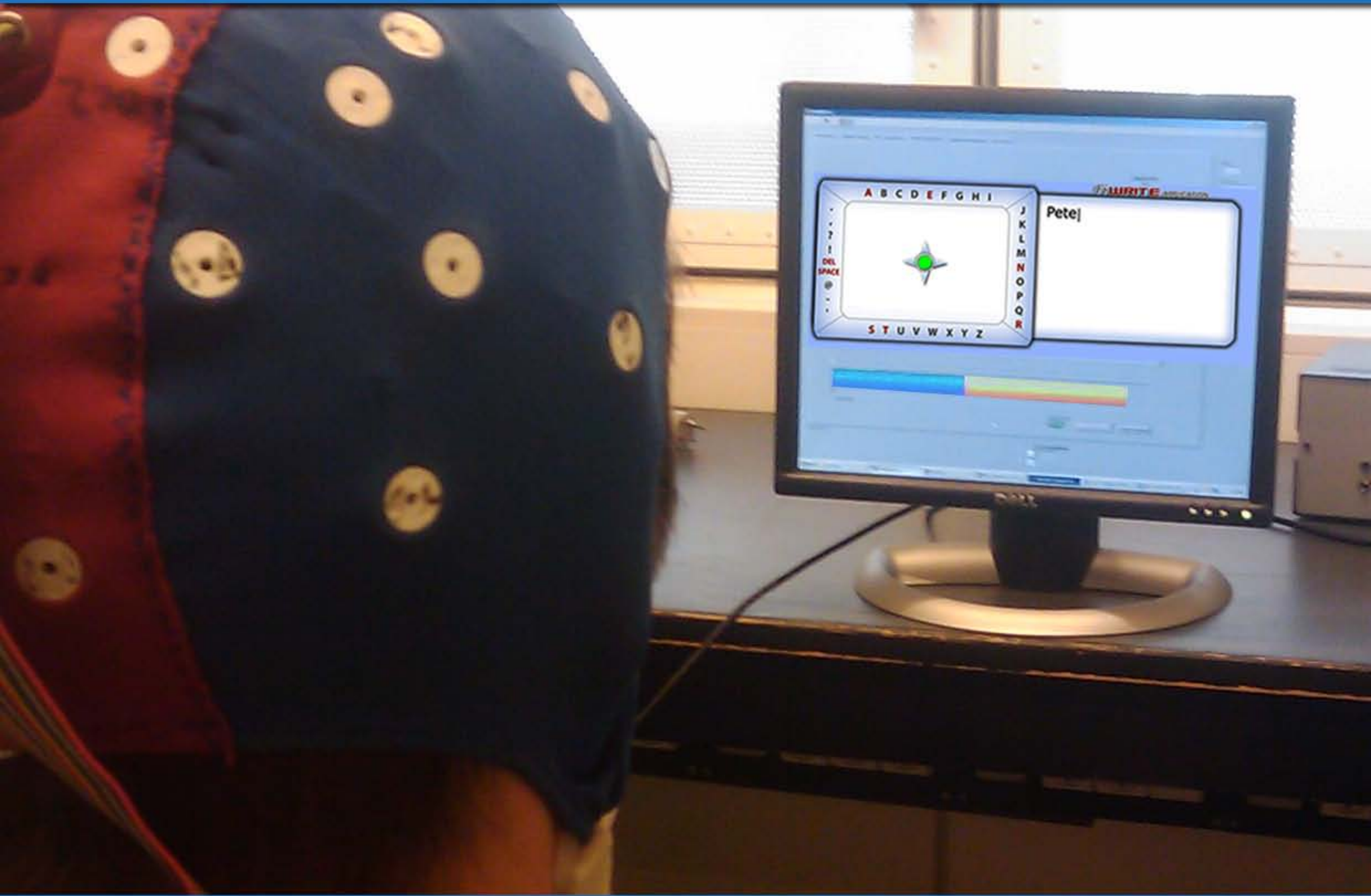
Fi-WRITE

A Text Input Technique Based on 4 Inputs For Brain-Computer Interface



Master Thesis

Sutharsan Narayanasamy (sw10)
Michael Tcherkewitch Mikkelsen (INF8)

Aalborg University - 2008
Department of Computer Science

AALBORG UNIVERSITET

Department of Computer Science

Selma Lagerlöfs vej 300

Phone 96358080

Fax 98159889

http://www.cs.aau.dk

TITLE:
Fi-WRITE - A Text Input Technique Based on 4 Inputs for Brain-Computer Interface

THEME:
Human Computer Interaction

PROJECT UNIT:
SW10 & INF8, Spring 2008.
February 1st – October 31st 2008.

PROJECT GROUP:
d614a

GROUP MEMBERS:
Michael Tcherkewitch Mikkelsen
Sutharsan Narayanasamy

PROJECT SUPERVISOR:
Jan Stage

NUMBER PRINTED: 4

NUMBER OF PAGES: 48

NUMBER OF APPENDICES: 2

SYNOPSIS
This master thesis concerns the development of a text writing technique with only 4 inputs (Fi-WRITE).

The primary motivation for this thesis comes from the need for an efficient text writing ability for people suffering from motor disorders. This text input technique therefore targets both the Brain-Computer Interface field, where inputs are given by imagination instead of muscular movements, as well as the non-BCI field.

Based on this challenge, an overall research question was designed. This question was then divided into two sub-questions, each marked by two phases which resulted in a research article. Through a number of experiments, the two articles together give an answer to our overall question. For the BCI part, we used the BCI system being developed at Aalborg University, while for the non-BCI goal we used a computer keyboard.

Our results showed that our texting technique, Fi-WRITE, was applicable in both BCI and non-BCI fields. Our best result in non-BCI usage was measured to 12.68 WPM, while in BCI we were not able to measure the efficiency.

PREFACE

This master thesis is written in the spring 2008, by an Informatics student and a Software Engineering student of the Department of Computer Science at Aalborg University.

This report gives an overview of the complete master thesis, by presenting the essential parts from the project. Our master thesis resulted in two articles, which are included in the appendix section. Since this report is based on the articles, it will have repetitions of some parts.

When the words *we* and *our* are used, they refer to the authors and *he* refers to *he/she*.

We would like to thank all the test subjects who participated in the experiments throughout the project period. We would also like to thank our supervisor, Jan Stage, for a great support in this thesis. Finally, we would like to thank the BCI-team at Aalborg University for the cooperation in the final experiment.

The Fi-WRITE application is included on the enclosed CD.

References
All the references are presented as [ ] with a number inside. The numbers represent the placement of the reference in the reference list.

Michael Tcherkewitch Mikkelsen (INF8)

Sutharsan Narayanasamy (SW10)

TABLE OF CONTENTS

1 INTRODUCTION

For text writing on computers, full-sized keyboards have been by far the most common and efficient text input devices. Standard computer keyboards are designed so each letter in the alphabet is mapped into each key on the keyboard. This allows the user to write text in high speed with the only physical requirement being to have hands. However, a keyboard is not efficient or suitable in all situations.

This could be for the following reasons

- The device which provides the text writing could be too small to include a fully-sized keyboard.
- The text input is supported by a non-keyboard device only.
- The interaction between human and computer cannot be realized using muscular movements.

In all these situations, using a keyboard will be inadequate. It is therefore important to develop a technique in which text can efficiently be entered in the situations mentioned above.

A more concrete situation where there is a need for an efficient text writing technique is in the field of Brain-Computer Interface (BCI). In BCI, the interaction between a human and a computer is obtained by the human mind and not other muscular movements. BCI has several aims, but one important motivation is helping people with severe disorders. According to WebMD [1] 1-2 of 100000 worldwide are suffering from these kind of disorders. In cases where people become totally paralyzed, the communication with their surroundings becomes significantly weakened. A text writing technique can return to these people their ability to communicate with the rest of the world.

At Aalborg University, a group of researchers (from now on referred to as AAU BCI-team) are developing a BCI system. This BCI system allows 4 inputs only. This is equivalent to using a keyboard with only 4 keys. The AAU BCI-team use non-invasive BCI in their development, and receive the electroencephalography (EEG) signals through an electro-cap, and not through a chip directly implanted into the brain, as invasive BCI does. The system with which we will be working provides online data analysis, so the information from the electro-cap can give feedback to the user while he is using it.

Thus, in this thesis, the aim is on developing a text writing technique targeting BCI. On the basis of the above mentioned problem, the overall research question for this thesis is formulated as:

Can we realize an efficient text input technique based on a limited number of inputs with a high level of usability, which is usable in both BCI and non-BCI context, and what is the efficiency of this?

To evaluate whether the technique has high level of usability, we will let the test subjects rate the usability by a scale from 1 to 9, where 1 is poor and 9 is high. This scale was chosen to give the subjects the possibility to rate neutrally. Thus, the definition:

***Usability**, is measured by a user based usability evaluation. Users will give a score from 1 to 9 in a number of questions [2] by comparing Fi-WRITE with other text writing applications. If the total average score passes 6, which is in the top third, we will rate it as high usability.*

Efficiency can be defined in many ways. In the article: *Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?* [3], the efficiency is defines as:

***Efficiency**, which is the relation between (1) the accuracy and completeness with which users achieve certain goals and (2) the resources expended in achieving them. Indicators of efficiency include task completion time and learning time.*

We will follow the same definition and use the task completion time as the only indicator of efficiency. To measure the task completion time we will use WPM (words per minute). A word is defined to have the length 4.4 characters [4].

The overall research question consists of two sub-questions, which each intend to give an in-depth answer to the overall question. The two sub-questions also define the two phases of this project: BCI and non-BCI.

1.1 RESEARCH QUESTION 1

The first research question in this thesis is based on the development part of the technique. The primary focus is designing an efficient text input technique with high usability, based on 4 inputs only. For this, the idea is to design three prototypes and evaluate them in a number of experiments and thereby find the best one with regard to usability and efficiency. All these lead to the formulation of the first research question:

Can we design a text writing technique based on a limited number of inputs, which has a high level of usability and efficiency?

The intention is to design a number of text input techniques and compare them against each other to find the overall best one. Efficiency and usability will be evaluated for the selected technique. This research question concerns realizing the text input technique, thus this goal will be attained by experiments in non-BCI environment. To evaluate the efficiency and usability we will use the same definitions as mentioned above.

1.2 RESEARCH QUESTION 2

The second phase involves Brain-Computer Interface. The desire is to evaluate the use of our technique, Fi-WRITE, in Brain-Computer Interface perspective and examine the efficiency of it. This leads to the second research question:

Is it possible to use BCI to enter characters in Fi-WRITE and what is the efficiency of this?

To evaluate the efficiency, we will examine the completion time of tasks. The same definition as mentioned earlier will be used, thus we will use the measurement unit WPM. The intention is to compare our results with similar results under the same conditions. For this evaluation we will use the BCI system being developed at Aalborg University.

2 THE RESEARCH ARTICLES

This chapter presents the two articles which form the basis of this thesis. The articles give individual answers to the corresponding research questions, while contributing to a joint answer to the overall research question for this project. The first article describes the development of the technique, while the second article puts the technique into BCI context. The articles can be found in the Appendix.

- Article 1: Fi-WRITE: A Text Input Technique Based On 4 Inputs

- Article 2: A Study of Fi-WRITE, A Text Input Technique With 4 Inputs, in Brain-Computer Interface Context

2.1 ARTICLE 1: FI-WRITE: A TEXT INPUT TECHNIQUE BASED ON 4 INPUTS

The aim of this article was to examine how to make an efficient input technique with only four inputs available. The article focused on the development and evaluation of such a technique, without having Brain-Computer Interface directly involved with it. But the goal of having the technique support BCI usage in a later stage required having some BCI requirement in mind when designing the prototypes for the final technique. One major requirement was an enforced delay time of 3.5 seconds between the inputs. The BCI system at AAU was not able to give continuous inputs. Therefore the first part of this article concerns the design phase, where three prototypes were designed and implemented. In the first phase of this article we describe the development of the three prototypes. Beside the development of the technique, the articles describe a number of experiments. In order to achieve our goals, it was necessary to perform three experiments:

1. Rotation Experiment
2. Comparison Experiment
3. Evaluation Experiment

Two of the prototypes were based on rotations, so the first experiment concerned finding a suitable rotation speed for these prototypes. The subjects, who participated in this experiment, were asked to write the same text in each rotating prototype using three different predefined speeds. Having found a suitable rotation speed, the next experiment was the comparison experiment where all three prototypes were compared. The comparison was based on task completion time, workload and an overall comparison in regards to each other. Twelve subjects were asked to write a sentence which we defined. They were introduced to the prototypes, to make them feel comfortable with the techniques, before the actual experiment was conducted. The workload was measured by using NASA TLX [5] [6], where each subject had to answer and compare different values regarding workload after testing each prototype. The overall comparison part was evaluated by using a questionnaire that was administered at the end of the experiment. The purpose was to compare each prototype in relation to each other by giving the best prototype a score of 1, the next-best a score of 2 and a score of 3 to the poorest one. All three factors (task completion time, workload and comparison) were equally important. For a combined value, the individual proportional differences (the differences in percentages) are what matters. Since, in all three factors, the lowest value is the best, the total scores are simply calculated by multiplying the three values. In this way it is only the difference in percent which matters, and the fact that the value in speed is a lot higher than the value in TLX makes no difference in the result. This calculation can only be done if no values are 0. The results are seen in Table 1.

	Task completion time (seconds)	Workload (TLX)	Overall comparison (1-3)	Total score
Prototype A	127	84,28	1,54	1098
Prototype B	154	106,61	2,29	2501
Prototype C	109	109,47	2,17	1729

Table 1: Results from prototype comparison experiment

The results in the table show that Prototype A received the lowest total score, followed by Prototype C and then Prototype B. For this reason, Prototype A was evaluated to be the overall best prototype, even though it was not the fastest one to use.

After the prototype-selection experiment, we optimized Prototype A according to the HCI principles of predictability, consistency and familiarity [2].

Finally, to evaluate the usability and to examine the efficiency of the prototype after improvement, a third experiment was conducted. In this experiment the subjects were asked to complete the given task as fast as possible, both with and without the delay, followed by an evaluation of the usability of the technique. The efficiency results, measured in WPM, showed that with the delay, the best performance was 1,64WPM and without the delay it was 12.68WPM. The usability was evaluated by a questionnaire which requested the subject to rate the usability in regards a number of HCI principles. On a scale from 1-9, the final technique, Fi-WRITE, was rated as 6.39.

Following the development, the next step was to compare the three prototypes with each other in order to select the most successful prototype. Having selected the best prototype, the next step was to develop that prototype further and make the usability higher. Finally, the article concludes with a usability evaluation.

2.2 ARTICLE 2: A STUDY OF FI-WRITE, A TEXT INPUT TECHNIQUE, IN BRAIN-COMPUTER INTERFACE CONTEXT

The second article concerns the challenges of people with severe motor disorder. In this article the aim is to examine the use of our 4-input based text writing technique, Fi-WRITE, when used with a Brain-Computer Interface, and to evaluate the efficiency of it. The article therefore consists of a laboratory experiment which examines the efficiency by having eight subjects write the same sentence. The intention was to examine the shortest task completion time when using Fi-WRITE with the BCI system, and thereby measure WPM.

BCI is still in its development stage, so an experiment involving BCI requires access to a BCI laboratory. Thus, our BCI experiment was conducted using the BCI laboratory located at Aalborg University, in cooperation with a BCI research team at the university. The communication with the BCI system could not be achieved directly, so a large effort went into making a bridge application which could manage the communication between the parts. The article describes the experiment in the laboratory following the development of this application.

Eight subjects participated in this experiment, which produced unexpected results. Only half of the eight subjects were able to control Fi-WRITE using the BCI-system. Furthermore, the four subjects who could control the system lost the control after a short time. The results clearly showed that Fi-WRITE was usable in BCI context. By usable we mean that it is possible to enter characters using the mind. But it was not possible to conclude anything regarding efficiency because the subjects could not complete the task. The accuracy rate was high, but only for the first 110 to 160 seconds; then the four subjects lost control of the system and were unable to use it further. We defined “lost control” as 20 consecutive undesired inputs. There could be several reasons for this surprising outcome, thus the final part of this articles gives some ideas on what could have affected the results. It should be noted that the BCI system at Aalborg University was still under development and our experiment was the first of its kind with their system.

3 RESEARCH METHOD

In this chapter, we describe the research method used for answering the research questions. We did four experiments in this study and the laboratory experiment method was used for each of them.

3.1 DESCRIPTION OF LABORATORY EXPERIMENTS

Laboratory experiments are characterized by taking place in artificial environments, created for the research experiment. Laboratory experiments do not necessarily have to take place in dedicated “real” laboratories, but can be conducted in various controlled environments, such as in an office, in a hallway or in a simulator. In this way it is possible to eliminate external disturbances that might affect the experiment. One of the great advantages of laboratory experiments is the opportunity to focus on specific phenomena of interest. Laboratory experiments are also highly replicable and facilitate good data collection. However, having limited relation to the real world can also be a disadvantage [7] [8].

3.2 LABORATORY EXPERIMENTS IN OUR STUDY

The four experiments in our study are divided into two categories; *the group-room experiment* and *the BCI-Lab experiment*.

THE GROUP-ROOM EXPERIMENT: EXPERIMENT 1, 2 & 3

The first three experiments in this study had the aim of 1; compare three rotating speeds in the two rotating prototypes and find the most suitable speed for each, 2; compare the three prototypes in order to find the one which was the overall best, and 3; evaluate the final prototype in regards to usability and efficiency. All three experiments had the goal of exploring the performances on one or more prototypes in non-BCI context.

Due to the final goal of evaluating Fi-WRITE in a BCI system, we took certain precautions to make sure the prototypes were tested in circumstances as close to the BCI circumstances as possible. Besides only having four inputs, we also implemented a delay within the system and made the input method such that the users would have a lower accuracy, to come closer to the data provided by the AAU BCI-team. To do so, we used a method from a previous study [9].

EXPERIMENT 1

The purpose of the first experiment was to compare three rotating speeds in each of the two rotating prototypes and find the most suitable speed for each. Before the actual experiment, a mini-experiment was conducted with two subjects. Those two subjects were asked to define three rotation speeds which were to be used in the actual experiment.

Eight male students from the Department of Computer Science at Aalborg University participated as subjects in this experiment. During this experiment, we discovered which of the three rotation speeds the subjects preferred for each of the two prototypes. The results from this evaluation enabled us to do the next experiment.

EXPERIMENT 2

The purpose of the second experiment was to compare the three prototypes in order to find the best prototype. The experiment was conducted with eight male students from the Department of Computer Science at Aalborg University as participants. The selection of the prototype was based on three factors: task completion time, workload and user evaluation.

EXPERIMENT 3

The purpose of the third experiment was to examine the final prototype in order to evaluate its efficiency and usability. The experiment was conducted with four male students from the Department of Computer Science at Aalborg University as participants.

Furthermore, the efficiency of the prototype without the delay was measured in this experiment. The evaluation of usability was based on the user feedback, while the efficiency was based on the task completion time.

THE BCI-LAB EXPERIMENT: EXPERIMENT 4

The purpose of this experiment was to evaluate the efficiency of Fi-WRITE when using it with BCI. The environment setup in this experiment was very different from in the first experiment, even though they both used laboratory experiment. This experiment took place in a BCI laboratory, using BCI equipments.

The eight subjects who participated in this experiment had no prior experience using BCI. Before the actual experiment, the subjects were asked to do a training session in order to calibrate the BCI system. The training was based on defining four pressure levels by imagination. Each pressure level had 25 trials. Following the training, each subject was asked to perform the experimental task. The task was the same as in the other experiments, but this time using imagination instead.

3.3 LIMITATION WITH OUR EXPERIMENTS

The results in experiment 1 were limited by the number of speeds used for evaluation as well as the lack of providing the subjects the opportunity to choose any speed. Experiment 2 was limited by the number of prototypes provided for evaluation. The usability results from experiment 3 are limited by the fact that the prototype was not compared to other systems. In experiment 4, the results might be limited by lack of experience from the subjects in using brain interaction. Furthermore, the results in this experiment are limited by the short training time before the actual experiment.

The number of participants is a limitation which is asserted in all three experiments. Additionally, only students from Department of Computer Science have participated in the four experiments.

4 CONCLUSION

The purpose of this thesis was to develop a text input technique using 4 inputs. The technique should have a high level of efficiency and usability while being usable in the field of Brain-Computer Interface as well as non-BCI fields. Based on these goals, an overall research question for this master thesis was defined:

Can we realize an efficient text input technique based on a limited number of inputs with a high level of usability, which is usable in both BCI and non-BCI context, and what is the efficiency of this?

The answer to this overall question was obtained by dividing the question into two sub-questions. In the following section, we review the two research question and their results. Based on these results, we can answer the overall question.

4.1 RESEARCH QUESTION 1:

Can we design a text writing technique based on a limited number of inputs which has a high level of usability and efficiency?

To answer this question, we conducted three experiments, which are described in article 1. The results from the first experiment were only used for the comparison of our three prototypes in experiment 2. The comparison was targeted against a selection of the best prototype, based on three factors: efficiency (WPM), workload (NASA-TLX) and an overall user evaluation. The results from experiment 2 enabled us to evaluate the usability and efficiency of the final technique, which was the aim of experiment 3.

To get an answer to the usability part of the question, the test subjects were asked to rate the usability of our final technique, Fi-WRITE, given a set of usability questions. On a scale from 1 (being lowest) to 9 (being highest), the overall usability of Fi-WRITE was rated as 6.39. From this we can conclude that the Fi-WRITE technique was rated to be in the top 70% of the scale regarding usability. Thus, we can conclude, based on our earlier definition of high usability, that the high level of usability condition was fulfilled.

The question of efficiency was also answered in experiment 3. The purpose of measuring efficiency was to see how fast the given task could be completed using our technique. This was measured in words per minute and compared to the results described in the related work section of article 1. The experiment showed that our fastest subject performed 12.68 WPM without the delay. When comparing this result with the results from our related studies with non-BCI techniques, Fi-WRITE was ranked among the best performances. Therefore, we can conclude that Fi-WRITE has a high level of efficiency.

4.2 RESEARCH QUESTION 2:

Is it possible to use BCI to enter characters in Fi-WRITE and what is the efficiency of this?

The answer for this question is obtained from experiment 4, which is described in article 2. In this article, we evaluate whether it is possible to use our technique with a BCI-system and measure its efficiency in

WPM. Having concluded on Fi-WRITE’s efficiency in non-BCI usage, it was interesting to evaluate its efficiency when controlling the input with the mind. The results from experiment 4, which was conducted in the BCI-laboratory at Aalborg University, showed that only half of the eight subjects were able to control Fi-WRITE using their minds. We could characterize the eight subjects into two groups; *able to use* and *unable to use*. *Unable to use* was defined by 20 consecutive undesired inputs.

The *able to use* group did manage to write desired characters using their minds, but in a limited time only. All four subjects from this group lost control after having entered 2 to 4 letters. Like *unable to use,* we defined ‘lost control’ as 20 consecutive undesired inputs.

The results showed that it was possible to enter characters using the imagination, thus we can conclude that it was possible to write characters in Fi-WRITE using BCI. The efficiency, on the contrary, could not be concluded because none of the subjects were able to complete the task. But an estimating of the efficiency in terms of the four ’able to use’ subjects only and limited time where they were able to use it, was calculated to be 0,33WPM in average, and 0,39WPM for the one who entered 4 correct characters. This is a weak performance compared to other BCI text writers. Even our fastest subjects WPM were between 13% and 77% slower than the BCI text writers in our related word.

4.3 THE OVERALL RESEARCH QUESTION:

Can we realize an efficient text input technique based on a limited number of inputs with a high level of usability, which is usable in both BCI and non-BCI context, and what is the efficiency of this?

We can conclude that it was possible to realize a text input technique with a limited number of inputs compared to the keyboard. We successfully designed and developed a text writing technique using 4 inputs only. Besides this, the goal of implementing high level of usability and efficiency into the technique can also be concluded to be successful according to our definitions. The technique should be useable in both BCI and non-BCI context. Based on the results from the experiments, Fi-WRITE could be used with both a keyboard and with BCI. Thus, this part can also be concluded as successful. Finally, the overall question asks for the efficiency when used in these contexts. In non-BCI use, we reached a maximum speed of 12.68 WPM among the novice users, which compared to the results from our related studies was better than all the other text writers. According to our definition of efficiency, it was not possible to conclude on the efficiency in the BCI context.

4.4 LIMITATIONS

The results from this study, especially in experiment 3 where the usability and efficiency in non-BCI use was evaluated, are limited by the number of subjects used. The selection of subject group and the number of prototypes developed could have affected the results as well. Additionally, the results from experiment 4 were limited by the training time allowed in the experiment.

4.5 FUTURE WORK

As a continuation of this study, it would be interesting to do another experiment in the BCI laboratory with significantly increased training time. In this context, one or more subjects could be trained over a period of time to examine the difference. Another interesting idea would be to examine whether a greater number of test subject would have changed our conclusion on usability and efficiency in non-BCI contexts. Finally, we would be interested in an evaluation of Fi-WRITE implementation in mobile phones.

5 REFERENCES

[1] WebMD, Amyotrophic Lateral Sclerosis (ALS) - Topic Overview, http://www.webmd.com/brain/tc/amyotrophic-lateral-sclerosis-als-topic-overview
[2] Dix, A., Finlay, J. E., Abowd, G. E., Beale, R. (2004) Human Computer Interaction 3rd edition, Harlow : Pearson Education
[3] Frøkjær, E., Hertzum, M., Hornbæk, K. (2000) Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?, CHI '2000
[4] Celikel, E., Dalkilic, M. E., Dalkilic, G. (2005) Word-based fixed and flexible list compression, International Symposium on Computer and Information Sciences
[5] Hart, S. G., Staveland, L. E. (1988) Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, NASA Ames Research Centre
[6] Human Performance Research Group (1987) NASA TASK LOAD INDEX (TLX), NASA Ames Research Centre
[7] Wynekoop, J.L., Conger, S.A. (1990) A Review of Computer Aided Software Engineering Research Methods. Proceedings of the IFIP TC8 WG 8.2 Working Conference on The Information Systems Research Arena of The 90's , Copenhagen, Denmark
[8] Kjeldskov, J., Graham, C. (2003) A Review of Mobile HCI Research Methods, University of Melbourne
[9] Mikkelsen, M. T., Narayanasamy, S. (2008) Designing a Feedback Application for Brain Computer Interface, Aalborg University

6 APPENDICES

ARTICLE 1:

FI-WRITE: A TEXT INPUT TECHNIQUE BASED ON 4 INPUTS

Fi-WRITE: A Text Input Technique Based On 4 Inputs

Michael Tcherkewitch Mikkelsen
tm@cs.aau.dk

Sutharsan Narayanasamy
qmar@cs.aau.dk

ABSTRACT

The purpose of this paper is to describe the development of a text writing technique, based on 4 inputs. Our desire was to make the technique have a high level of usability and efficiency and thereby compare it to other related techniques. For this reason, three text writing techniques were developed and compared in an experiment to select the best of the three techniques. Based on a set of factors the most suitable technique was selected and then optimized, on basis of HCI guidelines. The results showed that the system's usability was high. The comparison to other techniques showed that our technique had a higher word per minute ratio than all comparable sources. Our fastest subject reached 12.68 WPM.

INTRODUCTION

One of the most important tasks that a computer is useful for is writing. In the early days writing was usually done using a hand tool such as a pencil or a pen, but with time electronic text writing has become a method used by most people. Whether it is for small notes, e-mailing or academic reports, text writing has taken its place among many people's daily tasks.

The Human Computer Interaction (HCI) device, the computer keyboard, makes the writing task much easier and straight forward since each key on a typical keyboard represents each letter in the alphabet. Thus, by having a keyboard in front, the writing on computers does not pose any difficulties or major challenges. But there are many cases where a fully sized computer keyboard and the applications designed for keyboard usage are unsuitable. Among those are Brain Computer Interface (BCI), SMS writing and writing text on a Game console using a game controller. In these cases the use of a computer keyboard is not an option, but text writing still might be.

BCI is an example where the keyboard becomes completely obsolete. In BCI the interaction between the human and computer is dependent on the human's brain instead of muscular movements. In other words, the application gets its input from the brain as if it were given by the keyboard. But the major difference, besides using brains instead of muscles, is that in BCI the outputs are minimal, often 2 or 4, such that text writing applications designed for keyboards become ineffective or completely unusable.

Mobile phone technology is a field where the speed of the text writing as well as the size of the input mechanism are both of the essence. Mobile devices are expected to get smaller and smaller with more and better features while still maintaining the ability to quickly write text messages. There are several new text writing methods, for example *TiltText [1]* and *ChordTap [2]*. While these text writing techniques both have a high word per minute ratio, they still require all 12 keys on the mobile phone for input. These methods optimize the writing speed, but as with many others, have not been able to decrease the number of input possibilities, and therefore the number of keys, on the mobile phone.

Another example is the gaming field using game consoles. Game consoles, such as PlayStation and X-Box, come with game controllers. Even though these game controllers have a fairly high number of inputs in regards to what's necessary for playing games, they are not designed for text writing, like the computer keyboard. So writing text using game controllers is done with ineffective mechanisms [3]. The most common way is a simple 5-input based input technique, where a marker is navigated around an onscreen keyboard, using up, down, left, right and a select button[3].

In this study the aim is to develop a text writing technique that can later be adapted to BCI context. Our primary goal is to examine the efficiency and usability in the technique. Based on these goals, we defined the research question for this study:

Can we design a text writing technique based on a limited number of inputs, which has a high level of usability and efficiency?

To evaluate whether the technique has high level of usability, we will let the test subjects rate the usability by a scale from 1 to 9, where 1 is poor and 9 is excellent. This scale was chosen to give the subjects the possibly to rate neutrally. Thus, the definition:

Usability*, is measured by a user based usability evaluation. Users will give a score from 1 to 9 in a number of questions [4] by comparing Fi-WRITE with other text writing application. If the total average score passes 6, which is in the top third, we will rate it as high usability.*

Efficiency can be defined in many ways. In the article *Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?* [5], the efficiency is defines as:

Efficiency*, which is the relation between (1) the accuracy and completeness with which users achieve certain goals and (2) the resources expended in achieving them. Indicators of efficiency include task completion time and learning time.*

We will follow the same definition and use the task completion time as the only indicator of efficiency. To measure the task completion time we will use WPM (words per minute).

This paper therefore only presents the development part and testing in non-BCI environments. Three techniques will be designed and the most efficient one will be further developed as the final version. In order to test the efficiency of the techniques, we have implemented them into flash applications.

RELATED WORK

Previous studies have examined text writing techniques, with the main focus on the speed of the technique. Here they are split into two groups, BCI and non-BCI techniques.

BCI

BCI has become an interesting and important field and several BCI applications have been developed in recent times.

In the article Berlin Brain-Computer Interface [6] a study of several BCI text writers was investigated. This study evaluated offline and online signal processing, as well as both invasive and non-invasive BCI. Invasive BCI means that the electrodes measuring the brain signals are inserted into the brain, while in non-invasive BCI the electrodes are outside the skin. In the online, non-invasive BCI a speed of two to three letters per minute was the fastest a person could write. This research focused a lot on keeping a high level of HCI. Another method also described in this article has a writing speed of three letters per minute, though this requires several months of training before it can be achieved.

In another study [7], in which no emphasis was placed on HCI but only on the speed on the BCI writing, the fastest speed was 1.62 characters per minute. Two out of three subjects could keep this average speed over five words. These numbers were from online, non-invasive BCI. The focus this time was effectiveness, not efficiency, so the errors and the selections to delete them did not count in the completion time. This article evaluated the text writer Hex-o-spell.

The final study [8] described a very simple text writer, in which the top speed was three letters per minute. However in some cases only 0.15 letters per minute was achieved. While writing, the user had to make four to five selections in order to write a letter, and, errors were often made. With only a 65-90% accuracy rate, in some cases the user took almost 7 minutes to write a letter.

NON-BCI

The focus of this study is developing a text writing application with four inputs. Several applications and methods exist with the aim of reducing the number of keys used for interaction with applications. Most of these applications are found in the mobile device environment.

The current generation of mobile phones typically has a numeric keypad with 12 keys where multiple letters are mapped to a single key. The two most widely used methods for entering text using a numeric keypad are the multi-tap method ("this" = 8 44 444 7777) and the Tegic Communications' T9 method ("this" = 8447). Novice users of these methods achieve text entry rates of 5-10 words per minute [9].

Some alternatives to the T9 technology for writing on mobile phones have been developed; one of them is the TiltText [1]. The TiltText works like multi tap when writing; you have to decide on every letter. But instead of clicking a button several times, you simply tilt the phone. The censors in the phone detect which way the phone is tilted and based on that writes a specific letter. Novice users write eight WPM using this technique.

Another technique is the ChordTap [2]. In the ChordTap three chord keys are mounted on the phone. While writing the user pushes the chords to determine which letter to write, on the key which gets pressed, this technique is a lot like the TiltText, though the methods are very different. Novice users can write 10 words per minute using this.

Several other writing technique have been developed, simple ones like the a 3 key text entry [12], where the user moves back and forth in one line of letters and can click select. The slightly more sophisticated, like the 5 key text writer, where the user moves in a 2D plan, and can move up, down, left, right and select. Finally more complex ones like the 4-key EgdeWrite, where the user clicks 4 buttons in a certain order to get the desired letter. The order matches the letter as much as possible, for example: clicking bottom left, top left, bottom right and top right forms an N and writes the letter n. Novice user can write up to 12.5 WPM using EgdeWrite.

AN OVERVIEW OF TEXT WRITERS

Table 1 gives an overview of the reviewed text writing techniques. Here are the number of key inputs (keys), keystrokes per character (KSPC) and word per minute (WPM) listed.

	Keys	KSPC	WPM
Non-BCI			
Multi tab on cell phones [10]	9	2,13	7.3
T9 on cell phones [11]	11	1.0072 [13]	9
3 key text entry [12]	3	10,53[1]	9,10
5 key text entry [12]	5	3,13	10.62
Edge Write [12]	4	3,52	12.5
ChordTap [2]	12	2	10
TiltText [1]	9	1	8
BCI			
Spelling device (BCI) [6]	2	6-7	0.45
Hex-o-spell (BCI) [7]	1	2	1.73
Spelling device (BCI) [8]	2	6-7	0.68

Table 1: An overview of text input techniques

DESIGN PROCESS

The aim of this study was to develop a text writing technique only. Testing the technique in practical use required implementing the technique into an application. This section describes three different prototype applications, each of which includes a technique for text writing. The prototypes acted as candidates for the final application. Since the vision for the technique, thus also for the application, was to support compatibility with the AAU BCI-team's equipment, some requirements needed to be met when designing the prototypes.

[1] They use a technique called *key-repeat*

REQUIREMENTS:

- ***4 inputs only:*** In order to get the application working with AAU BCI, the application must be fully functional with four different inputs only. This means that the application must not require a fifth input to write a text.
- ***Input delay:*** The application must support insertion of a 3.5 seconds delay time between each input, which means that continuous input is impossible.

Besides these major requirements from the AAU BCI-team, we specified the following requirements:

- ***A-Z only**:* The application will support the English language only, which includes the letters A to Z. There will be no support for writing numerals in the application.
- ***At least 5 symbols and Space and Delete*:** The application should have at least seven commonly used symbols. Furthermore an option for deleting mistakes and separation of words with regards to spacing should be available.
- ***Frequently Used Letters*:** Users should be able to enter more frequently used letters faster than rarely used letters [14].
- ***Low KSPC*:** Keystroke per character should be kept as minimal as possible.
- ***Delay Indication:*** An indication of the time where input is not possible should be available.

Furthermore, because single letters are easier to read if they are uppercase and words are easier to read if they are lowercase, all the individual letters should be displayed in uppercase for selection and displayed in lowercase in the text box. Starting letters and letters after periods, question marks and exclamation marks should also be in uppercase [15].

PROTOTYPES

Once the requirements were determined, we were ready to design the prototypes. This section describes the development of the three prototypes, named Prototype A, Prototype B and Prototype C for easy reference. Most of the programming languages will satisfy our needs but Flash is a great tool for developing graphical applications. For this reason we decided to use Flash 8 (Action Scripts 2.0 as the programming language) for the implementation of the prototypes.

PROTOTYPE A

In this prototype all the characters are grouped and distributed into the four edges of a square: left, right, top and bottom.

We added a circle in each corner as an indicator to the user that the program is ready. The circles were colored either red or green. These colors were selected because they are often used in other places where green means go and red means stop, for example in traffic lights. In this prototype the order of the letters also followed the Familiarity guideline [4]. The order of the letters corresponds with the alphabet, both in the startup screen and in the sub-screens. The letters are placed clockwise according to the alphabet, except in the bottom box where they go from left to right, since that is the direction in which the English language is read. The same goes for the right box in the sub-screens; there it is from the top down.

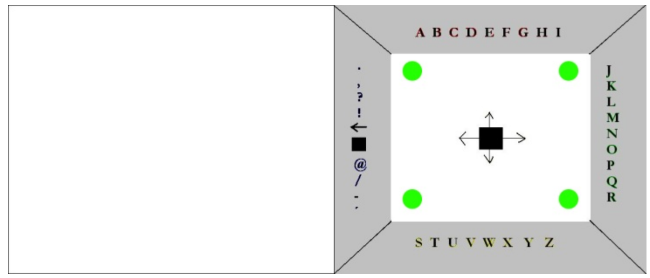


Figure 1: Prototype A – Startup screen.

As seen in Figure 1, the letters "A" and "E" are placed alphabetically clockwise from the top to bottom edge, while the symbols, space and delete are placed on the left side. A black square is placed in the middle, with an indication of the movable directions. There is a text field on the left side in the screen where the chosen characters appear.

A character selection is performed by moving the black square in the center to the corresponding edge. Choosing an edge leads to a second screen, as seen in Figure 2.

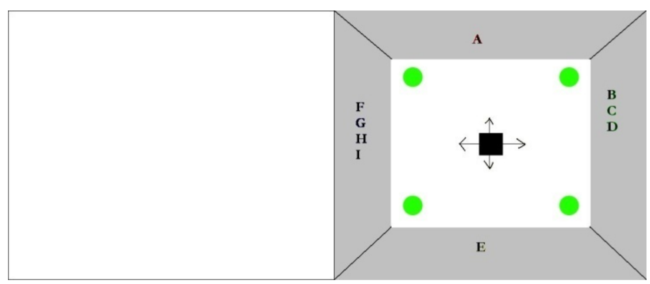


Figure 2: Prototype A – Example of second screen, where two letters are placed separately.

In this screen, all the characters from the chosen edge are distributed into four new edges. Again, starting from the top, the characters are distributed alphabetically. It should be noted that the letter “A” and letter “E” are placed alone. This means that choosing one of these edges will enter the corresponding letter on the screen. This is always the case in the second screen; two letters are placed alone, while the rest are in a group of three or four.

Choosing an edge with more than one character will lead to a new screen with four new edges.

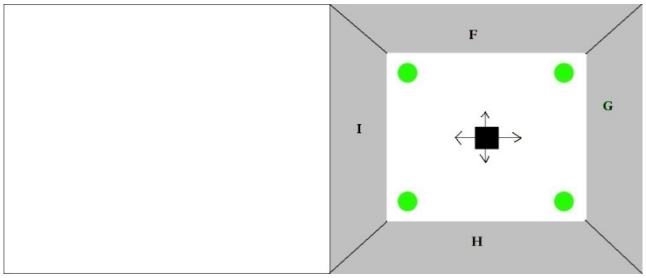


Figure 3: Prototype A – Example of third screen, where each letter is attached to each edge.

Because the maximum number of letters attached to each edge is reduced to four in the second screen, the next screen will have an edge for each remaining letter, which is seen in Figure 3. However, if an edge with only three letters is chosen, each letter will still have its own edge and a "back" option will appear in the left edge.

Figure 4 shows the screen after choosing an edge with one letter; the prototype returns to its startup screen while the chosen letter, in this example a "G", is entered in the text field to the left.

The symbols are reached in exactly the same way as letters – space and delete are reached with 2 steps.

Any character is reachable in either two or three steps in this prototype. The characters that are reachable in only two steps are not randomly chosen. The most frequently used characters in each edge are placed so they are reachable in two steps [14].

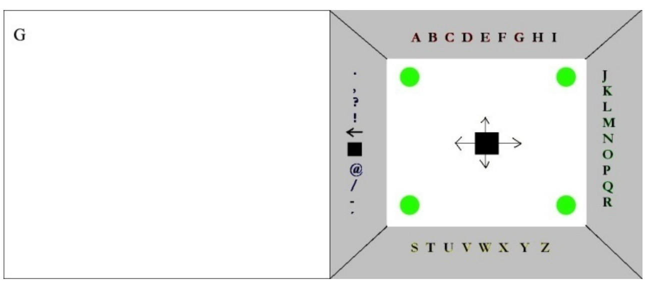


Figure 4: Prototype A – A letter has been entered.

Characters

This prototype provides a total of 36 characters including 26 letters (A-Z), delete (back arrow) and space (black square) options and the following eight symbols:

Period (.) comma (,) question mark (?) exclamation mark (!) at sign (@) slash (/) dash (-) apostrophe (')

PROTOTYPE B

This prototype is designed as a flower with seven leaves. The characters are distributed into six of the leaves, with each leaf holding six characters, as seen in Figure 5. The remaining leaf does not hold anything as it is meant to be used in the second screen only, or is reserved for reaching delete more quickly. There are four arrows in the center in the directions left, right, up and down, which are not rotating. These arrows are used for selection of a leaf.

This prototype also has the letters ordered alphabetically clockwise, like Prototype A. Even though this prototype rotates, we decided to put a red/green indicator in it, like in Prototype A. This was done so the subjects could try a rotation prototype both with and without an indicator, since there is not an indicator in Prototype C.

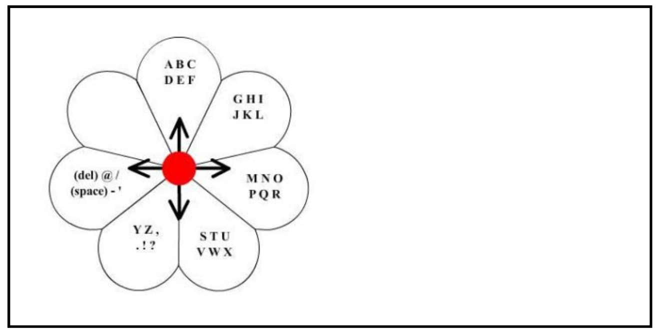


Figure 5: Prototype B – Startup screen.

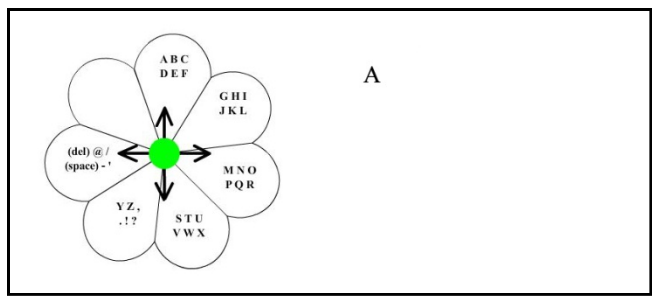


Figure 6: Prototype B – A letter has been entered.

When the prototype is started, the leaves rotate clockwise. Choosing a character requires first choosing the leaf where it is placed. With four arrows, each leaf is reachable in either left, top, right or down position, thus if one position misses it can be chosen by the following arrow instead. Choosing a leaf leads to the next screen where the six characters from the chosen leaf are distributed into a new flower with seven leaves. In this screen the seventh leaf includes a back option. Choosing a leaf in this screen will enter the corresponding character in the text field.

In this prototype a character is always written using only two steps.

Characters

This prototype provides a total of 36 characters including 26 letters (A-Z), delete and space options and the following eight symbols:

Period (.) comma (,) question mark (?) exclamation mark (!) at sign (@) slash (/) dash (-) apostrophe (')

PROTOTYPE C

Figure 7 shows a screenshot of Prototype C. In this prototype there are four rotating rings in four sizes. The characters are distributed into the rings with the largest ring containing 16 characters, the next ring 12, the next 8 and the smallest ring containing 4 characters. The most commonly used letters and space are placed in the smallest rings.

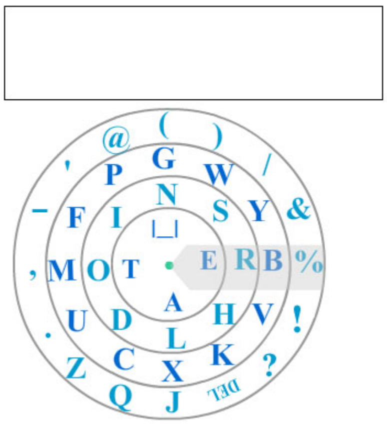


Figure 7: Prototype C - Startup screen

A gray horizontal field is seen in the middle right of the startup screen. This field acts as a marking field for the characters. The rings rotate at different speeds so that exactly one character from each ring will be in the marked area at one time. The rings are represented by the 1, 2, 3 and 4 keys on a computer keyboard, with 1 representing the smallest ring and 4 representing the largest. The text field is placed above the rings.

A character is chosen by pressing 1, 2, 3 or 4 when the desired character is in the gray field. For example if a letter from ring 2 needs to be written, the 2 key should be pressed when the desired letter is in the gray field.

Figure 8 shows an example where the 1 key is pressed. The letter "T" from ring 1 is entered in the text field.

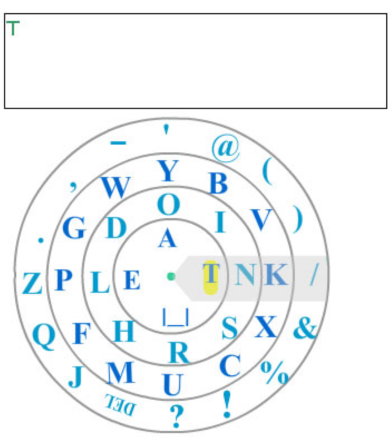


Figure 8: Prototype C - Entering a letter

Characters

This prototype provides a total of 40 characters including 26 letters (A-Z), delete and space options and the following twelve symbols:

Period (.) comma (,) question mark (?) exclamation mark (!) at sign (@) slash (/) dash (-) apostrophe (') and (&) percent (%) left parenthesis (( ) right parenthesis ( ))

THEORETICAL PERFORMANCES

This section presents the theoretical performances for all the three prototypes. The theoretical performance, which is depended on the delay time, is measured in word per minute. Different delays give different results and also affect which prototype is theoretically the fastest. Figure 9 shows a graphical illustration of the words per minute for all three prototypes with different delays, and 4.4 letters per word [16].

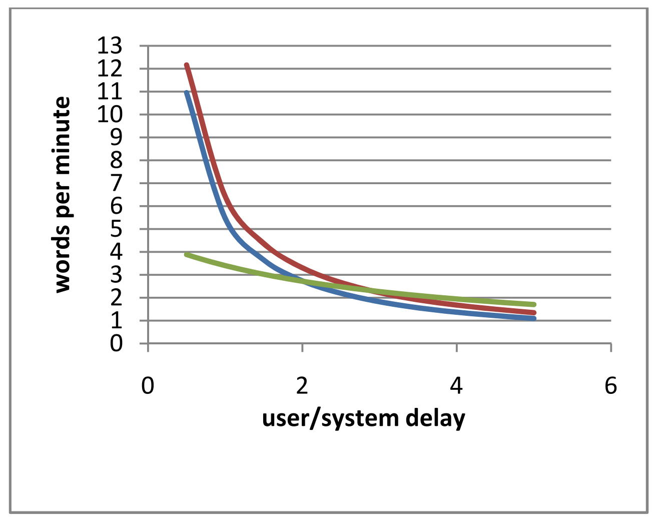


Figure 9: Words per minute for the three prototypes based on variable delay

This was calculated without taking errors or the users' reaction- and planning time into account. This was basically done this way for each letter: $\frac{60}{frequence(steps+delay)\cdot 4.4}$. This is simple in prototypes A and B, where the prototype goes back to the startup screen after each letter. But in prototype C it continues from the current position. So we just took all possible positions of the prototype into account and calculated the average of the number of steps from all positions to each letter.

However, this calculation is not 100% correct. To calculate a more accurate word per minute value for Prototype C, we would need to take into account the next letter possibility values for all letters and also for space. To do this we would have to find out the probability that a word starts with an "a", with a "b", etc. and the probability that a word ends with an "a", with a "b", etc. Even doing that, we would not be 100% correct in our calculations, since there are several possible positions of the prototype with all letters in the three inner rings. Calculating this would be impossible to do theoretically since it depends on the users' reaction time, error rate and when the error occur.

METHODS

This section describes the experiments carried out in this study. The experiments were:

Rotation Experiment: This first experiment included only the rotating prototypes, Prototype B and Prototype C. The purpose of this experiment was to find appropriate rotation speeds.

Comparison Experiment: In this experiment the aim was to compare the three prototypes to select the prototype with the best results.

Evaluation Experiment: This third experiment was conducted to evaluate the selected prototype in order to meet high level of usability and efficiency.

First, the rotation speeds in the two rotating prototypes, Prototype B and Prototype C, were tested in the two rotating prototypes. This experiment tested the prototypes using three different speeds each in order to find the most appropriate speed. Before the actual experiment, the three rotation speeds needed to be defined. To do that, we conducted a preliminary mini-experiment with two subjects to find the three rotation speeds which would be tested in the rotation experiment. The rotation experiment was followed up by a comparison experiment where the three prototypes were compared to each other. This experiment intended to find the most effective prototype of the three prototypes. The three prototypes were compared in different aspects, and both the performance and user feedback affected the selection of the winner. When the best prototype was found, it was optimized to meet the final requirements and some HCI rules to make the application user friendly. Finally, a third experiment was conducted to further improve the text writing application.

In the following section, we describe the conditions, which were similar for all the three experiments.. The method laboratory experiment [17] has been used in all experiments.

INPUT METHOD

The input method used in the experiments is the same as described in P9 [18]. A keyboard was used to imitate the input with delays equaling inputs from the BCI equipment.

TASK

In the first two experiments the subject was instructed to write the sentence: peter is sleeping. The sentence we used should be as close to the theoretical average as possible. With 4.4 letters per word, and two spaces, the subjects should have written 15.2 characters. We found that the number of steps to an average letter is 2.49 steps, based on the most frequent letters and the steps to each letter. The number of inputs in the task should therefore have been close to 38. The sentence we chose to use has 40 inputs, which is 5.3% more than an average sentence. We chose to use it anyway, since it was the best complete sentence we could find that did not include any complicated words.

SETTING

The testing area was a single room setup [19]. Our group room was used for this purpose. With a single room setup, we could get a good first hand impression of what the participant is doing. The setting is depicted in Figure 10.

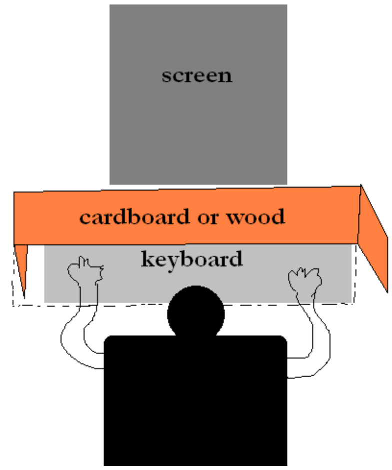


Figure 10: Experiment setting

PARTICIPANTS

The subjects were all male students from the Computer Science department at Aalborg University. None of the subjects had experience in similar text writing techniques like the input method used in the experiment.

TEST MONITORS

Both the authors of this paper filled the role of test monitor.

MATERIALS

The following materials were used in the experiments: a desktop computer running the prototypes, a laptop computer with the questionnaire, a video camera, two stacks of paper and a metal sheet.

PROCEDURE

Upon entering the test area, the subject was placed in front of a computer. An introduction was given to the subject before each experiment. The introduction was intended to generally describe how the experiment would be conducted, the purpose of the experiment, how to give inputs to the prototype and how the prototype functioned.

DATA COLLECTION

During the test we used a stopwatch to record the completion times for each prototype for each subject. We also recorded everything going on with a video camera. A questionnaire was used to get feedback from the subjects on user experience.

ROTATION EXPERIMENT

This first experiment was mainly conducted as a comparison test, although a small part of the test was exploratory [19]. In this experiment, three rotation speeds were tested and compared with regards to speed and stress. The following sections will describe all the aspects of the rotation experiment which was conducted using Prototype B and Prototype C.

SPEED SELECTION

Before the rotation experiment was to be conducted, we needed to define three speeds. To do this, we used two subjects who were also used in the input method test [18].

The selection was conducted in an unstructured manner, to get the highest level of accuracy. Because we only used two persons in the test and the subjects already knew each other, we used both subjects at the same time. The subjects were introduced to our two rotating prototypes, one at a time.

The subjects were allowed to change the rotation speed of the prototype with which they were working, to try out all the different speeds/. The goal for the two subjects was to decide on three speeds each within a manageable range: a speed they considered the best, a speed that was a bit slower, but still workable and a speed that was a bit faster but still usable. Once the two subjects each decided on the three speeds, the three averages of the two subjects' choices were selected.

TASK

The task in the experiment was to write a simple sentence defined by us. The sentence was:

*"**peter is sleeping**"*.

The subject was asked to write this sentence in three different speeds in each prototype.

The interaction with the subjects during the tests was kept at an absolute minimum. The only time we purposely interacted with a subject during the test was if they made a spelling error and did not notice it. After the second character following the error, the test monitor told the subject to check the spelling.

PARTICIPANTS

Twelve students participated as subjects in this test. We used Latin squares [20] to determine the order in which the subject would try the rotation speeds. .

PROCEDURE

The experiment was split into three phases where the last two phases included three sessions each:

1. Introduction session
2. Prototype X
    - Introduction session
    - Task session
    - Feedback session
3. Prototype Y
    - Introduction session
    - Task session
    - Feedback session

In the introduction phase, the subject was given an overall introduction to the experiment and the purpose of the experiment. He was also introduced to the input method.

Both the prototype phases were divided into three sessions. The first session included an introduction to the prototypes where a non-rotating version of the prototype was presented. The second session included the actual task, where the sentence was to be written using three different speeds. When this was completed, a feedback session with a questionnaire was held. The feedback session was intended to evaluate the user experience of the experiment.

DATA COLLECTION

During the test, both the subjects and the screens of the computers on which they were working were recorded by a camera. Later the recording was burned onto DVDs for further studying.

ROTATION EXPERIMENT RESULTS

The task completion time and stress factor for each prototype is analyzes in this section. Table 2 shows the average time for completing the tasks in each of the three rotation speeds in the two prototypes, based on the twelve subjects.

Prototype	Speed	Task completion time
B	1	191 seconds
	2	188 seconds
	3	198 seconds
C	1	128 seconds
	2	111 seconds
	3	113 seconds

Table 2: The average task completion time

Table 3 shows the average stress factor for each speed in the two prototypes, based on the twelve subjects.

Prototype	Speed	Stress factor (1-5 scale):
B	1	1,85
	2	3,00
	3	3,00
C	1	2,17
	2	2,25
	3	2,42

Table 3: The average stress factor

Since both the stress factor and the time to finish the task are each judged to be better as the values become smaller, a multiplication of the two values gives a joined score for the three speeds in both prototypes. This is shown in Table 4.

Prototype	Speed	Score:
B	1	350
	2	564
	3	595
C	1	278
	2	249
	3	273

Table 4: Joined speed score

From Table 4 we can see that speed 3 was evaluated as the most suitable speed in Prototype B. And speed 3 as the most suitable speed in Prototype C.

COMPARISON EXPERIMENT

In this experiment three prototypes were tested and evaluated, based on different factors, with the purpose of finding out which prototype was the best of the three.

The following section describes the specific aspects of the comparison test.

TASK

The task in the experiment was to write the same simple sentence used in the rotation experiment. The sentence used was:

*“**peter is sleeping**”*.

The subject was asked to write this sentence in three different speeds in each prototype.

The only direct interaction with the subject during the test was following an unchecked spelling error. If the subject did not notice the error, the test monitor told the subject to check the spelling.

PARTICIPANTS

Twelve students participated as subjects in this test. We used Latin squares [20] to determine the order in which the subject would use the prototypes.

PROCEDURE

This experiment was split into five phases where the three prototype phases included four sessions each:

1. Introduction
2. Prototype X
    - Introduction session
    - Training session
    - Task session
    - TLX session
3. Prototype Y
    - Introduction session
    - Training session
    - Task session
    - TLX session
4. Prototype Z
    - Introduction session
    - Training session
    - Task session
    - TLX session
5. Feedback

In the introduction phase the subject was given an overall introduction to the experiment and the purpose of it. He was also introduced to the input method and NASA-TLX [21] [22]. In the prototype phases, the subject was first introduced to the actual prototype followed by a quick training session of one minute to give an idea of how the prototype works. In the next session, the task session, the subject was asked to write the sentence. In the final session for each prototype, a NASA-TLX test was given to the subject to evaluate the workload level of each prototype. In the last phase, a questionnaire was given to evaluate the user experience.

DATA COLLECTION

During the test we used a stopwatch to record the completion times for each prototype for each subject. In case something went wrong with the stopwatch, we also recorded everything with a video camera, so we could watch the test again if needed.

COMPARISON EXPERIMENT RESULTS

In this experiment we had two sets of data, one from the experiment and questionnaire and one from the TLX test.

DATA FROM THE TEST AND THE QUESTIONNAIRE

In Table 5 the average task completion time for the twelve subjects in each of the three prototypes is shown. Using the task completion time an average words per minute was calculated. A word is defined to have the length 4,4 characters [23].

Prototype	Task completion time:	Words per minute
A	127 seconds	1.42 WPM
B	154 seconds	1.17 WPM
C	109 seconds	1.65 WPM

Table 5: The average task completion time

Table 6 shows the user evaluation averages, grading 1 to 3, where 1 is best.

Prototype	Best	Fastest	Easiest to use	Most manageable	Most relaxing
A	1.58	1.83	1.50	1.42	1.42
B	2.33	1.92	2.42	2.42	2.42
C	2.08	2.25	2.08	2.17	2.17

Table 6: User evaluation averages

Table 7 shows the user evaluation data based on a number of questions where they should compare the prototype against each other..

Prototype	Fastest	Easiest to use	Most manageable	Most relaxing
A	4	8	8	7
B	3	3	2	3
C	5	1	2	2

Table 7: High rating number score

The subjects were asked which prototype they felt was fastest to use. The results this is compared to the actually performance. This comparison is seen in Table 8

Prototype	Actually best performance	Sense of best performance
A	2	4
B	0	3
C	10	5

Table 8: Actually performance vs. sense of performance

An interesting thing to notice is that even though Prototype C clearly outperformed the other prototypes in completion time, the subjects themselves felt that this was not the case.

We also calculated a weighting value for each user evaluation category; we scored each category every time they matched up with the overall best prototype. This means that if a subject rated the overall best prototype in the order 3-2-1 and rated the fastest prototype in the order: 3-2-1, then the user evaluation category "fastest" would get a point. All the points for all categories were added up, and a user evaluation value was created. Because the "fastest" category and "overall best" category could have just 1 prototype in common, we made an additional rating. If a category and the "overall best" category had one prototype in common, that category would get one point and if they had three prototypes in common, that category would get three points.

User evaluation value with one points are shown in Table 9

	Fastest	Easiest to use	Most manageable	Most relaxing
Value:	2	4	6	4

Table 9: User evaluation with one point

User evaluation value with one and three points is shown in Table 10.

	Fastest	Easiest to use	Most manageable	Most relaxing
Value:	12	19	22	17

Table 10: User evaluation with one and three points

In Table 11 we have multiplied the user evaluation average and the value, first with only one point.

	Fastest	Easiest to use	Most manageable	Most relaxing	Combined
Prototype A	4	6	9	6	24
Prototype B	4	10	15	10	38
Prototype C	5	8	13	9	35

Table 11: User evaluation and rating with one

In Table 11 we can see that Prototype A has the lowest score, and here the lowest score is the best. Prototypes B and C have roughly 50% higher scores than Prototype A.

If we multiply the user evaluation average and the value with one and three points, the prototypes have the scores listed in Table 12.

Prototype	Fastest	Easiest to use	Most manageable	Most relaxing	Combined
A	22	29	31	24	106
B	23	46	53	41	163
C	27	40	48	37	151

Table 12: User evaluation and rating with three

This calculation gives the same result, which is that Prototype A gets a 33% better score than prototypes B and C. Also in both cases neither Prototype B nor Prototype C gets a better score than Prototype A in any of the individual categories.

DATA FROM THE NASA-TLX TEST

Three demands and three user experiences were evaluated in the TLX test and the average scores are shown in Table 13. In the scale 5-100, where 100 represents the highest demands, strain and frustration and worst performance:

Prototype	Mental demands	Physical demands	Time demands	Performance	Strain	Frustration
A	39,6	32,5	29,2	25,4	32,9	41,7
B	41,7	34,2	45,0	30,0	44,6	51,7
C	57,9	31,3	38,8	25,0	53,3	47,5

Table 13: TLX evaluation averages

The user was also asked to rate each category based on its importance for each prototype. The results of these ratings are shown in Table 14 in the scale 0-5, where 0 equals less important while 5 equals most important.

Prototype	Mental demands	Physical demands	Time demands	Performance	Strain	Frustration
A	2,8	2,3	1,5	3,2	2,6	2,7
B	3,0	1,2	3,7	2,1	2,2	2,9
C	3,3	1,0	3,9	2,3	2,2	2,3

Table 14: TLX rating averages

In this NASA-TLX test [22] we had two set of values; a grading for each category and a score for each of those categories. We multiplied the values, so the scores in the categories which had the highest rating became the most important scores, the results of this can be seen in Table 15.

Prototype	Mental demands	Physical demands	Time demands	Performance	Strain	Frustration
A	112	73	44	80	85	111
B	125	40	165	63	97	151
C	193	31	152	58	116	107

Table 15: TLX evaluation x rating

To make this clear, the average scores from all the categories are shown in Table 16.

	Average score
Prototype A	84
Prototype B	107
Prototype C	109

Table 16: TLX evaluation x rating average

From this, it is seen that Prototype A has the lowest (best) score, with a score that's roughly 20% lower than prototypes B and C.

SUMMARY

Prototype A was clearly the best of the three prototypes based on the tests we ran. Therefore we decided to further develop Prototype A.

To get a visual overview of the prototype ratings, we have multiplied the user evaluation and TLX values with a constant. Doing it in this way will not affect the ratios between the numbers in each category, but instead give a joined score, as seen Figure 11.

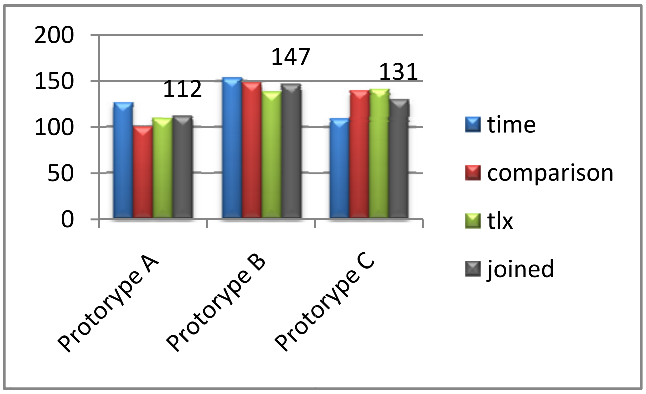


Figure 11: An overview of all prototypes in the experiment

IMPROVING PROTOTYPE A

After Prototype A was selected in the prototype test, we made several improvements and named the technique Fi-WRITE (*F*our *I*nput *Wri*ting *Te*chnique). Some improvements were only on the graphical side, while others were functional. Most of the improvements we made were based on the source *Human Computer Interaction 3rd Edition* [4]. We changed two things based on the two principles of predictability and consistency. First we changed the display so that all the most frequently used letters and alpha-numeric characters were always in the same spot in the sub-screens. We also changed the "back" function so that it is always to the right and in the first sub-screen one could always select right to get to the "back" function., The most frequently used letters and alpha-numeric characters are marked with a red color on the first screen to make it easier to remember which letters are only two steps away. Three changes were made based on the familiarity principle. We gave the program an overall blue color and a white field for the letters, to make the users think of the program Word, by Microsoft. Figure 12 shows a screenshot of the graphical optimizations.

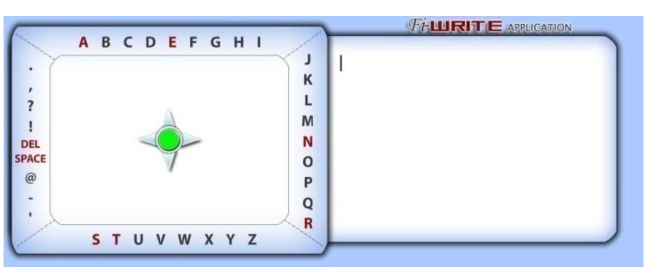


Figure 12: Screenshot of the Fi-WRITE application after optimization

EVALUATION EXPERIMENT

The third and final experiment in this article was to evaluate Fi-WRITE further. The purpose of this experiment was to evaluate the usability and effectiveness. The experiment was conducted as an exploratory test .

In the experiment, we included a test with no delay, where the subjects were given the task of writing an sentence of 3 words as fast as possible.

In the following section we describe all aspects of the test.

TASK

In this experiment the subjects were asked to write the sentence ***Peter is walking down the street, while jumping up and down, using both legs.***.

For the part without the delay the task was:
was walking down

PARTICIPANTS

Four students from the Computer Science Department at Aalborg University participated as subjects in this experiment.

PROCEDURE

Introduction texts were read to each subject before the test began. The text was intended to generally describe how the test would be run, the purpose of the test, how to give inputs to the prototype and how the prototype functions.

Upon entering the test area, the subject was placed in front of a computer. A text was then read out loud to the subject. After the prototype was tested, the subject was asked to fill out a questionnaire [4] about the prototype. The scale on the questionnaire was 1-9. We used an odd number for the maximum value because then there would be a neutral number, in our case five. When this questionnaire was completed, the test monitor asked the subject to go into details in all the areas of the questionnaire that he graded the prototype as below average. The test was complete after the interview.

First the subject would get asked to write the sentence: ***peter is sleeping***, but would get stopped after one minute. This was done to make the subjects more familiar with the writing technique prior to the experiment.

In the experiment, the subject was asked to write a sentence which took 192 inputs. This was done, in order to compare the experiment results from both before and after the graphical improvement.

In addition to the normal input method, we tested how fast each participant could write a 3-word sentence requiring 40 inputs.

This was done to be able to compare our program's word per minute ratio with other applications' word per minute ratio.

DATA COLLECTION

During the test we used a stopwatch to record the completion times for each prototype for each subject. We also recorded everything going on with a video camera.

EVALUATION EXPERIMENT RESULTS

Table 17 shows the questionnaire with answers for this experiment. The questionnaire was originally in Danish, but is here translated into English.

Test subjects →		**1**	**2**	**3**	**4**	**Average score**
	Bad / good	8	6	6	8	7
	Frustrating / calming	4	6	3	7	5
	Dull / stimulating	6	4	3	4	4,25
	Hard- / easy to use	6	7	9	9	7,75
The program is basically	Ineffective / effective	6	5	6	8	6,25
	Hard- / easy to read	9	3	9	9	7,5
	Bad- / good font	8	2	5	9	6
The letters are	Bad- / good order	7	5	7	7	6,5
The layout	Bad / good	8	5	5	8	6,5
The colors	Bad / good	8	4	2	8	5,5
Using it for the first time	Hard / easy	6	4	2	9	5,25
Learning to use it optimally	Hard / easy	6	6	7	7	6,5
Remembering where the letters are located	Took a long- / short time	5	3	5	8	5,25
Correction of spelling errors	Difficult / easy	6	9	8	4	6,75
The program is basically	Unreliable / reliable	7	7	8	8	7,5
System malfunctions appear	Often / rarely	8	9	9	9	8,75
Average score		6,75	5,31	5,88	7,63	6,39

Table 17: Usability rating results

The scale was from 1 to 9, where 9 was the most positive.

Our subjects took an average of 34.5 seconds to write the three words. However, our subjects were not very familiar with the program, due to the fact that they had only been using it for roughly 10 minutes when testing how fast they could write. The fastest subject took 27 seconds and the slowest took 45 seconds. Since we had slightly more experience using the system than our subjects (though not more experience with writing without delay), we decided to also test this on ourselves. Our fastest time was 14 seconds. All the times from the speed-test can be seen in Table 18.

Subject	**Time 1**	**Time 2**	**Time 3**	**Time 4**	**Time 5**	**Time 6**	**Time 7**	**Fastest time (sec)**
1	1.00	0.51	0.48	0.45	0.45	……	……	45
2	0.36	0.41	0.36	0.33	0.31	……	……	31
3	0.55	0.48	0.51	0.35	0.35	……	……	35
4	0.45	0.32	0.30	0.29	0.28	0.27	……	27
5	0.25	0.22	0.18	0.19	0.15	0.14	0.15	14

Table 18: Time it took to write three words without delay

During the test, we stopped the subjects and asked them what they thought of the location of the box in which the text appeared. All subjects felt it was natural that the box would be to the right of the selection part.

All subjects were also asked if they would rather have seen the letters in the bottom frame as "Z Y X W V U T S" instead of "S T U V W X Y Z", to correspond with the clockwise placement of the letters. Half the subjects didn't care one way or the other and the other half liked them going from left to right, and not clockwise. Furthermore, the subjects were asked if they liked the way the letters were ordered, or if they would rather have seen the letters ordered in another way, for example: q, w, e, r, t, y… and so on, like the order on a keyboard. But all subjects liked the way the letters were placed on the prototype.

Two of the subjects pointed out that it was hard to tell the difference between " , " and " ' ", the comma and the apostrophe.

There are some interesting issues to note from the questionnaire. Subject three rated the use of colors very low, due to the fact that the subject was colorblind. We used dark gray and dark red colored letters but the subject could not tell those two colors apart. Also the subject rated "using it for the first time" low, which he said was because of the colors.

Subject three found the program fairly frustrating and the same time found it dull. Those two ratings were related; he found it to be dull because of the delay time we implemented and because of that delay time, he also found it frustrating. Also, subjects one and two mentioned that the delay in the program made it irritating.

Subject two felt that it was a bad font and that it was hard to read, but could not tell us why or which font or kind of font he felt would be more suitable for the program.

Subject two did not like the color, because the subject simply does not like the color blue, and the color blue is the main color in our interface.

All four subjects had problems remembering where the different letters were in the beginning. One subject rated "remembering where the letters are located" higher than the others. He said he felt that during the test, he quickly learned the location of a lot of letters, and thought that with very little training, remembering the letters location would be easy.

It should be noted that in six of the categories, the difference between the highest and the lowest score were five or more, which normally makes the result inaccurate. However, if the highest and lowest score in all 16 categories are removed, the joined average score will only increase by 1.7%, so the changes are not significant.

During the experiment we measured the writing speed two different times, once in the beginning and once in the end. We found that the subjects in average took 3.18 seconds longer writing a word in the end of the experiment, than in the beginning.

From the overall speed in the experiment we found that the subjects in average were 1.16 seconds faster per word than before the improvements were made.

SUMMARY

The subjects liked the order of the letters. The position of the textbox was good. Some subjects had problems telling comma and apostrophe apart. The time to write three words, without delay was 34.5 seconds (6 WPM) in average and 27 seconds (7.7 WPM) for the fastest. The fastest time we had in the project group was 14 seconds (12.68 WPM). One subject could not tell dark gray and dark red apart. One third did not like the font. Due to delay, the subjects found the program both dull and frustrating. The average program score on a scale from one to nine was 6.39. The writing speed of the subjects dropped during the experiment and the writing speed of the system were 1.16 seconds per word faster than prior to the improvements.

FINAL CHANGES

After the optimization test, we only made some minor changes to the application. We made the following changes to make the application more user friendly: the sentence always starts with an uppercase letter, after a “. “, “! “ or “? ”, the next letter would always be uppercase and if an ”i” stands alone or is followed by an apostrophe, it will turn into an uppercase “I”.

DISCUSSION

The results from the rotation experiment showed that the completion of the task was fastest on a suitable speed only – not too slow or too fast. Having the speed increased did not mean that the completion time became better, instead subjects made more errors. This experiment was limited by the fact that the speed could only be one of the three different speeds that we defined. The results might have been better if subjects had the option to choose the speed by trying any speeds.

The results from the comparison experiment showed clearly that Prototype C was the fastest one to use. But due to the fact that it was rotating all the time, the users found it stressful to use. In contrast, Prototype A, which ranked second place in speed, was rated to be the most comfortable and effective one to use. The users felt it was significant that they could control the speed of writing themselves.

The best performance for Fi-WRITE (as we titled the winning prototype), with a delay of 3,5 seconds, was 1,42WPM while the best performance without the delay was 12,68 WPM.

COMPARING WITH OTHER RESULTS

Both our performance results, one with a delay and one without a delay, are compared with other similar results. Our best result with the delay, 1.41WPM, is compared with BCI text writers while our best result without the delay, 12.68WPM, is compared to non-BCI text writers. In Table 19 clearly shows that Fi-WRITE is among the best of the related text writers.

	Keys	KSPC	WPM	Fi-WRITE
Non-BCI				
Multi tab on cell phones [10]	9	2,13	7.3	73.7% faster
T9 on cell phones [11]	11	1.0072 [13]	9	40.9% faster
3 key text entry [12]	3	10,53[2]	9,10	39.3% faster
5 key text entry [12]	5	3,13	10.62	19.4% faster
Edge Write [12]	4	3,52	12.5	1.4% faster
ChordTap [2]	12	2	10	26.8% faster
TiltText [1]	9	1	8	58.5% faster
Fi-WRITE	4	2,4888	12.68	
BCI				
Spelling device [6]	2	6-7	0.45	254.6% faster
Hex-o-spell [7]	1	2	1.73	17.9% slower
Spelling device [8]	2	6-7	0.68	134.7% faster
Fi-WRITE	4	2,4888	1.41	

Table 19: A comparison of FI-WRITE results with other related results

Edge Write, also with 4 inputs, is the text writer whose results are closest to our results without delay. An interesting issue is that Fi-WRITE beats the standard text writing techniques for mobile phones, multi-tab and T9.

In the comparison with BCI text writers, it should be noted that the BCI text writer results were based on experiments using the brain as the input method, Hex-o-spell was the only one that performed better than Fi-WRITE with delay.

PERSPECTIVES OF USE

In this section we discuss the possibility of using Fi-WRITE in fields other than BCI where it might improve performance. The other fields of interest are mobile phones where Fi-WRITE could be used when writing SMS, and in game consoles when entering text, for example in the high score list.

FI-WRITE IN MOBILE PHONES

Most mobile phones use about 12 buttons when writing SMS’s with the T9 technique, and 10 buttons when writing without T9. Implementing Fi-WRITE on mobile phones would only require four buttons. The evolution in the mobile industry aspires to minimizing the devices as much as possible while keeping the performance as high as possible. Therefore, implementing Fi-WRITE into mobile phones could be an advantage because the number of buttons, thus the space used for them, can be reduced. However, when writing with the Fi-WRITE, the user has to look at the display, and if the user cannot click up, down, left and right without looking at the keys, the user has to look back and forth between the keys and the display between every input.

[2] They use a technique called *key-repeat*

Furthermore, on new generation mobile phones, such as HTC Diamond [24] and iPhone [25], a tilt sensor /accelerometer is included. Using this technology input is given to the software by tilting the phone in left, right, up or down directions. Thus, this technology can take great advantage of Fi-WRITE as well.

Lastly, there is a chance that it is easier to remember the location of all the keys when having fewer keys, and also that it is easier to remember because up, down, left and right are positions to which everyone can relate and have used before. Because of this, the user will probably have an easier time when writing SMS's in a car, for example. Also in a car, the display could get shown in the windscreen, so the driver does not have to look down at the mobile phone when writing. However, writing SMS's this way will have to be tested, to find out if the driver can do this and concentrate on driving at the same time..

FI-WRITE IN GAME CONSOLES

Interaction with game consoles are often supported by enclosed game controllers only. Thus the inputs in the games are given by using the game controller, which typically has four buttons for the four directions and a few buttons for actions. Many games include text writing for entering names etc., for example in high score lists. But due to the limited number of inputs, entering text using the game controllers lacks an effective mechanism [3]. Implementing Fi-WRITE in games for game consoles might improve the effectiveness of writing text on game consoles.

CONCLUSION

The purpose of this study was to develop a text input technique with 4 inputs only. The technique should have a high level of usability and a high level of efficiency. To match a high level of usability, Fi-WRITE was optimized based on a number of HCI principles and evaluated in the experiments. Users rated the usability of Fi-WRITE to be 6.39 on a scale from 1 to 9, which is clearly in the high end of the scale. Thus, on the basis of our definition for usability we can conclude that high level of usability was successfully fulfilled.

To measure the efficiency, we used words per minute as the dominant factor. Compared to other results from related studies, Fi-WRITE proved to be an efficient input technique.

It is not possible to conclude anything from the comparisons between our results and the results from BCI studies, since their text writing used a totally different input method, which is the brain signal. Thus, a conclusion of whether Fi-WRITE is effective compared to other BCI text writers is not possible in this study. On the other hand comparing our results with non-BCI is more realistic. Fi-WRITE reached a performance of 12,86WPM, which is better than the best result, by Edge Write, from our related studies. Hereby it can be concluded that Fi-WRITE is efficient compared to the results from the related work section.

LIMITATIONS

The results in this study are limited by the delay time we had to implement in the prototypes. The number of subjects used in the experiments also limited the results. Like that only students from the Computer Science Department at AAU were used in the experiments. Furthermore the number of prototypes developed could have limited the results as well.

FUTURE WORK

As continuation of this study is could be interesting to improve Fi-WRITE further. This includes a predictive language model and future work in the program which can be done in regards to T9, auto completion and manipulation.

Predictive Text

Predictive text is an input technology based on a dictionary that reduces the keystroke per character. T9 is such a technology [26]. This could be implemented in Fi-WRITE. The number of possible words would be higher than on a mobile phone, because the letters are distributed on three "buttons", compared to a mobile phone which uses eight buttons. The predictive text technology also contains an auto completion feature, which also makes the writing faster.

Personalization of Fi-WRITE

Another interesting improvement could be letting the users localize the Fi-WRITE by adding or replacing characters in it. An additional feature for this purpose could be letting the user rearrange the characters as desired.

Finally, the user would also be able to generate new auto corrections. For example, if the user writes "wkr" and selects a space, the program changes it to "with kind regards".

ACKNOWLEDGMENTS

We want to thank all the participants who made the experiments possible.

REFERENCES

[1] Wigdor, D., Balakrishnan, R. (2003) TiltText: Using Tilt for Text Input to Mobile Phones, ACM

[2] Wigdor, D., Balakrishnan, R. (2004) A Comparison of Consecutive and Concurrent Input Text Entry Techniques for Mobile Phones, ACM

[3] Wilson, A. D., Agrawala, M. (2006) Text Entry Using a Dual Joystick Game Controller, ACM

[4] Dix, A., Finlay, J. E., Abowd, G. E., Beale, R. (2004) Human Computer Interaction 3rd edition, Harlow : Pearson Education

[5] Frøkjær, E., Hertzum, M., Hornbæk, K. (2000) Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?, CHI '2000

[6] Krepki, R., Curio, G., Blankertz, B., Müller, K.-R. (2006) Berlin Brain–Computer Interface—The HCI

communication channel for discovery, Academic Press, Inc.

[7] Blankertz, B., Krauledat, M., Dornhege, G., Williamson, J., Smith, R. M., Müller, K.-R. (2007) A Note on Brain Actuated Spelling with the Berlin Brain-Computer Interface, Springer-Verlag Berlin Heidelberg

[8] Wolpaw, J., Birbuamer, N., McFarland, D. J., Pfurtscheller. G., Vaughan, T. M. (2002) Brain–computer interfaces for communication and control, Elsevier Science Ireland Ltd

[9] Hsu, B.-J., Mahajan, M., Acero, A., Multimodal Text Entry on Mobile Devices, Microsoft Research

[10] Oniszczak, A., MacKenzie, I. S. (2004) A Comparison of Two Input Methods for Keypads on Mobile Devices, ACM

[11] Clawson, J., Lyons, K., Clarkson, E., Starner, T. Mobile Text Entry: An empirical study and analysis of mini-QWERTY keyboards, ACM

[12] Wobbrock, J., Myers, B., Rothroch, B. (2006)Few-key Text Entry Revisited: Mnemonic Gestures on Four Keys, ACM

[13] MacKenzie, I. S. (2002) KSPC (Keystrokes per Character) as a Characteristic of Text Entry Techniques, Springer-Verlag

[14] Lewand, R. E. (2000) Cryptological Mathematics, Mathematical Association of America

[15] Dix, A., Finlay, J. E., Abowd, G. E., Beale, R. (1997) Human Computer Interaction 2nd edition, Prentice Hall

[16] Celikel, E., Dalkilic, M. E., Dalkilic, G. (2005) Word-based fixed and flexible list compression, International Symposium on Computer and Information Sciences

[17] Wynekoop, J.L., Conger, S.A. (1990) A Review of Computer Aided Software Engineering Research Methods, The Information Systems Research Arena of The 90's

[18] Mikkelsen, M. T., Narayanasamy, S. (2008) Designing a Feedback Application for Brain Computer Interface, Aalborg University

[19] Rubin, J. (1994) Handbook of Usability Testing, Wiley

[20] Williams, E. J., (1949) Experimental designs balanced for the estimation of residual effects of treatments, Australian Journal of Scientific Research

[21] Hart, S. G., Staveland, L. E. (1988) Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, NASA Ames Research Centre

[22] Human Performance Research Group (1987) NASA TASK LOAD INDEX (TLX), NASA Ames Research Centre

[23] Celikel, E., Dalkilic, M. E., Dalkilic, G. (2005) Word-based fixed and flexible list compression, International Symposium on Computer and Information Sciences

[24] HTC cooperation, Specifications, http://www.htc.com/www/product/touchdiamond/specification.html

[25] Apple inc., Technical Specifications, http://www.apple.com/iphone/specs.html

[26] Nuance Communications, Inc, Learn To Use T9®, http://www.t9.com/learn

ARTICLE 2:

A STUDY OF FI-WRITE, A TEXT INPUT TECHNIQUE WITH 4 INPUTS, IN BRAIN-COMPUTER INTERFACE CONTEXT

A Study of Fi-WRITE, A Text Input Technique, in Brain-Computer Interface Context

Michael Mikkelsen
tm@cs.aau.dk

Sutharsan Narayanasamy
qmar@cs.aau.dk

ABSTRACT

In this paper, we describe a study of a text writing technique based on 4 inputs, Fi-WRITE, when using it with Brain-Computer Interface. The aim was to examine the use as well as the efficiency of Fi-WRITE when used with BCI. For this, we conducted an experiment in a BCI laboratory. The experiment resulted in only half the subjects being able to use Fi-WRITE with BCI, and they could only do so for 100-160 seconds. Because the subjects were unable to complete the task, we could not conclude on the efficiency of Fi-WRITE from the this experiment. However, we were able to conclude that the system worked in BCI context.

1. INTRODUCTION

The computer has become an indispensable part of our lives. The way the human and the computer communicate in relation to Human Computer Interaction (HCI) has been a very important issue when designing software. The most widely used interaction devices between humans and computers are keyboards and mice.

Researchers around the world have been seeking new and more expressive ways of communication with computers and one of the hot topics today is the interface between the computer and the human mind [1].

The idea is to use the human brain to directly control the interaction. This technology, which is known as Brain-Computer Interface (BCI), is based on input from the brain in the form of electroencephalogram (EEG) or other electrophysiological measurements of brain function. The research on BCI began in earnest in the early 1970s, but the first working human experiment was only seen in the mid 1990s. Since then, a number of BCI systems have been developed, with differences in the way the brain signals are used as input, the way signals are processed and the way the output is produced. [2] [3]

The fact that BCI depends on a human's brain activities, in contrast to HCI which requires a human's muscular movements, is an advantage in many aspects. For instance, BCI is suitable for people suffering from severe motor disabilities, like neuropathic disease, that partially or wholly paralyzes the muscular activities. An example is people suffering from Amyotrophic Lateral Sclerosis (ALS) also known as Motor Neuron Diseases, which is a progressive, usually fatal, neuromuscular disease that at some point prohibits people from using any conventional augmentative communication methods, despite still being able to understand everything happening in their surrounding environment [4].

The incident of ALS is 1-2 per 10000 worldwide [5]. Some 4600 people in the United States are newly diagnosed with ALS each year, and nearly 30000 people residing in the U.S. at any time are living with it [6] [7].

Thus, a BCI system can benefit this group of people in keeping their communication with the rest of the world.

Another example where BCI is useful is in the aero space industry, for controlling purposes. BCI can be implemented as a supplement to the existing method of controlling an aero plane, letting the pilot perform more tasks at the same time than he is able to do with his hands.

At the Department of Health Science and Technology at Aalborg University, a group of researchers are developing a BCI system. These researchers (from now on referred to as AAU BCI-team) will implement a four output solution in their system, which means that the brain signal will be translated into four different outputs as if the signal were given by the computer keyboard. The system uses the non-invasive method, which derives the signals through an electro cap, whereas the invasive method implants the electrodes into the brain. The AAU BCI-team process the signals online, which is live processing of data, as opposed to offline processing where the signals are stored for future processing.

Figure 1 depicts the way in which BCI is processed. The signals from the brain are picked up using the electrodes in the brain cap that is applied to the scalp. The signals are then processed and classified before they are sent to the computer system as inputs. A feedback application running on the computer transforms the input into the desired output on the screen.

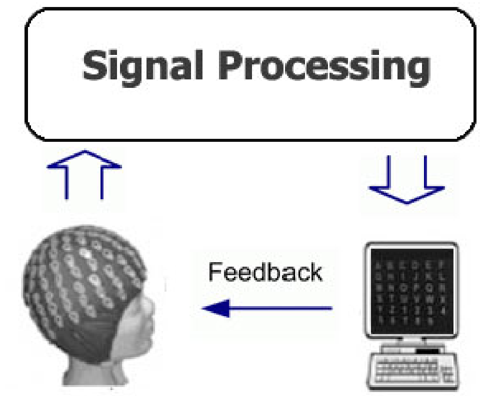


Figure 1: Brain-Computer Interface system

Text writing applications, games, medical applications, and wheelchair controllers are among the available feedback applications. Many BCI systems are limited by the number of inputs supported, thus most of PC applications become completely unusable in BCI context.

This study aims to evaluate our text input technique based on 4 inputs, Fi-WRITE, using the BCI system developed at AAU. The primary goal was to examine whether it is possible to enter characters in Fi-WRITE by the imagination. Furthermore we wanted to examine the efficiency of Fi-WRITE when used with BCI. Based on this we defined the research question for this study:

Is it possible to use BCI to enter characters in Fi-WRITE and what is the efficiency of this?

Efficiency can be defined in many ways. In the article *Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?* [8], the efficiency is defines as:

Efficiency, which is the relation between (1) the accuracy and completeness with which users achieve certain goals and (2) the resources expended in achieving them. Indicators of efficiency include task completion time and learning time.

We will follow the same definition and use the task completion time as the only indicator of efficiency. To measure the task completion time we will use WPM (words per minute), where a word is defined to have the length 4,4 characters [9].

To realize the above mentioned goals, an experiment in the BCI Laboratory at Aalborg University will be conducted.

2. RELATED WORK

Because BCI has become an interesting and important field, several BCI applications have been developed in recent times. We describe some of them in this section.

In the article Berlin Brain-Computer Interface [10] a study of several BCI text writers was investigated. This study evaluated both invasive and non-invasive BCI, as well as both offline and online signal processing. In the online, non-invasive BCI system, a speed of two to three letters per minute was the fastest a person could write. This research focused on keeping a high level of HCI. Another method also described in this article has a writing speed of three letters per minute, though this requires several months of training before it can be achieved.

In another study, *A Note on Brain Actuated Spelling with the Berlin Brain-Computer Interface[11]*, in which no emphasis was placed on HCI but only on the speed on the BCI writing, the fastest speed was 1.62 characters per minute. Two out of three subjects could keep this average speed over five words. These numbers were from online, non-invasive BCI. The focus this time was effectiveness, not efficiency, so the errors and the selections to delete them did not count in the completion time. This article evaluated the text writer Hex-o-spell.

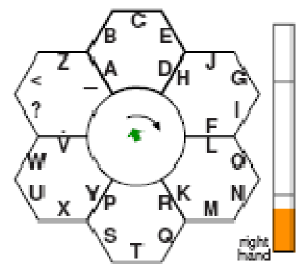


Figure 2: Hex-o-spell text writer

Furthermore, the study *Brain–computer interfaces for communication and control* [10] described a very simple text writer, in which the top speed was three letters per minute. However in some cases only 0.15 letters per minute was achieved. While writing, the user had to make four to five selections in order to write a letter, and errors were often made. With only a 65-90% accuracy rate, in some cases the user took almost 7 minutes to write a letter.

Here, we describe some of the most used and the most acclaimed text writer applications.

An often used BCI text writer application is the Hex-o-spell. Besides being used a lot, this text writer is also highly acclaimed. It works with only one input and consists of six hexagons with all the letters and symbols inside. An arrow is constantly spinning, and when the arrow reaches the desired hexagon, the user thinks in order to select the hexagon. The same goes for the next level, when there is only 1 letter in each hexagon.

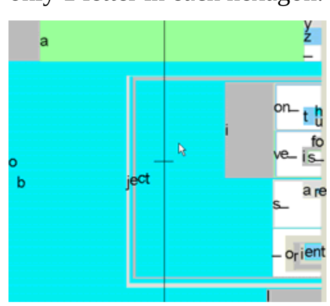


Figure 3: The text writing application - Dasher

The Dasher text writing application is highly acclaimed. Two or four inputs can be used. When using Dasher, all the letters are on the right, and the user moves the curser from the left to the right, and moves the cursor up or down to select the desired letter. Some letters are easier to hit than others, based on the previous letters and words.

Finally, a very commonly used application is a two input text writer. This is used in a lot of places but does not have a specific name. The alphabet is divided into two groups, and the user has to select one of the groups each time. So basically, on each selection, the user removes half the possibilities, until there is only one letter left.

3. Fi-WRITE

Fi-WRITE is a text writing application which we, the authors of this paper, initially designed for BCI usage. Using 4 inputs only, it gives the user the ability to write some text.

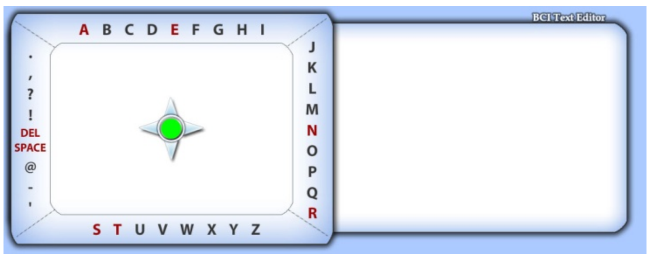


Figure 4: A screenshot of Fi-WRITE startup screen

As seen in Figure 4, the application is split into two sections; the selection of letters and symbols is in the left side while the text box where the chosen letter appears is on the right side. In the left section, the letters and symbols are distributed into 4 edges: left, right, top and bottom. The letters are arranged clockwise in alphabetical order from the top, with the symbols, space and delete on the left.

Choosing a character

All the characters, whether it is a letter, symbol, space or delete, are reached in the same way. They are chosen by selecting the edge where the character is placed. Choosing one edge splits the group into four new groups attached to each edge in the following screen.

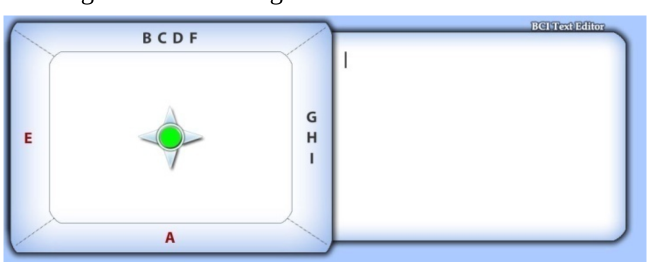


Figure 5: Fi-WRITE letter section

See Figure 5. If the desired character is placed alone in an edge in the second screen, it can be selected by choosing the correct edge. If it is again grouped with other characters, the step is repeated. Every character can be reached by either two or three steps.

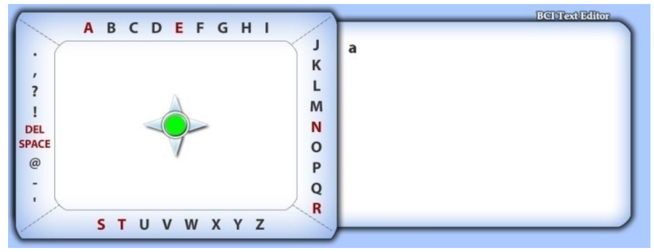


Figure 6: Fi-WRITE - a letter is entered

Those characters that can be reached with two steps are indicated with the color red. The circle in the middle indicates if a new command is reachable. It can be either red or green. Red means wait, while green means that the characters or character-groups are selectable.

The application, which is implemented using Action Scripts 2.0 in Flash, is based on the English alphabet and 7 of the most often used symbols[1], along with space and delete.

4. METHOD

As mentioned, the evaluation of Fi-WRITE was conducted using the BCI system developed at AAU. This experiment was done in the BCI laboratory at AAU. The goal is only to do an evaluation of Fi-WRITE and not the BCI system.

For Fi-WRITE to operate with the BCI system, we were required to make a bridge application in JAVA.

4.1 THE BRIDGE APPLICATION

In order for Fi-WRITE to work, it needs inputs from a keyboard. In relation to BCI, Fi-WRITE will receive input from the online data analysis provided by the AAU BCI-team. The online data analysis part is implemented using LabVIEW. A third application was needed for our flash application, Fi-WRITE, to communicate with the LabVIEW application. For this purpose, we developed a bridge application in Java which can interact with both LabVIEW and Flash applications.

The bridge application enables communication with LabVIEW by using the TCP/IP connection. Using a robot class, it generates the inputs. Figure 7 depicts the communication flow between the three applications. When a connection is established, LabVIEW sends either 1, 2, 3, 4 or 5, which represents the directions left, right, up, down, and empty to Java. When Java receives these, they are received as 58, 59, 60, 61 or 62. In the third step, Java sends the corresponding direction output to the computer, which in our case is received by the Flash application. In the next step, Java sends LabVIEW a flag (2), which indicates that Java is ready for the next input.

This process is meant to be repeated each time the subject tries to make an input with his mind.

[1] This was decided and chosen by the authors of this paper.

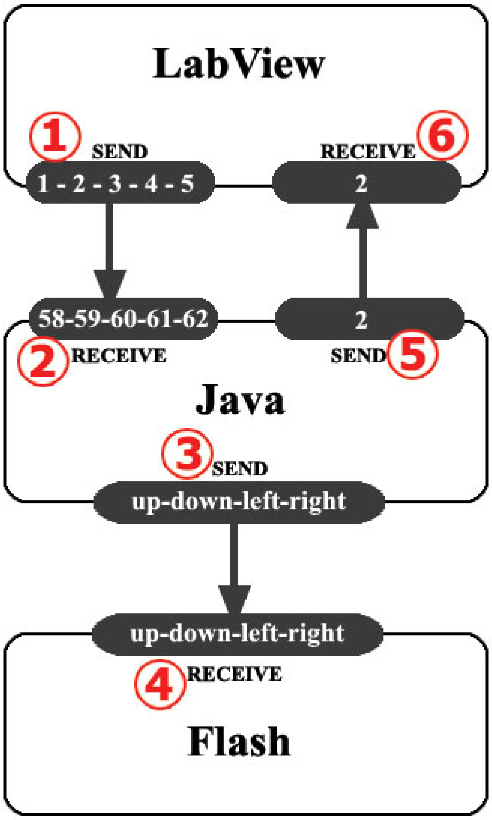


Figure 7: The communication between LabView, Java and Flash

5. BCI EXPERIMENT

We used the BCI laboratory at Aalborg University for the evaluation of our application. In the following section, we describe in detail all of the aspects regarding the experiment.

5.1 PARTICIPANTS

The experiment was conducted with eight male students from the Department of Computer Science at Aalborg University. The average age of the subjects was 26.

5.2 MATERIALS AND SETTING

The testing area was a laboratory, which was set up as a single room setup [12]. A video camera was setup to record the experiments. As seen in Figure 8, the subject was seated in an armchair looking at a monitor, while controlling the Fi-WRITE through an Electro-cap.

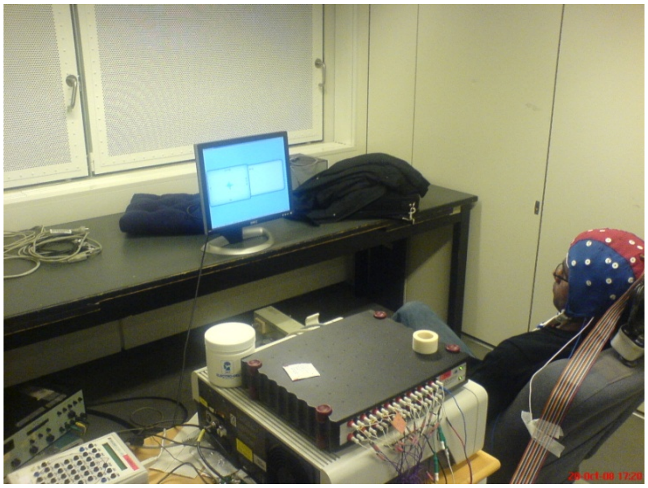


Figure 8: Experiment setting

Electro-Cap: The electro-cap was placed on the subject's head.

Gel: For the electro-cap to get a better reading of the EEG signals, a gel was injected between the electro-cap and the subjects head.

Electrodes: To validate the results, electrodes were placed on the subjects legs. These electrodes measured muscle activity, to check if the subject was thinking about an action, or actually performing it.

Monitor: A monitor on which our program interface was shown was placed in front of the subject.

PC: Fi-WRITE, along with the LabVIEW application, was running on a Windows PC that was connected to the monitor.

Armchair: When performing an experiment in the BCI laboratory, the subject was required to be seated in an armchair.

Video camera: A video camera was used to record everything the subject did during the experiment. The camera was placed behind the subject to the right and in roughly a 30 degree angle; it was filming the screen directly and the subject partly from behind.

Pedal: During the experiment, the subjects had their right foot placed on a pedal which could not move.

5.3 INPUT DELAY

The input delay time was 10 seconds according to the AAU BCI-team. Due to this no continues input can be realized, but instead the subjects were only able to give an input every 10 seconds.

5.4 TRAINING SESSION

Before each subject was tested in the experiment, the subject had to go through a training session. This is a very important part of the experiment, because during the training, the BCI equipment gets a lot of information from the brain, which is turned into a calibration file for the individual subject after the training. This way the system is adapting to the individual user's way of thinking. Each training session consisted of 100 trials, The users had to imagine the four different pressure levels, as if he was pressing an accelerator pedal in a car, 25 times each. The trials appeared in a random order, with an indication of which pressure level to imagine.

5.5 EXPERIMENTAL TASK

The subjects were asked to write the following sentence in Fi-WRITE, using only their brain:

"Peter is sleeping"

5.6 PROCEDURE

The experiment had one test observer and one test monitor, who also acted as an observer when the actual experiment was started. The experiment was divided into the following four phases:

1. Overall introduction
2. The training session
3. The experimental task
4. Filling out the questionnaire

Overall introduction

In the overall introduction, the subjects were welcomed, whereupon the test monitor presented the agenda for the experiment and introduced the environment and the equipment. All interaction during the experiment followed a specific manuscript, unless the subject asked unexpected questions.

The training session

Before the actual training session, the subject was introduced to how the training session was going to take place and what exactly was needed to be practiced. When he understood the concept, he was asked to do the training.

The experimental task

The experiment's task phase also had a short introduction, to ensure that the subject understood what the experiment was about. When accepted he was

Filling out the questionnaire

After completion of the experimental task, a questionnaire was handed out to the subject, where a number of questions regarding the subject's performance was to be rated. The questionnaire can be found in the Appendix.

5.7 DATA COLLECTION

As mentioned, there were two observers during the experiment. Both the observers kept track of errors and task completion time. The data from both observers was compared to ensure it was equal. In case the data were different or anything was missing, the recording from the video camera was used to validate the data.

6. RESULTS

First we give a brief presentation of the major outcomes in the test, after which we present the subjects' accuracies in the experiment, how many letters were written by the different subjects and the efficiency as measured in WPM. Next, we present an overview of which words were written during the test and sum up all the results from the test.

Out of the eight subjects who participated in the experiment, four of them were unable to give a voluntary stabile input to the system, meaning they could not control what the system did. All four subjects who were able to control the system were only able to do so for a limited period of time, which was not the same for all subjects. The outputs from the eight subjects are shown in Table 1. From this table, it is seen that only 2-4 letters were entered correctly. Both the correct inputs and incorrect inputs were measured for the control time only.

#	Letters written	Correct input	Wrong input
1	Petppppppppp	12	2
2	Pet(del)(del)(del)(del)(del) (del)(del)(del)(del) (del)	9	3
3	Pessssssssss	9	2
4	Pppppyppnpppp	0	0
5	Peteppppppkppp	13	1
6	Bbbbbbbb	0	0
7	Ssssssssss	0	0
8	Sssssssnsszssnsss	0	0

Table 1: The letters written during the test

The four subjects who were unable to control the system gave a constant input to the system, which resulted in the Fi-WRITE writing the same letter over and over again. When 20 consecutive wrong inputs were reached, we evaluated the subject as having lost control or being unable to use the system.

In Table 2, different average accuracies based on the correct inputs are listed. The accuracies are all calculated by number of correct inputs multiplied by 100, then divided by the total number of inputs.

Aspects of accuracies	Accuracy
The overall accuracy of the system	17.7%
The overall accuracy of the first three minutes of each experiment	29.2%
The accuracy of the four subject who were able to use the system	35.6%
The accuracy of the first three minutes of each experiment, but only of the four who could use the system	58.4%

The accuracy or the four subject who were to use the system, up until they were unable to control the system	84%

Table 2: The accuracy of the system

As can be seen in Table 2, the accuracy of the subjects who could control the system was fairly high, before they lost control. An overview of each subjects' accuracy can be seen in Figure 9. This chart shows the average accuracy for each subject while they had control over the system.

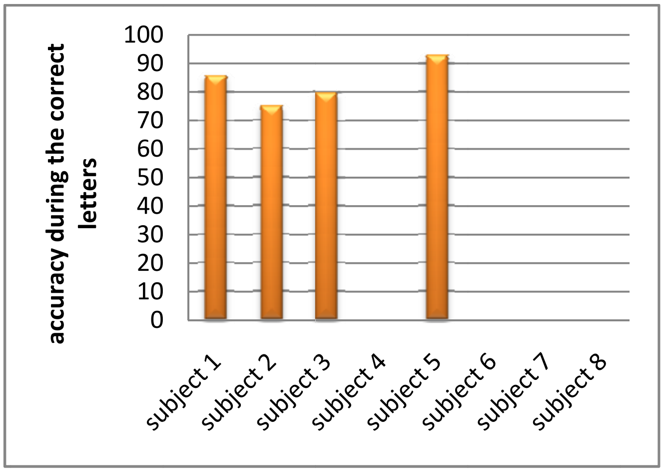


Figure 9: Accuracy during the correct letters

Of the subjects who could control the system, the longest a subject was able to keep the control of it was for two minutes and 20 seconds, while the shortest was 1 minute and 50 seconds. The results for these subjects are displayed in Figure 10.

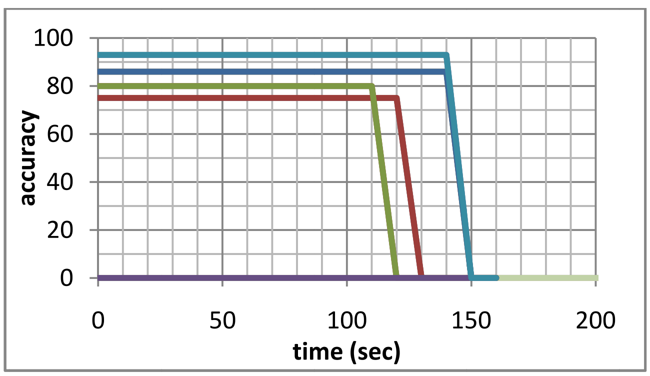


Figure 10: Accuracy timeline of all subjects

Since the period the subjects could keep up this accuracy was limited, the number of correct letters being written was very low; it can be seen in Figure 11. Here it should be noted that this is the number of correct letters they reached before losing control, not the number of correct letter displayed when the experiment ended.

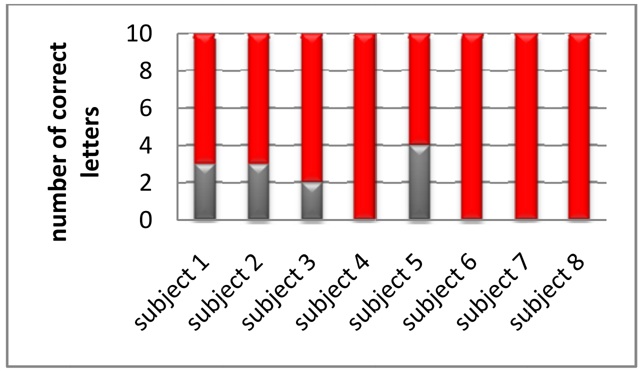


Figure 11: Number of correct letters during the entire experiment

As can be seen from comparing Figure 9 and Figure 11, it is not only a matter of accuracy to be able to write the highest number of letters. Even though subject two's accuracy was lower than subject three's (see Table 2), subject two was able to write one more letter than subject three. This was due to two factors; subject two was able to use the program longer before losing control and subject two simply made his errors at a better time during his writing. If a subject is able to simply go back after an error, it costs two steps less than if a wrong letter has to be input. For example, if two subjects wish to write an "e" and nothing else, both subjects must go to the "A-I" group. If both subjects make an error here, and subject one goes up instead of left while subject two goes to the right instead of the left, in this case subject one needs five steps in order to complete the task, while subject two only needs three steps.

Table 2 shows the WPM the subjects were able to write using the system. The WPM value is calculated based on the characters per minute. Even though a subject has a WPM value, it does not mean that he wrote a whole word. It is simply used as a measurement of writing speed.

	Char/min	WPM
The average overall speed of the system	0,30	0,07
The average overall speed of the first three minutes of each experiment	0,50	0,11
The average speed of the four subject who were able to use the system	0,61	0,14
The average speed of the first three minutes of each experiment, but only of the four who could use the system	1,00	0,23
The average speed of the four subject who were able to use the system, up until they were unable to control the system	1,44	0,33

Table 3: The speed of the system

Subject	Time	Letters	Errors	Accuracy	WPM	Letters entered
1	140	3	2	86%	0.29	Petppppppppp
2	120	3	3	75%	0.34	Pet(del)(del)(del)(del)(del)(del)(del)(del)(del)(del)
3	110	2	2	80%	0.27	Pessssssssss
5	140	4	1	93%	0.39	Petepppppkppp

Table 4: A complete overview of the successful subjects

When only evaluating the subjects who could control the system, and only while they had control of the system, we found that our system had a word per minute ratio of 0.32, meaning the subjects in average could write 0.32 words for every minute they were using the system.

The WPM of all the subjects (see Figure 12) follows the overall same tendency as the accuracy (see Figure 10: Accuracy timeline of all subjects).

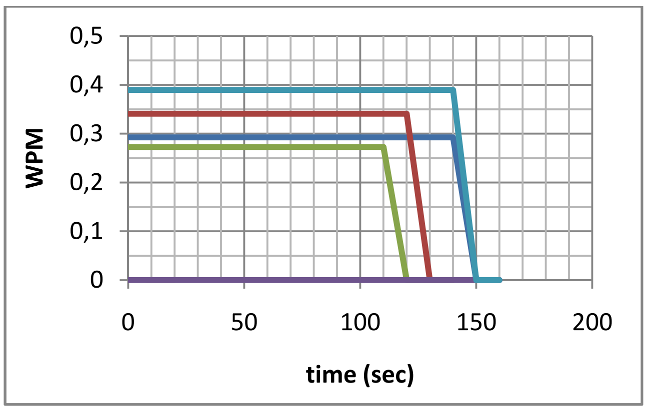


Figure 12: WPM timeline of all subjects

Even though they follow the same overall tendency, the two graphs are not the same. The subject with the lowest accuracy during the experiment had the second highest WPM in the same experiment.

The four subjects who could not control the system and the four subjects,who lost control of the system still gave an input to the system, but it was not the input they wanted. An example of the input of a person losing control of the system is this: *petepppppkppp,* which can be seen in Table 1. Here the person was only able to write four letters voluntary and after that he kept giving the input for "right". The additional "k" in the sentence, was created because the subject on one trial was able to change the input to "up". However, the subject did not give the input intentionally and was not able to give it again later.

The four subjects who were able to control the system in the beginning all lost control later. When they lost control, the system got the same input and the Fi-WRITE wrote the same letter over and over again, or in one case kept deleting letters.

Summary

Only half our subjects could use the system and the subjects who could use it could only do so for 1 minute and 50 seconds to 2 minutes and 20 seconds. The overall speed of the system was 0.07 WPM, but when only counting while it worked, the speed was 0.32 WPM. The subjects' average accuracy was 17.7%, while the subjects who could use the system, had an accuracy between 75% and 93% before they lost control. The four subjects who could use the system were able to write 2, 3, 3 and 4 correct letters. **Error! Reference source not found.** presents a brief overview of the results from the subjects who could control the system.

7. DISCUSSION

As mentioned earlier, the results were surprising both to us and to the AAU BCI-team. First of all the input delay time was increased to 10 seconds, this was totally different from the 3.5 seconds which was informed in the development phase. Beside this, only four of the eight subjects were able to use the system. The most unexpected finding was that the subjects were suddenly unable to control the system.

The average word per minute for the four subjects who could control the system was 0.32 while they had control of it. Compared to the word per minutes that other articles presented, our results are between 10% and 50% lower. Furthermore, our subjects could not keep writing for very long, only between 100 and 140 seconds.

The BCI-system being developed at Aalborg University is unique; no one in the world had developed a BCI system with exactly the same conditions. The BCI system was being developed simultaneously with the development of Fi-WRITE. Even though the system was not 100% completed at the testing stage, it was validated to be used in our experiment. The AAU BCI-team advised us not to be too optimistic about the outcome and told us to only expect about 30% accuracy. So it was surprising, especially for them, that some of the subjects managed to get an accuracy of 100% in the beginning of the task.

The results led us into the speculation of why the experiment resulted as it did with the following questions:

1. *Why did only four of eight subjects manage to control the system?*
2. *Why did the subjects lose control after some few minutes?*

This could be caused by several factors that we are unable to verify, but one important factor is the training timeframe. We only used 25 trials for training each pressure level. This gives a total of 100 trials which corresponds to 12-13 minutes. In related BCI experiments, subject had been trained in many hours over a period of several months [13]. For this reason we increased the trials to 50 instead of 25 in an additional mini-test conducted using one of the subjects which could not control the system in his first try. This did not change anything in the results. Instead, the subject felt doing 200 trials successively put too much strain on him and found it difficult to focus at the end.

Another factor is concentration. The brain is very sensitive, thus several issues can influence one's imagination without being aware of it. The fact that the BCI-system records everything in the experiment's task interval required 100% concentration to produce useful results. Thus if concentration is weakened, this will impact the results negatively.

The level of complexity is another factor. The subject had to define four different pressure levels by imagining. These four levels should be exactly the same each time when doing the trials. The subject was not provided any feedback of matching imaginations in the training session. This requires competence.

A fourth factor could be the mapping. In the training session, the subjects were asked to calibrate the system by imagining four different pressure levels, as if pressing an accelerator pedal in a car. Each pressure level was equal to a direction in Fi-WRITE. In other words, the subject should find which direction he needs to go to enter the desired character, and imagine the corresponding pressure level. This is clearly a lack of mapping, thus the subject might have had difficulties in transforming pressure levels to directions. The information that the measurements in the BCI system are based on pressure levels was given us very late, after the completion of Fi-WRITE development, so we had no chance to redesign our technique to match the mapping within the given timeframe.

These factors all assume that the BCI equipment was working and performing as expected. But one factor might also be lack of performance from the BCI equipments. This could be the amplifier, the electro-cap or some other equipment, or the algorithm used to analyze the EEG data. All these factors regarding the BCI equipments are outside our field and therefore hard to verify in which part it could have gone wrong.

In order to find the exact problem, another experiment is needed, with the above mentioned factors considered more precisely. A more positive outcome might result from taking one or more of these issues into account.

8. CONCLUSION

The purpose of this study was to evaluate the effectiveness of Fi-WRITE when used in BCI context. The evaluation was made with eight subjects in a BCI laboratory, where four of the subjects managed to produce useful data in the beginning phase only. Although the results were good in a short period for four of the subjects, the overall outcome of the experiment was not comparable with other BCI results.

From the results in this study, it is not possible to conclude on the effectiveness of Fi-WRITE. The only clear conclusion from this study is that Fi-WRITE did work in BCI context. Although it was limited, it was actually possible to write some desired characters in Fi-WRITE based on the imaginations.

The results from the Fi-WRITE evaluation are limited by several factors. The timeframe given to the training session might have been too small. Another limitation is that we only used a specific group, students from the Department of Computer Science at AAU, for the evaluation. Furthermore, we were also limited by the number of test subjects and the lack of experience they had with brain interaction tasks. Finally, the increased delay time affected the results.

As a continuation of this project, it would be interesting to do another experiment in the BCI laboratory, both to get a more successful result and to examine the reason for the unexpected outcome of this experiment. More of the factors which limited this experiment could be improved, as well as taking the possible factors for the negative results into account.

9. ACKNOWLEDGMENTS

Furthermore we want to thank all the participants who made the experiments possible. And last, but not least, we thank the AAU BCI-team for the cooperation.

10. REFERENCES

[1] Howard, T. (1996) Beyond the Big Barrier, Personal Computer World

[2] Seabrook, R. H.C. (1994) The Brain-Computer Interface: Techniques for Controlling Machines, Anne Arundel Community College

[3] Dornhege, G., Millan, J. D. R., Hinterberger, T., McFarland, D. J., Müller, K.-R. (2007) Toward brain-computer interfacing, MIT Press

[4] National Institute of Neurological Disorders and Stroke, Amyotrophic Lateral Sclerosis Fact Sheet, http://www.ninds.nih.gov/disorders/amyotrophiclateral sclerosis/detail_amyotrophiclateralsclerosis.htm

[5] WebMD, Amyotrophic Lateral (ALS) – Topic Overview,

http://www.webmd.com/brain/tc/amyotrophic-lateral-sclerosis-als-topic-overview

[6] Healthcommunities.com, Inc., Amyotrophic Lateral Sclerosis (ALS)/Lou Gehrig's Disease, http://www.neurologychannel.com/als/index.shtml

[7] Wickelgren, I. (2003) Tabbing the Mind, American Association for the Advancement of Science

[8] Frøkjær, E., Hertzum, M., Hornbæk, K. (2000) Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?, CHI ‘2000

[9] Celikel, E., Dalkilic, M. E., Dalkilic, G. (2005) Word-based fixed and flexible list compression, International Symposium on Computer and Information Sciences

[10] Krepki, R., Curio, G., Blankertz, B., Müller, K.-R. (2006) Berlin Brain–Computer Interface—The HCI communication channel for discovery, Academic Press, Inc.

[11] Blankertz, B., Krauledat, M., Dornhege, G., Williamson, J., Smith, R. M., Müller, K.-R. (2007) A Note on Brain Actuated Spelling with the Berlin Brain-Computer Interface, Springer-Verlag Berlin Heidelberg

[12] Rubin, J. (1994) Handbook of Usability Testing, Wiley

[13] Wolpaw, J., Birbuamer, N., McFarland, D. J., Pfurtscheller. G., Vaughan, T. M. (2002) Brain–computer interfaces for communication and control, Elsevier Science Ireland Ltd