BC(eye): Combining Eye-Gaze Input with Brain-Computer Interaction

Roman Vilimek[1] and Thorsten O. Zander[2]

[1] Siemens AG, Corporate Technology, User Interface Design
Otto-Hahn-Ring 6, 81730 Munich, Germany
[2] Technische Universität Berlin, Chair of Human-Machine Systems, Team PhyPA
Franklinstr. 28-29, 10587 Berlin, Germany
roman.vilimek@siemens.com, thorsten.zander@mms.tu-berlin.de

Abstract. Gaze-based interfaces gained increasing importance in multimodal human-computer interaction research with the improvement of tracking technologies over the last few years. The activation of selected objects in most eye-controlled applications is based on dwell times. This interaction technique can easily lead to errors if the users do not pay very close attention to where they are looking. We developed a multimodal interface involving eye movements to determine the object of interest and a Brain-Computer Interface to simulate the mouse click. Experimental results show, that although a combined BCI/eye-gaze interface is somewhat slower it reliably leads to less errors in comparison to standard dwell time eye-gaze interfaces.

Keywords: Brain-Computer Interaction, BCI, multimodal, eye tracking, eye-controlled applications.

1 Motivation

With the idea of "eyes as output" Richard Bolt introduced eye-gaze input to facilitate human-computer interaction already in 1982 [1]. Since then, numerous studies were conducted on how to utilize the user's eye movements for working with graphical user interfaces (GUIs). These studies have shown that eye tracking can be a successful means of controlling the mouse cursor and more (cf. for instance [2]). Since the first activities in this field, gaze control has become an accepted input modality. It has proven to be very intuitive, fast and especially useful in hands-free operation scenarios [3-5].

However, whereas moving the mouse cursor with eye movements is quite intuitive, it is not that easy to find a good mechanism for performing the click operation. Ideas like using eye blinks for activation were rejected already in the first studies on this subject as it is impossible for the user to exercise precise enough control over the blink reflex [6]. Most solutions are based on dwell times, i.e. the user has to fixate an item for a pre-defined period of time in order to activate it. This technique has to face the inherent problem of finding the optimal dwell time. If the dwell time is too short,

C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2009, LNCS 5615, pp. 593–602, 2009.

click events will be carried out unintentionally leading to errors. If it is too long, fewer errors will be made but more experienced users will get annoyed.

In our study we developed and evaluated a completely different approach: using a Brain-Computer Interface (BCI) to confirm object selections made by eye tracking. Brain-computer interaction and eye-gaze input can be regarded as complementary modalities in the respect that they compensate for each other's disadvantages. The combination overcomes the BCI drawback of having problems differentiating between more than two commands because only one activation thought needs to be tracked reliably. If this activation works properly, a new solution to perform click operations in gaze-based interfaces can be established by providing an explicit, nevertheless not overt visible, command under complete user control.

2 Eye-Gaze Input and BCI

This section outlines general properties of eye-tracking and BCI as input modalities. In detail, it describes the advantages and some of the major challenges when using these technologies individually and describes how a combination of the two in a multimodal interface may leverage their potential.

2.1 Gaze-controlled User Interfaces

Eye-gaze interaction can be a convenient and – with certain restrictions – a natural addition to human-computer interaction. The eye gaze of humans is basically an indicator for a person's attention over time [7]. For human-computer interaction this means that the mouse cursor and visual focus usually correspond to each other which implies an intuitive substitution of the conventional mouse control by eye movements.

However, this rule does not always apply. The design of gaze-based systems has to consider unintentional fixations and sporadic dwellings on objects that typically occur during visual search or when people are engaged in demanding mental activity (cf. [8]). This fact is known as the "Midas Touch" problem: Although it may be helpful to simply look at an object and have the corresponding actions occur without further activity, it soon becomes annoying as it gets almost impossible to let the gaze rest anywhere without issuing a command [9]. The problem directly points to the challenge of defining the mouse click operation in gaze-controlled environments.

In past research dwell-time based solutions proved to be the best technique that can establish an even faster interaction process than using a mouse [10, 11]. However, choosing a dwell time duration is always a trade-off between speed and accuracy. Furthermore, a well defined feedback informing the user about the current state of the activation progress is crucial, but can be difficult to design [12]. Even with adaptive algorithms, like e.g. shortening the dwell time period with growing user experience, one major problem remains: The system can not know whether the user fixates a command button for a long time because he wants to trigger an action or because the description is difficult to read, he reflects about the corresponding system action, he tries to understand the meaning of a complex icon… it is simply not possible to find a perfect relation between gaze duration and user intention. Thus, it will be beneficial to replace this implicit way of issuing a command with a more direct and controllable user action.

A different solution for the Midas Touch problem consists of providing the user with a manual control key for activation, a so-called gaze button (for details cf. [13]). This activation key has the same functionality like a mouse button and allows to click an object which has been selected by gaze tracking. The gaze button offers a greater amount of control for the users and reduces activation errors. It is especially useful for people with muscle diseases like muscular dystrophy who may not be able to fine-control mouse movements but still can perform some simple movements. The disadvantage is, of course, that systems with a gaze button are no longer hands-free.

2.2 Brain-Computer Interfaces

Brain-computer interfaces provide a unidirectional channel from human to computer without the involvement of any muscular activity. Contrary to most standard EEG analyses, BCIs isolate feature-patterns from online EEG data. Thoughts or intentions of activity as well as conscious or unconscious information processing evoke specific neuronal activity that can be detected as specific patterns in EEG. These patterns are extracted from the very noisy EEG signal by filtering techniques and methods of machine learning. As BCIs work in real time, no averaging over many trials is possible, so the challenge is to find the relevant pattern in only one single trial. Once the signal of interested is detected it can be used in two ways. Either BCIs allow the user to deliberately control system properties by brain signals. These so-called active BCIs, which enable the users to perform direct commands, are typically operated by forming the intention of a motor movement like imagining to move the right hand. Or BCIs work by recognizing specific mental states of the user like high workload peaks with the goal that the intelligent systems can adapt to the user's current needs.

At present most BCI research focuses on solutions for the medical care sector where significant contributions were made in assisting people with massively restricted motor abilities [14, 15]. These applications can be regarded as specialized high-end solutions for a relatively small number of users. Access to a mass market will be possible most likely for gaming devices. Having to wear head-mounted equipment and a relatively low degree of accuracy may be a less important factor when establishing completely new game experiences.

However, most applications relying on active control suffer from the small number of available commands. BCIs typically can only differentiate between two commands as they analyze whether an imaginary movement is reflected in the right or left hemispherical primary motor cortex. Hence, the highest potential for BCIs lies in the realm of multimodal environments where one or two explicit commands can suffice and may significantly increase the overall system performance.

In order to get the EEG data needed for using a BCI, electrodes are positioned on the user's scalp. This is quite time-consuming; typically it takes about 20-30 minutes for 32 electrodes. Additional time is required to adjust the BCI itself. As the EEG signal varies considerably not only between users but also within one person at different times a classifier needs to be trained before every usage. Because of these side conditions, current BCIs are incapable for usage outside the laboratory.

But researchers and industry are already working on new solutions. With the advance of more efficient algorithms, the training effort reduces steadily. Several groups are concentrating on the development of dry electrodes and mobile EEG systems that allow usage in a broader range of environments (cf. [16]). Last but not least the high interest of the gaming industry in brain-controlled devices boosts development. Thus, although there are still difficulties, BCIs are a promising technology for HCI applications [17].

2.3 Combining Eye-Gaze Input and BCI

In the research reported here we decided to evaluate Brain-Computer Interaction as a supportive modality to eye-gaze input. Mouse movements, i.e. the selection of a target object on a GUI, are mapped to eye movements. A mouse click, i.e. the activation of a selected object, will only be carried out if the user fixates on an object and imagines a special movement of both hands at the same time.

Obviously, there also exist other options in multimodal environments that are more stable, like speech or gesture control. However, a BCI offers the advantage of hands-free operation and does not demand any additional muscular activity or overt command. The combination of eye-gaze input and a BCI is especially suitable for demanding working environments like sterile operating conditions, when wearing protective clothing or if the working condition severely restricts the range of body movement. In contrast to voice control, other people in the same room do not get distracted and there is no interference with human-to-human verbal communication.

The eye movements themselves, however, are quite a challenge for an EEG based BCI. The eye is a powerful dipole that disturbs the detection of the much weaker brainwaves. This means for the activation thought a pattern needs to be found that poses less weight on the frontal electrodes. The recognition algorithm needs to be able to deal with the noise produced by eye movements.

In this investigation, we did not limit the scope to the context of assisting physically challenged people but have tried to learn more about the potential of a multimodal BCI/eye-gaze interface. This has several implications. First, anybody should be able to use the system after a short training session. Therefore, contrary to most experiments on BCIs, our participants had no working experience with a BCI before. Second, the replacement of dwell times by BCI need to prove to be a better solution to the Midas Touch problem by yielding lower error rates in the selection tasks. Task completion times should also be lower or at least comparable. Finally, using the new interface must at least be as convenient as a the gaze-based interface. Thus, the workload associated with BCI/eye-gaze interface may not be higher and using it should be preferred in comparison to conventional eye-tracking interfaces.

3 Experimental Evaluation of BC(eye)

This experiment compares a BCI-based activation of targets in an eye-controlled selection task against two conventional dwell time solutions with different activation latencies. The study aims to determine the degree to which BCI can match or even exceed dwell time activation in respect to effectiveness, efficiency, and demands on cognitive resources in “clicking” the target stimulus.

Task difficulty in the selection task was varied by showing either simple visual stimuli with only a few random characters or by presenting more complex visual stimuli featuring a higher number of characters. Two different dwell times, short and long, were chosen for a better representation of the range of typical interaction situations with gaze-controlled applications. Assuming that signal extraction and pattern recognition of current BCIs still need a substantial minimal presence duration of the activation thought and that processing these signals takes additional time, it does not seem very likely that subjects will be able to complete tasks with the BCI faster. The question of interest here is whether they are significantly slower with a BCI than with dwell times.

The activation thought via BCI is a conscious, explicit command – in contrast to the implicit commands of dwell time solutions. Thus, the error rate in the BCI condition should be substantially lower, especially for difficult selection tasks.

3.1 Methods

Participants. Ten participants (five female, five male) took part in the present study. They were monetarily compensated for their participation. Their ages ranged from 19 to 36 years. Before engaging in the experiment subjects were screened for shortness of sleep, tiredness, and alcohol or drug consumption. All participants reported normal or corrected-to-normal vision.

Tasks. The participants had to perform a search-and-select task. They were presented with stimuli consisting of four characters in the "easy" condition and seven characters in the "difficult" condition. The reference stimulus was displayed in the middle of the screen. Around this item twelve stimuli were shown in a circular arrangement, eleven distractors and one target stimulus, which was identical to the reference stimulus. The radial arrangement of search stimuli ensured a constant spatial distance to the reference stimulus. All search strings consisted of consonants only. The distractors shared a constant amount of characters with the target. Examples of the search screens are shown in Figure 1.

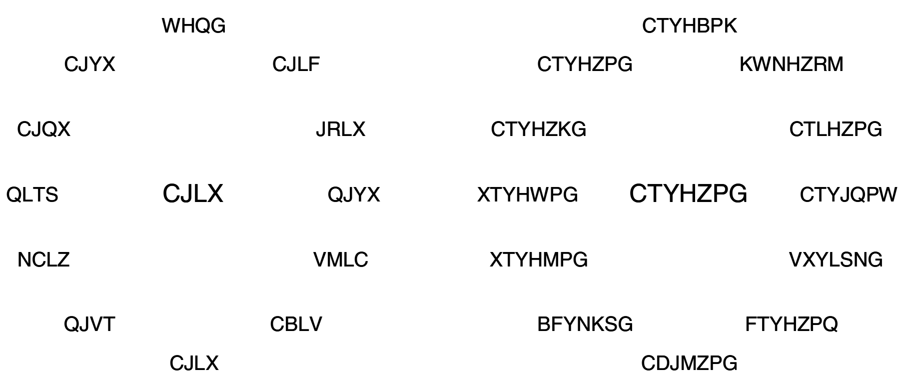

Fig. 1. Examples for easy (left) and difficult (right) search tasks

Subjects had to select the target stimulus by either fixating it for the given dwell time or by thinking the activation thought. It was not possible to use standard suggestions for dwell time durations from literature (e.g. [4]), because the difficulty levels of the search task are not directly comparable to search tasks on a GUI in terms of absolute time needed for identification. Rather the tasks were chosen to be easily kept in working memory in the “easy” condition and to almost exceed its storing abilities in the “difficult” condition. To make sure that the dwell times match stimulus complexity, different versions were tested in pre-experiments. The selection criterion was that the short version is still well controllable and that the long activation latency is not perceived as slowing down the user. The short dwell time was 1.000 milliseconds, the long dwell time 1.500 milliseconds.

For BCI activation, the participants had to imagine closing both hands to fists and then to turn them against each other like when wringing out a cloth by twisting it tightly. They were told not to involve any overt muscular activity.

Apparatus. Brain data were registered using a 32 channel EEG system (Brain Products, actiCap). Electrodes were positioned according to the 10-20 system covering all relevant areas. Signal processing was focussed on the sensomotric areas C3 and C4. Grounding was established with electrode Fz. Eye movements were tracked with an infrared camera equipped remote eye tracker (SensoMotoric Instruments, iView X RED). Lighting conditions were held constant during the experiment.

Design and Procedure. Two different levels of search difficulty (easy, hard) and three levels of activation technique (method of activation: dwell time short, dwell time long, BCI) were varied in a 2×3 within subjects factorial design. Participants went through the levels of the factor activation technique in separate blocks. The order of these blocks was counterbalanced. Subjects completed 30 trials per condition. The experiment itself took about 1 hour, the whole test procedure about 2.5 hours. Effectiveness was measured in terms of errors in task completion. Efficiency was defined as the time needed to complete a search task. Mental workload was assessed with the unweighted version of the NASA Task Load Index (Raw Task Load Index, RTLX) [18].

After making sure that all EEG electrodes were in place and working, additional EMG electrodes were attached to the participants’ arms to monitor for muscular activity. Before and during the technical preparations subjects received a general overview on the procedure of the experiment and their tasks. A complete and summarized presentation of the test setting was given afterwards. To finalize the preparation phase, subjects practiced using the BCI command and engaged in training the BCI classifier with a task that was very similar to the later search task. If the training was successful, a short calibration of the eye tracker followed and the experiment started. Each trial was terminated after 15 seconds if the participants were not able to locate the target stimulus. These trials were excluded from further analysis. The NASA TLX was filled in after each condition. At the end the participants had the opportunity to discuss their experiences with the experimenter and were asked to rate the activation techniques according to their preferences.

3.2 Results and Discussion

Time needed for task completion and accuracy (data on errors) were averaged across all subjects for each selection method and level of search difficulty. Trials with errors were not included in the analysis of response time. First, an analysis of variance was conducted on the results. The alpha level for significance was chosen to be .05. In a second step, the data of the easy and difficult condition were pooled for each selection method. This allows to take a closer look in pairwise comparisons between BCI vs. long dwell time and BCI vs. short dwell time. To avoid any problems associated with multiple testing, differences will be regarded as significant with an alpha level of .025 for these comparisons.

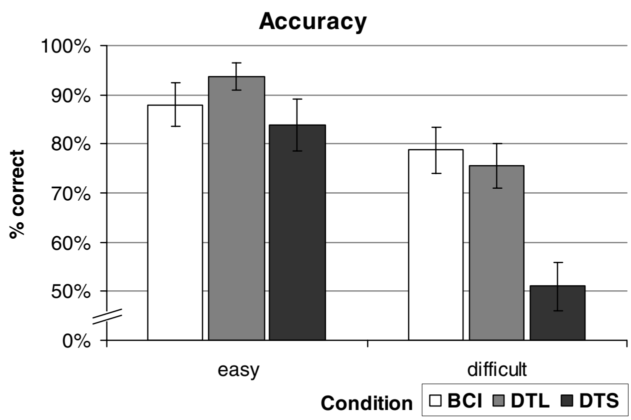

Fig. 2. Percentage of correct selections: Brain-Computer Interface (BCI), long dwell times (DTL) and short dwell times (DTS)

The accuracy data are summarized in Figure 2. The "easy" condition yielded 88.0% correct selections when using the BCI. In 93.8% of all tasks correct answers were produced in the "dwell time long" (DTL) condition, in the "dwell time short" (DTS) condition 83.8%. Fewer correct selections were made in the "difficult" condition. Remarkably, the BCI leads to the best results with 78.7% correct selections, although the difference to the long dwell time, 75.6%, is only marginal. The short dwell time condition, however, lead to a strong negative effect on performance as the percentage of correct answers dropped to 51.1%. This change in the result pattern in the difficult condition is reflected in a significant search condition × activation technique interaction ($F(2,18) = 13.30$, $p < .001$). An analysis of the main effects confirms general differences between the activation techniques ($F(2,18) = 12.47$, $p < .001$) and that the difficult search condition leads to more errors ($F(1,9) = 38.37$, $p < .001$).

The pooled BCI accuracy average is 83.3% correct selections, the corresponding values for dwell time long and dwell time short are 84.7% and 67.4%. Pairwise t-tests reveal that the better performance of the BCI compared to "dwell time short" is significant ($t(9) = 3.66$, $p = .005$). The small differences between BCI and "dwell time long" is not reliable ($t(9) = 0.33$, $p = .75$). As expected, the BCI allows users to activate (click) GUI items more precisely than a dwell time solution with short latencies. Long dwell times are suited for precise object activation but do not prove to be substantially better than BCI based selection.

Task completion was fastest in both search conditions with short dwell times (easy: 3.98 s; difficult: 5.38 s). Next was dwell time long (4.79 s; 7.37 s), leaving BCI the slowest method of activation (5.90 s; 8.84 s). This general difference between the input methods is statistically confirmed ($F(2,18) = 56.25$, $p < .001$). The results are depicted in Figure 3.

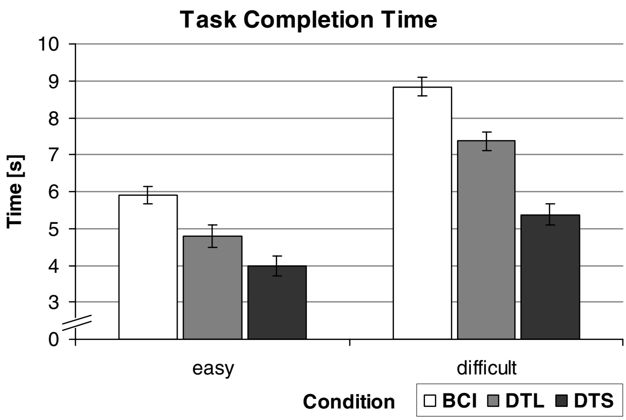

Fig. 3. Task completion times: Brain-Computer Interface (BCI), long dwell times (DTL) and short dwell times (DTS)

Looking at these data also shows that the difficult search task leads to longer search times, which is only of minor interest ($F(1,9) = 102.38$, $p < .001$). The significant search condition × activation technique reflects the larger differences between the selection techniques in the difficult compared to the easy condition ($F(2,18) = 7.46$, $p < .01$). The pairwise comparisons support the view that BCI selection was slowest (BCI: 7.37 s; dwell time long: 6.08 s; dwell time short: 4.68 s). These differences are significant (BCI – DTL: $t(9) = 4.31$, $p = .002$; BCI – DTS: $t(9) = 13.57$, $p < .001$).

Overall the TLX results show no differences in workload between the activation techniques. On a scale ranging from 0 to 10 with higher values standing for higher workload, BCI yielded 4.7, DTL 4.6 and DTS 4.6 ($F(2,18) = 0.18$, $p = .84$). Judging on this basis BCI does not come at the cost of higher cognitive demands. In the preferences ratings at the end of the experiment, 9 out of our 10 participants preferred using the combined BCI/eye-gaze interface over the standard gaze-based interface.

4 Conclusions and Outlook

Taken together, the state of technology allows to perform more accurate activations with BCI than with dwell time solutions with short latencies. Also quite remarkable is the strong user voting preference for using a BCI instead of dwell times for the activation of selected objects. However, using BCI is still somewhat slower. Nonetheless, although statistically significant, the magnitude of the difference between BCI and the dwell time solutions is remarkably small. Therefore, BCI has successfully proven to be a real competitor for dwell time activation already at the current state of technological development. This clearly indicates that it is a forthcoming technology for multimodal interfaces indeed.

Furthermore, integrating brain-computer interaction into a multimodal system opens up the option of using it as a means of direct input on the one side while simultaneously monitoring the user's workload on the other side [19]. Thus, behind any work on BCIs also stands the vision of building more ergonomic work places with future UI technology [20].

Acknowledgments. We would like to thank Matthias Rötting and especially Matti Gärtner and Christian Kothe for their support in conducting the experiments. We also want to thank Sonja Pedell for helpful comments on an earlier draft of this document.

References

1. Bolt, R.A.: Eyes at the Interface. In: Proceedings of the 1982 Conference on Human Factors in Computing Systems, pp. 360–362. ACM Press, New York (1982)
2. Engell-Nielsen, T., Glenstrup, A.J., Hansen, J.P.: Eye Gaze Interaction: A New Media - Not Just a Fast Mouse. In: Itoh, K., Komatsubara, A., Kuwano, S. (eds.) Handbook of Human Factors / Ergonomics, pp. 445–455. Asakura Publishing, Tokyo (2003)
3. Nilsson, S., Gustafsson, T., Carleberg, P.: Hands Free Interaction with Virtual Information in a Real Environment. In: Proceedings of COGAIN 2007, Leicester, UK, pp. 53–57 (2007)
4. Jacob, R.J.K.: What You Look at Is What You Get. IEEE Computer 26, 65–66 (1993)
5. Murata, A.: Eye-Gaze Input Versus Mouse: Cursor Control as a Function of Age. Int. J. Hum-Comput. Int. 21, 1–14 (2006)
6. Hutchinson, T.F.: Eye-Gaze Computer Interfaces: Computers That Sense Eye Positions on the Display. Computer 26, 65–67 (1993)
7. Kahneman, D.: Attention and Effort. Prentice Hall, Englewood Cliffs (1973)
8. Yarbus, A.L.: Eye Movements During Perception of Complex Objects. In: Riggs, L.A. (ed.) Eye Movements and Vision, pp. 171–196. Plenum Press, New York (1967)
9. Jacob, R.J.K., Legett, J.J., Myers, B.A., Pausch, R.: Interaction Styles and Input/Output Devices. Behaviour & Information Technology 12, 69–79 (1993)
10. Sibert, L.E., Jacob, R.J.K.: Evaluation of Eye Gaze Interaction. In: CHI 2000: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 281–288. ACM Press, New York (2000)
11. Jacob, R.J.K.: The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look at Is What You Get. ACM Transactions on Information Systems 9, 152–169 (1991)
12. Beinhauer, W., Vilimek, R., Richter, A.: Eye-Controlled Applications - Opportunities and Limits. In: Proceedings of Human-Computer Interaction International 2005. Lawrence Erlbaum Associates, Mahwah (2005)

13. Salvucci, D.D., Anderson, J.R.: Intelligent Gaze-Added Interfaces. In: CHI 2000: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 273–280. ACM Press, New York (2000)
14. Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., Perelmouter, J., Taub, E., Flor, H.: A Spelling Device for the Paralyzed. Nature 398, 297–298 (1999)
15. Leeb, R., Friedman, M.-P.G.R., Scherer, R., Slater, M., Pfurtscheller, G.: Self-Paced (Asynchronous) BCI Control of a Wheelchair in Virtual Environments: A Case Study with a Tetraplegic. Computational Intelligence and Neuroscience 7, 1–8 (2007)
16. Gargiulo, G., Bifulco, P., Calvo, R.A., Cesarelli, M., Jin, C., van Schaik, A.: A Mobile EEG System with Dry Electrodes. In: IEEE Biomedical Circuits and Systems Conference. IEEE, Baltimore (2008)
17. Birbaumer, N.: Breaking the Silence: Brain-Computer Interfaces (BCI) for Communication and Motor Control. Psychophysiology 43, 517–532 (2006)
18. Byers, J.C., Bittner, A.C., Hill, S.G.: Traditional and Raw Task Load Index (TLX) Correlations: Are Paired Comparisons Necessary? In: Mital, A. (ed.) Advances in Industrial Ergonomics and Safety, pp. 481–485. Taylor & Francis, London (1989)
19. Cutrell, E., Tan, D.: BCI for Passive Input in HCI. In: Proceedings of ACM CHI 2008. ACM Press, New York (2008)
20. Zander, T.O., Kothe, C., Welke, S., Rötting, M.: Enhancing Human-Machine Systems with Secondary Input from Passive Brain-Computer Interfaces. In: Proceedings of the 4th International BCI Workshop & Training Course. Graz University of Technology Publishing House, Graz (2008)