university of groningen

faculty of behavioural and social sciences

<A Motor imagery-based Brain Computer Interface>

Jamilla de Koning

Master Thesis - Master programme Clinical Neuropsychology

[*3361780*]
[*August*] [*2022*]
Department of Psychology
University of Groningen
Supervisor /Examiner: dr S. Enriquez-Geppert
Second reviewer: dr. L. Pillette

A thesis is an aptitude test for students. The approval of the thesis is proof that the student has sufficient research and reporting skills to graduate, but does not guarantee the quality of the research and the results of the research as such, and the thesis is therefore not necessarily suitable to be used as an academic source to refer to. If you would like to know more about the research discussed in this thesis and any publications based on it, to which you could refer, please contact the supervisor mentioned.

Abstract

Brain-Computer Interfaces (BCIs) are receiving more and more attention and interest in research. The applications of these systems range from medical and assistive (in stroke rehabilitation or paralyzed patients) to non-medical (videogames and virtual reality) (Lotte et al., 2015). One of the most used methods to control a BCI is motor imagery (MI). BCI performance is influenced by a variety of factors including classification algorithms, users' states/traits and the way users are trained to perform a BCI task. The current study aims to test any difference between two newly created instructions modes (video vs written) and analyze important users' characteristics (motivation and vividness of visual imagery) which are known to influence BCI performance. BCI performance was measured both with classification accuracy and subjective performance measures. 27 subjects were recruited for the study that consisted of two sessions in which kinesthetic motor imagery tasks (hand imagery and feet imagery) were performed. Additionally, participants had to fill out questionnaires regarding motivation, vividness of visual imagery, mind-body techniques, creative activities, and experience with electronic devices. Two RM-ANOVA were run to find any difference in BCI performance between the two instruction methods, additionally, correlation analyses were run between two independent variables motivation and vividness of visual imagery and BCI performance (classification accuracy and subjective performance). Results showed that no difference was present between the two instruction methods in BCI performance. Moreover, higher vividness of visual imagery was related to higher subjective performance. These findings might be useful for future research to find the most appropriate method of instructing participants and to design a BCI system adaptable to the needs of each participant.

Keywords: Brain Computer Interfaces, motor imagery, motivation, vividness of visual imagery, classification accuracy, subjective performance, instructions

A motor imagery based Brain Computer Interface

Think about a sports competition like free climbing. An athlete might be waiting for their turn and mentally rehearsing what they are going to do when the competition starts (e.g. going through all the single steps to get to the top of the wall, the types of movement they have to perform). They also might imagine how the movement feels (e.g. muscle contracting, the temperature of the wall when touching it with their hands, the type of surface they touch with their hands). This example depicts the construct of motor imagery (MI), which is defined as the mental process of imagining a movement of part of the body without making an overt movement (de Vries & Mulder, 2007).

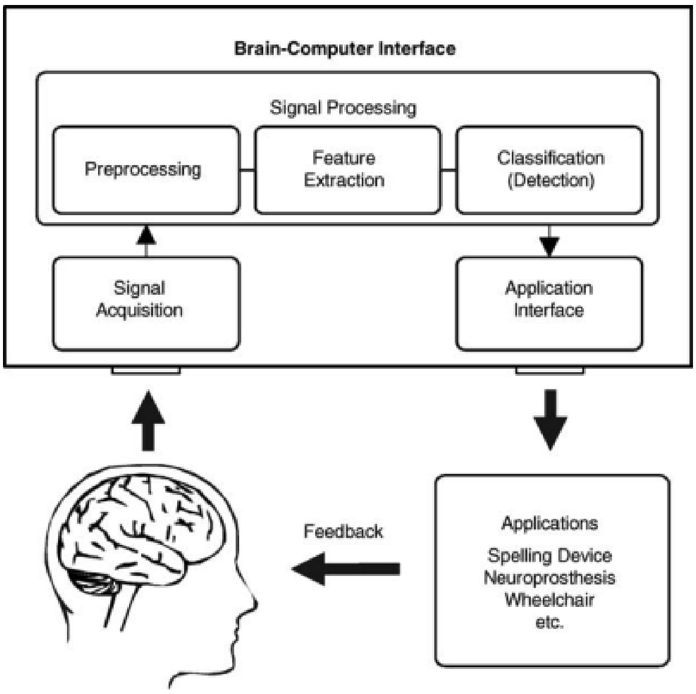
Motor imagery is part of a bigger concept that is mental imagery, which also includes other cognitive mental tasks like auditory imagery or mental calculations. It is known that when one performs an action certain brain regions are activated (Ehrsson et al., 2003) and these same regions are similarly engaged during mental imagery. Specific brain patterns in the motor and somatosensory cortices change because of motor imagery or movement (Nam et al., 2018). More specifically, changes in Mu oscillations, measured with an electroencephalogram (EEG) above the sensorimotor areas, are seen, which are called event-related desynchronization (ERD) and event-related synchronization (ERS) (Nam et al., 2018). These are respectively a decrease and increase in Mu frequency band amplitude in sensorimotor areas after movement or motor imagery. It has been suggested that Mu rhythm can be used as a predictor of motor imagery success (Chen et al., 2021)

Motor imagery is one of the most used methods to control Brain-Computer Interfaces (BCIs). The term BCI was introduced in 1973 by the researcher J. Vidal, who described it as the "utilization of brain signals in a man-computer dialogue" (Vidal, 1973) and who first tried to implement an electroencephalogram-based BCI. Recent research has given a more descriptive definition for such a system, also because of its widespread use;

BCIs are communication systems that make it possible for the user to send commands to a computer using their brain activity (Lotte et al., 2013). As it was just mentioned, to operate a BCI, the users have to control their brain state, for example using MI tasks (Jeunet et al., 2016b). The acquisition of this brain activity is done by using either invasive (intracortical electrodes, ECoG) or non-invasive (EEG, fNIRS, MEG, fMRI) systems. The brain activity is then processed and identified by the machine using signal processing and machine learning (Jeunet et al., 2016b). In this paper we focus on EEG as this is the most used system to acquire brain activity and the best suited for application (see Figure 1 for a BCI system overview).

Figure 1

BCI system overview



Note. Components of a BCI system from signal acquisition to possible BCI applications. From “Human Brain-Computer Interface”, by G. Pfurtscheller, C. Neuper and N. Birbaumer, in A. Riehle and E. Vaadia (Eds.), *Motor cortex in voluntary movements: a distributed system for distributed functions* (p. 367), 2005, Boca Raton: CRC Press.

The applications of EEG-based BCIs are wide. They range from medical and assistive applications, for example in motor rehabilitation after a stroke or to make a paralyzed patient communicate with the environment, to non-medical ones, for example in video games (Roc et al., 2021). One of the most interesting and relevant applications of MI-BCIs is in the case of stroke or paralyzed patients. In paralyzed patients, the usual pathways from the brain to the nerves and muscles are damaged or disrupted, consequently, patients are unable to move parts of their body (Ropper et al., 2019). BCIs make it possible to circumvent these usual pathways (Peng et al., 2022). Even though the applications of these systems are wide, we encounter the problem of “BCI illiteracy”. Research shows that between 15 and 30% of users are unable to use a MI-BCI system (Thompson, 2019). When taking into account the part of users that is not illiterate, the average performance they reach is usually still rather low (around 75% of accuracy) (Jeunet et al., 2015), considering that 70% of accuracy is the minimum performance required for communication (Ahn & Jun, 2015). Due to this problem, there is a need to find a way to increase performance. However, as the decision to set a specific minimum performance accuracy required for communication is currently up to the researchers (Thompson, 2019), ‘labelling’ participants that are BCI illiterate might be complex and might vary per study depending on the set threshold.

There are a variety of reasons why the phenomenon of BCI illiteracy can arise. The first is that classification algorithms are still improvable (Jeunet et al., 2017). Second, participant training might also be responsible for MI-BCI performance variations and needs to be investigated more. In order to understand the user training process we need to understand which factors impact the ability to control a BCI (Jeunet et al., 2017), for example, inter-individual differences in the participants, ranging from mood to motivation, to visuo-motor coordination (Nijboer et al., 2008; Hammer et al., 2014). These factors could be placed in the third possible reason leading to BCI illiteracy. A cognitive model was created to gather together the most recent findings on intrinsic (i.e., participants' traits and states) and extrinsic factors (i.e., factors influencing traits and states) that have been found to influence BCI performance, which is measured in this case with classification accuracy (Jeunet et al., 2017). Up until now, no comprehensive model has been presented for MI-BCI user training (Roc et al., 2021). In the following sections, the two factors influencing the ability to control a BCI are presented in more depth.

User training: instructions

The instruction method for MI tasks the participant has to perform is an essential component of the user training. Currently, there is no standardized method of instruction yet and little is known about the most suitable mode of instruction. A possible reason is that, as studies do not report how participants were instructed, thereby preventing comparisons between studies, it might be challenging to determine the most effective method of instruction. Moreover, instructions currently used in MI-based brain-computer interfaces training may not be appropriate (Roc et al., 2021) as these do not incorporate recommendations from psychology and instructional design (Jeunet et al., 2016a). Three types of instructions have been described depending on when these are given during the training procedure, namely: 1) general instructions (e.g. presentation of the BCI system), 2) instructions on what mental tasks are used (e.g. motor imagery or mental calculation) and 3) guidance on how to perform the tasks (Roc et al., 2021). A step-by-step protocol was presented for participants to learn how to control a MI-BCI (Rimbert et al., 2020): (a) make the user aware of the sensations that are usually felt when performing a movement (pressure, heat, muscle contraction etc), (b) user starts to progressively reduce muscular activity so that in the end they are only left with kinesthetic imagination of the movement, (c) user is able to perform the Kinesthetic Motor Imagery (KMI) whenever they want. This protocol shows the importance of KMI, which is described as "the ability to imagine performing a movement by means of having an impression of the muscle contraction and sensation during an actual movement" (Yang et al., 2021). Another type of imagery to control BCIs is visual-motor imagery (VMI). KMI is preferred to VMI to improve MI-BCI control (Neuper et al., 2005) because brain wave patterns can be more easily detected with kinesthetic imagery than with visual-motor imagery. Alongside KMI, specificity and familiarity with the MI task are important for the training (Roc et al., 2021). For example, rather than asking users to imagine a general and nonspecific movement of their right hand, it is preferred to ask them a more specific one like squeezing a stress ball in their hand. One of the goals of the present study was to present, test and compare two novel instruction methods (written or video) based on the just mentioned findings to find a suitable way to instruct users before using a MI-based BCI.

Inter-individual differences

The ability to control BCIs could also be influenced by different inter-individual factors. These can be divided into users' traits and states. States like higher motivation (Kleih-Dahms et al., 2021; Nijboer et al., 2008), better mood (Nijboer et al., 2008), higher attention level (Grosse-Wentrup and Schölkopf, 2012), lower fear of the BCI system and of incompetence (Nijboer et al., 2008; Nijboer et al., 2010) have been found to positively influence BCI performance. Traits such as attention span (Hammer et al., 2012), visual-motor coordination (Hammer et al., 2014), learning style, tension and self-reliance (Jeunet et al., 2015) have been related to BCI performance. Moreover, the user's relationship with the BCIs and in general with electronic devices can affect performance; users that feel anxious when using electronic devices and not in control of the device have more problems controlling BCIs (Jeunet et al., 2016b). Additionally, some people are unable to mentally visualize places, people or objects which is can be called 'aphantasia' (Merriam-Webster, n.d.). This phenomenon has been found to be correlated with MI-BCI performance (Leeuwis et al., 2021).

Finally, not only the user's traits and states could have an impact on BCI performance, but also extrinsic factors, described as design artefacts and cognitive activities or exercises which are thought to optimize users' performance (Jeunet et al., 2017). For example, activities like playing an instrument could lead to higher BCI performance (Randolph, 2012). Furthermore, there is evidence that practising meditation can predict and improve the ability to control a BCI (Jiang et al., 2021; He et al., 2015).

Measure of BCI performance

MI-BCI research mostly measures performance with classification accuracy scores, meaning how well the classification algorithms can pick up a specific brain pattern and correctly classify it as such. This measure depends on different aspects like classification algorithms used, training time etc., and is a combination of the user and machine performances (Roc et al., 2021). For this reason, along with the usual measure of accuracy being classification accuracy, it is valuable to look at a more user-centered and subjective method of measuring performance, such as self-predicted or subjective performance. As a matter of fact, participants are able to self-predict their BCI performance, even in the absence of feedback (Ahn et al., 2018).

Purpose of this study

The present study aimed to contribute to the literature on user training by looking at two modalities of instructions (video vs written) and exploring any possible differences between the two. Secondly, important users' characteristics which are known to influence and predict BCI performance were investigated. Predicting a user's (in)ability to control a MI-BCI system could avoid a loss of time and energy for experimenters and users (Jeunet et al., 2015). By doing so, we aimed to help create a BCI system that could be adapted to a specific user, according to their characteristics

In the current study, BCI performance was measured both with classification accuracy and subjective performance. By doing so, we are able to analyze both a measure of performance based on machine learning (classification accuracy) and one that focuses on the subjective experience of the user.

Finally, we investigated four hypotheses. First, H1) we expected to find differences in subjective performance scores between the video and written instruction groups. Similarly, for the second hypothesis, H2) we expected to find differences in classification accuracy scores between the two instruction groups. Third, H3) we hypothesized that motivation is positively related to classification accuracy and subjective performance. Finally, H4) we expected vividness of visual imagery to be positively related to classification accuracy and subjective performance scores.

Method

The current study is part of ongoing research, called Testing Instructions and Classifiers study (TIC study), which aimed at creating and testing new instruction methods (video vs written) and testing different existing classifiers. This research was done at the Clinical Neuropsychology department of the University of Groningen, it was approved by the Ethical Committee of the Psychology department of the University of Groningen and was conducted according to the Declaration of Helsinki. Participants were recruited either via a platform for first-year psychology students (SONA) of the Psychology faculty or via the researchers' close environment (e.g. friends). Participants recruited via SONA received compensation with SONA credits, while the rest did not receive any compensation.

Materials

Questionnaires

Motivation was assessed using the Questionnaire of Current Motivation (QCM) for BCI, which is the adapted version utilized by Kleih and Kübler (2013) of the original (Rheinberg, Vollmeyer, & Burns, 2001). The QCM divides motivation into four components: mastery of confidence, fear of incompetence, interest and challenge. These are measured with 18 items in total. The participants rated these items on a 7-point Likert scale from 1 (*strongly disagree*) to 7 (*fully agree*). Items 6, 9, 12 were reversed for scoring. Cronbach's alpha for the scale was found to be satisfactory for Session 1 (α = .6) and relatively high for Session 2 (α = .75). For scoring, one mean score of all items for Session 1 and one for Session 2 were calculated for each participant. A higher score in the scale indicated higher motivation.

Subjective performance of motor imagery was measured using two items created by the researchers, namely “I felt completely confident in performing these tasks” and “I performed these tasks exactly as it was instructed to me”. These items were scored with a 4-point Likert scale from -2 (*strongly disagree*) to 2 (*strongly agree*). Reliability for Session 1 was good (α = .81) as well as for Session 2 (α = .77). For the sake of the analysis, the ratings were recoded to values between 1 to 4. A mean score for Session 1 and Session 2 was calculated for later analysis.

Vividness of visual imagery is described as the “Intensity to which people can visualize settings, persons, and objects in mind” (Leeuwis et al., 2021, p. 3). This was measured with the Vividness of Visual Imagery questionnaire (Marks, 1973), which is also a questionnaire used for The questionnaire comprises 4 different images (a friend, a rising sun, a shop, a country scene) about which the participants have to imagine a series of different scenes (n=16) and rate how vivid the image they created in their mind was.

Ratings were made on a 5-point Likert scale from -2 (*no image at all, you only know that you are thinking of the person*) to 2 (*perfectly clear and as vivid as normal vision)*. For the sake of analysis, the ratings were recoded to values between 1 and 5. For every participant, a mean score was calculated on this questionnaire. The higher the score the higher the participant's vividness of visual imagery was. The scale was found to be highly reliable (α = .91).

Three more variables were measured using the Experiences, Motivation, Expectations and Attitudes regarding BCI/Neurofeedback (ExMEA-BCI/NF) questionnaire, which was created from a Groningen-France collaboration of researchers.

The first questionnaire measured experience with electronic devices, which was adapted from the German 'Kontrollüberzeugungen im Umgang mit Technik' [KUT] questionnaire (Beier, 2004). This concept was measured with nine items (e.g. "Electronic devices are often unclear and difficult to use", "When I encounter difficulties or problems in using an electronic device, I will try to solve them"). These items were rated on a scale ranging from 0 (*completely disagree*) to 100 (*completely agree*). Items 2, 5, 7, 8 and 9 were reversed. To score the scale, a mean score for each participant and standard deviation was calculated. A higher score on this scale represents more experience and confidence with the use of electronic devices. Chronbach's alpha for this nine-item scale was .897.

The second questionnaire measured experience with mind-body techniques. The participant was asked the frequency of practice of mind-body techniques (e.g. Meditation/Mindfulness, Prayer, Yoga etc.) in the last year. For this, eight types of techniques had to be rated on a 5-point Likert scale from 1 (*never*) to 5 (*daily or almost daily*). The alpha reliability score of the scale from the current sample was .72, indicating good reliability. For scoring, the mean score and standard deviation of the scale for each participant was calculated.

The third and last questionnaire measured creative activities. Here participants rated the frequency at which they engage in seven given creative activities (e.g. practice of musical instruments, playing computer games) on a 5-point Likert scale ranging from 1 (*daily*) to 5 (*never*). Mean score of the scale and standard deviation were calculated.

These questionnaires were all implemented on Qualtrics.

Instructions

The two instructions modes (written vs video) used in this research were newly created by the research group and in collaboration with the International BIG Shared MT-BCI protocol Instruction subgroup. Written instructions consist of four different texts for the four different mental tasks. The whole text includes 1222 words. The four video instructions, one per mental task, are on average 3.79 minutes long. The video is narrated while a person (female) is shown as a participant doing the task (third person perspective). For the hand and feet imagery tasks, the videos are in a first person perspective, showing close-up shots of a hand and feet. For more details, the reader can refer to Appendix A and B.

Procedure

The day before the participants took part in the actual experiment, two questionnaires were sent by email to fill out at home: one with demographic information (including informed consent and description of the study) and the Vividness of Visual Imagery Questionnaire (VVIQ) and the other which comprehends two, namely Questionnaire of Current Motivation (QCM) and Experiences, motivation, expectations and attitudes regarding BCI/Neurofeedback (ExMEA-BCI/NF). Participants were asked to come to the EEG lab twice, to identify any potential learning effects. Three days after the first assessment, participants repeated the same experiment (see Figure 2 for the workflow of the study).

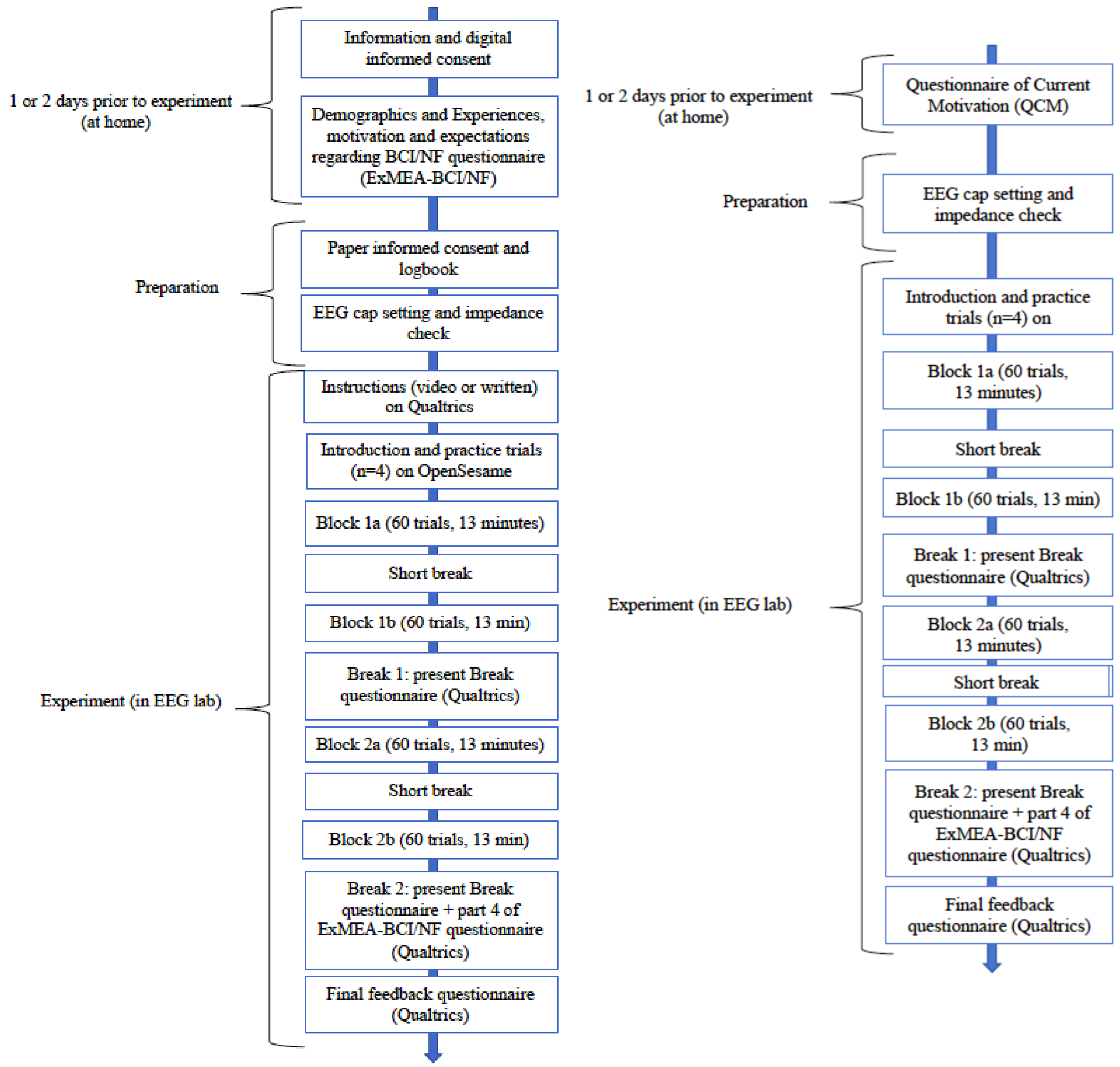
In the first session, in the preparation room, the EEG cap was set and the researchers explained the materials used for the cap setting (e.g. conductive gel, blunt syringe, reference nose electrode, disinfectant gel to clean the skin on the forehead and nose before setting the cap). Impedance check was done by keeping impedances below 10 Ohms. Next, in the experiment room, the participants sat on a comfortable chair with an armrest facing a computer screen. Then, they randomly received either video or written instructions about how to perform different mental imagery tasks. Afterwards, the experimenters, while repeating the impedance check, showed how movements, outside noise and blinking could contaminate EEG by showing this on the screen in front of the participants. Participants were then instructed to sit still and blink as little as possible. Then, participants were given an introduction to the experiment and the stimuli. Afterwards, participants had to explain their tasks. Then, four practice trials (one for each mental task) followed. Finally, the actual experiment was run. Here the participants had to perform four mental imagery tasks, which will be described in the following section. In the second session, participants went through the same experimental process without receiving instructions.

Experimental Tasks: Mental Imagery

Four mental tasks were performed, two motor tasks and two cognitive tasks: (a) kinesthetic imagination of squeezing a small ball with one dominant hand, (b) kinesthetic imagination of curling the feet towards the knees (feet flexion), (c) listening to specific a song (auditory imagery), (d) choosing a three-digit number and subtracting a specific number from it continuously (mental calculation). The experiment consisted of two conditions, namely: ‘Choice’ (26 minutes) and ‘Random’ (26 minutes), which were each presented in two blocks, so four blocks in total. For the choice blocks, the participants could choose which of the four tasks to perform but had to perform 15 trials for each MI task in total (60 in total for each block), for the random ones they were required to perform the mental task (60 per block in total) presented on the screen. Between the two conditions, participants filled out other questionnaires about the tasks they just performed and their experiences/feelings during and after the task (including subjective performance). The same questionnaires were presented at the end of block two. The total time of the experiment was 270 minutes, 170 minutes for the first session and 120 for the second, including EEG cap setting and questionnaires. Practice trials and the experiment were presented through OpenSesame (Mathôt et al., 2012).

Figure 2

Workflow of The Study for Session 1 (left) and Session 2 (right).



EEG data recording

EEG data was recorded with a 64 Ag/AgCI electrodes Waveguard™ connect cap (with wet electrodes), using the Twente Medical Systems International B.V. (TMSi) REFA amplifier. For EEG data acquisition, OpenVibe (Renard et al., 2010) was used. The EEG channels were selected following the 10-12 international system. For this, we used 32 channels, which were: Fp1, Fp2, AF3, AFz, AF4, F7, F3, Fz, F4, F8, FT7, FC5, FC1, FC2, FC6, FT8, T7, C3, C1, Cz, C2, C4, T8, TP7, CP5, CP1, CP2, CP6, TP8, P3, Pz, P4. The ground electrode was AFz and the reference electrode was set on the nose. The sampling rate was 512 Hz. Impedances were kept below 10 Ohms.

Offline EEG Analysis

Preprocessing

For the analysis, only motor imagery was used, namely hand imagery and feet imagery. First, raw data was pre-processed using EEGlab functions (Brunner et al., 2013), which is based on Matlab (2019). The EEG data was low and high pass filtered for the frequency bands of interest with 40 Hz and .1 Hz to remove drifts caused by skin conductance and noise from the power line. Then, the data was downsampled to 250 Hz and re-referenced to the nose electrode (Nz). Following, an independent component analysis (ICA) with RUNICA was used to take out and then reconstruct the scalp activity without components related to eye artifacts. Data was then epoched from 0 to 7 sec. To ensure a clean data sample a rest-artefact rejection was run. In the final step, in case of bad impedances of single channels, meaning if bad impedances were present in around 70% of the total recording (including all conditions), interpolation was performed. The minimum number of epochs needed per task was set to 27 to make sure that the chosen classifier would have enough samples/epochs to train each class.

Feature Extraction and Classification

The preprocessed EEG data was then imported to Python (Van Rossum, & Drake, 1995). For feature extraction, Common Spatial Pattern (CSP) was used to extract relevant features. Next, for classification, Linear Discriminant Analysis (LDA) was used, as this is the most used and robust algorithm in BCIs applications (Nam et al., 2018). Specifically LDA has been proven successful in MI based BCIs (Guger et al., 2000), which is the focus of this paper.

Analysis

The dependent variables in all the hypotheses were classification accuracy (in percentages) and subjective performance scores for each participant.

To test the first hypothesis, namely whether there was a difference in subjective performance between the two methods of instruction (video vs written), two RM-ANOVAs were run, with the between-subject factor being *Instruction Method* (video vs written) and within-subject factor being Session (1,2). For the second hypothesis, which tested whether there was a difference in classification accuracy between the two methods of instruction (video vs written), two RM-ANOVAs were run, with the between-subject factor being instruction method (video vs written) and within-subject factor being Session (1,2). If significant results were found, a post-hoc test was performed. Effect sizes, indicated with partial eta squared (η^2) were interpreted as small ($\leq .01$), medium ($\geq .06$) and large ($\geq .14$). Assumptions of RM-ANOVA of normality, homogeneity, sphericity and outliers were checked with, respectively: Shapiro-Wilk's test of normality, Levene's test of homogeneity of variances, Mauchly's test and boxplots.

For the third hypothesis, whether *motivation* was correlated with classification accuracy and subjective performance (both in Session 1 and Session 2), a Pearson's correlation analysis was run. For the fourth, whether *vividness of visual imagery* was correlated to classification accuracy and subjective performance, another Pearson's correlation analysis was run. For the correlation analyses, if any assumption of Pearson's correlation was violated, a non-parametric correlation test was run. Effect sizes for correlations were interpreted as small (r = .1), moderate (r = .3) or large (r = .5). Assumptions for Pearson's correlation of normality, linearity and outliers were checked using, respectively: Shapiro-Wilk's test of normality and visually using a scatterplot for the last two assumptions.

For all the tests performed, a p-value of <.05 was used. All statistical analyses were carried out using SPSS (IBM Corp., 2020). A power analysis was conducted using G*Power version 3.1.9.6 (Faul et al., 2007). For the correlation analyses, since only less than 27 datasets were collected, only large effect sizes (r > .5) could be detected. For RM-ANOVAs, only medium to large effect sizes (Cohen's *d* > .08) could be found with N = 26.

Results

Sample Characteristics

A total of 27 participants participated in the study with a mean age of 20.74 years (SD = 2.7) and 77.8% females. Most participants were naive to MI (81.5%) and BCIs (96.3%). Additional personal information collected included: level of education, rated on a 6-point scale depending on the level; native language, rated on a 4-point scale; presence of neurological or psychiatric disorders, presence of tinnitus, prior use of BCI and of MI were rated with a "yes" or "no" choice (see Table 1). Participants were excluded for different reasons: when they did not fill out a questionnaire, because of errors during recording, preprocessing and processing of EEG data, and finally when the number of epochs was less than 27. The final N for each variable can be seen in Table 2.

Table 1

Sample Characteristic

Personal variables	Total sample (N = 27)
Age (in years)[b]	20.74 (2.71) 18-28
Gender	
Female	21 (77.8%)
Male	6 (22.2%)
Education level[a]	
Primary school	
Secondary school	16 (59.3%)
(Technical) secondary school diploma	2 (7.4%)
University degree	8 (29.6%)
Doctorate degree	
Other	1 (3.7%)
Native language[a]	
English	3 (11.1%)
German	8 (29.6%)
Dutch	12 (44.4%)
Other	5 (14.8%)
Tinnitus[a]	
Yes	1 (3.7%)
Previous use of BCI[a]	
No	26 (96.3%)
Previous use of MI[a]	
No	22 (81.5%)

[a] Values are frequency in the sample and valid % in parenthesis. [b] Values are means with standard deviation in parentheses.

Preliminary analysis

Mean scores of classification accuracy seem to slightly increase from Session 1 (M = 64.21%, SD = 19.98) to Session 2 (M = 66.34%, SD = 15.09) with a min of 35% and a max of 94.87%. Nine subjects out of 25 (36%) reached the minimally needed accuracy for BCI control of 70% (Ahn & Jun, 2015). Subjective performance in Session 1(M = 3.36, SD = .46) is very comparable with Session 2 (M = 3.37, SD = .45). It can be noted that the average subjective performance in the two sessions was very high since the scale ranged from 1 to 4, meaning that the participants on average felt confident about their performance in the MI tasks.

Vividness of visual imagery was scored with 3.93 (SD = .57), thus on average the majority of participants experienced a visual imagery that was between moderately clear and vivid and perfectly clear. Motivation scores stayed similar between Session 1 (M = 4.56, SD = .57) and Session 2 (M = 4.62, SD = .69). Descriptive statistics including mean, standard deviation, maximum and minimum score, and group size of the variables of interest can be seen in Table 2.

Table 2

Descriptive Statistics for Variables Scales

	N	Minimum	Maximum	Mean	Std. Deviation
Classification accuracy S1[a]	25	38.46	94.87	64.21	18.98
Classification accuracy S2[a]	24	35.00	94.87	66.34	15.09
Vividness of visual imagery[b]	27	2.75	4.88	3.93	.57
Subjective performance S1[c]	27	2.25	4.00	3.36	.46
Subjective performance S2[c]	26	2.50	4.00	3.37	.45
Motivation S1[d]	27	3.61	5.61	4.56	.57
Motivation S2[d]	25	3.00	5.61	4.62	.69

Note. S1 = Session 1; S2 = Session 2

[a]The scale ranges from 0 to 100. [b]The scale ranges from 1 to 5. [c]The scale ranges from 1 to 4. [d]The scale ranges from 1 to 7.

Hypothesis 1

By looking at descriptive statistics (see Table 3 and Appendix E) we see that the mean score of subjective performance for the two instructions methods is higher in the video method than the written, both in Session 1 and 2. This was however not confirmed by the RM-ANOVA, but it is worth mentioning for further future analysis.

The normality and homogeneity assumptions were met, as assessed by Shapiro-Wilk's test of normality and Levene's test of homogeneity of variances, respectively. There were five outliers in the sample, as assessed visually using a boxplot (see Figure D1). When excluding these from the analysis, the results pointed at the same conclusions (see Appendix D for the same analysis run without the outliers). The sphericity assumption was checked with the Mauchly's test. Since this test is used when the within-subject factor has three or more levels (Field, 2013) the test gave no value, indicating that the sphericity assumption was met. The RM-ANOVA showed that there was no statistically significant main effect of time in subjective performance scores, ($F(1,24)= .0$, $p = 1.00$, partial $\eta^2 = .0$), so there was no difference in subjective performance at the two time points (Session 1 and Session 2). Moreover, the main effect for instruction method (video vs written), showed no significant difference in subjective performance scores between the two instruction groups ($F(1, 24) = .85$, $p=.37$, $\eta^2 = .03$).

Hypothesis 2

Descriptive statistics can be seen in Table 3. For Session 1, the mean classification accuracy for the video instructions was 67.17% (SD = 19.73) and for the written ones was 58.91% (SD =18.88). In Session 2, the mean classification accuracy for the video instruction was 70.86% (SD = 17.41) and 61.82% (SD = 11.33) for written instructions. This shows that the mean classification accuracy score was higher in the video method of instruction than for the video. This result was not confirmed by the RM-ANOVA.

Assumptions for RM-ANOVA of normality, sphericity, outliers were checked using Shapiro-Wilk's test of normality, Mauchly's test and boxplot, respectively. Normality assumptions were fulfilled apart from the classification accuracy variable for Session 1 (see appendix C). However, since RM-ANOVA is a quite robust measure against normality violations and this assumption was only slightly violated the analysis was still run. No main effect was found for time ($F(1,22) = 1.06$, $p = .32$), nor for instruction method (video vs written) ($F(1, 22) = 2.09$, $p=.16$, $\eta^2 = .09$). However, a medium effect size for the main effect of instruction method was present, suggesting clinical relevance.

Table 3

Descriptive Statistics in RM-ANOVA of subjective performance and classification accuracy in Session 1 and 2 for the two instruction method groups (video and written)

	Instruction method	Mean	SD	N
Subjective performance S1[a]	Video	3.40	.43	13
	Written	3.33	.51	13
Subjective performance S2[a]	Video	3.48	.45	13
	Written	3.25	.44	13
Classification accuracy S1[b]	Video	67.17	19.73	12
	Written	58.91	16.88	12
Classification accuracy S2[b]	Video	70.86	17.41	12
	Written	61.82	11.33	12

Note. S1 = Session 1; S2 = Session 2.

[a]The minimum and maximum possible values are 1 and 4, respectively. [b]The minimum and maximum possible values are 0 and 100, respectively.

Hypothesis 3

To assess the relationship between motivation and subjective performance, assumptions for Pearson's correlation of normality, outliers, linearity were first checked. These were all met. In Session 1, a moderate positive correlation was found, however this was not statistically significant ($r = .32$, $p = .054$, one-tailed). In Session 2, a small correlation between motivation and subjective performance was found ($r = .27$, $p = .10$, one-tailed) which turned out to be not statistically significant.

A correlation analysis was also computed to test the relationship between motivation and classification accuracy. Since the variable classification accuracy in Session 1 was not normally distributed and the linearity assumption was not met, a Kendall's tau-b correlation was run. There was a weak positive association between the two variables, which was not significant ($\tau_b = .01$, $p = .47$, one-tailed). In Session 2, the linearity assumption was not met, consequently, a Spearman's rank-order correlation was run. The assumption of monotonic relationship was checked and met. There was a moderate negative and significant correlation between motivation and classification accuracy in Session 2 ($r_s = -.37$, $p < .05$, one-tailed).

Hypothesis 4

In Session 1 and 2, the assumptions of normality, linearity and outliers were met, as assessed by Shapiro-Wilk's test and a scatterplot respectively, thus Pearson's correlations were run. In Session 1, a large positive significant correlation was found between these two variables ($r = .60$, $p < .001$, one-tailed). In Session 2, a moderate positive correlation was also found between vividness and subjective performance ($r = .39$, $p = .03$, one-tailed).

To test whether vividness of visual imagery and classification accuracy were positively correlated assumptions of normality, linearity and outliers assumptions were checked. The first two were both not met for Session 1, consequently a Kendall's tau-b correlation was run. There was a small positive, but not statistically significant association between vividness and classification accuracy in Session 1 ($\tau_b = .003$, $p = .49$, one-tailed). In Session 2, the linearity assumption was not met, hence Spearman's rank-order correlation was run. A weak negative correlation was found between vividness of visual imagery and classification accuracy, this was however not statistically significant ($r_s = -.11$, $p = .30$, one-tailed).

Exploratory analysis

An additional analysis was done to examine the sample more closely, dividing the sample in two groups of people with lower and higher classification accuracy using a median split. Descriptive statistics of the variables of interest for each group were calculated (Table 4). However, visually there were no differences between the two groups (see Figure 3).

Additionally, independent-samples t-tests were run on each variable for the two groups. Assumption of normality for every variable was checked and met. Assumption of outliers was checked visually with boxplots and was met for all variables except for the variable experiences with electronic devices, in which three outliers were present. For this reason, the outliers were removed to check whether these affect the results. Removing these led to the same conclusion (see Appendix F, Table F2), consequently the outliers were left in for the analysis. The assumption of homogeneity of variance was checked with the Levene's Test for equality of variances. For subjective performance, mind-body techniques and creative activities the assumption of homogeneity of variance was violated, thus a Welch t-test was run to determine if there were any differences in the variables between lower and higher scorers. No statistically significant differences were found between the two groups in the six variables (for test results see Table F1).

Table 4

Mean and standard deviation of median split groups low and high

	Lower	Higher
	M(SD)	M(SD)
Subjective performance S1+S2[a]	3.29(.50)	3.40(.31)
Motivation S1+S2[b]	4.51(.59)	4.63(.53)
Vividness of visual imagery[c]	3.85(.56)	3.93(.58)

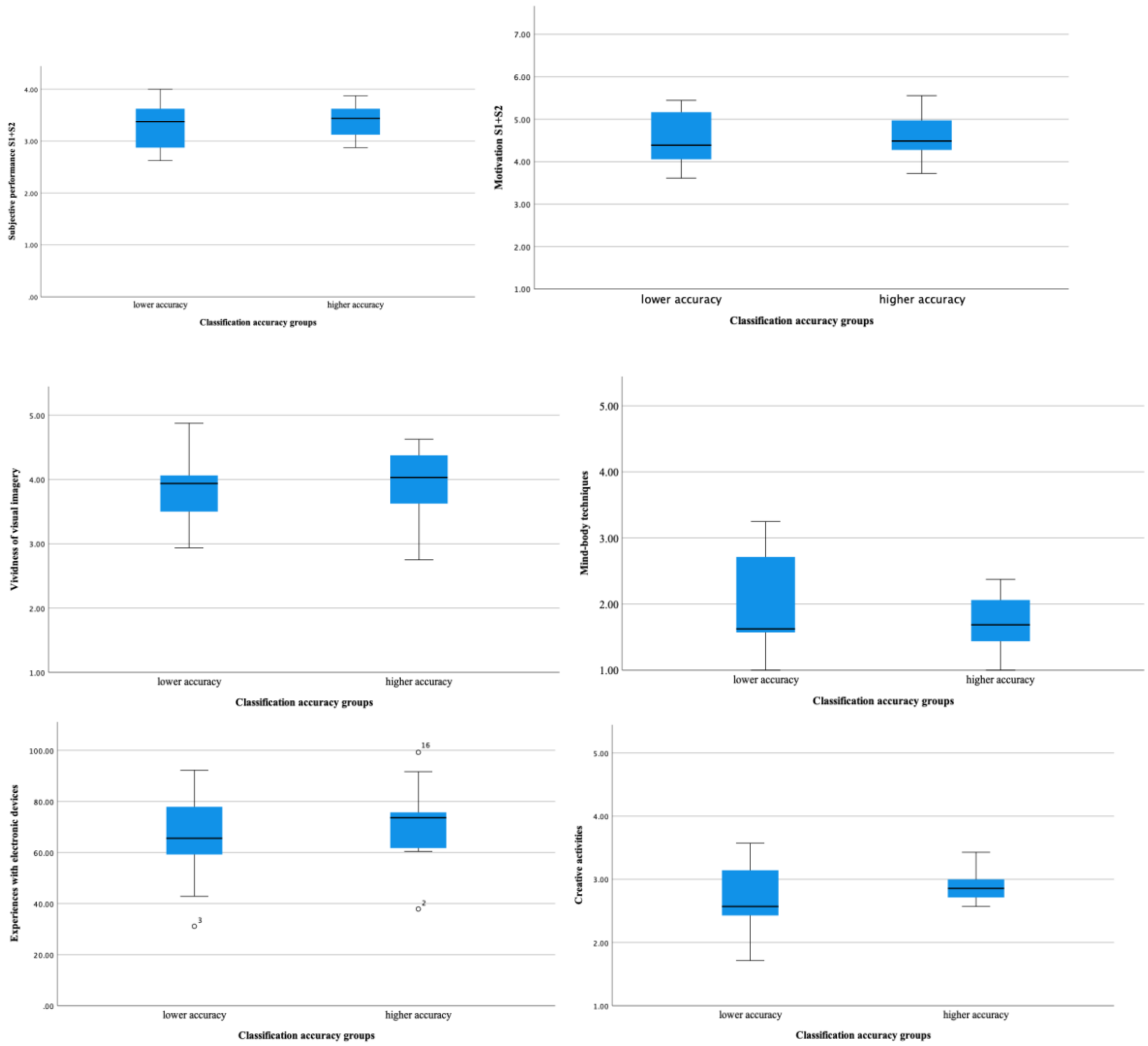
Mind-body techniques[c]	1.96(.73)	1.73(.40)
Experience with electronic devices[d]	67.03(17.62)	71.26(15.60)
Creative activities[c]	2.67(.55)	2.87(.25)

[a]The scale ranges from 1 to 4. [b]The scale ranges from 1 to 7. [c]The scale ranges from 1 to 5.

[d]The scale ranges from 0 to 100.

Figure 3

Boxplots of the Six Different Variables in Session 1 and 2 According to Accuracy Group (Lower and Higher).



Note. S1 = Session 1; S2 = Session 2

Discussion

The current study looked at various factors that influence BCI performance (measured with both classification accuracy and subjective performance). Participants took part in two sessions in which they performed MI tasks, including motor imagery, auditory imagery and mental calculation without receiving feedback on their performance. For the analysis, only motor imagery was examined. This study also investigated novel instruction methods, which could be a valuable addition to the field in which no standardized instruction method has been created or compared. No differences in BCI performance scores (measured with subjective performance and classification accuracy) were found between the two instruction methods. Correlation analysis showed that vividness of visual imagery was positively correlated with subjective performance in both Session 1 and Session 2. Motivation was found to be negatively correlated with classification accuracy in Session 2.

In the following sections, the different hypothesis, results and their implications will be discussed in more detail.

Instructions and BCI performance

Participants in the study received either video or written instructions about the mental tasks that had to be performed based on recent findings including KMI, familiarity and specificity of the MI tasks (Neuper et al., 2005; Roc et al., 2021). These instructions were created to test whether the modality in which participants receive instructions can affect BCI performance and to find the most effective way to instruct users before using a MI-based BCI.

The results are not in line with our first and second hypotheses, meaning that users scored similarly on the subjective performance measure and classification accuracy independently of the instruction method. This does not necessarily mean that it is not a valuable finding. These results highlight that the modality of instruction might not be important for BCI performance: each instruction medium (video and printed) brings its own advantages and disadvantages (Alexander, 2013), thus there is no difference between watching and reading instructions. In fact, as Roc et al. (2021) explain, users learn in different ways and the experience of the learner should determine how much guidance should be provided during their training. Roc et al. also suggest that in some cases we could let the users choose how they want to be trained according to their needs. Consequently, knowing a person's learning style and their preferred method of instruction could lead to the development of a BCI system tailored to every subject. In fact, in order to improve the efficiency of BCI training, training protocols may need to be tailored to the user's characteristics (Lotte et al., 2013). It is important to note that, in the second hypothesis, while we could not find a statistically significant result, we did find a clinically relevant result between the two instruction methods for classification accuracy. Alexander (2013) has shown that participants who received video instructions made fewer errors and their accuracy was better in the task than users who received written instructions. These results could be considered for further analysis in future studies testing modalities of instruction.

Inter-individual differences

Research in the field of users' states affecting BCI performance has shown that motivation is one of them (Hammer et al., 2012; Nijboer et al., 2008). In the present study, motivation was found not to be positively correlated with subjective performance, thus as motivation in participants increased, confidence in their own performance did not increase with it. Moreover, motivation was found to be negatively correlated with classification accuracy only in Session 2, however this result could not be confirmed because of lack of power from the low sample size in this analysis. Kleih et al. (2011) theorize that motivation facilitates learning an SMR-BCI task, meaning that higher motivation leads to higher BCI performance. Learners' motivation in general is an important mediator of performance

(Shute, 2008). Interestingly, it is known that receiving feedback can motivate learning (Shute, 2008). In contrast with past studies, participants in the present study did not receive feedback after performing MI tasks. As a result, participants may have been less motivated to perform well.

The fourth hypothesis is in part supported. As expected, vividness was positively correlated with subjective performance in both Session 1 and 2, meaning that the better the users could visualize setting, people and objects, the more confident they felt about having performed well in the motor imagery tasks. Vividness was not correlated with classification accuracy both in Session 1 and Session 2, a result not in line with our hypothesis. Only two studies reported a positive relation between BCI performance and vividness of visual imagery (Leeuwis et al., 2021; Vuckovic & Osuagwu, 2013). However, to better predict BCI performance another measure of visual imagery might be used. Especially if the end goal in MI-BCI research is to have users who are able to perform MI in order to control the BCI system, a measure to assess and predict MI performance, instead of visual imagery, might be more effective. For example, an objective measure to assess motor imagery could be used instead of a self-reported subjective measure. For example, Madan and Singhal (2013) presented a test objectively measuring motor imagery ability. This could be implemented in MI-BCI research so that the users' skills in motor imagery could be assessed in advance, which consequently could help with prediction of BCI performance. According to the power analysis for correlation analyses, only a large effect size ($r > .5$) could be detected in our sample, consequently the results of the correlation analyses must be carefully looked at. For future studies, a larger sample size might be needed to assess these relationships.

Additionally, this paper explored the differences between poorer and higher performers in the sample. Studies have tried to find the reason for this phenomenon by identifying the characteristics that differentiate poor performers (Ahn et al., 2018). In the current study, when dividing the sample in lower and higher performers (measured with classification accuracy), no differences were found in a number of variables: subjective performance, motivation, vividness of visual imagery, creative activities (i.e. knitting, playing an instrument, playing computer games etc.), mind-body techniques (i.e. meditation, prayer, yoga, etc.) and experience with electronic devices. Since the two groups were formed using a median split, thus creating a precise cut-off score, some participants who scored just below the cut-off were still considered poor performers, while they scored very closely to the higher performers' group. For this reason, it might be more valuable to compare two extreme groups including highest and lowest performers so that more distinct cut-off scores can be selected. This could help find characteristics that distinguish the two groups that could explain the gap in their performance.

Measures of BCI performance

It is known that, in order to successfully control a BCI, a minimum of 70% of classification accuracy is required (Ahn & Jun, 2015), however, in the current sample only 36% of the participants reached this value. Classification accuracy is the most used measure of BCI performance (Lotte & Jeunet, 2018). However, other measures of BCI accuracy are present and sometimes classification accuracy is not considered the best option as it presents many limitations. Lotte and Jeunet (2018) suggest that this measure should be used together with other metrics to interpret more accurately users' performance and learning.

Research has shown that users can self-predict BCI performance (Ahn et al., 2018), hence this paper explored subjective performance as a subjective and user-centered measure of BCI performance. Subjective performance measured how confident the users were that they performed the MI tasks well. However, because measures based on self-reports are susceptible to a variety of factors (MacIntyre et al., 2018), we combined the subjective measure with a more objective one namely classification accuracy. We created a valuable combination of measurements to better understand all the mechanisms involved and not only those based on machine learning classification algorithms.

Limitations

This study had a number of limitations. The generalizability of the correlation analyses results is limited by a lack of power due to having a small sample size. However, because of the length of the experiment, in such a study it was more challenging to find more participants. A second limitation is that our sample was recruited via SONA, which is a pool for first-year psychology students that awards them credits for participation. This is also the focus of most papers, however, because of this focus, these results are usually not applicable to BCI end-users (Kleih & Kübler, 2018), which are those individuals that will end up using the BCI in their daily life and that may differ from healthy individuals due to, for example, cognitive impairments or perceptual difficulties (Kleih and Kübler, 2018).

Another limitation is that the study included two sessions that were longer than one hour, which could have been tiring for participants. Tiredness is also known to negatively affect motivation which affects attention and finally performance (Jeunet et al., 2017).

Finally, no feedback was given to the participants regarding their performance on the MI tasks. Because the participants could not see their performance and progress, the experiment could have been less challenging or motivating. In fact, as Lotte et al. (2013) state, it is important that the user feels competent when receiving feedback so that motivation, effort and learning efficiency are increased. The use of feedback is also important for participants to regulate brain waves that are classifiable by the BIC system (Ahn et al., 2018). This is possible when the user can check whether they are performing well or not, meaning when they receive feedback.

Future directions and recommendations

Based on the limitations and findings of this paper, we suggest a number of future directions. Shorter sessions might be preferred in future studies in order to keep participants motivated, focused and attentive to have a better BCI performance. Alternatively, if longer sessions are needed, more breaks can be added. Furthermore, a more representative sample should be considered so that results can be transferred to end-users too. This way, not only healthy participants are involved but also patients who will be the primary users of and will benefit the most from the BCI for example because of their motor impairments (e.g. paralyzed patients, stroke).

Finally, even though only vividness of visual imagery was found to be correlated with subjective performance and the other studied variables did not affect BCI performance, it is recommended to monitor these factors with a more adequate sample. As Jeunet et al. (2015) state, if we use certain experimental designs or cognitive trainings, we could influence those traits and states that are known to influence MI-BCI performance. The factors that have been found to influence BCI performance are many more than the ones considered in this study. These should all be placed in a model to find the most accurate way to characterize a user and its needs. Once we have a complete model of factors influencing BCI performance including users' states, traits and training, BCI systems could be adapted to different users.

References

Ahn, M., Cho, H., Ahn, S., & Jun, S. C. (2018). User's Self-Prediction of Performance in Motor Imagery Brain-Computer Interface. *Frontiers in human neuroscience, 12*, 59. https://doi.org/10.3389/fnhum.2018.00059

Ahn, M., and Jun, S. C. (2015). Performance variation in motor imagery brain-computer interface: a brief review. *Journal of neuroscience methods, 243*, 103–110. https://doi.org/10.1016/j.jneumeth.2015.01.033

Alexander, K. P. (2013). The usability of print and online video instructions. *Technical Communication Quarterly*, *22*(3), 237-259.

Beier, G. (2004). *Kontrollüberzeugungen im Umgang mit Technik : ein Persönlichkeitsmerkmal mit Relevanz für die Gestaltung technischer Systeme*. Dissertation.de.

Brunner, C., Delorme, A., & Makeig, S. (2013). Eeglab - an Open Source Matlab Toolbox for Electrophysiological Research. Biomedizinische Technik. *Biomedical engineering, 58* Suppl 1, /j/bmte.2013.58.issue-s1-G/bmt-2013-4182/bmt-2013-4182.xml. https://doi.org/10.1515/bmt-2013-4182

Chen, Y. Y., Lambert, K., Madan, C. R., & Singhal, A. (2021). Mu oscillations and motor imagery performance: A reflection of intra-individual success, not inter-individual ability. *Human movement science, 78*, 102819. https://doi.org/10.1016/j.humov.2021.102819

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). US, Hillsdale: Lawrence Earlbaum Associates.

de Vries, S., & Mulder, T. (2007). Motor imagery and stroke rehabilitation: a critical discussion. *Journal of rehabilitation medicine, 39*(1), 5–13. https://doi.org/10.2340/16501977-0020

Ehrsson, H. H., Geyer, S., & Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *Journal of neurophysiology, 90*(5), 3304–3316. https://doi.org/10.1152/jn.01113.2002

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. https://doi.org/10.3758/BF03193146

Field, A. (2013). Discovering statistics using IBM SPSS statistics. Sage.

Grosse-Wentrup, M., & Schölkopf, B. (2012). High γ-power predicts performance in sensorimotor-rhythm brain-computer interfaces. *Journal of neural engineering, 9*(4), 046001. https://doi.org/10.1088/1741-2560/9/4/046001

Guger, C., Ramoser, H., & Pfurtscheller, G. (2000). Real-time EEG analysis with subject-specific spatial patterns for a brain-computer interface (BCI). *IEEE transactions on rehabilitation engineering, 8*(4), 447-456.

Hammer, E. M., Halder, S., Blankertz, B., Sannelli, C., Dickhaus, T., Kleih, S., ... & Kübler, A. (2012). Psychological predictors of SMR-BCI performance. *Biological psychology, 89*(1), 80-86.

Hammer, E. M., Kaufmann, T., Kleih, S. C., Blankertz, B., & Kübler, A. (2014). Visuo-motor coordination ability predicts performance with brain-computer interfaces controlled by modulation of sensorimotor rhythms (SMR). *Frontiers in human neuroscience, 8,* 574.

He, B., Baxter, B., Edelman, B. J., Cline, C. C., and Ye, W. W. Noninvasive Brain-Computer Interfaces Based on Sensorimotor Rhythms. *Proceedings of the IEEE*, vol. 103, no. 6, pp. 907-925, June 2015, doi: 10.1109/JPROC.2015.2407272.

IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp

Jeunet, C., Jahanpour, E., & Lotte, F. (2016a). Why standard brain-computer interface (BCI) training protocols should be changed: an experimental study. *Journal of neural engineering, 13*(3), 036024. https://doi.org/10.1088/1741-2560/13/3/036024

Jeunet, C., N'Kaoua, B., & Lotte, F. (2016b). Advances in user-training for mental-imagery-based BCI control: Psychological and cognitive factors and their neural correlates. *Progress in brain research, 228*, 3–35. https://doi.org/10.1016/bs.pbr.2016.04.002

Jeunet, C., N'Kaoua, B., & Lotte, F. (2017). Towards a Cognitive Model of MI-BCI User Training. GBCIC.

Jeunet, C., N'Kaoua, B., Subramanian, S., Hachet, M., & Lotte, F. (2015). Predicting mental imagery-based BCI performance from personality, cognitive profile and neurophysiological patterns. *PloS one, 10*(12), e0143962.

Jiang, X., Lopez, E., Stieger, J. R., Greco, C. M., & He, B. (2021). Effects of Long-Term Meditation Practices on Sensorimotor Rhythm-Based Brain-Computer Interface Learning. *Frontiers in neuroscience, 14,* 584971.

Kleih-Dahms, S. C., Botrel, L., & Kübler, A. (2021). The influence of motivation and emotion on sensorimotor rhythm-based brain-computer interface performance. *Psychophysiology*, *58*(8), e13832. https://doi.org/10.1111/psyp.13832

Kleih, S. C., & Kübler, A. (2013). Empathy, motivation, and P300 BCI performance. *Frontiers in human neuroscience, 7,* 642. https://doi.org/10.3389/fnhum.2013.00642

Kleih, S. C. & Kübler (2018). Why user-centered design is relevant for Brain-Computer Interfacing and how it can be implemented in study protocols. In C. S. Nam, A.

Nijholt, & F. Lotte (Eds.*), Brain-computer interfaces handbook: technological and theoretical advances* (1st ed., p. 586). CRC Press.

Kleih, S.C., Riccio, A., Mattia, D., Kaiser, V., Friedrich, E.V.C, Scherer, R., Müller-Putz, G., Neuper, C., & Kübler, A. (2011) Motivation influences performance in SMR-BCI. n G. Müller-Putz, R. Scherer, M. Billinger, A. Kreilinger, V. Kaiser, & C.Neuper (Eds.), *Proceedings of the 5th International Brain-ComputerInterface Conference*, Technical University of Graz, Graz, Austria,September 22–24, 108–111.

Leeuwis, N., Paas, A., & Alimardani, M. (2021). Vividness of visual imagery and personality impact motor-imagery brain computer interfaces. *Frontiers in Human Neuroscience, 15*, 634748.

Lotte, F., Bougrain, L. and Clerc, M. (2015). Electroencephalography (EEG)-Based Brain–Computer Interfaces. *Wiley Encyclopedia of Electrical and Electronics Engineering*. https://doi.org/10.1002/047134608X.W8278

Lotte, F., & Jeunet, C. (2018). Defining and quantifying users' mental imagery-based BCI skills: a first step. *Journal of Neural Engineering, 15*.

Lotte, F., Larrue, F. and Mühl, C., (2013). Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design. *Frontiers in Human Neuroscience, 7*.

MacIntyre, T. E., Madan, C. R., Moran, A. P., Collet, C., & Guillot, A. (2018). Motor imagery, performance and motor rehabilitation. *Progress in brain research, 240*, 141–159. https://doi.org/10.1016/bs.pbr.2018.09.010.

Madan, C. R., & Singhal, A. (2013). Introducing TAMI: an objective test of ability in movement imagery. *Journal of motor behavior, 45*(2), 153–166. https://doi.org/10.1080/00222895.2013.763764

Marks D. F. (1973). Visual imagery differences in the recall of pictures. *British journal of psychology, 64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*(2), 314-324. doi:10.3758/s13428-011-0168-7

Merriam-Webster. (n.d.). Aphantasia. In *Merriam-Webster.com dictionary*. Retrieved August 4, 2022, from https://www.merriam-webster.com/dictionary/aphantasia

Nam, C.S., Nijholt, A., & Lotte, F. (Eds.). (2018). *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances* (1st ed.). CRC Press. https://doi.org/10.1201/9781351231954

Neuper, C., Scherer, R., Reiner, M., & Pfurtscheller, G. (2005). Imagery of motor actions: differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. Brain research. *Cognitive brain research, 25*(3), 668–677. https://doi.org/10.1016/j.cogbrainres.2005.08.014

Nijboer, F., Birbaumer, N., & Kübler, A. (2010). The influence of psychological state and motivation on brain–computer interface performance in patients with amyotrophic lateral sclerosis–a longitudinal study. *Frontiers in Neuropharmacology, 4*, 55.

Nijboer, F., Furdea, A., Gunst, I., Mellinger, J., McFarland, D. J., Birbaumer, N., & Kübler, A. (2008). An auditory brain–computer interface (BCI). *Journal of neuroscience methods, 167*(1), 43-50.

Peng, Y., Wang, J., Liu, Z., Zhong, L., Wen, X., Wang, P., Gong, X., & Liu, H. (2022). The application of brain-computer interface in upper limb dysfunction after stroke: A systematic review and meta-analysis of randomized controlled trials. *Frontiers in Human Neuroscience, 16*. https://doi-org.proxy-ub.rug.nl/10.3389/fnhum.2022.798883

Pfurtscheller, G., Neuper, C., & Birbaumer, N. (2005). Human brain-computer interface. In A. Riehle & E. Vaadia (Eds.), *Motor cortex in voluntary movements: a distributed system for distributed functions* (1st ed., pp. 367-401). CRC Press.

Randolph, A. B. (2012). Not all created equal: individual-technology fit of brain-computer interfaces. In 2012 45th Hawaii International Conference on System Sciences (pp. 572-578). IEEE.

Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., & Lécuyer, A. (2010). OpenViBE: An open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence: Teleoperators and Virtual Environments, 19*(1), 35–53. https://doi.org/10.1162/pres.19.1.35

Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern-und leistungssituationen. *Diagnostica, 47*, 57-66. http://dx.doi.org/10.1026//0012-1924.47.2.57

Rimbert, S., Bougrain, L., & Fleck, S. (2020). Learning how to generate kinesthetic motor imagery using a BCI-based learning environment: A comparative study based on guided or trial-and-error approaches. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 2483-2498). IEEE.

Roc, A., Pillette, L., Mladenovic, J., Benaroch, C., N'Kaoua, B., Jeunet, C., & Lotte, F. (2021). A review of user training methods in brain computer interfaces based on mental tasks. *Journal of neural engineering, 18*(1), 10.1088/1741-2552/abca17. https://doi.org/10.1088/1741-2552/abca17

Ropper, A. H., Samuels, M. A., & Klein, J. (2019). Adams and victor's principles of neurology. McGraw-Hill Education.

Shute, V. J. (2008). Focus on formative feedback. *Review of educational research*, *78*(1), 153-189.

Thompson M. C. (2019). Critiquing the Concept of BCI Illiteracy. *Science and engineering ethics*, *25*(4), 1217–1233. https://doi.org/10.1007/s11948-018-0061-1

Vidal J. J. (1973). Toward direct brain-computer communication. *Annual review of biophysics and bioengineering*, *2*, 157–180. https://doi.org/10.1146/annurev.bb.02.060173.001105

Van Rossum, G., & Drake Jr, F. L. (1995). Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam.

Vuckovic, A., & Osuagwu, B. A. (2013). Using a motor imagery questionnaire to estimate the performance of a brain–computer interface based on object oriented motor imagery. *Clinical Neurophysiology*, *124*(8), 1586-1595.

Yang, Y.J., Jeon, E.J., Kim, J.S., & Chung C. K. (2021) Characterization of kinesthetic motor imagery compared with visual motor imageries. *Sci Rep 11*, 3751. https://doi.org/10.1038/s41598-021-82241-0

Appendix A

Written instructions

To start, you will be introduced to the hand imagery task.

As you have read in the information sheet you will be asked to perform motor imagery. Motor imagery involves imagining a movement without actually executing it overtly. Please pay close attention to the instructions.

Find a comfortable sitting position and straighten your back. Breathe deeply but calmly throughout the experiment.

We are going to practice a specific movement using our hands.

Please place your hands on the table or on the armrests of the chair you are sitting on with the palms facing upwards. You will perform a grabbing movement with your hands using a stress ball. Please place the stress ball in your dominant hand. Pay close attention to the feeling of the stress ball in your hands. While grabbing the stress ball, all of your fingers should be in direct contact with it. Now, you will squeeze the stress ball tightly and hold it for about one second. Then slowly open your hands back to an open resting position. Again, squeeze (...) and release. Please do not hesitate to get in touch with the experimenter now, should you have any questions. (...) If you are confident with this movement, repeat this several times. While doing this, pay close attention to any kind of sensations you might be experiencing. Pay special attention to the perception of pressure, heat, skin perception, and muscular tension in your arm or hand.

Now, you can put the stress ball aside. Next, you will attempt to replicate the movement in your head as accurately as possible. Please start by performing the movement without the stress ball but reduce the muscular activity progressively until you are left with nothing but **the** imagination. Try to imagine the same sensations you noticed earlier. Recreate the experience with as much detail as possible in your mind. Pay attention to the muscular movement, feeling of the pressure, and sensation of temperature you have experienced during the actual movement.

Again, if you have any questions, please contact the experimenter. (...) If you are confident to move on, perform this task entirely in your imagination. During the imagination keep your eyes open! Repeat the movement at least five times in your mind.

We will now explain the concept of mental calculation and the task to perform.

As you have read in the information sheet you are asked to perform mental calculations. Mental calculation involves arithmetical calculations without the help of tools or supplies. Please pay close attention to the instructions.

Find a comfortable sitting position and straighten your back. Breathe deeply but calmly throughout the experiment.

We will practice some specific mental calculations.

You will continuously subtract a one-digit number (between 3 and 9) from 100. An example for this is: 100 - 7 = 93, 93 - 7 = 86 and so forth, until you reach a negative number.

Now you will do such a calculation. Continuously subtract 4 from 100 while speaking out loud or quietly mumbling to yourself until you reach a negative number. If any mistakes arise in this or any further trial, please just continue with the last number that you remember. The process of mental calculation is more important than the actual result of the calculation. Please now attempt these subtractions.

Now you will do another calculation. However, this time you will calculate only in your mind. Hence, do not move your lips. Continuously subtract 9 from 100 until you reach a negative number. Again, do not worry about any calculation mistakes but please just continue with the last number that you remember. During the calculation, please keep your eyes open! Please calculate now.

We will now explore the foot motor imagery task.

As you have read in the information sheet, you will be asked to perform motor imagery. Motor imagery involves imagining a movement without actually executing it overtly. Please pay close attention to the instructions.

Find a comfortable sitting position and straighten your back. Breathe deeply but calmly throughout the experiment.

We will now practice a specific movement using our feet.

Please remove your shoes for better mobility. (...) Place both of your feet flat on the ground in a comfortable position. To begin, you will lift the front part of your feet as high as possible. Your heels will remain on the ground. Curl your toes up as far as you can, as if you wanted to get them as close to your knees as possible. You will hold this stance for about one second and then slowly ease your feet back to the resting position. Please do not hesitate to get in touch with the experimenter now, should you have any questions. (...) Please repeat this movement several times.

While doing this, pay close attention to any kind of sensations you might be experiencing. Pay special attention to any tension in your leg and feet muscles, as well as possible temperature changes or simply the sensations of pressure on your feet.

Now you can rest your feet. Next, you will attempt to replicate the movement in your head as accurately as possible. Please start by performing the movement but reduce the muscular activity progressively until you are left with nothing but the imagination. Try to imagine the same sensations you noticed earlier. Recreate the experience as detailed as possible in your mind. Pay attention to every muscle movement and sensation of pressure and temperature you have experienced while actually performing the movement.

Again, if you have any questions, please contact the experimenter. (...) Now perform this task entirely in your imagination. During the imagination keep your eyes open! Repeat the movement at least five times in your mind.

We will now explain how to perform the auditory imagination task.

As you have read in the information sheet you are asked to perform Auditory Imagery. Auditory Imagery involves the imagination of listening to a familiar tune. Please pay close attention to the instructions.

Find a comfortable sitting position and straighten your back. Breathe deeply but calmly throughout the experiment.

We will now practice a specific auditory imagination. Think of a song that you may have had stuck in your head before. This song can be entirely instrumental or include lyrics. In a moment you may play this tune yourself on one of your devices. You could use YouTube, Spotify, Apple Music, or any other platform you wish to use. It is important that you <u>do not watch the music video</u>, as we would like you to solely focus on the melody, the background instrumentals, and possible lyrics. You will later attempt to reproduce the listening experience in your head. Pick a part of a song that you can later replay in your head. You can now start to play your song of choice.

Now, without moving your lips or humming, try to listen to this tune in your head. Try to remain still as you do this. Hereby, try to imagine listening to this tune, instead of singing along with it. Reconstruct the melody, the tune and possibly the lyrics as you listen to the song in your head. During the imagination keep your eyes open.

Please perform this imagination now over the next 20 seconds.

If you have any questions or problems with this imagery please contact one of the researchers present.

Appendix B

Figure B1

Hand imagery video instructions

Note. Image from video instructions for hand imagery with first person perspective

Figure B2

Mental calculation video instructions

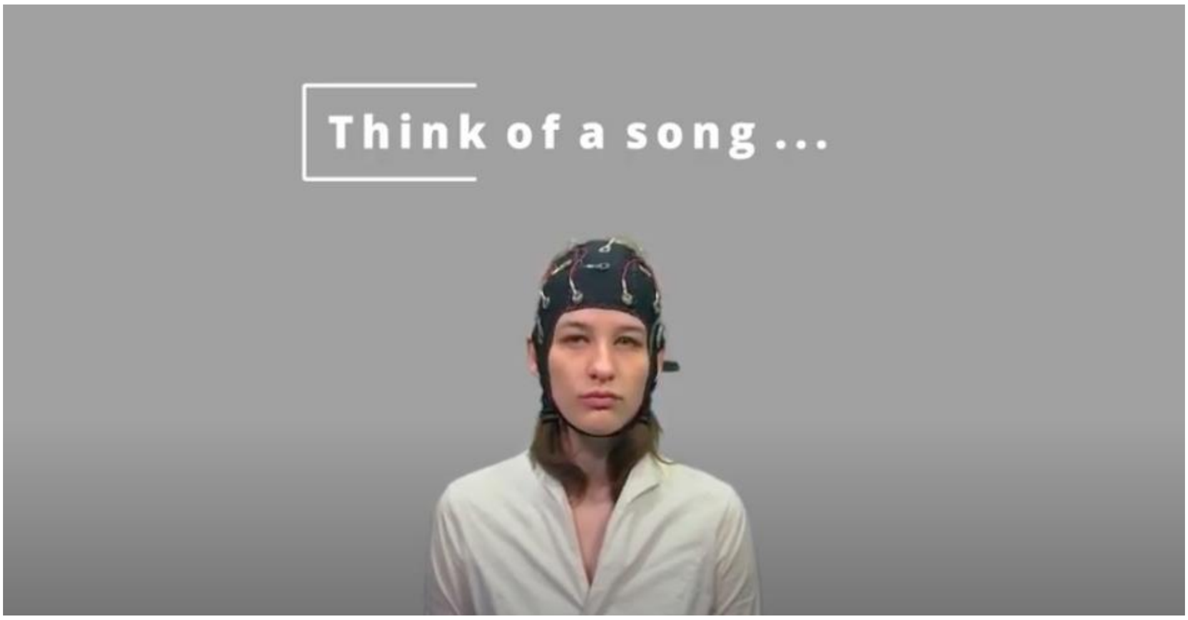
Note. Image from video instruction for mental calculation

Figure B3

Note. Image from video instruction for feet imagery with first person perspective

Figure B4

Auditory imagery video instructions



Note. Image from video instruction for auditory imagery

Appendix C

Table C1

Normality check for the Variables of Interest

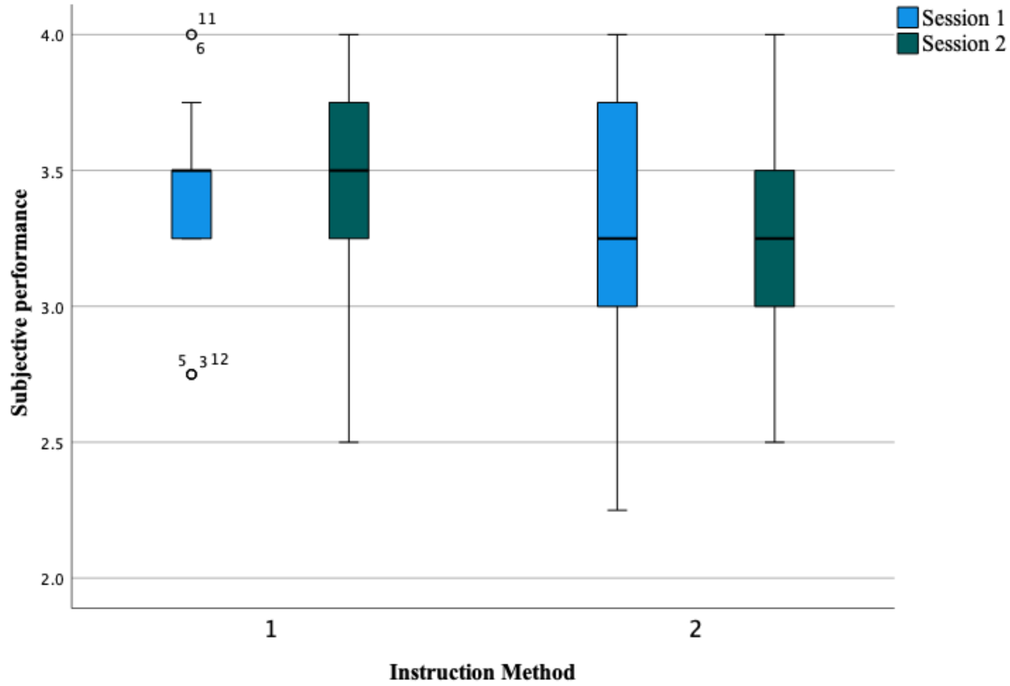
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Motivation S1	.101	27	.200*	.958	27	.332
Motivation S2	.123	25	.200*	.951	25	.264
Subjective performance S1	.175	27	.033	.940	27	.121
Subjective performance S2	.155	26	.109	.927	26	.066
Vividness of visual imagery	.107	27	.200*	.972	27	.654
Classification accuracy S1	.173	25	.053	.891	25	.011
Classification accuracy S2	.101	24	.200*	.973	24	.751

*. This is a lower bound of the true significance

a. Lilliefors Significance Correction

Figure C1

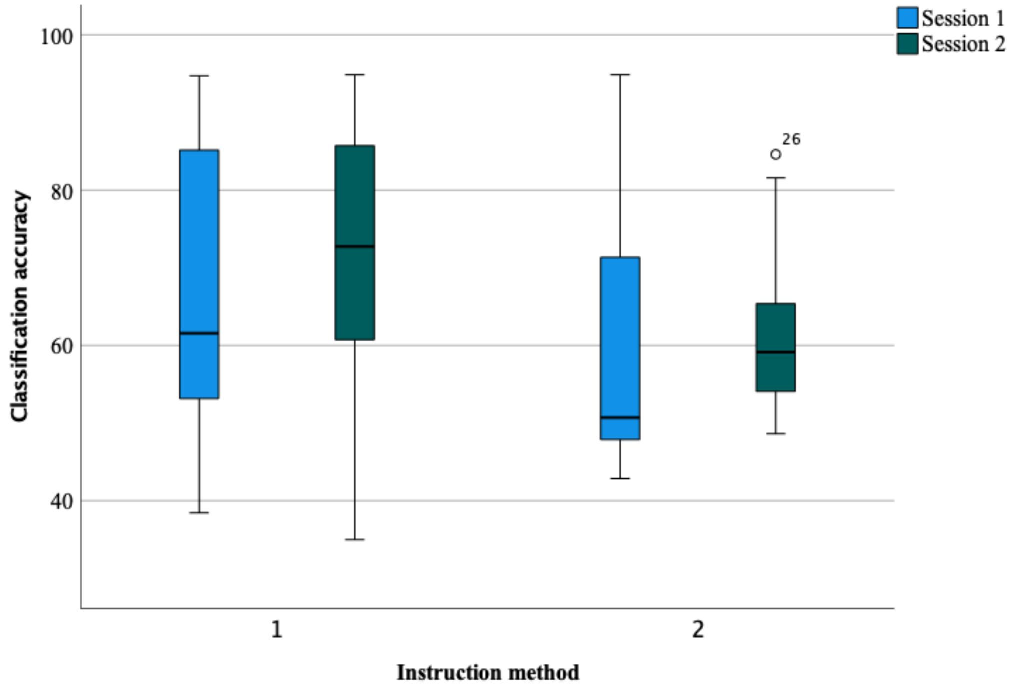
Outliers check for RM-ANOVA for hypothesis one



Note. Y-axis represents subjective performance scores.

Figure C2

Outliers check for RM-ANOVA for hypothesis two



Note. Y-axis represents classification accuracy scores.

Appendix D

RM-ANOVA for hypothesis one without outliers

Table D1

Tests of Within-Subjects Effects

Measure: Subjective Performance

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
session	Sphericity Assumed	.001	1	.001	.010	.923	.001
session * instructionmethod	Sphericity Assumed	.072	1	.072	.982	.334	.049
Error(session)	Sphericity Assumed	1.395	19	.073			

Table D2

Tests of Between-Subjects Effects

Measure: Subjective Performance

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	462.769	1	462.769	1892.864	.000	.990
instructionmethod	.661	1	.661	2.705	.116	.125
Error	4.645	19	.244			

Figure D1

Plot of Estimated Marginal Means RM-ANOVA with Classification Accuracy and Instruction Method.

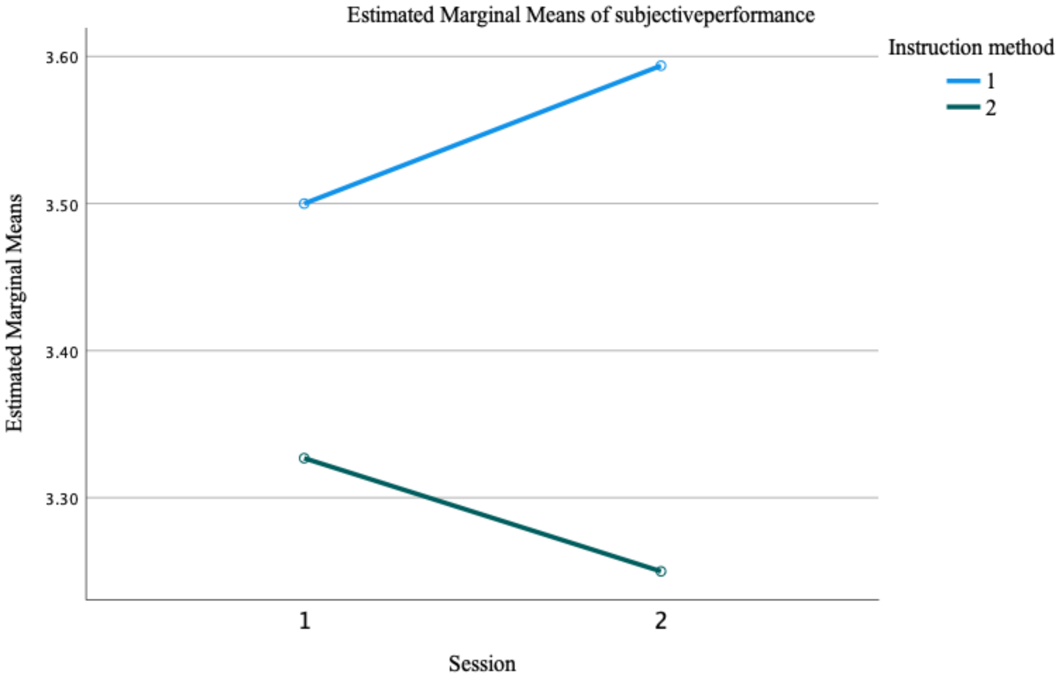


Table D3

Descriptive Statistics in RM-ANOVA of Subjective Performance in Session 1 and 2 for the Two Instruction Methods (Video and Written) Without the Outliers.

	Instruction method	Mean	Std. Deviation	N
Subjective performance session 1	Video	3.50	.13	8
	Written	3.33	.51	13
	Total	3.39	.42	21
Subjective performance session 2	Video	3.59	.23	8
	Written	3.25	.44	13
	Total	3.38	.41	21

Appendix E

Figure E1

Plot of Estimated Marginal Means RM-ANOVA with Subjective Performance and Instruction Method

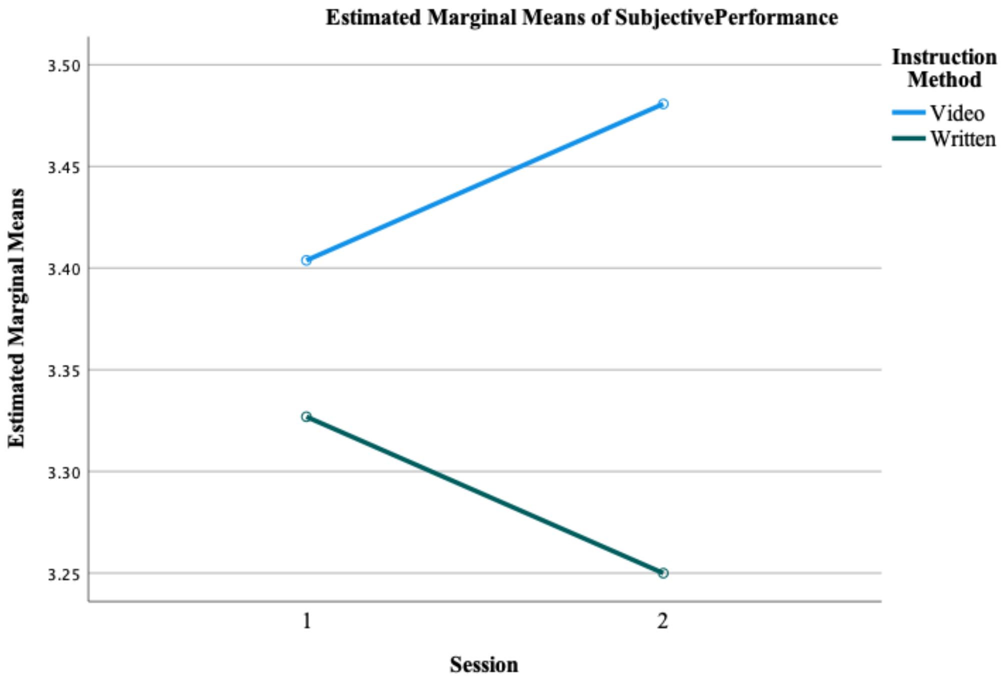
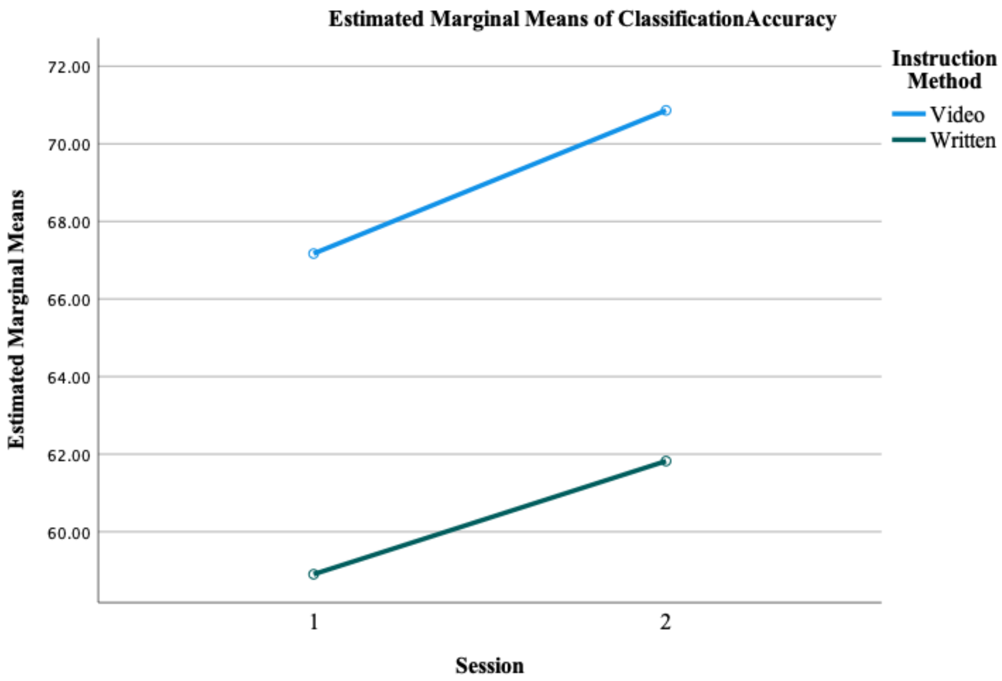


Figure E2

Plot of Estimated Marginal Means RM-ANOVA with Classification Accuracy and Instruction Method



Appendix F

Table F1

Independent t-tests for lower and higher classification accuracy groups for different variables of interest

		Levene's Test for Equality of Variances				
		F	Sig.	t	df	Sig. (2-tailed)
mean_subjS1S2	Equal variances assumed	6.884	.015	-.643	23	.527
	Equal variances not assumed			-.654	20.339	.520
QCM_S1S2	Equal variances assumed	.279	.602	-.548	23	.589
	Equal variances not assumed			-.550	22.983	.587
Vividness of visual imagery	Equal variances assumed	.013	.909	-.357	23	.725
	Equal variances not assumed			-.356	22.708	.725
Mind-body techniques	Equal variances assumed	6.396	.019	.973	23	.341
	Equal variances not assumed			.995	18.935	.332
Experiences with electronic devices	Equal variances assumed	.410	.528	-.634	23	.532
	Equal variances not assumed			-.637	22.967	.530
Creative activities	Equal variances assumed	8.847	.007	-1.140	23	.266
	Equal variances not assumed			-1.172	16.868	.257

Table F2

Independent t-test for experience with electronic devices for the two classification accuracy groups (lower and higher) without outliers

		Levene's Test for Equality of Variances				
		F	Sig.	t	df	Sig. (2-tailed)
Experiences with electronic devices	Equal variances assumed	2.838	.108	-.334	20	.742
	Equal variances not assumed			-.347	18.890	.732