Identifying workload levels with a low-cost EEG device using an arithmetic task

Bachelor's Project Thesis

Merel Wiersma, s2386038, M.Wiersma.10@student.rug.nl,
Supervisor: dr. J.P. Borst

Abstract: EEG-based systems are widely used because of their high temporal resolution, but they're not affordable for everyone. One of the newest low-cost EEG devices that was released is the Emotiv Insight, which is a mobile five channel headset promoted to produce clean, robust signals anytime, anywhere. The objective of this research was to test its abilities concerning distinguishing different workload levels, since monitoring workload can help us for example at work create a safer and better working environment, causing higher productivity and motivation. Previous research revealed that in event related potential (ERP) measures, the P300 reflects attentional and working memory processes. Therefore, we have manipulated workload levels by varying long term memory retrieval and working memory updates. The ERP results showed small significant differences around the P300 for absence compared to the presence of working memory updates at the AF4 channel. But for the other channels and the results concerning long term memory retrievel, no significant differences were found around the P300. Therefore, the conclusion of this research is that the Emotiv Insight was not capable of distinguishing between different workload levels.

Keywords: Emotiv Insight headset; ERP measures; Long term memory; P300; Workload; working memory

1 Introduction

Over 80 years already it is known that electrical activity of the brain can be recorded externally by electrodes on the surface of the scalp (24). This resulted in electroencephalography-devices which are widely used because of their portability, high temporal resolution and relatively low costs compared to other non-invasive methods such as MEG and fMRI (40). Nevertheless, the traditional EEG-device is still not affordable for everyone.

Figure 1.1: The Emotiv Insight

Currently, more and more low-cost EEG devices appear on the market. One of the newest low-cost EEG devices that was released is the Emotiv Insight(Emotiv Systems Inc., San Francisco, CA, USA), see Figure 1.1. This is a mobile five channel headset promoted to produce clean, robust signals anytime, anywhere. Although this sounds very promising, the question remains what can be measured with only five channels. The objective of this research is to test its abilities concerning mental workload, because mental workload is an important and relevant subject in our modern knowledge society, where information overload is a fact of life. Monitoring mental workload can help us to create better and safer working environments, causing higher motivation and productivity.(27; 28) But this is only feasible with a robust, low-cost EEG-device.

1.1 Mental Workload

"Mental workload" has been defined in different ways, but a common definition is that mental

workload is the ratio between task demands and a person's capacity, where workload is high when task demands are close to exceeding capacity.(15; 26). In other words, mental workload describes the level of mental resources utilized when a person is performing a task(22).

Mental workload can be measured in different ways, but research revealed that EEG is the most sensitive or promising indicator of mental workload compared to physiological variables like different eye and heart related measurements(3; 4; 6; 8). Another way of measuring workload is with subjective rating scales like RSME(41), SWAT and NASA-TLX (25). But this is distracting for the subject (23; 39) and possibly suffers from biases because it is always measured at one point (13). A last alternative for measuring workload is measuring performance measures like accuracy, but this is also only one measurement each trial. EEG on the other hand, provide a continuous record of mental workload over time.

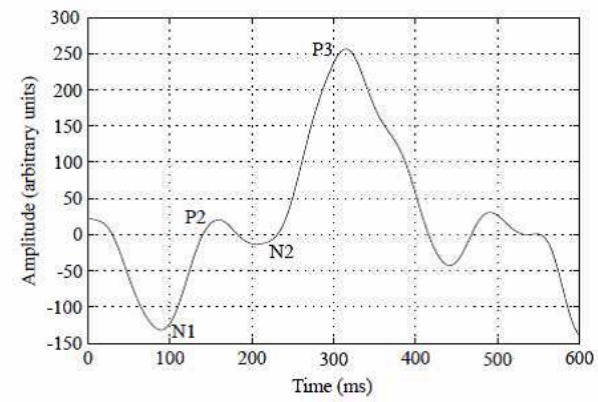

Figure 1.2: Overview early ERP attributes

One aspect of EEG are event related potential measures. ERPs are scalp-recorded fluctuations in the brain's electrical activity, elicited by a stimulus event(35). ERPs are calculated by averaging EEG epochs time-locked of this particular event, for this research solving equations. According to previous research, the P300 component of ERP measures reflects attentional and working memory processes(29; 30). Many studies found that when memory or workload increases, the P300 decreases(1; 12; 16; 19; 31; 32; 38). For Example, Allison used a shooting game in which workload was manipulated by adjusting the number of enemies. The ERP amplitude diminished as game difficulty increased.

Other early components like the N100, N200 (1; 18; 36) and the P1 (31) respond to workload or task difficulty as well. In Figure 1.2 you can see an overview of these ERP components.

Although it is clear that it is possible to measure workload with a full-scale EEG device, the question is whether this is also possible with a simpler device. In three papers, research has been done with a low-cost EEG device, the Emotiv Epoch, concerning mental workload levels. This mobile 14 channel device is more expensive than the Emotiv Insight but still cheaper than a traditional EEG device. The results showed significant differences between workload levels (2; 17; 37).

1.2 Arithmetic task

An often used task to measure mental workload levels is the well-known n-back task (5; 7; 9; 14; 33; 34; 37). For this task, you have to recall an item you saw n items back. However, for this task it is not clear which resources in the brain are employed exactly, because people have different strategies to remember the items. The same holds for other tasks for which the participant has to remember multiple items, like a grid memory task and a forward or backward digit span task (4). Furthermore, these tasks are artificial, especially designed for research, not something people do or learn by themselves.

In another study subjects had to perform a silent reading task under different difficulty levels (17). But during a silent reading task, subjects can suffer mind wandering. So, again we are not sure what exactly is happening in the brain.

Arithmetic tasks on the other hand, can include specific areas in the brain and are less artificial since people learn mathematics on school and use them in their daily life. For example to calculate how much they have to pay a friend after eating together. Simple equations can also be solved quite fast, which prevents mind wandering from happening. Berka, for example, used an addition task requiring participants to employ working memory

and executive function resources. An obvious workload effect was shown (4). Echkard also performed a research in which mental workload was measured using a simple mathematical task. Again, a correlation between increasing the difficulty of the calculations and mental workload was shown (11).

For the current study, we designed an arithmetic task with different types of equations that do or do not require the subjects to employ their working and/or long term memory. So, to operationalize mental workload we varied working memory updates and long term memory retrieval.

Based on the previous researches described above, the expectation is that the Emotiv is able to measure workload levels. However, the differences may be less significant than with the Epoch and classical EEG devices, due to the fact that it has less channels. The difference we think should be possible to measure is the decrease of the P300 when the workload is high, meaning working memory updates and or long term memory retrievals are required.

2 Methods

The experiment consisted of two phases: a training phase and a test phase. During the training phase, the subject learned to know the task by solving one equation of each type. In the test phase, during which the EEG data were collected, 50 equations of each type had to be solved by the subject. The "types" differed based on the requirement of long term memory retrieval and/or working memory updates. One type required none of the above.

2.1 Participants

Thirteen students from the University of Groningen participated for a compensation of 8 Euros. The mean age of the participants was 23. All participants were able to speak fluent English and had normal visual capabilities. Halfway through testing the last participant the battery of the Emotiv Insight broke. The EEG data of participant 9 was too disturbed and therefore not useful. Participant 10 did not understand the task fully correctly for the first 72 trials and was not able to solve all equations. Therefore, the results of this participant have been removed as well. So for the stimulus-locked analysis ten participants remained, four of them women and six men.

For the response-locked analysis, two more participants (both a male and a female) had to be removed, due to too many eye-movements in the EEG data. In section 3, these different analyses will be explained in further details.

2.2 Materials

The Emotiv Insight was used to record electroencephalographic data. This wireless Bluetooth device has five electrodes, located and named as in Figure 2.1.

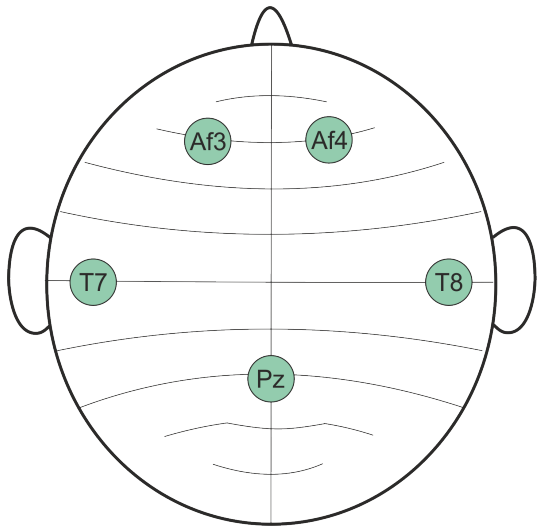

Figure 2.1: Overview and location of the electrodes of the Emotiv Insight

To collect the electroencephalograph data, the Emotiv was connected to the Emotiv software Testbench using bluetooth. Using virtual COM-ports, Testbench and OpenSesame(21) were connected. The experiment was implemented and the markers for the EEG data were defined in Opensesame. EEGlab(10), which is an interactive Matlab toolbox, was used to filter and analyse the data afterwards.

2.3 Design

The task consist of five types of equations that will be explained below. First an example equation will be given, followed by a detailed description.

- Type 1: x = 100, x = y with y in 0 - 100
  This type of equation can be solved without workload, since the answer is y, which is in this case 100.

- Type 2: x + 12 = 48,x + y = z or x - y = z with y in 0 - 100, z in 0 - 100, x + y in 0 - 100 and x - y in 0 - 100
  To solve this type of equations, the subject has to employ his or her working memory to move y to the other side of the equation. In this case, 12 has to be moved to the other side of the equation and has to be subtracted from 48 after that. The mathematical rules needed to solve these equations are so easy that they are intuitive and therefore the use of our long term memory is minimal.

- Type 3: x = 4 / 2, x = y / z or x = y * z, with y in 1 - 10, z in 1 - 10, y/z in 1 - 100 and y * z in 1 - 100
  To solve this type of equation, the subject has to employ its long term memory to retrieve the answer, which is in this case 2.

- Type 4: 4x = 12, x/y = z or x * y = z with same conditions as above.
  This type is quite similar to type 3, except from the fact that a working memory update has to take place, because a digit had to be moved to the other side of the equation (in this case 4). After that, it is a multiplication or division just like type 3. In this case 12 / 3, which gives us the answer 4. So, both long term memory retrieval and working memory updates are required.

- Type 5: 4x + 5 = 17, ix + y = z or ix - y = z with i in 2 - 10, y in 2 - 100 and z in 1 - 100
  This last type is a combination of the above, requiring the most steps to be able to solve it. So both long term memory retrieval and working memory updates are required. In this case, 5 has to be transferred to the other side of the equation and must be subtracted from 17. The result, 12, has to be kept online in the working memory. After that, 4 can be transferred to the other side. Then, 12 has to be divided by 4 to bring us to the answer of the equation which is 3 in this case.

2.4 Procedure

The experiment began with a few personal questions the subject had to answer: their gender and their age. After that they saw an instruction screen. It explained the task, told them how they should answer by using the mouse and that it was more important to answer correctly then fast. This last thing was included, because only correctly answered trials can be used for the data analyses. But, since unusually slow trials are useless as well, the instructions told them that the equation would disappear after a while as well.

After the instructions, the training phase started. This phase consisted of one equation of each type, which sums up to five equations. Participant were asked whether they understood the task and were ready to start with the real experiment with a text screen. When he or she pressed the ok button, the real experiment began, which consisted of 250 equations, 50 of each type. These were generated arbitrarily - within the requirements described above - and presented in a random order. The 250 equations were divided into three blocks. Since 250 cannot be divided by three equally, the first block consisted of 84 equations, and both the second and third of 83 equations.

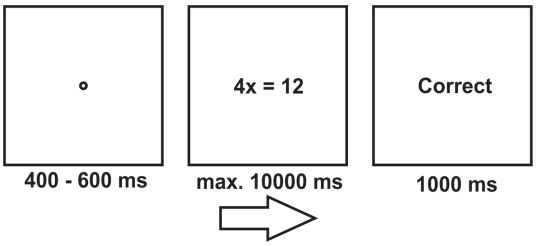

Figure 2.2: Overview of the arithmetic task

Each trial consisted of a fixation dot, equation and feedback, see Figure 2.2. The fixation dot was shown for a random duration between 400 and 600 milliseconds, to prevent expectation. The equation was shown for a maximum of ten seconds. The sub-

ject had to indicate he or she knew the answer by clicking on the screen. Afterwards, an onscreen keyboard appeared for a max of two seconds, so that the subject could perform the whole experiment by mouse only. This ensured that the subject moved as little as possible. After the participant had submitted an answer, feedback was given which was either the word correct or wrong. This was done because, as said before, only correctly answered trials could be used for the analysis of the data.

2.5 EEG recording and Analyses

During the whole task the participants were wearing the Emotiv Insight. For most participants a little bit of salty gel was needed to establish a stable connection between the five electrodes and the scalp. Behind the left ear, the reference electrode was located. The data were re-referenced over all channels. Frequencies below 1 and above 30 Hz were filtered out of the data. After that, the data were epoched based on both the moment of the stimulus and the moment of the response. The stimulus-locked data were epoched 200 ms before the stimulus and 1500 ms after it. The response-locked data were epoched 1500 ms before the response and 200 milliseconds after it. The stimulus-locked data were baselined based on the 200 milliseconds before the stimulus appeared on the screen. For the response-locked data the baseline was based on the 200 ms after the response was given by the participants.

During wrong answered trials we cannot tell what happened in the brain of the subject, so these were removed. Then, with help of EEGlab (10), the trials with extreme values were marked. All trials are visually inspected and the outliers were removed. A lot of data were disturbed and had therefore to be removed, which came down to about half of the trials for each participant. We decided at least 100 trials had to remain for containing the data in the analyses for the results. We used this threshold because a lower one would make the data less trustworthy and if we required more trials, not enough participants would remain.

3 Results

Both behavioural data and EEG data were collected during the experiment. As said in Section 2.5, the EEG data has been analysed in two ways, response-locked and stimulus-locked. Stimulus-locked means that the zero is the moment at which the stimulus appear on the screen. Response-locked means that zero is the moment at which the subject filled in the answer. For both the analyses, different workload levels based on working memory updates and long term memory retrievals will be researched and visualized in this section.

3.1 Behavioural results

Type	Reaction time	Answered correctly
1	851 ms	99.9%
2	3629 ms	97.9%
3	1723 ms	99.2%
4	2224 ms	99.2%
5	3976 ms	97.4%

Table 3.1: Table of the behavioural data

To see if the task we used was reliable, we took a look at the behavioural data of the participants, see Table 3.1 for an overview. As expected, the first type was answered the fastest, since for this type of equation the answer was right there. No transitions had to be made and no information had to be retrieved or kept online, so neither the working memory nor long term memory was used. Type 3 was quite simple as well, since the answers were already stored the brain. Only the retrieval of the answer had to take place. Type 4 was similar to type 3, but it included a working memory update. We see this in the reaction time which was indeed a little slower. Type 2 required only working memory updates. These answers could not be retrieved, but had to be calculated. The reaction times indicate that this is much slower than retrieving the answer from the long term memory. Type 5 contained the most steps to solve and was indeed answered the slowest. This was a combination of type 2 and type 4.

The second column of Table 3.1 tells us the percentage of equations that were answered correctly. For all types this amount is quite high, telling us that the participants did understand the task and the equations were not too difficult.

3.2 ERP results

To take a first look at the ERP data, a stimulus-locked and a response-locked graph have been made for each channel, see Figures 3.1 and 3.2. In each graph all five different types of equations are shown in a different color. In all stimulus-locked graphs the P300 component of the ERP can be recognized, except for the T7 graph. Unexpectedly, type 1 - which requires the lowest amount of workload - shows, compared to the other types, an increased P300 instead of a diminished one. Furthermore, type 5 - which requires the highest amount of workload - shows an average P300, while we expected it to be the most diminished compared to the other types. Type 2 has also an increased P300 compared to the other types, even though we thought it would be more diminished because of a relative high workload. Type 3 and 4 show a bigger difference in P300 for channels Pz and T8 than expected, since these types were quite similar. These results suggest that the P300 increased with a higher workload and diminished with a lower one which is exactly the opposite of what we expected.

In the response-locked graphs, all types look to have an arbitrary graph, no ERP components can be identified. For channel AF3, type 1 has a big dip at the left side of the graph, but this could not be explained. There is also no link between the different channels and the different types visible.

At each point in time, for both stimulus-locked analysis and response-locked analysis, a t-test has been done to test if there is a significant difference between the equations that required the subject to employ their long term memory and the equations that did not. The same was done for the equations that required subjects to employ working memory and the equations that did not. The significance level that was used is 0.05. The t-test that has been executed was paired, since all participants performed all conditions - they answered all different kinds of equations.

In Figure 3.3 the equations that did (blue) and did not (red) require long term memory retrieval are set out for each channel. The grey areas show which parts of the graph differ significantly. The translucent red and blue areas show the standard errors. These reflect what the results could look like with more participants. We see that the lines as wel als the translucent areas are quite overlapping for all graphs, which means that the are quite similar, even when the population would have been bigger. The absence of difference is also shown by the significant areas, which are only a few and very little. Furthermore, no significant differences are shown by the results around the P300. The same can be seen in Figure 3.4, which shows the differences between the equations that did (blue) and did not (red) require working memory updates. In this graph slightly more significant areas are shown, but these are still small and only one of them -in channel AF4- is at the P300 peak. Furthermore, both the red and blue lines and areas are again quite similar.

Similar graphs for the response-locked analysis can be found in in Figures 3.5 and 3.6. A lot of significant areas are shown by this graphs, but these are mostly small and not consistent between the different graphs. Furthermore, these significant areas are spread across the whole graphs which makes it impossible to find any logical patterns.

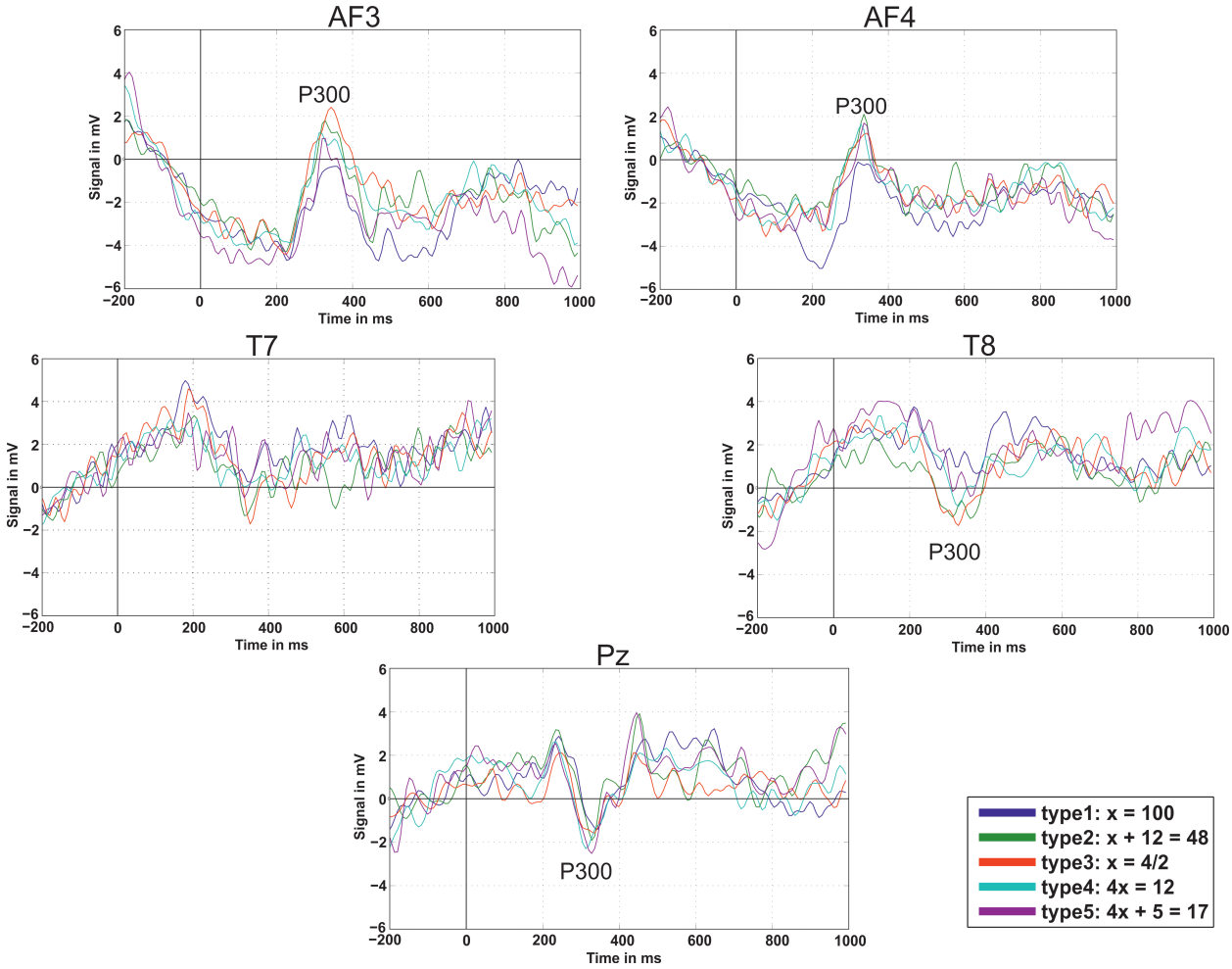

Figure 3.1: Stimulus-locked graphs for all channels

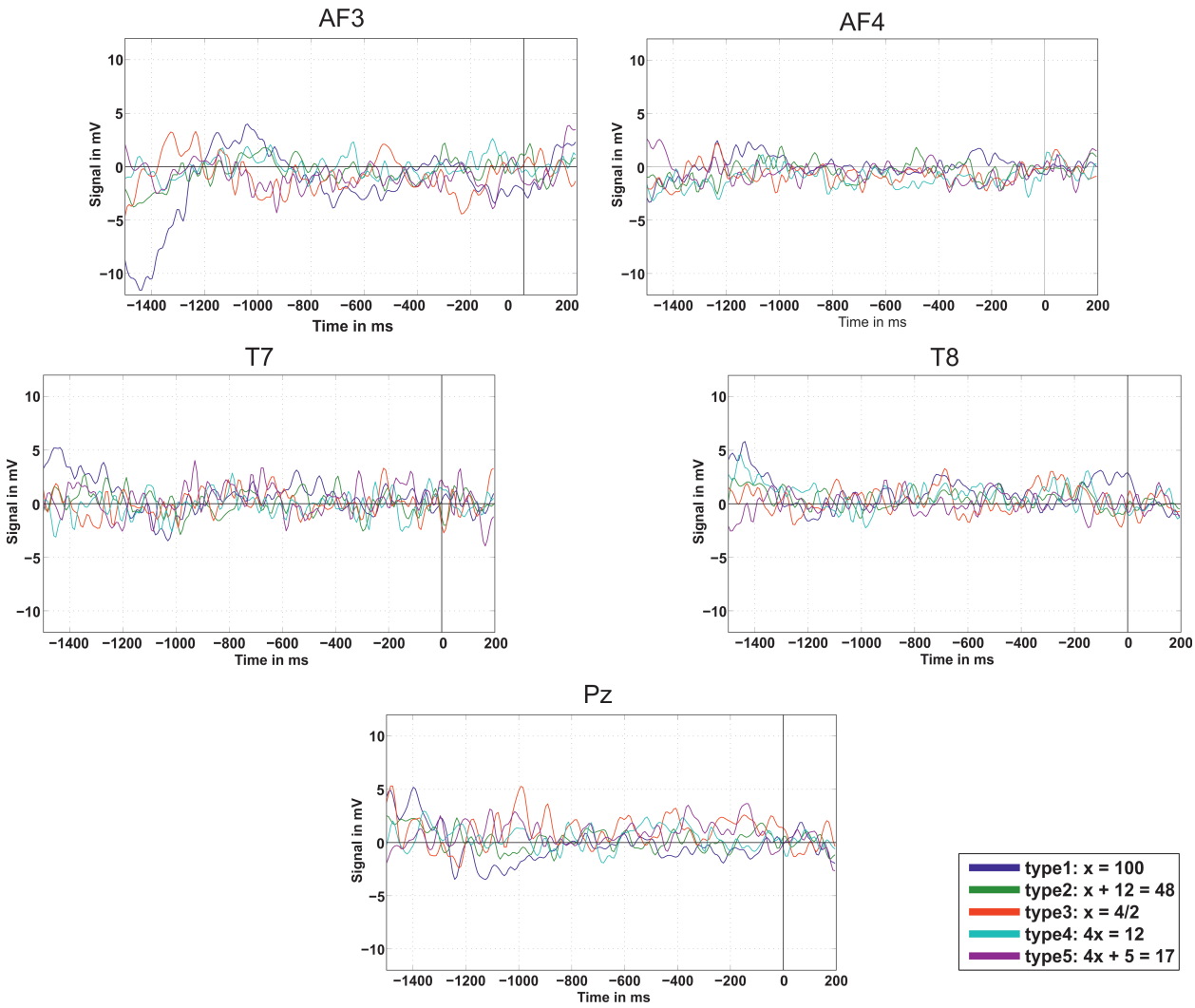

Figure 3.2: Response-locked graphs for all channels

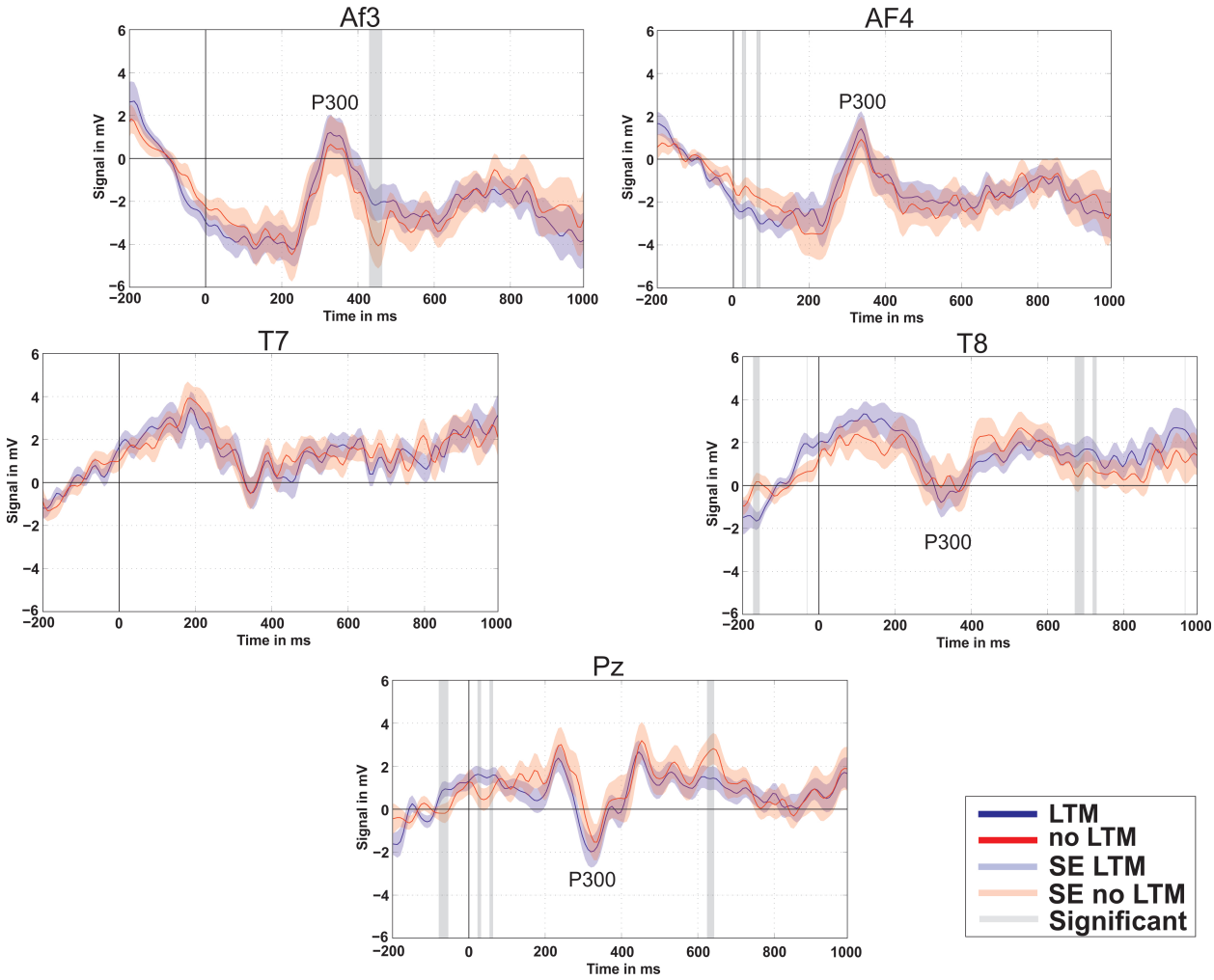

Figure 3.3: Stimulus-locked long term memory graphs of all channels with the standard errors and significant parts

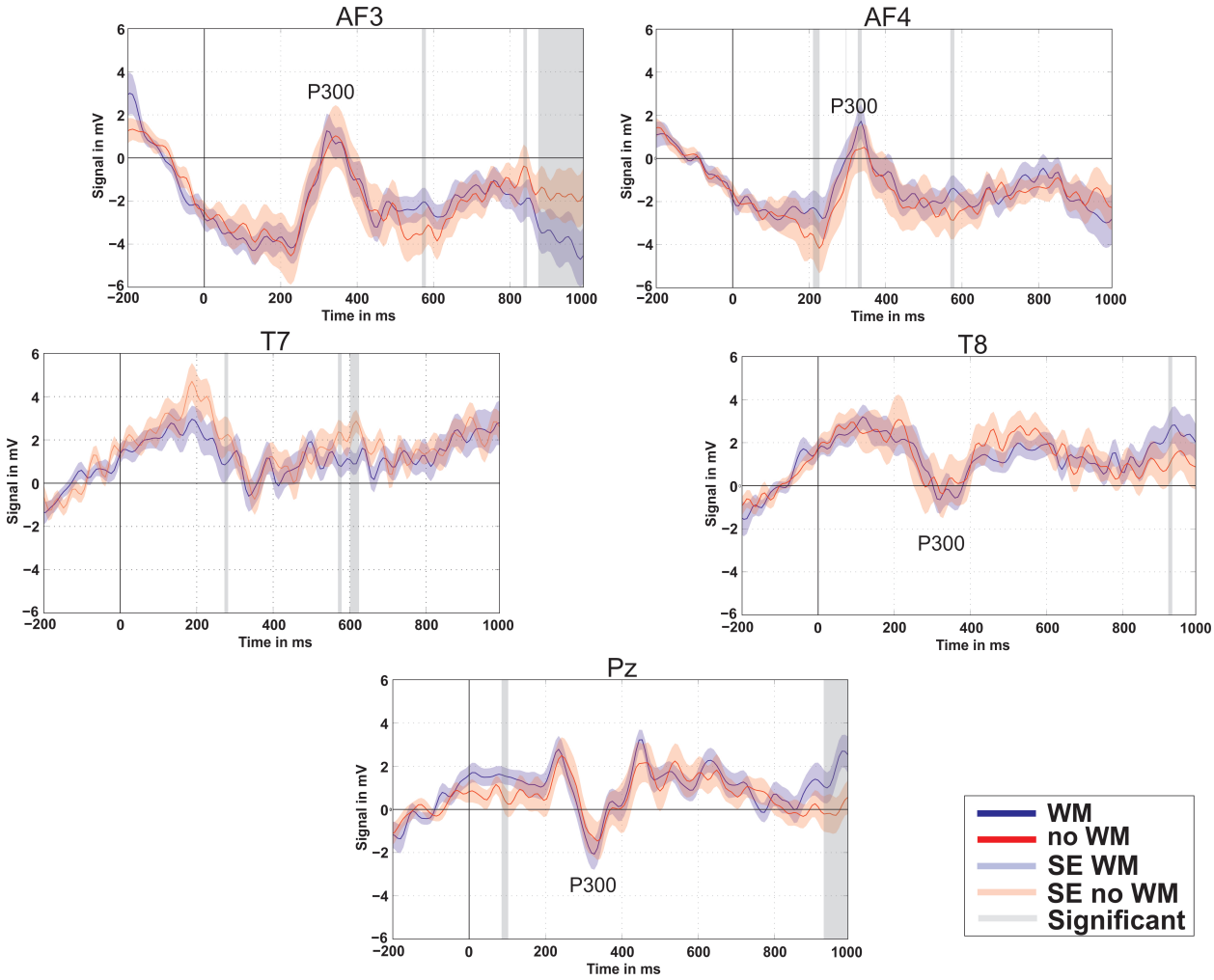

Figure 3.4: Stimulus-locked working memory graphs of all channels with the the standard errors and significant parts

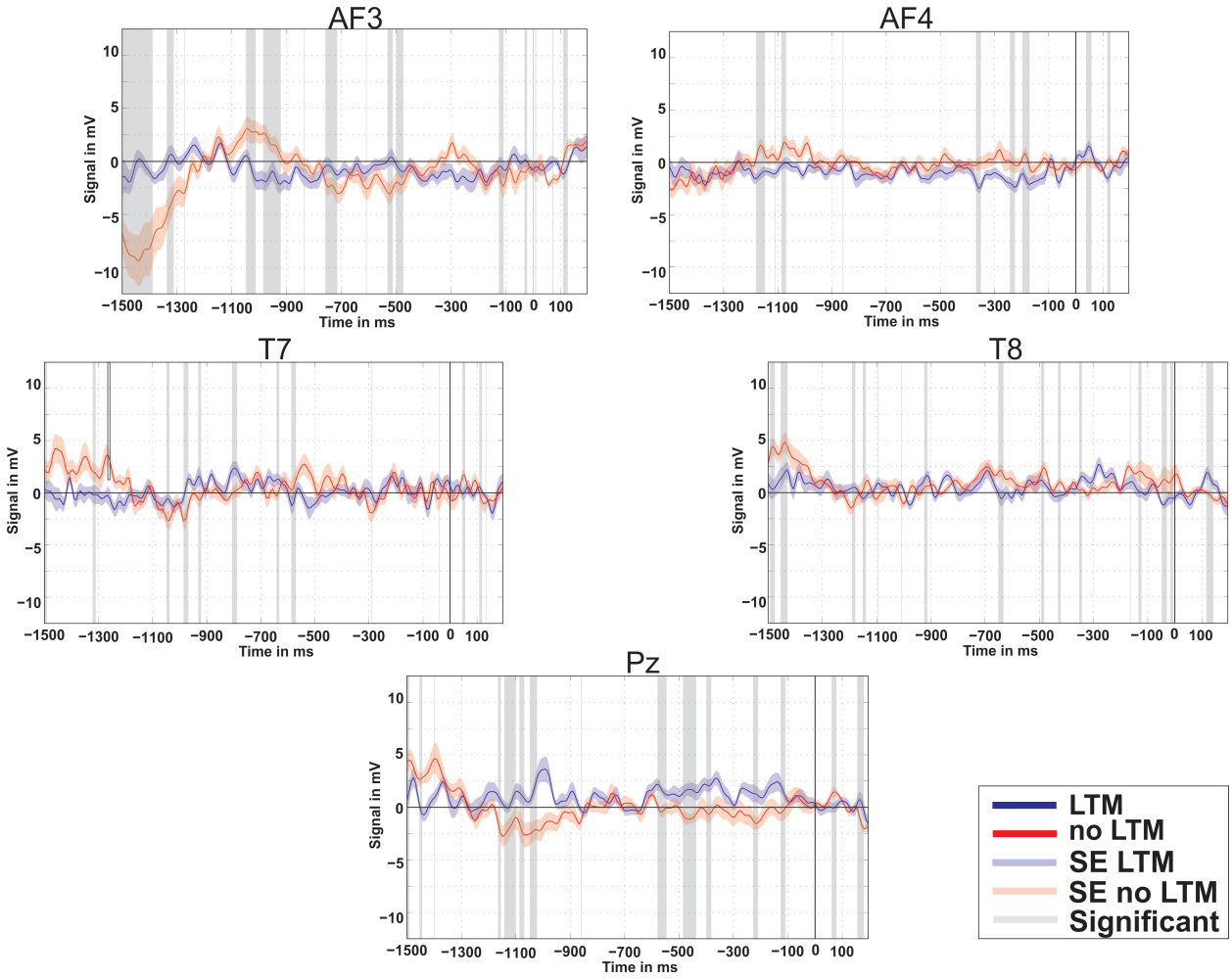

Figure 3.5: Response-locked graphs of the five channels concerning long term memory retrieval with the standard errors and significant parts

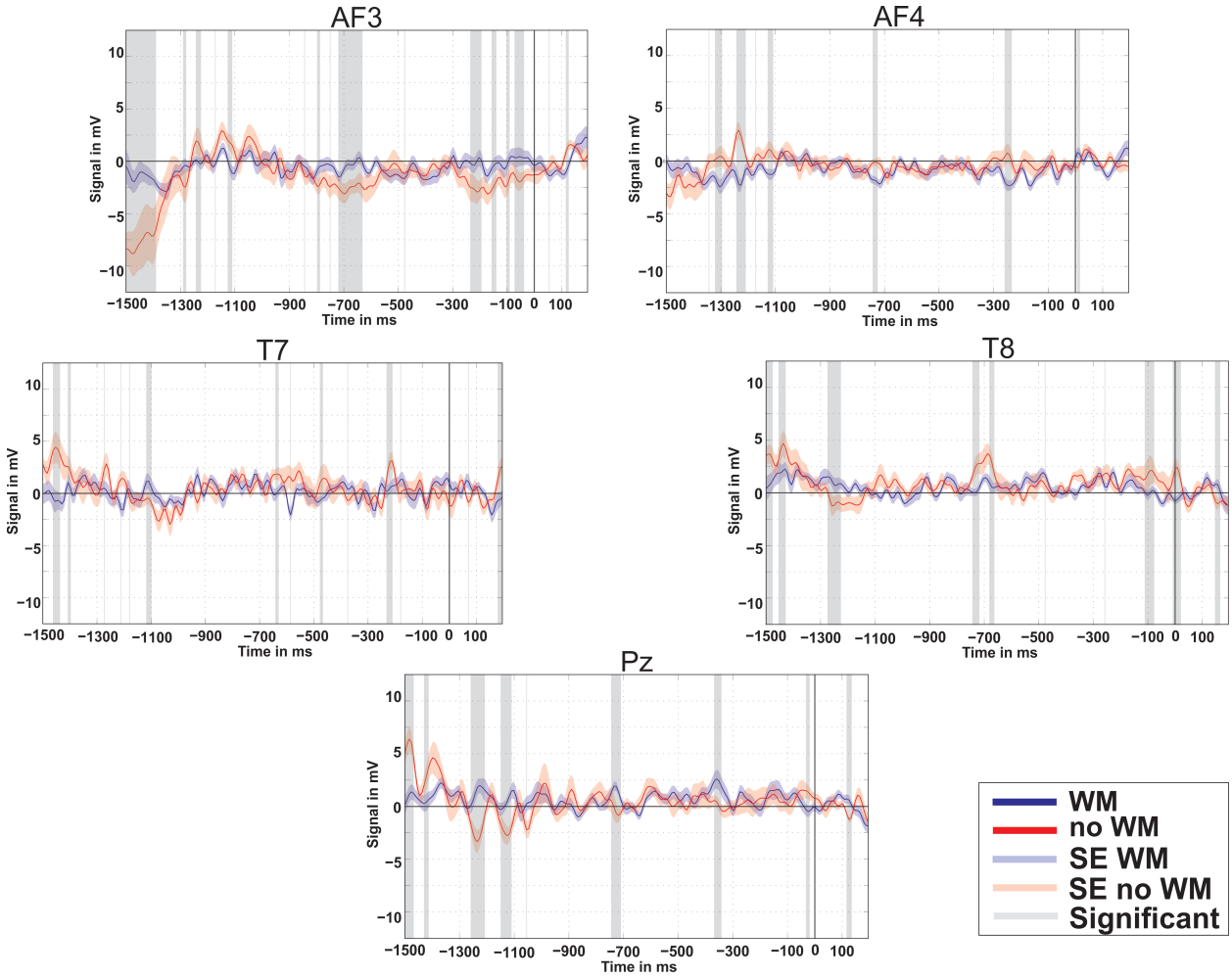

Figure 3.6: Response-locked graphs of the five channels concerning working memory updates with the standard errors and significant parts

4 Discussion

To test the abilities of the low-cost EEG device the Emotive Insight, a mental workload experiment was conducted. From previous research we have learned that early ERP attributes, especially the P300, reflect mental workload (1; 12; 16; 18; 29; 30; 31; 32; 36; 38). Using an arithmetic task, we research two types of mental workload, one requiring long term memory and the other requiring working memory. These types were analysed both response-locked and stimulus-locked. The results as showed in Section 3 will be discussed below.

4.1 Mental workload effects

In our results of the stimulus-locked analysis for the long term memory retrieval, we saw that the graph with the equations that involved long term memory retrieval and the ones that did not were quite similar, see Figure 3.3. No significant differences around the P300 were found. Based on the standard error, we can conclude that even with a bigger population no significant differences would have been found.

The results of the response-locked analysis concerning long term memory retrieval shows significant areas all over the place, which makes it impossible to draw any conclusions. The behavioural data on the other hand, did show differences between the situations in which the participants had to employ their long term memory and the ones they did not, see 3.1. Based on these results we can conclude that the Emotiv insight is not capable to distinguish between the situations in which long term memory retrievals were required and the situations in which they were not.

Our stimulus-locked results that concern the requirement of the working memory, do not show a convincing significant difference around the P300 either. Only the AF4 channel shows a little significant difference around the P300, but this one is very small and in contradiction with the literature. In the behavioural data in Subsection 3.1 we saw that equations that included working memory updates were the most difficult ones. In the graph for channel AF4 we see that the line representing the requirement of working memory updates shows an increased P300 compared to the line representing the ones that did not, instead of a diminished P300 like stated by the literature(1; 12; 16; 19; 31; 38; 32).

The results of the response-locked analysis for the working memory updates show significant areas all over the place as well. So again, no components or patterns could be found in this data. Based on this results we can draw a similar conclusion as for the long term memory retrieval. The Emotiv Insight was not able to distinguish between the situations in which working memory updates were required and the situations in which it was not.

Based on the above we can conclude that no convincing differences could be found between the different types of mental workload. The behavioural data on the other hand did show the expected differences. So, based on these results we can conclude that the Emotiv Insight is not capable of measuring significant differences in mental workload.

4.2 Implications

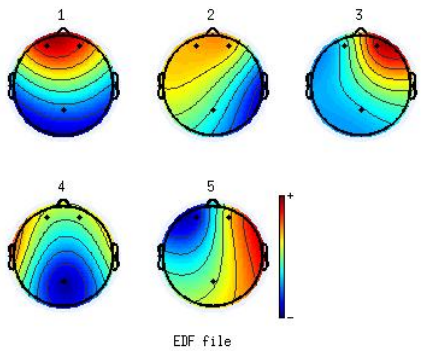

Figure 4.1: Example output ICA algorithm

One of the reasons that we could not find significant differences in mental workload could be the small number of participants. This shortage was due to the fact that the battery of the Emotiv broke down after twelve and a half participant. The standard errors were quite overlapping, suggesting that there would not have been more significant differences with a bigger population, but with more subjects the measurement will be more precise, so the standard errors will go down. This means that the translucent areas around the graphs will

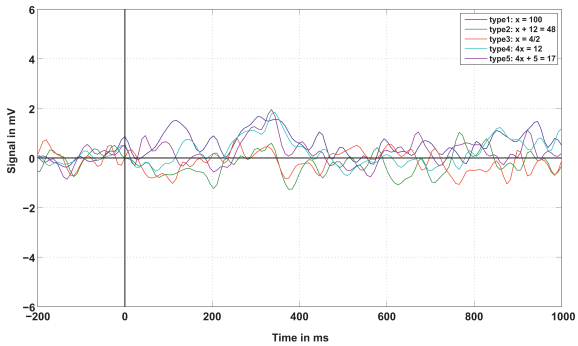

Figure 4.2: Stimulus-locked AF3 graph filtered by ICA

be smaller with more participants, so a significant difference could appear.

That we could not even find an ERP component or any form of a recurring pattern in the response-locked data could be due to the fact that the response times differed quite a lot. Close to the stimulus this does not change too much since all trials started at the same moment so the difference is small, but close to the response, the difference is bigger.

The Emotiv Insight is also quite sensitive, the electrodes did not always show a perfect connection with the scalp, the signal was easy disturbed and the graphs showed in Section 3 are, despite the filtering, quite disturbed as well. Furthermore, eye-movements were visible in the EEG data of many the trials. We have tried to filter this artefact out of the data using ICA (independent component analysis). ICA is an algorithm that is able to filter out obvious artefacts (20). It shows you an overview of the artefacts it has found, and at which part of the brain most of the activity of this artefact was located, see Figure 4.1 for an example. For most subjects, the eye-movements seemed to be recognized by the algorithm. For the example in Figure 4.1, one seems to be the eye-movement artefact. But, because we only had five channels, it filtered out too much of the signal and could therefore not be used, see Figure 4.2. So, the Emotiv Insight was in our experience not as robust as has been promoted.

4.3 Conclusion

Based on our results can be concluded that the Emotiv Insight was not able to distinguish between different workload levels. There were no convincing significant differences between the workload levels around the P300 and the only small difference that appeared was in contradiction with previous studies. The behavioural data on the other hand did show the expected differences. So, the Emotiv Insight might not be as suitable for measuring different workload levels as we hoped for.

Further research is needed, because a daily device to measure mental workload would still be very useful. Another component of EEG data is frequency band analysis. This could be tried as well. Also, more participants might help to get better and more trustworthy results. However, we think it might be better to test another low-cost EEG device in follow-up research.

References

[1] B.Z. Allison and J. Polich. Workload assessment of computer gaming using a single-stimulus event-related potential paradigm. *Biological Psychology*, 77(3):277–283, 2008.

[2] E. Anderson, K. Potter, L. Matzen, J. Shepherd, G. Preston, and C. Silva. A user study of visualization effectiveness using eeg and cognitive load. *Comput. Graph. Forum*, 30(3):791 – 800, 2011.

[3] N.R. Bailey, M.W. Scerbo, F.G. Freeman, P.J. Mikulka, and L.A. Scott. Comparison of a brain-based adaptive system and a manual adaptable system for invoking automation. *Human Factors The Journal of the Human Factors and Ergonomics Society*, 48(4):693–709, 2006.

[4] C Berka, D. J. Levendowski, M. N. Lumicao, A Yau, G Davis, V. T. Zivkovic, R. E. Olmstead, P. D. Tremoulet, and P. L. Craven. *EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory Tasks.* Aerospace Medical Association, 2007.

[5] T.S. Braver, J.D. Cohen, L.E. Nystrom, J. Jonides, E.E. Smith, and D.C. Noll. A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage*, 5:49–62, 1997.

[6] J.B. Brookings, G.F. Wilson, and C.R. Swain. Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, 42(3):361–377, 1996.

[7] A. Brouwer, M.A. Hogervorst, J.B.F van Erp, T. Heffelaar, P.H. Zimmerman, and R. Oostenveld. Estimating workload using eeg spectral power and erps in the n-back task. *Journal of Neural Engineering*, 9(4):14, 2012.

[8] J.C. Christensen, J.R. Estepp, G.F. Wilson, and C.A. Russell. Psychophysiological responses to changes in workload during simulated air traffic control. *The effects of day-to-day variability of physiological data on operator functional state classification*, 59(1):57–63, 2012.

[9] A. Conway, M. Kane, M. Bunting, D. Hambrick, O. Wilhelm, and R. Engle. Working memory span task: A methodological review and users guide. *Psychonomic Bull.*, 12:769 – 786, 2005.

[10] A. Delorme and S. Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

[11] H.H. Eckhard and M.P. Jamer. Pupil size in relation to mental activity during simple problem-solving. *Science, New Series*, 143(3611):1300 – 1190 – 1192, 1964.

[12] J.L. Evans. P300 as a measure of processing capacity in auditory and visual domains in specific language impairment. *Brain Research*, 1389:93, 2011.

[13] P.A. Hancock and N. Meshkati. *Human Mental Workload*. North-Holland, Amsterdam, The Netherlands, 1988.

[14] J.M. Jansma, N.F. Ramsey, R. Coppola, and R. Kahn. Specific versus nonspecific brain activity in a parametric n-back task. *Neuroimage*, 12:688 – 697, 2000.

[15] B.H. Kantowitz. *Human Factors Psychology, 3. Mental Workload*. Elsevier, The Netherlands, Amsterdam, 1987.

[16] T. Kida, Y. Nishihira, A. Hatta, T. Wasaka, T. Tazoe, Y. Sakajiri, H. Nakata, T. Kaneda, K. Kuroiwa, S. Akiyama, M. Sakamoto, K. Kamijo, and T. Higashiura. *Resource allocation and somatosensory P300 amplitude during dual task: effects of tracking speed and predictability of tracking direction*. Elsevier, Ireland, 2004.

[17] A. Knoll, Y. Wang, F. Chen, J. Xu, N. Ruiz, J. Epps, and P. Zarjam. Measuring cognitive workload with low-cost electroencephalograph. In *Human-Computer Interaction-INTERACT 2011*, pages 568–571. Springer Berlin Heidelbergs, 2011.

[18] A.F. Kramer, L.J. Trejo, and D. Humphrey. Assessment of mental workload with task-irrelevant auditory probes. *Biological Psychology*, 40(1):83 – 100, 1995.

[19] D.E.J. Linden. The p300: Where in the brain is it produced and what does it tell us? *The Neuroscientist*, pages 563–576, 2005.

[20] S. Makeig, A.J. Bell, T.P. Jung, and T.J. Sejnowski. Independent component analysis of electroencephalographic data. In *Advances in Information Processing System 8*, pages 145–151, Cambridge, Massachusetts, London, England, 1995. The MIT Press.

[21] S. Mathot, D. Schreij, and J. Theeuwes. Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2):314–324, 2012.

[22] N. Moray. *Mental Workload: Its Theory and Measurement*. Plenum, New York, NY, USA, 1979.

[23] W.F. Moroney, D.W. Biers, and F.T. Eggemeier. Some measurement and methodological considerations in the application of sub-

jective workload measurement techniques. *International Journal of Aviation Psychology*, 5(1):87–106, 1995.

[24] E. Neidermeyer and F. Lopes da Silva. *Electroencephalography: Basic principals, clinical applications, and related fields.* Lippincott Williams and Wilkins, Philadelphia, Pennsylvania, United States, 2005.

[25] T.E. Nygren. Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors The Journal of the Human Factors and Ergonomics Society*, 33(1):17–33, 1991.

[26] R. D. O'Donnell and F. T. Eggemeier. *Workload assessment methodology.* John Wiley AND Sons, Oxford, England, 1986.

[27] R. Parasuraman. Neuroergonomics: research and practice. *Theor. Issues in Ergon. Sci.*, 4(2):5–20, 2003.

[28] R. Parasuraman and M. Rizzo. *Neuroergonomics: The brain at work.* Oxford University Press, Oxford, United Kingdom, 2008.

[29] J. Polich. Updating p300: an integrative theory of p3a and p3b. *Clin Neurophysiol.*, 118(10):2128–2148, 2007.

[30] J. Polich and A. Kok. Cognitive and biological determinants of p300: an integrative review. *Biological Psychology*, 41(2):103 – 146, 1995.

[31] N. Pratt. Effects of working memory load on visual selective attention: behavioral and electrophysiological evidence. *Frontiers in Human Neuroscience*, 5:57, 2011.

[32] M. Raabe, R.M. Rutschmann, M. Schrauf, and M. Greenlee. Neural correlates of simulated driving: auditory oddball responses dependent on workload. *Foundations of Augmented Cognition*, pages 1067–1076, 2005.

[33] J. Ragland, B. Turetsky, R. Gur, F. Gunning-Dixon, T. Turner, L. Schroeder, and R. Chan. Working memory for complex figures: An fmri comparison of letter and fractal n-back tasks. *Neuropsychology*, 16:370 – 379, 2002.

[34] R. Ravizza, M. Behrmann, and J. Fiez. Right parietal contributions to verbal working memory: Spatial or executive? *Neuropsychologia*, 43:2057 – 2067, 2005.

[35] M.D. Rugg. *Encyclopedia of Neuroscience: Event-Related-Potentials (ERPs).* Academic Press, 2009.

[36] P. Ullsperger, G. Freude, and U. Erdmann. Auditory probe sensitivity to mental workload changes - an event-related potential study. *International Journal of Psychophysiology*, 40(3):101 – 9, 2001.

[37] S. Wang, J. Gwizdka, and W.A. Chaovalitwongse. Using wireless eeg signals to assess memory workload in the n-back task. *IEEE Transactions on Human-machine Systems*, pages 424–435, 2015.

[38] S. Watter, G.M. Geffen, and L. Geffen. The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, 38(3):998–1003, 2001.

[39] G.F. Wilson and C.A. Russell. Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human Factors The Journal of the Human Factors and Ergonomics Society*, 49(6):1005 – 1018, 2008.

[40] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan. Brain-comupeter interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767 – 791, 2002.

[41] F.R.A. Zijlstra. *Efficiency in Work Behavior: A Design Approach for Modern Tools.* Doctoral dissertation, Technical University of Delft, The Netherlands, 1993.