Research Article

Text Entry by Gazing and Smiling

Outi Tuisku,[1] Veikko Surakka,[1] Ville Rantanen,[2] Toni Vanhala,[1,3] and Jukka Lekkala[2]

[1] *Research Group for Emotions, Sociality, and Computing, Tampere Unit for Computer-Human Interaction (TAUCHI), School of Information Sciences, University of Tampere, Kanslerinrinne 1, 33014 Tampere, Finland*

[2] *Sensor Technology and Biomeasurements, Department of Automation Science and Engineering, Tampere University of Technology, P.O. Box 692, 33101 Tampere, Finland*

[3] *ICT for Health, VTT Technical Research Centre of Finland, Tekniikankatu 1, P.O. Box 1300, 33101 Tampere, Finland*

Correspondence should be addressed to Outi Tuisku; outi.tuisku@sis.uta.fi

Received 8 November 2012; Accepted 28 February 2013

Academic Editor: Kerstin S. Eklundh

Face Interface is a wearable prototype that combines the use of voluntary gaze direction and facial activations, for pointing and selecting objects on a computer screen, respectively. The aim was to investigate the functionality of the prototype for entering text. First, three on-screen keyboard layout designs were developed and tested ($n = 10$) to find a layout that would be more suitable for text entry with the prototype than traditional QWERTY layout. The task was to enter one word ten times with each of the layouts by pointing letters with gaze and select them by smiling. Subjective ratings showed that a layout with large keys on the edge and small keys near the center of the keyboard was rated as the most enjoyable, clearest, and most functional. Second, using this layout, the aim of the second experiment ($n = 12$) was to compare entering text with Face Interface to entering text with mouse. The results showed that text entry rate for Face Interface was 20 characters per minute (cpm) and 27 cpm for the mouse. For Face Interface, keystrokes per character (KSPC) value was 1.1 and minimum string distance (MSD) error rate was 0.12. These values compare especially well with other similar techniques.

1. Introduction

Recently, there have been several attempts to develop alternative human-computer interaction (HCI) methods that utilize eye tracking in combination with another human behavior related measurement. One line of investigation has been to measure signals that originate from human facial expression systems [1–4]. One reason for using facial muscle behavior in HCI has been the fact that the human facial system is versatile when used for communication purposes and whose functionality could serve as a potential solution in HCI systems as well [3]. Pointing and selecting as well as text entry are the most common tasks in HCI, and thus being able to carry them out with acceptable performance can be considered important in order to consider an HCI solution fit for use.

The potential of the human facial system has been already utilized in the context of eye tracking research. For example, eye blinks have been used for selecting objects when gaze direction has been used for pointing [5, 6]. The choice of using use eye blinks results from the fact that video-based eye trackers that image the eyes are able to recognize whether the eyes are opened or closed [5, 7]. The relation to the facial muscle system comes from the fact that eye blinks result from the activation of *orbicularis oculi* facial muscle [8]. While video-based eye trackers track the eyes, computer vision methods can be used to track the eyelids directly [9]. Eye blinks which are used for selection purposes could be mistaken for unintentional eye closure, which in turn can cause unwarranted selection incidents.

The use of facial actions other than blinking can offer more functional solutions in combination with gaze pointing. Other facial muscles that can be used in HCI are, for example, *corrugator supercilii* (i.e., activated when frowning) or *zygomaticus major* (i.e., activated when smiling). In 2004, a real time HCI method was introduced where voluntary gaze direction was used for pointing and voluntarily produced facial muscle activations were used for object selection [3].

Remote eye tracker was used to record the gaze direction and facial electromyography (EMG) was used to measure frowning related facial muscle activations. An overall mean pointing and selection task time of 0.7 seconds was reporter using a relatively simple experimental setup [10]. In comparison to the computer mouse, the results showed that as measured with pointing task times, use of the mouse proved significantly faster than the new technique in the shortest pointing distance. However, with medium and long distances, there were no statistically significant differences between the mouse and the new technique. In a follow-up study, the technique was extended so that two EMG channels (i.e., frowning and smiling related electrical activity) could be used with gaze direction and to then determine which of the two would function better as a selection technique [4]. Findings revealed that smiling functioned faster than frowning; overall mean pointing task times were 0.5 seconds and 0.8 seconds for smiling and frowning, respectively.

San Agustin et al. [2] compared the use of two pointing techniques and two selection techniques. These were the mouse and the gaze for pointing and the mouse click and voluntarily produced changes in facial EMG (i.e., frowning or jaw clenching) for selecting objects. They tested all the four possible pointing and selection combinations with simple tasks. Results showed that the overall mean task time was 0.4 seconds. Gaze combined with facial EMG was the fastest one of the pointing and selection combinations with the task time of approximately 0.35 seconds. Chin et al. [1] used facial EMG for correcting the inaccuracy of the eye tracker as well as for selecting objects. If the cursor was not inside the object when the user gazed at the object, facial muscle activations were used to move the cursor on the object. For example, left and right jaws clench resulted in the cursor to move left and right, respectively. Finally, user selected the target by clenching the whole jaw. This technique resulted in a mean task time of 4.7 seconds. These findings show that combination of gaze direction and facial activity measurement can function very well and effectively for pointing and selecting in HCI. This depends, of course, on the test setup and activations required for facial behavior.

Recently, further studies in developing the technique that combines the use of voluntary gaze direction and facial muscle activations have resulted in a prototype called Face Interface. It is a wearable device that is built in the frames of protective glasses and it consists of both a video-based eye tracker and capacitive sensors for measuring facial activity [11–14]. Capacitive sensors measure the movement of facial tissue instead of the electrical activity of the muscles that the EMG measures. Capacitive measurement has the advantage that it requires no contact with the facial skin, and thus, no preparation of the skin is needed. In addition, it makes a significant difference in respect to the wearability of the prototype. The aim is that the user just wears the prototype and starts to interact with computer.

In the first version of Face Interface, a wired commercial USB web camera was used for imaging the eye. There was no compensation for head movements, and thus, a chin rest was used for preventing involuntary head movements [13]. In the second version, a scene camera was added to

help in compensating head movements so that it was used to image the display that was used as a reference for the head-movement compensation. The prototype was also made wireless [12, 14]. In earlier experiments, the functionality of the prototype has been tested using simple Fitts' law style pointing and selecting tasks, so that the task was first to select a home square and then to select a target circle using different pointing distances, pointing angles, and sizes of the targets [10, 15, 16]. In the first experiment, the participants used frowning as the selection technique. It was found that the Face Interface prototype was functional in pointing and selecting tasks. The overall mean pointing task time of 2.5 seconds was reported [13]. In the second experiment participants used either lowering (i.e., frowning) or raising the eyebrows as the selection technique. The results showed an overall mean pointing task time of 2.4 seconds by means of frowning and 1.6 seconds by means of raising the eyebrows. Further, an important finding was that the objects were difficult to point and select on the left and right edges of the display which suggested a design guideline that objects should be larger on the edges of the display than in the middle of the display to make them easier to select [14]. Both of the above studies revealed that the larger targets were easier to select than smaller targets which is similar to other gaze-based or manual pointing studies [17].

All the above new studies have been done with controlled setups that allow calculation of performance metrics using Fitts' law analyses [2–4, 13, 14]. However, little by little testing new interaction techniques need to be extended to more applied (i.e., closer to "real world") tasks. Text entry is such a task that has been widely studied with systems that use only gaze input [18]. In this case, pointing is done by gazing and selection by various dwell time algorithms. Evidently, text entry studies fits well to the research of the Face Interface prototype because it has the same functions (i.e., pointing and clicking). A more direct route to apply these functions for text entry is to use an on-screen keyboard.

Most of the on-screen text entry studies with gaze-based methods have used a traditional QWERTY layout. That is mainly because QWERTY has the advantage that it is familiar for most users [19]. Based on previous research with Face Interface, however, it is clear that a QWERTY layout keyboard with equally sized buttons would not be optimally functional [14]. This is due to the fact that accuracy of eye tracking varies depending on gaze direction, and gazing with the eye closer to the extremities of its rotational range makes the tracking less accurate. This is because the eye tracking camera is often placed in front of the eye and gaze is more precisely tracked based on features on the eye (e.g., pupil) when the gaze is straight towards the camera and the features thus cover more of the image. Also, near the extremities of eye's rotational range the eyelid may occlude part of the pupil making eye tracking less accurate than when the pupil is fully visible [12]. As stated, earlier results showed that pointing was more accurate in the middle of the computer display than at the edges of the screen. This made the target selection at the edges of the screen difficult if the targets were same sized all over the display. In the case of QWERTY on-screen keyboard arrangement, for example, some frequently used letters are

situated on the edges of the keyboard (e.g., letter "a" is situated in the left side of the layout) which makes them difficult to point and select [20]. Thus, different types of keyboard layouts are needed to create a better functionality for the Face Interface prototype. An encouraging first result that writing speed could be faster with different keyboard arrangement comes, for example, in a study by Špakov and Majaranta [21]. They designed an optimized keyboard arrangement for on-screen keyboard. Most frequently used letters were placed in the topmost row of the keyboard in order to make them easy and fast to select. Their results showed that the mean writing speeds were 11.1 wpm for QWERTY and 12.18 wpm for the optimized letter placement. Similar results have also been found in different text entry studies [22, 23]. In addition to Face Interface, there are also other wearable eye trackers where an eye camera has been placed in front of the user's eye and they could benefit from new kind of on-screen keyboard layout as well [24, 25]. Also high-end remote eye trackers could benefit from a new kind of layout for an on-screen keyboard [20].

In most eye typing studies, the layout design (e.g., key size and placement) of the keyboard has not been explicitly considered. GazeTalk is one exception [26–28]. The GazeTalk consisted of 11 cells, including a text field and 10 buttons. The size of the buttons was approximately 8 cm × 8 cm and the size of the text field was 8 cm × 16 cm. In the ten buttons, six letters were visible at a time, as well as space, backspace, possibility to select letters from alphabet listing, and eight most likely words. The six visible letters changed after every typed character based on a predictive algorithm, so that if user had typed, for example, "ca", the visible letters that would be predicted as the most probable letters are, for example, t, r, n, u, l, and b. The eight most probable words dynamically changed during typing, similarly as did the six letters. If the visible letters were not what a user wished to type, the next character had to be selected from the alphabet listing. The buttons were selected using a dwell time algorithm. In a longitudinal study with the GazeTalk, the maximum text entry speed for Danish and Japanese text, after thousand typed sentences, was approximately 9.4 words per minute (wpm) and 29.9 cpm, respectively.

Dasher is one alternative for text entry that uses only one modality (e.g., mouse or gaze) [29]. It is a zooming interface which user operates with continuous pointing gestures. At first, the letters are placed in the alphabetical order in a column in the right-hand side of the screen. User should then move the cursor to the place where the desired letter is, for example, by looking at the letter. The area where the desired letter starts to grow and most probable next letters will come closer to the current cursor position. Letter is selected when it crosses a horizontal line in the center of the screen. User should navigate through the letters simply by looking at them. At first glance, the letters may seem to be unorganized because the sizing of the letters is different based on the probability of the next character so that the more probable the next character the larger its size is. This is the problem that may arise when using Dasher for the first time and some people may not understand the logic of Dasher even with long

practice. After two and half hours of practice with Dasher, users were able to write text at an approximate rate of 17 wpm. After the first 15-minute session, the average writing speed was approximately 2.5 wpm [30].

As stated above, the keyboard layouts have not been explicitly studied. In most cases, the layout was just pre-designed and then an experiment was run to test the typing speed that could be achieved with it. It is possible that keyboard layout design can be an important factor especially with new alternative interaction techniques. It should be noted that the subjective experiences of the keyboard layout could be a more important factor than the writing speed when deciding which keyboard layout to use because the user might not realize while writing which layout is faster in terms of writing speed. Thus, it is important to compare the possible layouts with each other in one fair method where the places of the letters would not interfere with the evaluation of the layout. It has been common to compare new pointing devices to computer mouse [2, 3]. However, users have a very long experience with mouse pointing, and this makes the comparison of new techniques to mouse somewhat unfair because it is the case that the mouse is very likely to be better than any new techniques. Further, usually it has been the case that the words to be written have been randomized while the letter placement of the keyboard has stayed static [31, 32]. This type of randomization is to certify partly the quality of experimental arrangement and to rule out possible biases resulting from, for example, the use of same order of the words to be written. This does not, however, change the fact when people have used some interaction technique like mouse for over some years the comparison with a new interaction technique will always result in favor of the older technique. When comparing new interaction techniques to traditional ones, the use of randomization could be considered to balance out the advantage the use of traditional pointing device in most cases has. So, randomizing the places of the letters in the keyboard could result in a more fair comparison between a new technique and a traditional one. The randomization can help in balancing out the huge advantage the mouse has over new interaction techniques. Further, randomization can be also a suitable method when comparing different input devices with each other because the places of letters that are learned with one pointing device do not provide advantage to the next condition where another device is used. In terms of text entry rate, it was shown already in the 1980s that random placing of letters do not have an effect on the writing speed, if the random letter arrangement is compared to alphabetically arranged letters [33].

When testing new techniques, objective measures of functionality are imperative. However, it is equally important to measure how the participants rate (i.e., experience) the functionality of these new techniques. There are several possibilities to measure subjective ratings of the techniques. One possibility is to use the semantic differential method which is a combination of associational and scaling procedures [34, 35]. With this method participants can rate their experiences using a set of bipolar scales that can vary, for example, from bad to good or from boring to fun. Ratings along these scales can be done by self-assessment manikin (SAM) or

a modification of it. In HCI studies, these types of scales have been frequently used to analyze experiences about new interaction techniques [3, 4, 13, 14].

Using bipolar rating scales, Surakka et al. [3] reported that the use of their technique was rated as faster than the use of the mouse. On the other hand, mouse was rated as easier and more accurate to use. San Agustin et al. [2] reported similar findings as their participants rated the combined technique as faster but less accurate to use than the mouse pointing. Recently, the participants rated the usage of the Face Interface prototype as enjoyable, easy, fast, efficient, and accurate [12]. When comparing the use of different facial activations as the selection technique with bipolar rating scales, the studies have found no difference in ratings between frowning and smiling [4], and frowning and raising eyebrows [14]. In order to deepen the systematically collected ratings, also interviews can be done to get additional information [36].

To summarize, so far the research on Face Interface has been concentrated on simple pointing and selection tasks to test the functionality of the prototype. The natural continuation would then be to extend the task with Face Interface to on-screen text entry. The previous research with Face Interface showed that there are some parts on the widescreen display that were difficult to point and select [14]. To remediate this, three different keyboard layouts were designed and tested. They were all designed so that the sizes of the keys were larger in the edges of the keyboard than in the middle of the keyboard. Another feature that was used was the randomization of letters each time a word was entered to balance out the advantage mouse interaction has over all new interaction methods.

Two experiments were run to investigate the keyboard layout and to compare writing with Face Interface to writing with the mouse. The aim in the first experiment (i.e., layout selection experiment) was to compare three different on-screen keyboard layouts, so that it would be equally easy to select any character from the keyboard. The three designed layouts were then pilot tested with ten participants to see which of the layouts would be most promising to be used in the future. After the experiment, participants rated the used keyboard layouts and a short interview was conducted. In the second experiment (i.e., text entry experiment) the aim was to compare entering text with Face Interface to entering text with computer mouse. The on-screen keyboard layout that was selected as the most prominent to be used with Face Interface in the first experiment and was used in the second experiment. In both of the experiments, the task of the participants was to enter one word at a time, and after the experiment participants rated their experiences on six bipolar scales and were shortly interviewed.

2. Face Interface

Face Interface is an eye-glass like wireless wearable device that combines the use of wearable video-based eye tracker and a capacitive sensor to detect the movement of facial skin resulting from the activation of facial muscles. The third generation Face Interface device is shown in Figure 1.

FIGURE 1: The Face Interface prototype.

The prototype device was built on the frames of protective glasses. The head-worn device includes two cameras, one for imaging the eye and the other for imaging the computer screen, an infrared (IR) light emitting diode for illuminating the eye and to provide the corneal reflection, sensors and electronics for detecting facial movements using a capacitive method, and a Class 2 Bluetooth radio (RN-42 by Roving Networks) for serial transmission of the measured capacitance signal. The used cameras were low-cost, commercial complementary metal oxide semiconductor (CMOS) cameras. The eye camera was a greyscale camera with a resolution of 352×288 pixels that was modified to image IR wavelengths, and the scene camera was a color camera with a resolution of 597×537 pixels. The frame rate for both of the cameras was 25 frames per second. The eye camera is placed near the user's left eye and the IR light source was placed right next to it. The scene camera was placed in front of the user's forehead [12, 14]. The facial movement sensors are based on capacitance measurement with a programmable controller for capacitance touch sensors (AD7147 by Analog Devices). The capacitive sensors in the frames were placed in front of both eyebrows and cheeks, and one was placed in front of the forehead.

In addition to the head-worn device, a separate carry-on unit to house some components responsible for the wireless operation was included. The unit included a power supply, four AA batteries, and two wireless analogue video transmitters that used the common free frequencies at 2.4 GHz. The PC computer was connected to a receiving station consisting of two video receivers with a power supply, and two frame grabbers for the video signals. The capacitive signal was received with computer's Bluetooth functionality.

Computer vision library OpenCV version 2.1 [37] was utilized to extract features from the image streams of both eye and scene cameras. Pupil detection was based on the corneal reflection method. The algorithm that was used for pupil detection and corneal reflection detection was the same

that Rantanen et al. [12] introduced. Calibration of the eye tracker was done in a similar manner as in the OpenEyes project [38]. Head movements in relation to the computer screen were compensated using a computer vision algorithm. The screen detection algorithm aimed to find the frames of a dark rimmed computer display from the scene camera image and, thus, no separate markers (e.g., colored dots on the borders to be detected) were needed. The algorithm was based on three observations. First, there were one or two highly contrasted edges that separated the display surface from the surrounding background. The screen is typically brightly illuminated and thus lighter than the surroundings. Further, many monitors have a black frame that surrounds the display surface. Thus, there is a sharp contrast between the illumination of the display surface and the surrounding space (e.g., monitor frame or background), and there may also be another edge with high contrast between the dark monitor edge and the background. Second, both the display surface and the monitor frame are typically rectangular, which means that they have four straight corners. Third, the corners of the outer border of the monitor frame are relatively close to the corners of the display surface. These three features were used to rank potential screen candidates to select a best one [39]. For example, a candidate with a dark rimmed border was preferred to one without.

Previously, with Face Interface, only frowning and raising the eyebrows have been used as the selection technique [12–14]. Earlier by using facial EMG, Surakka et al. [4] had compared the use of frowning and smiling as the selection technique and found out that the smiling was a faster selection technique than frowning when voluntary gaze direction was used as the pointing technique. Further, Rantanen et al. [39] found that smiling as the selection technique does not interfere with the accuracy of the eye tracker. On the basis of these findings, Face Interface was updated so that smiling activity can be tracked with capacitive sensors.

3. Layout Selection Experiment

3.1. Methods

3.1.1. Participants. Ten (7 male, 3 female) able-bodied volunteers participated in the experiment. Their mean age was 29.5 years (range 21–44 years). All of them had normal or corrected-to-normal (i.e., with contact lenses) vision. All were native Finnish speakers. To avoid any bias, participants had no knowledge of the design of the layouts.

3.1.2. Apparatus. The Face Interface prototype was used as the pointing and selection device. A $24''$ widescreen display was used and the viewing distance was approximately 60 cm. A computer with Windows XP operating system was used to run the experiment. The software for online processing of the data from the prototype was implemented with Microsoft Visual C++ 2008 [12]. The software translated the obtained information to cursor movements and selections on the computer screen.
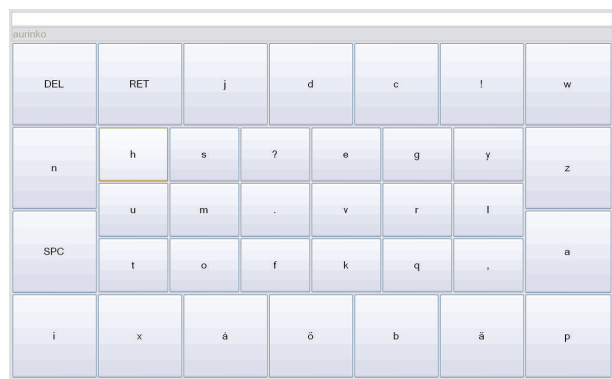

Figure 2: Layout 1.

Figure 3: Layout 2.

Three different keyboard layouts were designed (see Figures 2–4). Each of them consisted of 36 keys, although, the keys were laid out differently. In every keyboard RET, key represented the Enter key, SPC key represented the space key, and DEL key represents the delete key. In Layout 1 (see Figure 2), the keys in the middle of the screen were made smaller because earlier research has shown that it is easier to select smaller keys in the middle of the screen than on the edges of the screen. In Layout 2 (see Figure 3), keys only on the edges of the keyboard were made larger and smaller keys were used the middle of the keyboard. Finally, in Layout 3 (see Figure 4), the sizes of the keys were gradually increasing from the middle of the keyboard to the edges of the keyboard. Keyboards were implemented in .NET environment using Visual Basic 2008 programming language.

The keyboard layouts (i.e., button placement) were kept static but the places of the letters were randomized every time the participant had entered the requested word and pressed the Enter key. This approach was chosen in order to prevent the possible learning effects of the keyboard layout. Keys were highlighted when the participant's gaze was inside a key. When participant had pressed the key, a "click" sound was played to indicate the selection. Cursor was not visible. The characters that the user typed appeared in the white text box at the top of the keyboard. The grey text box under it showed the word to be written.

Figure 4: Layout 3.

3.1.3. Experimental Task. The task was to write one word (as in other studies [31, 40]) "aurinko" (i.e., sun in English) ten times with each of the three keyboard layouts. The word "aurinko" was chosen from the list of 1000 most common Finnish words and it was chosen because of three reasons: (1) it was a quite long word compared to other common ones such as "ei" (no in English) or "silmä" (eye), (2) because it was a noun, and (3) because each of the characters appeared only once.

The users entered characters by looking at the desired character and smiled in order to select the character. When participants had entered the word "aurinko" once, she or he was instructed to press the Enter key (i.e., the key that had label "RET" on it). After the participant had hit the Enter key, the places of the letters were randomized, and participant was required to look the letters needed to write word "aurinko" again. This procedure was repeated until the participant had written the word ten times. After that, the keyboard disappeared. In total, one participant wrote the word "aurinko" 30 times in total, ten times with each of the keyboard layouts.

3.1.4. Procedure. When a participant arrived in the laboratory, the laboratory and the equipment were introduced to him or her. The participant was asked to sign an informed consent form. The participant was told that the purpose of the experiment was to evaluate three different layouts of on-screen keyboards using gaze direction as the pointing technique and smiling as the selection technique. Then, the prototype was introduced to the participant. The participant wore the prototype and saw live videos from the eye camera and from the scene camera. She or he was instructed to try different head orientations to see how large head movements were possible while still keeping the display visible in the scene camera image. Next, the participant was instructed to try and perform clicks by smiling. After a few successful clicks were produced, the eye tracker was calibrated.

Before conducting the experiment, there was a practice session consisting of 5 trials which were precluded from the actual experiment. In the practice session a keyboard with same equal sized keys was used and participants wrote the word "elokuva" (i.e., movie in English) five times. The

participants were told to perform the tasks as fast and as accurate as possible. Then, there was a short relaxation period before the actual experiment. Then the eye tracker was calibrated and the actual experiment started. The order of the used keyboard layouts was counterbalanced. The eye tracker was re-calibrated during the experiment when needed (i.e., approximately 0.2 times per participant on average). After the participant had completed the task with one keyboard layout, she or he was allowed to rest for a while, if necessary.

At the end of the experiment, participant rated the keyboard layout that was used on three different scales: enjoyableness, clarity, and functionality. For the enjoyableness and clarity ratings, they saw the pictures of each of the keyboard layouts lined up in the computer screen in a randomized order. They were asked to select the layout that was most enjoyable and most clear out of the three layouts. If they could not decide, they were allowed to select the "I don't know" option. The order of enjoyableness and clarity ratings were counterbalanced. For the functionality ratings, participants were allowed to interact with each of the keyboards as long as they wished and they were allowed to write anything they liked. After trying out every layout, they were asked which of the three layouts were the most functional in their opinion. Short (semistructured) interview was conducted after participants had rated the layouts in the three specified scales. Completing the whole experiment took approximately an hour per participant.

3.1.5. Metrics. Text entry rate was measured in cpm. The measure of cpm was chosen instead of the often used wpm measure because only one word at a time was to be written. Similar approach was also chosen by Helmert et al. [31]. Error rates were measured in two different ways: the minimum string distance (MSD) error rate and keystrokes per character (KSPC). The MSD error rate was measured with the improved MSD error rates as suggested by Soukoreff and MacKenzie [41]. MSD error rate is calculated by comparing the transcribed text (i.e., the text that was written by the participant) with the presented text, using minimum string distance. The key strokes per character (KSPC) value indicates how often the participants cancelled characters [41]. In a best case scenario KSPC = 1.00, which indicates that each key press has produced a correct character. However, if a participant makes a correction during text entry (i.e., presses Delete key and chooses another letter), the value of KSPC is larger than one. Thus, KSPC measures the accuracy of the text input process. Note that MSD error rate only compares the transcribed text to the presented text, whereas KSPC takes into account the errors produced.

3.2. Results. Data for statistical analyses were extracted from the moment of entering the first character to the selection of Enter character at the end of the word.

3.2.1. Text Entry Rate. Text entry rate for each of the on-screen keyboard layouts is presented in Figure 5. The overall mean text entry rate ± standard error of the means (S.E.M.s.) for Layout 1 was 14.5 ± 1.7 cpm, 14.9 ± 1.4 cpm for Layout 2,
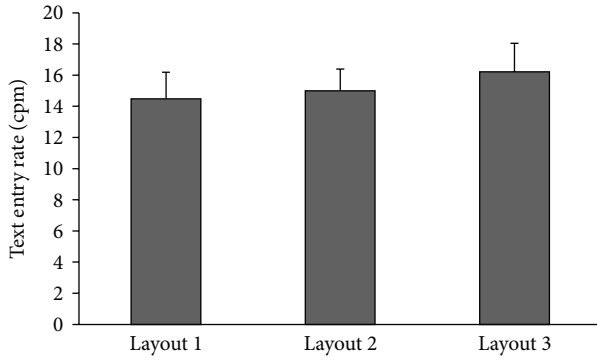
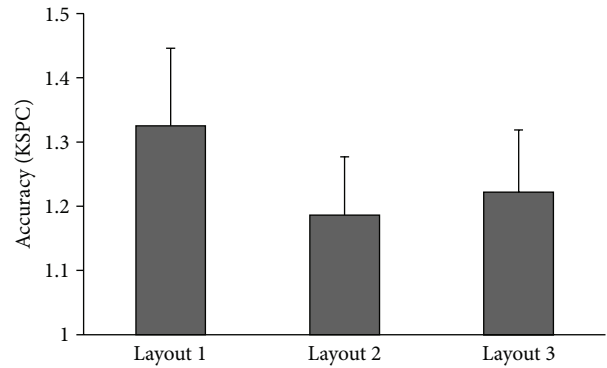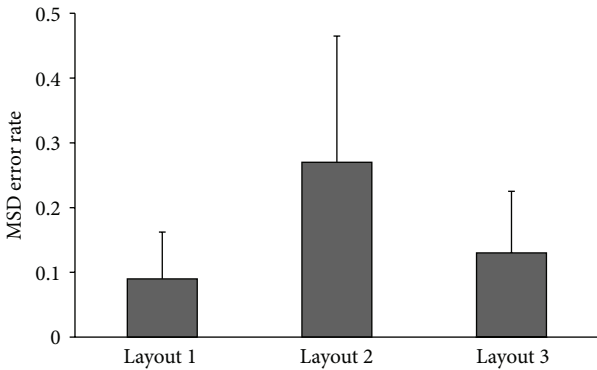FIGURE 5: Text entry rate by cpm.

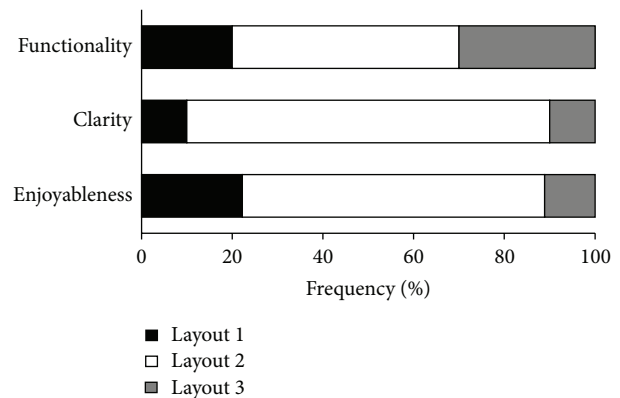FIGURE 7: KSPC values.

FIGURE 6: MSD error rate.

FIGURE 8: Subjective ratings.

and 16.2 ± 1.8 cpm for Layout 3. A one-way repeated measures analysis of variance (ANOVA) did not reveal a statistically significant effect of the layout.

3.2.2. Error Rates. MSD error rates by the layout and task number are presented in Figure 6. The overall mean for MSD error rate ± S.E.M was 0.09 ± 0.07 for Layout 1, 0.27 ± 0.19 for Layout 2, and 0.12 ± 0.09 for Layout 3. A one-way ANOVA did not reveal a statistically significant effect of the layout.

KSCP values by the layout and task number are presented in Figure 7. The overall mean for KSPC ± S.E.M. was 1.3 ± 0.12 for Layout 1, 1.18 ± 0.09 for Layout 2, and 1.22 ± 0.09 for Layout 3. A one-way ANOVA showed a statistically significant effect of the layout $F(2, 18) = 4.2$, $P < 0.05$. The post hoc pairwise comparisons were not statistically significant. Based on the results shown in Figure 7, it seemed, however, that the Layout 2 was the most promising in terms of effectiveness.

3.2.3. Subjective Ratings. The results of the ratings of the layouts can be seen in Figure 8. From the Figure 8, it is clear that the participants clearly preferred Layout 2. One participant considered the layouts equally enjoyable.

Overall, participants liked to use the Face Interface technique for text input. After the experiment they gave spontaneous comments such as "this was cool" or "that was fun." One participant commented that the technique feels natural to use and other participant commented that

the use of smiling for selections felt fun. One participant commented that entering text with the prototype felt very easy. Comments about each of the three keyboard layouts were mainly positive. About Layout 1, we got comments such as "it is easy to see many letters at the same time, because the keys are smaller in the middle [of the keyboard]." This was actually noted by two participants. Further, it was commented that Layout 1 is nicely geometrically shaped which makes it enjoyable, and clear. Of Layout 2, participants gave comments such as "the layout was nice to look at", as well as it was said to be "calm". It was also mentioned that it was easy to seek the letters from this layout. Further, one participant said that the equal sizing of the keys in the layout made entering the text enjoyable. Also, it was mentioned that the text entry felt to be the fastest with Layout 2. One participant commented that with Layout 2, it was easiest to find the letters because of the layout was so pleasant looking. One participant said that he preferred Layout 2 because it did not contain small keys, as the other layouts did. It was also noted that from this layout, it was easiest to select the keys (because of their size). Of Layout 3 comments such as "even though it is not really pretty to look at, nor does it seem to be very clear, it is still the most functional layout" were given. The participants motivated their answers by saying that the keys seemed to be just on the right place, and the sizing of the keys was really good. One participant mentioned that it was easy to find the

letters from Layout 3 because it "directed" one's gaze naturally (i.e., because of the sort of a spiral-shaped layout).

Some negative comments were given as well. For example, Layout 1 was found to be "hideous" by one participant and "awful" by other participant. On Layout 3, it was mainly mentioned that it was "fuzzy" and it did not seem to have clear logic behind how the keys were placed. Layout 2 was actually the only layout that received only positive comments.

4. Text Entry Experiment

First experiment indicated that the Layout 2 would be the most promising one to be used with Face Interface. Thus, based on the results of Experiment 1, a second experiment was run with Layout 2.

4.1. Methods

4.1.1. Participants. Twelve (4 male and 8 female) able-bodied volunteers participated in the experiment. Their mean age was 24 years (range 18–37). All the participants had normal vision. They were native Finnish speakers and they were novices in using any gaze-based system for controlling computers and same for using Face Interface. However, they were pretty experienced users of computer mouse; that is, their average experience in using mouse was 14.5 years (range 9–20).

4.1.2. Apparatus. The same prototype device was used as in Experiment 1. The keyboard that was used, was Layout 2 selected from Experiment 1 study. Again, the places of letters changed during the experiment to prevent the possible advantage that mouse may have if the letter placement is known (e.g., QWERTY). The mouse that was used was an optical mouse (Logitech Mouse m100). Mouse speed was set to medium level.

4.1.3. Experimental Task. The experimental task was the same as in Experiment 1.

4.1.4. Procedure. When the participant arrived in the laboratory, the laboratory and the equipment were introduced to him/her. The participant signed an informed consent form before the experiment. The participant was told that the aim of the experiment was to compare typing with mouse to the typing with Face Interface prototype using on-screen keyboard. The order of the used pointing device was counterbalanced so that half of the participants started with Face Interface and the other half started with the mouse. Participants performed the experimental task ten times, then there were a short pause and they performed the experimental task again ten times. After participant had conducted the experimental task with one pointing device she or he rated the experience with six nine-point bipolar scales. The scales were general evaluation (i.e., from bad to good), difficulty (i.e., from difficult to easy), speed (i.e., from slow to fast), accuracy (i.e., from inaccurate to accurate), enjoyableness (i.e., from unpleasant to pleasant), and efficiency (i.e., from
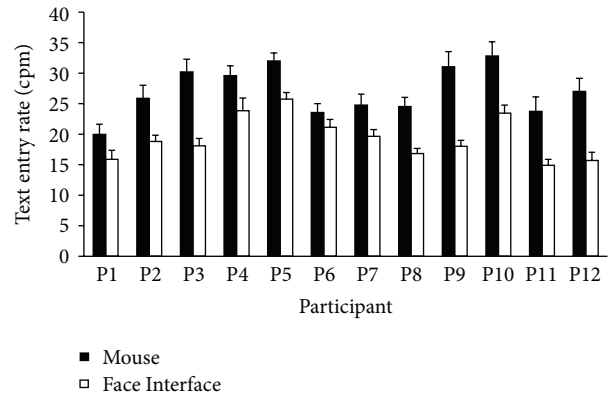

FIGURE 9: Text entry rate for both pointing devices divided by participant.

inefficient to efficient). The scales varied from −4 (e.g., bad experience) to +4 (e.g., good experience), and 0 represented a neutral experience (e.g., not slow nor fast). Then the same procedure was repeated with the other pointing device. After participant had completed the task with both of the pointing devices, a short interview was completed. Conducting the whole experiment took approximately 60 minutes.

4.1.5. Metrics. Same metrics were used as in Experiment 1.

4.2. Results

4.2.1. Text Entry Rate. Text entry rate for both pointing devices by participant is presented in Figure 9. The overall mean text entry rate ± S.E.M.s. for Face Interface was 19.4 ± 1.9 cpm and 27.1 ± 2.8 cpm for the mouse. A one-way ANOVA for the pointing device showed a statistically significant effect of the pointing device $F(1, 11) = 66.8$, $P < 0.001$. A one-way ANOVA for Face Interface showed a statistically significant effect of session $F(1, 11) = 4.9$, $P < 0.05$. For the mouse, the effect of session was not statistically significant.

Figure 10 shows minimum and maximum values for every participant by pointing device. The overall mean maximum values for Face Interface was 33.4 cpm and 48.3 cpm for mouse.

4.2.2. Error Rates. MSD error rate for both pointing devices by participant is presented in Figure 11. The overall mean MSD error rate ± S.E.M was 0.12 ± 0.09 for Face Interface and 0 ± 0 for mouse. A one-way ANOVA for the pointing device showed a statistically significant effect of pointing device $F(1, 11) = 18.6$, $P < 0.01$. One-way ANOVAs for the session were not statistically significant.

KSPC values for both pointing devices by participant are presented in Figure 12. The overall mean KSPC ± S.E.M was 1.1 ± 0.05 for Face Interface and 1.0 ± 0.002 for mouse. The one-way ANOVA showed a statistically significant effect of pointing device $F(1, 11) = 45.9$, $P < 0.001$. One-way ANOVAs for session were not statistically significant.
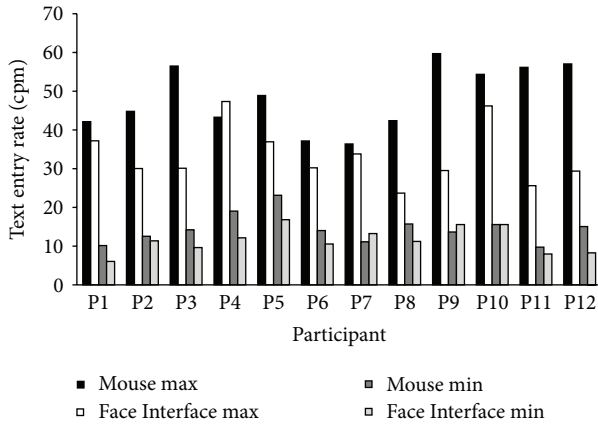
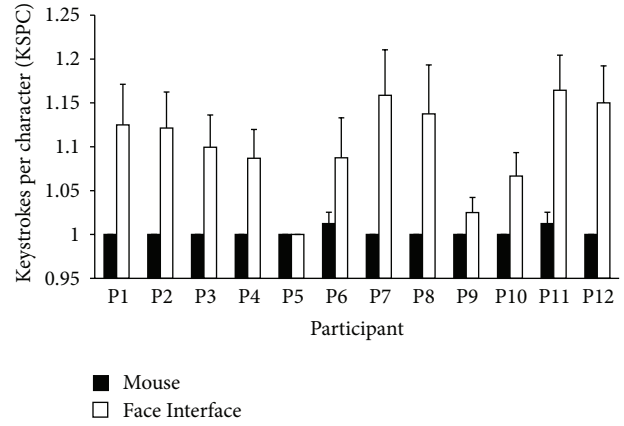Figure 10: Maximum and minimum values for mouse and Face Interface.
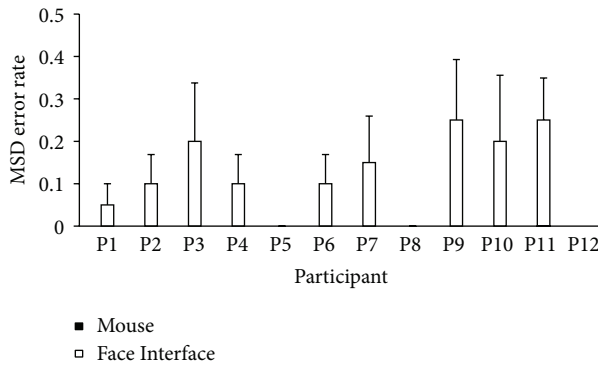

Figure 12: KSPC values by participant.

Figure 11: MSD error rate by participant.

Table 1: Mean ranks of ratings of both techniques.

	Face Interface	Mouse	Mann-whitney U	Significance
General evaluation	10.2	14.8	44.5	n.s.
Difficulty	8.6	16.4	25.0	$P < 0.01$
Speed	9.4	15.6	35.0	$P < 0.05$
Accuracy	7.8	17.2	15.5	$P < 0.01$
Efficiency	9.7	15.3	38.0	n.s.
Enjoyableness	10.3	14.8	43.5	n.s.

4.2.3. Subjective Ratings. Mann-Whitney U test was used for pairwise comparisons because it is commonly used for comparing two independent samples with each other. Mean ranks of subjective ratings are presented in Table 1.

After the experiment a short interview was conducted. First, participants were asked if they could see that the prototype would be used wider in future. All the participants answered yes to that question. They justified their answers by stating, for example, that the use of the prototype was quite easy and the prototype was interesting to use. Participants liked especially the fact that gaze was used for pointing because it was easy and even natural to use. They also

mentioned that the smiling felt natural as the selection technique but it required some time to get to use to it.

Participants were also asked that in what kind of task they would think that the Face Interface prototype could be used in the future. The answers varied between participants. However, some common points could also be found. Nine participants out of 12 answered that disabled people could use this prototype for communicating with other people. They saw the prototype as a promising concept for the disabled people who cannot move their hands. There were many other ideas as well. For example, Face Interface could be used as an alternative for remote control while watching television. Reasoning for using Face Interface as remote control could be that people are getting lazy, and thus, in future, even moving a remote-control with hand may require too much effort and thus, Face Interface could be a potential solution. It was also suggested that this kind of technique could be used in public when interacting with large interactive billboards or tourists could use it when interacting with a map in a strange city. The map could, for example, show sights that are near the place that user is looking and user could then select an attraction that she or he would like to know more. On a similar topic, a lecturer could use this to emphasis some specific point to students from the slides. Further, prototype could be used in a loud spaces, where talking would be impossible, for example, in a factory. Idea that Face Interface could be used while driving a car also came up. That is, there would be a transparent screen in the windshield and the driver could do something with it. One task that was mentioned that Face Interface could be used while playing video games or children could use it when playing.

On the other hand, there were fewer ideas about where Face Interface could not be used. Participants suggested that Face Interface could not be used in tasks that need really high accuracy or where the result is not shown. One such task that was mentioned was the use of PIN code in ATM's. Overall, participants found more tasks that were suitable for Face Interface and less that were not. Four of the participants did not come up with anything where Face Interface could not be used.

The last interview question was a word association task which was roughly based on the semantic differential method [34]. Task was to list words that came to their mind when using Face Interface. Participants listed many different kinds of words. To the glasses it was linked words such as *eye glasses, sun glasses, gaze,* and *eye pointer*. About the technique that words such as *fast to absorb, handy, new, useful, fun, advanced, science fiction, future, modern, interesting, futuristic, challenge,* and *21st century* were mentioned. Some negative words were mentioned as well, such as *requires focusing, difficult, and troublesome*. Again, more positive than negative comments were mentioned.

5. Discussion

The aim of the layout selection experiment was to find out which of the three designed keyboard layouts would be the most promising one to be used with Face Interface and in the actual typing experiment. Because the statistical analyses did not reveal significant differences in the text entry rate (cpm) and MSD error rate between the three different keyboard layouts, it suggests that text entry was neither significantly faster nor significantly erroneous with any of the keyboard layouts. In general, avoiding small keys near the edges of the screen seems to have successfully compensated for the previously found problems in selecting objects mainly at the corners of the display as well as at the left and right edges of the display [14]. With the current keyboard layout designs that problem was overcome. Subjective ratings showed that participants preferred the design of Layout 2. This preference was supported by the results of KSPC. Thus, both subjective and objective data led us to choose the Layout 2 for the subsequent typing experiment.

The text entry experiment showed that the overall mean text input speeds with Face Interface and mouse were 20 cpm and 27 cpm, respectively. The ANOVA revealed that entering text with the mouse was significantly faster than entering text with Face Interface. The slower text entry speed of Face Interface can be explained with the fact that participants did not have any previous experience with Face Interface; they only had a practice of approximately 5 minutes prior to the experiment. With mouse, however, they had experience over 15 years on average. Thus, from this point of view the found difference in text entry speed is relatively small. In addition to the radical differences in earlier experiences in using these two interaction methods there can be other explaining factors. For example, there is evidence that during eye typing users, especially novices, tend to gaze at the results of typing which can slow down the typing speed [26, 42]. Also the randomization might have had an effect for this, because when eyes are used as both: input and observation method, it might cause slowness for typing because participant cannot look for the next character before she/he has typed the current one. On the other hand, when using mouse (i.e., hand) as an input method, then eyes are free for searching the next character while cursor is still on the previous character.

Further, Figure 10 shows minimum and maximum values for mouse and Face Interface for text entry speed for every participant. One interesting finding for Face Interface was that one participant actually had a maximum text entry speed value of 47.4 cpm which was even higher than that participant's maximum speed with mouse, that is, 43.3 cpm. On the other hand, it can be seen from the Figure 10 that the slowest values for Face Interface and mouse are on a similar level. In other studies where text entry with gaze-based solutions has been compared to mouse, the results have shown, similarly as in the current experiment, that the mouse has been faster in terms of speed [27, 28]. For GazeTalk, for example, the writing speed with mouse was reported to be approximately 7.5 wpm (for Danish text) and 15 cpm (for Japanese text) [28].

We note especially that the comparison between mouse and Face Interface must be done cautiously because Face Interface uses two different modalities that are not traditionally (in contrast to hands) used neither for controlling computers nor manipulating objects. First, eye, for example, is primarily a perceptual organ [43]. People move their eyes involuntarily towards new stimuli which may cause problems when using eye gaze as the pointing method. It is known that there are some problems that have an effect when gaze direction is used for pointing, for example, inaccuracy of the eye trackers. Second, even if a person thinks that his or her gaze is fixated on a target; his eyes are actually actively moving [44]. This movement is known as fixation jitter [45] and it can cause inaccuracy in eye tracking. Third, it is in fact a rather new invention to use facial muscle movements as the selection method [3, 11, 46]. Some problems may occur, for example, from the fact that some people may find voluntarily control of facial muscles to be difficult [47]. Thus, from the fact that there are some possible difficulties in both techniques, it could be assumed that an integrated use of voluntary gaze direction and facial muscle movements for interacting with computers can be challenging at first. For example, there may be a delay in selecting the object when the cursor is inside an object because the pointing and selection are operated with two different modalities. Of course with such little amount of practice as in the current study, the smooth combination of these two modalities to a fast interaction techniques is not possible. Based on the above discussion, it seems quite natural that error rate with Face Interface is much higher than error rate with mouse.

Even though this experiment was a different from a traditional text entry experiment in a sense that the aim was to compare two pointing and selection techniques, some comparative results to other techniques can be given. Because text entry rate in most of the other studies have been reported using wpm value [27, 28, 30, 48], converting values from cpm to wpm in the current study gives an impression of the writing speed as compared to other systems. Of course, converting the cpm to wpm might not be the most reliable in this case because participants wrote only one word at a time in the current experiment. However, the wpm values give the possibility to compare the results to other studies. Wpm values in the present study were for Face Interface 4 wpm and 5 wpm for the mouse. It is noteworthy to mention that, for example, first time users of Dasher wrote text with an average speed of 2.5 wpm after the first session [29]. In

text entry studies with gaze gestures as the input method, a bit lower text entry speed has been achieved: approximately 2.3 wpm after the first session with EyeWrite [49]. Further, Porta and Turina [50] reported that their novice participants wrote one phrase that included 13 characters in 188.5 seconds which corresponds roughly to 4 cpm. For the GazeTalk [28], the grand mean text entry speed was 6.22 wpm for Danish text and 11.71 cpm for Japanese text. We note, again, that comparison to purely gaze based studies can be problematic because these do not require any other modality integration for functional user interface. In a different multimodal technique where the object was pointed by gaze and the selection was made utilizing signals from the brain, the text entry speed was found to be 9.1 cpm [51].

Error rate analysis using MSD revealed that participants made only few errors. This can be seen from the overall mean MSD error rate of 0.12 for Face Interface and 0.0 for the mouse. For the GazeTalk, the MSD error rate was 1.09 [28], and for the Dasher the MSD error rate was approximately 10. When interacting with on-screen keyboard using dwell time as the object selection method a MSD value of 1.28 has been reported [48]. The only possible comparison of the KSPC results is one with eye typing experiments. That is because KSPC as the measure is intended for such a text entry systems in which keys are pressed (i.e., keystrokes are created) which makes KSPC unsuitable for measuring the performance of, for example, for Dasher. The overall mean KSPC value for Face Interface was 1.1 and 1.0 for mouse. Majaranta et al. [48] reported the grand mean for KSPC of 1.09 in the first session within a longitudinal study which corresponds to this study because that session lasted approximately similar time period as the current experiment. Helmert et al. [31] reported KSPC values from 1.00 (dwell time of 700 ms) to 1.18 (dwell time of 350 ms) when one word at the time was written. The results of KSPC are very promising for Face Interface and they compare very well with other gaze based text entry techniques.

Quite naturally, the re-randomization of the letters after each written word had an effect to the text entry speed because participants had to find the correct characters time after time and the participants could not rely on their earlier experience about places of the keys as is the case, for example, in a QWERTY layout. However, the present aim was to study and compare two different pointing devices by specifically excluding the effects of letter placement. For this purpose, letter randomization was necessary. Even with this arrangement, the text entry speed in the current study compared well with other studies.

The ratings of the two techniques revealed that participants rated the use of mouse as more accurate, faster and easier than the use of Face Interface. In the ratings of general evaluation, efficiency and enjoyableness there were no statistically significant differences between the two techniques. In a way, this is a positive finding because it indicates that participants rated the use of the mouse and the prototype as equal with these three scales. Current ratings compare well with previous studies where similar techniques have been compared to mouse. For example, Surakka et al. [3] reported that participants rated the gaze combined with EMG technique as faster and less accurate and more difficult to use than the computer mouse. Further, San Agustin et al. [2] found that their participants rated the gaze pointing combined with facial EMG as faster but less accurate to use than the mouse pointing. Thus, current ratings are similar to those reported earlier. Tuisku et al. [14] reported ratings that were on the same level as in current study. An interesting finding from the interviews was that even though the participants used the Face Interface prototype only for short period of time and were novices in using it, they were still able to name many possibilities in which Face Interface could possibly be used in future. Thus, it seems that they were able to see the potential of Face Interface.

The present results showed that entering text with the prototype is possible. This experiment revealed also possible designs for on-screen keyboard layout. Even though the results showed that mouse was faster in terms of text entry speed, the results were promising for future text entry with Face Interface. Present results also confirmed that smiling can be used as the selection technique with Face Interface which can offer more possibilities to use it in the future. That is, the user is able to choose the selection technique he/she would like to use from the three possible options (i.e., frowning, smiling, and raising the eyebrows). It is noteworthy to mention that for the people who the use of speech is not possible, interactive conversation is seen as tolerable when it achieves a minimum rate of 3 wpm [52]. Face Interface met these minimum requirements even with the randomized keyboard. To further improve the technique, next steps would be to decide the letter placement in the keyboard based on the used language's most common characters and to run a longitudinal study in order to see the actual text entry rate that can be achieved with Face Interface. Although the current results with the prototype were not superior to the mouse the results are encouraging for further research and development of face based technologies.
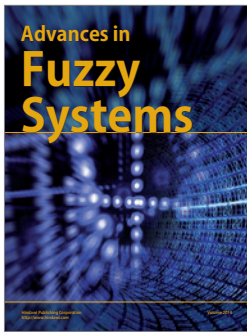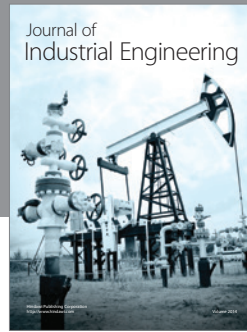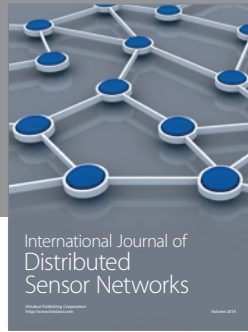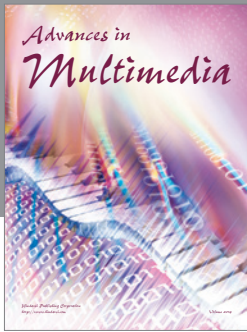
Acknowledgments

References

[1] C. A. Chin, A. Barreto, J. G. Cremades, and M. Adjouadi, "Integrated electromyogram and eye-gaze tracking cursor control system for computer users with motor disabilities," *Journal of Rehabilitation Research and Development*, vol. 45, no. 1, pp. 161–174, 2008.

[2] J. San Agustin, J. C. Mateo, J. P. Hansen, and A. Villanueva, "Evaluation of the potential of gaze input for game interaction," *PsychNology Journal*, vol. 7, no. 2, pp. 213–236, 2009.

[3] V. Surakka, M. Illi, and P. Isokoski, "Gazing and frowning as a new human-computer interaction technique," *ACM Transactions on Applied Perceptions*, vol. 1, no. 1, pp. 40–56, 2004.

[4] V. Surakka, P. Isokoski, M. Illi, and K. Salminen, "Is it better to gaze and frown or gaze and smile when controlling user interfaces?" in *Proceedings of the 11th International Conference on Human-Computer Interaction (HCI '05)*, CD-ROM, p. 7, July 2005.

[5] B. Ashtiani and I. S. MacKenzie, "BlinkWrite2: an improved text entry method using eye blinks," in *Proceedings of the ACM Symposium on Eye-Tracking Research and Applications (ETRA '10)*, pp. 339–346, March 2010.

[6] A. Sesin, M. Adjouadi, M. Cabrerizo, M. Ayala, and A. Barreto, "Adaptive eye-gaze tracking using neural-network-based user profiles to assist people with motor disability," *Journal of Rehabilitation Research and Development*, vol. 45, no. 6, pp. 801–818, 2008.

[7] H. Heikkilä and K. J. Räihä, "Simple gaze gestures and the closure of the eyes as an interaction technique," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*, pp. 147–154, March 2012.

[8] A. J. Fridlund, *Human Facial Expression: An Evolutionary View*, Academic Press, San Diego, Calif, USA, 1994.

[9] A. Królak and P. Strumiłło, "Eye-blink detection system for human-computer interaction," *Universal Access in the Information Society*, vol. 11, no. 4, pp. 409–219, 2012.

[10] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of Experimental Psychology*, vol. 47, no. 6, pp. 381–391, 1954.

[11] V. Rantanen, P. H. Niemenlehto, J. Verho, and J. Lekkala, "Capacitive facial movement detection for human-computer interaction to click by frowning and lifting eyebrows," *Medical and Biological Engineering and Computing*, vol. 48, no. 1, pp. 39–47, 2010.

[12] V. Rantanen, T. Vanhala, O. Tuisku et al., "A wearable, wireless gaze tracker with integrated selection command source for human-computer interaction," *IEEE Transactions on Information Technology in BioMedicine*, vol. 15, no. 5, pp. 795–801, 2011.

[13] O. Tuisku, V. Surakka, Y. Gizatdinova et al., "Gazing and frowning to computers can be enjoyable," in *Proceedings of the 3rd International Conference on Knowledge and Systems Engineering (KSE '11)*, pp. 211–218, October 2011.

[14] O. Tuisku, V. Surakka, T. Vanhala, V. Rantanen, and J. Lekkala, "Wireless face interface: using voluntary gaze direction and facial muscle activations for human-computer interaction," *Interacting with Computers*, vol. 24, no. 1, pp. 1–9, 2012.

[15] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction," *Human-Computer Interaction*, vol. 7, no. 1, pp. 91–139, 1992.

[16] S. A. Douglas and A. K. Mithal, "Effect of reducing homing time on the speed of a finger-controlled isometric pointing device," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*, pp. 411–416, April 1994.

[17] C. Ware and H. H. Mikaelian, "An evaluation of an eye tracker as a device for computer input," in *Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface (CHI '87)*, pp. 183–188, Ontario, Canada, April 1987.

[18] P. Majaranta and K. J. Räihä, "Twenty years of eye typing: systems and design issues," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '02)*, pp. 15–22, New Orleans, La, USA, March 2002.

[19] F. E. Sandnes and A. Aubert, "Bimanual text entry using game controllers: relying on users' spatial familiarity with QWERTY," *Interacting with Computers*, vol. 19, no. 2, pp. 140–150, 2007.

[20] K. J. Räihä and S. Ovaska, "An exploratory study of eye typing fundamentals: dwell time, text entry rate, errors, and workload," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, pp. 3001–3010, May 2012.

[21] O. Špakov and P. Majaranta, "Scrollable keyboards for casual typing," *PsychNology Journal*, vol. 7, no. 2, pp. 159–173, 2009.

[22] X. Bi, B. A. Smith, and S. Zhai, "Quasi-Qwerty soft keyboard optimization," in *Proceedings of the 28th Annual CHI Conference on Human Factors in Computing Systems (CHI '10)*, pp. 283–286, April 2010.

[23] I. S. MacKenzie and S. X. Zhang, "Design and evaluation of a high-performance soft keyboard," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '99)*, pp. 25–31, May 1999.

[24] D. Li, J. Babcock, and D. J. Parkhurst, "openEyes: a low-cost head-mounted eye-tracking solution," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '06)*, pp. 95–100, San Diego, Calif, USA, March 2006.

[25] W. J. Ryan, A. T. Duchowski, E. A. Vincent, and D. Battisto, "Match-moving for area-based analysis of eye movements in natural tasks," in *Proceedings of the ACM Symposium on Eye-Tracking Research and Applications (ETRA '10)*, pp. 235–242, March 2010.

[26] H. Aoki, J. P. Hansen, and K. Itoh, "Learning to interact with a computer by gaze," *Behaviour and Information Technology*, vol. 27, no. 4, pp. 339–344, 2008.

[27] J. P. Hansen, A. S. Johansen, D. W. Hansen, K. Itoh, and S. Mashino, "Command without a click: dwell time typing by mouse and gaze selections," in *Human-Computer Interaction (INTERACT '03)*, M. Rauterberg, M. Menozzi, and J. Wesson, Eds., pp. 121–128, IOS Press, Amsterdam, The Netherlands, 2003.

[28] J. P. Hansen, K. Tørning, A. S. Johansen, K. Itoh, and H. Aoki, "Gaze typing compared with input by head and hand," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '04)*, pp. 131–138, San Antonio, Tex, USA, March 2004.

[29] D. J. Ward and D. J. C. MacKay, "Artificial intelligence: fast hands-free writing by gaze direction," *Nature*, vol. 418, no. 6900, p. 838, 2002.

[30] O. Tuisku, P. Majaranta, P. Isokoski, and K. J. Räihä, "Now Dasher! Dash away!: longitudinal study of fast text entry by eye gaze," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '08)*, pp. 19–26, March 2008.

[31] J. R. Helmert, S. Pannasch, and B. M. Velichkovsky, "Influences of dwell time and cursor control on the performance in gaze driven typing," *Journal of Eye Movement Research*, vol. 2, no. 4, pp. 1–8, 2008.

[32] P. Majaranta, I. S. MacKenzie, A. Aula, and K. J. Räihä, "Effects of feedback and dwell time on eye typing speed and accuracy," *Universal Access in the Information Society*, vol. 5, no. 2, pp. 199–208, 2006.

[33] D. A. Norman and D. Fisher, "Why alphabetic keyboards are not easy to use: keyboard layout doesn't much matter," *Human Factors*, vol. 24, no. 5, pp. 509–519, 1982.

[34] M. M. Bradley, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[35] C. E. Osgood, "The nature and measurement of meaning," *Psychological Bulletin*, vol. 49, no. 3, pp. 197–237, 1952.

[36] R. Menzies, A. Waller, and H. Pain, "Peer interviews: an adapted methodology for contextual understanding in user-centred design," in *Proceedings of the 13th International ACM SIGAC-CESS Conference on Computers and Accessibility (ASSETS '11)*, pp. 273–274, Dundee, UK, October 2011.

[37] G. Bradski and A. Kaehler, *Learning Opencv: Computer Vision with the Opencv Library*, O'Reilly Media, Sebastopol, Calif, USA, 2008.

[38] D. Li, D. Winfield, and D. J. Parkhurst, "Starburst: a hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 1–8, San Diego, Calif, USA, June 2005.

[39] V. Rantanen, J. Verho, J. Lekkala, O. Tuisku, V. Surakka, and T. Vanhala, "The effect of clicking by smiling on the accuracy of head-mounted gaze tracking," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*, pp. 345–348, March 2012.

[40] P. O. Kristensson and S. Zhai, "Relaxing stylus typing precision by geometric pattern matching," in *Proceedings of the 10th international conference on Intelligent user interfaces (IUI '05)*, pp. 151–158, San Diego, Calif, USA, January 2005.

[41] R. W. Soukoreff and I. S. MacKenzie, "Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*, pp. 113–120, April 2003.

[42] R. Bates, "Have patience with your eye mouse! Eye-gaze interaction with computers can work," in *Proceedings of the 1st Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT '02)*, pp. 33–38, March 2002.

[43] S. Zhai, "What's in the eyes for attentive input," *Communications of the ACM*, vol. 46, no. 3, pp. 34–39, 2003.

[44] R. J. K. Jacob, "The use of eye movements in human-computer interaction techniques: what you look is what you get," *ACM Transactions on Information Systems*, vol. 9, no. 2, pp. 152–169, 1991.

[45] M. Ashmore, A. T. Duchowski, and G. Shoemaker, "Efficient eye pointing with a fisheye lens," in *Proceedings of Graphics Interface 2005 (GI '05)*, pp. 203–210, Ontario, Canada, May 2005.

[46] A. B. Barreto, S. D. Scargle, and M. Adjouadi, "A practical EMG-based human-computer interface for users with motor disabilities," *Journal of Rehabilitation Research and Development*, vol. 37, no. 1, pp. 53–64, 2000.

[47] R. W. Levenson, P. Ekman, and W. V. Friesen, "Voluntary facial action generates emotion-specific autonomic nervous system activity," *Psychophysiology*, vol. 27, no. 4, pp. 363–384, 1990.

[48] P. Majaranta, U. K. Ahola, and O. Špakov, "Fast gaze typing with an adjustable dwell time," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, pp. 357–360, Boston, Mass, USA, April 2009.

[49] J. O. Wobbrock, J. Rubinstein, M. W. Sawyer, and A. T. Duchowski, "Longitudinal evaluation of discrete consecutive gaze gestures for text entry," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '08)*, pp. 11–18, Savannah, Ga, USA, March 2008.

[50] M. Porta and M. Turina, "Eye-S: a full-screen input modality for pure eye-based communication," in *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '08)*, pp. 27–34, March 2008.

[51] X. Yong, M. Fatourechi, R. K. Ward, and G. E. Birch, "The design of point-and-click system by integrating a self-paced brain-computer interface with an eye-tracker," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 4, pp. 590–602, 2011.

[52] J. J. Darragh and I. H. Witten, *The Reactive Keyboard*, Cambridge University Press, New York, NY, USA, 1992.