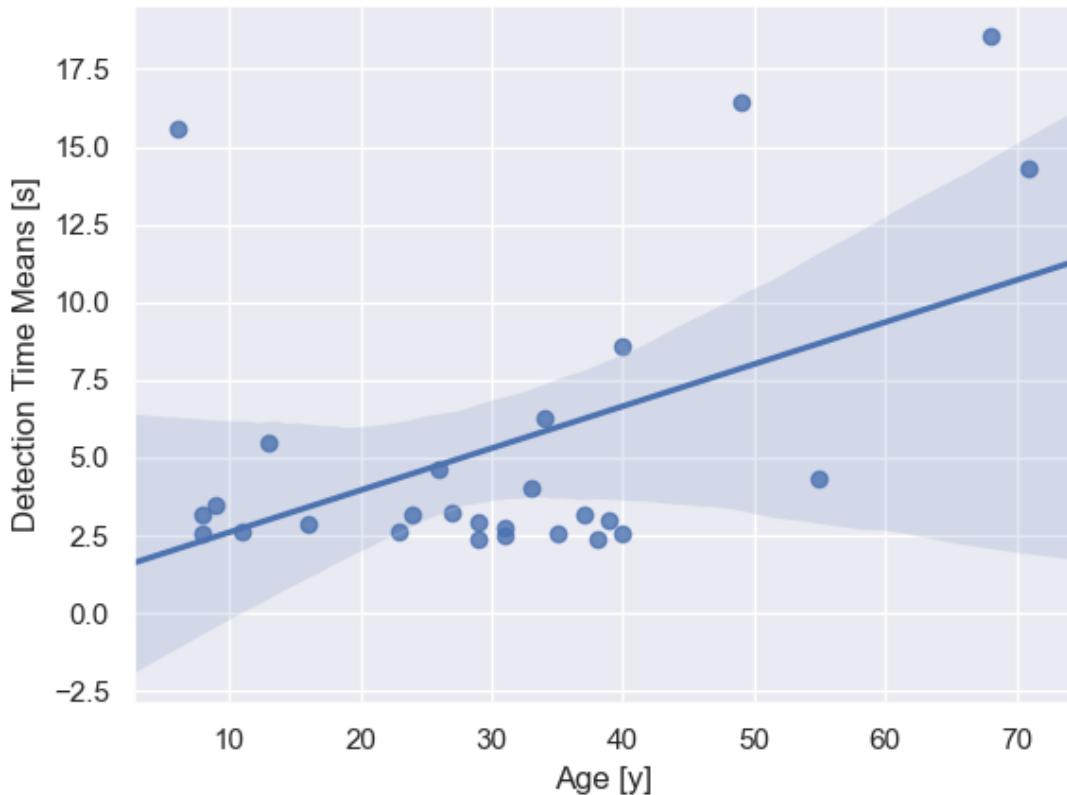


Figure 5.1: Regression graph with confidence interval

The general trend is that there in fact is a correlation between age and achieved detection times. However, taken the confidence interval into account, this result is neither clear nor definitive. If the experiment were to be repeated with the same parameters and all the results end up within the pictured confidence interval it might very well be that new trials would in fact yield a result, where the gradient of the regression graph might be negative. Nevertheless, there are two distinct key features in figure ??:

5 Findings

Consistency Both younger and older people tend to be less consistent with their achieved detection times. For the younger people below the age of 20 there aren't simply enough datapoints because two outliers already skewed the resultset too much. In fact the youngest tested person with an age of six performed almost as bad as the oldest person. A possible explanation is that this participant wasn't able to concentrate on the experiment, which required dedication and focus on the task at hand. The observation of this particular participant was that he started to move around on the chair and loose focus on the task, which resulted in longer detection times. A possible explanation is that the other children were already in 3rd grade upwards and didn't have any issue with focussing on a task for a few minutes. So in this case age does have an cognitive influence but not necessarily in a way which was anticipated.

For older people the consistency is also spreaded more than it is within the 20-40 years range, which achieved by far the most consistent results. The reason why older people tend to deliver less consistent results for many different trials can be also attributed to the ability to concentrate for prolonged periods. In fact, if one single cue took longer to detect, this had a domino effect because the concentration deteriorated even faster, resulting in considerably longer detection times later on.

Missing Data Points Knowing the results of the survey and the age distribution (figure ??) shows that not all participants, which are represented in the survey also participated in the experiment. The reason being that three participants in the above 40 age group had bad calibration scores and thus could not even participate in the experiment itself. Although this obviously cannot be added to the study as definitive result, it underlines the general performance of that group, where a less consistent overall performance can be expected.

5.2 Mathematical proof

Figure ?? already showed that within a 95% confidence interval given enough data points, the same study should also yield results within the same range if the parameters of the study are not altered. This section continues on the assumption that with a sufficiently large number of participants, the mean times density displayed in figure ?? will asymptotically approach the kernel density function overlayed on the histogram.

Under this assumption the probability of a *Type II Error*, which is represented by the β parameter can be determined. To do so, the following parameters have to be calculated first in the following part:

$$r_{xy} = \frac{S_{xy}}{S_x S_y} \quad (5.1)$$

Where r is the correlation coefficient in the range between $-1 < r < 1$. 1 in this case means perfect linear correlation and therefore a perfect positive correlation and -1 a perfect negative linear correlation. S_x and S_y are the standard deviations for the corresponding datapoints: $S_x =$ age and $S_y =$ detection time. The standard deviation is calculated with the following equation:

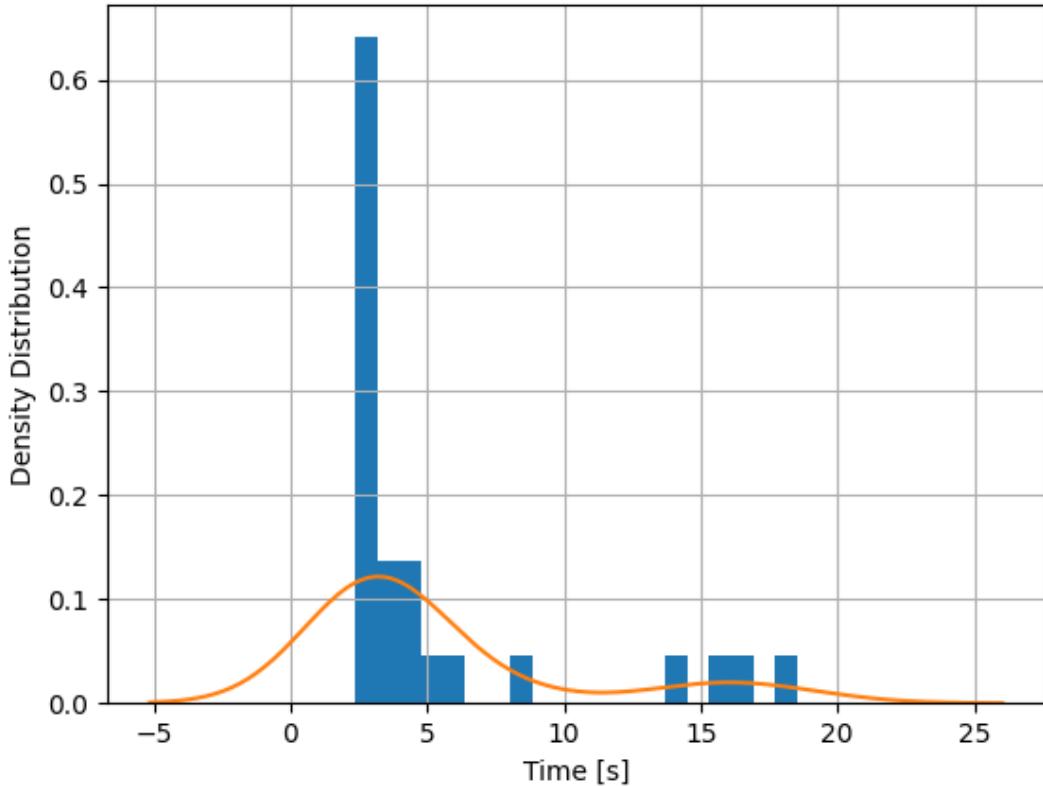


Figure 5.2: Regression graph with confidence interval

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}} \quad (5.2)$$

Where x is a single element of the sample and n the number of samples. For equation ?? the different x and y values have to be multiplied for the S_{xy} value, which is called *covariance*. The following equations will use the S_x notations with . With the numbers from the experiments, these numbers yield the following results:

$$S_x = 17.7 S_y = 9.11 S_{xy} = 36.9 \quad (5.3)$$

Using formula ?? with the results under ?? yields:

$$r_{xy} = 0.23 \quad (5.4)$$

This variable r will be later used as the effect size in the β calculation.

The next step will be calculating the z-score using the distribution shown in figure ?? . The equation for that z-score is as follows:

$$Z = \frac{x - \mu}{\sigma} \quad (5.5)$$

Where $x = 1 - \alpha/2$, μ is the mean of the sample and σ is the sample standard deviation. Using that values, the Z

Scoreresults to :

$$Z = 1337 \quad (5.6)$$

Which indicates a **strong/weak/inconclusive** relationship between age and detection times.

Section ?? already mentioned that in order to rule out *Type I Errors* and *Type II Errors* the α and β parameters have to be defined. In fact α was already set to 0.05 or 5% but β has to be calculated. The official formula is:

$$\beta = \text{something} \quad (5.7)$$

The contents of this formula are...

By using that formula a table for different β or rather $1 - \beta$ e.g. Powervalues can be calculated to evaluate the number of necessity trials necessary to prevent a *Type II Error*.

| Col1 | Col2 | Col2 | Col3 |
|------|------|-------|------|
| 1 | 6 | 87837 | 787 |
| 2 | 7 | 78 | 5415 |
| 3 | 545 | 778 | 7507 |
| 4 | 545 | 18744 | 7560 |
| 5 | 88 | 788 | 6344 |

Table 5.1: Table will be replaced...

Expand a little on results here.

The most apparent feature is in fact that there is a positive correlation. Up until the age of fourty, the results are fairly uniform in the 2-5 seconds range. People above this age however have a significantly higher mean in this study. One possible explantion can be given in regards to concentrativeness: Older people reported in the subsequent questionnaire that they sometimes had difficulties in focussing on the task. This coincided with longer detection times clearly observable during the experiment. Also one subject in the 40s-range reported that due to a long and stressful day at work, his ability to concentrate was significantly diminished. One cause, which can be ruled out beyond all doubt is that older people had a problem with the interface or the usage of the BCI itself. Almost none of the participants from very young age to the oldest ones reported difficulties in the task - see figure ?? : Difficulty.

Formally: with a 5% confidence intervall painted on the regression curve - which was required according to section ??, there still is a significant positive correlation between age and mean detection time. Therefore the hypothesis is disproven.

5.3 Other findings

The collected dataset allowed for other interesting insights into possible performance parameters of the participants, which were already briefly shown in the ???. This section briefly discusses these results and interpretes them in the context of this study.

5.3.1 Glasses

Whether the participant wears glasses or not seems to have an influence on the outcome on the performance. However, this might coincide with the fact that participants with glasses were over 10 years older in average: 30.4 years for the non-glasses group and 40.5 years for the glasses group. This is especially reasonable to assume since people tend to have more issues with eyesight the older they become.

5.3.2 Gender

According to figure ??, there seems to be a difference between male and female participants when it comes to BCI performance. This could be an interesting starting point for further research into the *why* of that anomaly. Since this study was not designed to provide an answer for this question, no further conclusion will be drawn, since it will be pure speculation. As for the influence on the dataset for this study, the survey results reflect that there is no huge age gap between the genders, which might reflect a systematic error on this side. In fact, the female participants were older with an average age of 37.3 years in comparison with the male participants with 33.2 years.

5.3.3 Target position

In regards to the timing for the different targets, which is shown in figure ??, the result is inconclusive. Although the question was raised during the conduction of the experiment, whether there it makes a difference in this setting, there weren't enough data-points to prove or show a significant difference in the timings between the targets with certainty. Nor was the experiment itself structured in a way which allowed for conclusive data collection. There is also no influence on the collected data for the main objective, since every participant was shown the same experiment. Hence any systematic error would have manifested evenly across all participants if it exists.

6 Conclusion

6.1 Results

Put this in here:

- Refer to Findings
- construct argument whether formally rejected or not

6.2 Future Work

There are several ways to continue with this research in subsequent studies, based on the outcome of this study. These potential studies reflect on the collected results and several key points, where this study did not dive into due to the different scope. Some of them are briefly outlined in the following paragraphs:

Gender Based on the gender gap in section ??, there might be a factor in the difference in genders, which has not been discovered yet. Since there was no apparent age gap between the two genders - the female participants were in fact older - which might explain the result, there might be another factor governing this. However, the results were inconclusive. Therefore a quantitative study to reveal a systematic influence might be worthwhile.

Quantitative Study This study could be expanded into a quantitative study in order to gain a deeper understanding of what the drivers are behind the influence of age. Especially since the ability to concentrate might have an influence here as well. A larger number of participants in the order of around a few thousand individuals very likely will draw a clearer picture of potential causes. The questionnaire could be improved as well to gain a better understanding of the process and learning curve of each individual.

Positions Briefly discussed in the results and findings was the subject of a potential influence of the target positions. Especially in regards to a potential required movement of the head in order to focus on new cues. This will very likely increase the t_1 time - see section ?. An approach, where the target positions are varied across a wide range of different positions might exhibit different results in regards to timing. This also has some relevance to theoretical applications in controlling machinery or large interfaces - see section ?? for details.

6 Conclusion

Learning One factor, which this study was not able to collect a sufficient amount of data on, was the learning progress of different participants. During the conduction of the experiment it was subjectively observed that the reaction times for each participant were faster in the later cues of the whole experiment compared to the earlier ones. However the data was inconclusive due to a high noise in the dataset for different individuals. A potential study could evaluate if for smaller subsequent cues the participant learns to use the interface on a cognitive level and has shorter average detection times on subsequent trials.

Other demographic factors Section ?? discussed the answers to different demographic factors, which might have had a systematic influence on the results. Apart from the necessity to wear glasses and the gender, there were no participants providing a different than negative answer. Nevertheless, it would be interesting to see whether colourblindness, intoxication or medication would result in different performance. Especially caffeine consumption as a popular neurological effective drug might be worthwhile to examine.

7 Acknowledgements

I want to thank Prof. Dr. Roland Greule and Dipl. Inf. Rüdiger Höfert to examine this study and always set the right impulses to guide the outcome of this thesis. Rüdiger Höfert in particular helped me out with the procurement of necessary equipment to conduct the study. I also want to express my gratitude to the Hamburg University of Applied Sciences to provide the necessary framework to obtain my degree and aid with the procurement of necessary equipment. I also want to express my appreciation for every single participant who took part in my experiments despite certain constraints due to the Covid19 pandemic. A special thanks goes to B.Sc. Moritz Bednorz, who helped me sharpen and clarify my knowledge in the field of medical engineering and design of experiment, which gave this thesis a coherent perspective on all domains. I will be eternally grateful to my father Dr. Thomas Neudecker to set the right impulses which made pursue an academic degree. He may rest in piece. My biggest thanks is dedicated to my girlfriend B.A. Alexandra Michels for her emotional support and creative input to some of the figures in this thesis.

A Material

A.1 Survey forms

The following forms were used during the experiment. They are written in the german language since every participant was native german.

Every document is in its original layout, hence the difference in terms of layout to the rest of this document.

Informationsbogen vor dem Experiment

Informationsblatt

In diesem Experiment wird ein sogenanntes *Brain-Computer-Interface* (BCI) verwendet, dieses misst ähnlich einem EKG (Elektrokardiogramm) die Gehirnströme. Dafür werden sogenannte *Visuell Evozierte Potenziale* (VEP) mittels spezieller visueller Stimuli erzeugt. Die Messergebnisse werden anschließend maschinell mit Hilfe eines mathematischen Modells verarbeitet und Steuerbefehle für einen Computer daraus generiert.

Das Experiment verursacht **keine** Schmerzen oder sonstige ungewohnte oder unangenehme Empfindungen.

Das Experiment verursacht **keinerlei** gesundheitliche Schäden durch Strahlenexposition oder das Einbringen von Fremdkörpern in den Körper.

Es können **keine** Gedanken gelesen und **keine** Daten gesammelt werden, die später einem Individuum und seinen politischen oder persönlichen Ansichten persönlichen Ansichten zuordenbar sind.

Es werden **keine** Daten gesammelt, die an eine dritte Partei über das Internet oder sonstwie übertragen werden.

Es findet **keinerlei** Beeinflussung des Gehirns durch den Sensor statt. Zu jedem Zeitpunkt des Versuchs ist das Datenfluss strikt unidirektional in Richtung des Computers.

Es wird **keine** Manipulation des Gehirns durchgeführt, die in irgendeiner Weise unmittelbar oder zukünftig eine der obengenannten Szenarien ermöglicht.

Alle Daten werden strikt anonym anhand der Teilnehmernummer verarbeitet und zu keiner Zeit werden anhand der Daten Rückschlüsse auf den Teilnehmer möglich sein.

Die gesammelten Daten belaufen sich auf zwei Fragebögen und eine maschinell generierte Log-Datei mit dem folgenden exemplarischen Inhalt:

Einverständniserklärung

Ich bin über den Inhalt des Experiments und die Verarbeitung der gesammelten Daten belehrt worden und stimme der Teilnahme und der anonymen Verarbeitung der gesammelten Daten zu.

Datum und Unterschrift: _____

Fragebogen vor dem Experiment

Einleitung

Dieser Fragebogen hat zum Ziel demografische und medizinische Daten zu sammeln, die im weiteren Verlauf der Studie zur Auswertung herangezogen werden können. Sämtliche Daten werden anonymisiert verarbeitet.

Teilnehmernummer: _____ (*Wird Ihnen mitgeteilt.*)

Fragebogen

1. Alter: _____
2. Biologisches Geschlecht: W M keine Angabe
3. Tragen Sie eine Brille? Ja Nein
4. Leiden Sie unter neurologischen Erkrankungen? Ja Nein
5. Nehmen Sie Medikamente zu sich, die sich auf das Gehirn auswirken? Ja Nein
6. Haben Sie in den letzten 12h Alkohol zu sich genommen? Ja Nein
7. Haben Sie eine Farbsehschwäche? Ja Nein

Bitte beantworten Sie die folgenden Fragen auf einer Skala von 1 (schlecht) bis 5 (gut):

8a. Wie fühlen Sie sich? 1 ———— 5

Fragebogen nach dem Experiment

Einleitung

Dieser Fragebogen hat zum Ziel Daten zur Versuchsdurchführung zu sammeln, die im weiteren Verlauf der Studie zur Auswertung herangezogen werden können. Die von Ihnen angegeben Daten in diesem Fragebogen werden für die spätere Auswertung der Studie herangezogen. Sämtliche Daten werden anonymisiert und vertraulich behandelt.

Teilnehmernummer: _____ (*Wird Ihnen mitgeteilt.*)

Fragebogen

Bitte beantworten Sie die folgenden Fragen auf einer Skala von 1 = wenig/angenehm bis 5 = viel/unangenehm:

- 1a. Haben Sie während des Versuchs Schwindel verspürt? 1 2 3 4 5
- 1b. Wie gut konnten Sie sich auf das Experiment konzentrieren? 1 2 3 4 5
- 1c. Beurteilen Sie den Tragekomfort des Sensors? 1 2 3 4 5
- 1d. Beurteilen Sie die wahrgenommene Zeit zum Einrichten des Sensors 1 2 3 4 5
- 1e. Beurteilen Sie wie schwer Ihnen diese Übung fiel: 1 2 3 4 5

Fragen zur Technologie

2. Stehen Sie dieser Technologie skeptisch gegenüber? Ja Nein keine Angabe
3. Würden Sie ein BCI selbst verwenden? Ja Nein keine Angabe
4. Sehen Sie einen Mehrwert in dieser Technologie? Ja Nein keine Angabe

A.2 Example logfile

This is an excerpt of a genuine logfile and shows all the relevant data which was gathered during a trial.

Since the data is anonymized, there is no way to identify a certain individual from this excerpt.

```

1626361044463;Model.NeuroTagMarkedAsTargetLogEntry;7
1626361044468;Model.SensorTelemetryLogEntry;94;95;94;93;93;85;97;94;C
1626361044468;Model.ExperimentEventLogEntry;started
1626361044559;Model.NeuroTagConfidenceLogEntry;7;0.03284512
1626361044858;Model.NeuroTagConfidenceLogEntry;7;0.02825086
1626361045158;Model.NeuroTagConfidenceLogEntry;7;0.04584821
1626361045458;Model.NeuroTagConfidenceLogEntry;7;0.0324579
1626361045758;Model.NeuroTagConfidenceLogEntry;7;0.1790879
1626361046075;Model.NeuroTagConfidenceLogEntry;7;0.2322663
1626361046374;Model.NeuroTagConfidenceLogEntry;7;0.2289219
1626361046658;Model.NeuroTagConfidenceLogEntry;7;0.2858692
1626361046958;Model.NeuroTagConfidenceLogEntry;7;0.2982492
1626361047258;Model.NeuroTagMarkedAsTargetLogEntry;2
1626361047259;Model.NeuroTagHitLogEntry;7
1626361049623;Model.NeuroTagMarkedAsTargetLogEntry;9
1626361049623;Model.NeuroTagHitLogEntry;2
1626361051955;Model.NeuroTagMarkedAsTargetLogEntry;8
1626361051955;Model.NeuroTagHitLogEntry;9
1626361054287;Model.NeuroTagMarkedAsTargetLogEntry;7
1626361054287;Model.NeuroTagHitLogEntry;8
1626361054287;Model.NeuroTagConfidenceLogEntry;7;0.07452609
1626361056636;Model.NeuroTagMarkedAsTargetLogEntry;1
1626361056636;Model.NeuroTagHitLogEntry;7
1626361058985;Model.NeuroTagMarkedAsTargetLogEntry;2
1626361058986;Model.NeuroTagHitLogEntry;1
1626361061335;Model.NeuroTagMarkedAsTargetLogEntry;7
1626361061335;Model.NeuroTagHitLogEntry;2
1626361061335;Model.NeuroTagConfidenceLogEntry;7;0
1626361063699;Model.NeuroTagMarkedAsTargetLogEntry;0
1626361063700;Model.NeuroTagHitLogEntry;7
1626361066049;Model.NeuroTagMarkedAsTargetLogEntry;8
1626361066049;Model.NeuroTagHitLogEntry;0
1626361066049;Model.NeuroTagConfidenceLogEntry;8;0.1005409
1626361068381;Model.NeuroTagMarkedAsTargetLogEntry;7
1626361068381;Model.NeuroTagHitLogEntry;8
1626361070729;Model.NeuroTagMarkedAsTargetLogEntry;6
1626361070729;Model.NeuroTagHitLogEntry;7
1626361073079;Model.NeuroTagConfidenceLogEntry;6;0.1839611
1626361073396;Model.NeuroTagConfidenceLogEntry;6;0.2086784
1626361073696;Model.NeuroTagConfidenceLogEntry;6;0.2204314
1626361073995;Model.NeuroTagConfidenceLogEntry;6;0.220725
1626361074278;Model.NeuroTagConfidenceLogEntry;6;0.2152889

```

A Material

```
1626361074578;Model.NeuroTagConfidenceLogEntry;6;0.1819418  
1626361074861;Model.NeuroTagConfidenceLogEntry;6;0.225923  
1626361075195;Model.NeuroTagConfidenceLogEntry;6;0.1962474  
1626361075495;Model.NeuroTagConfidenceLogEntry;6;0.2939416  
1626361075795;Model.NeuroTagMarkedAsTargetLogEntry;2
```

List of Figures

List of Tables

I hereby confirm that this thesis is my own work and that I have not sought or used inadmissible help of third parties to produce this work and that I have clearly referenced all sources used in this thesis. I have fully referenced and used inverted commas for all text directly or indirectly quoted from a source.

Place and date

Julius Neudecker