Enhancing Universal Access – EEG Based Learnability Assessment

Christian Stickel[1], Josef Fink[2], and Andreas Holzinger[3]

[1] VirtualTrends Int. Ltd. & Co. KG, R&D, Lerchenweg 3, 36137 Grossenlueder, Germany
christian.stickel@virtual-trends.de
2 University of Applied Sciences Frankfurt, FB2 Computer Science,
Nibelungenplatz 1, 60318 Frankfurt, Germany
jfink@fb2.fh-frankfurt.de
[3]Medical University Graz, Institute of Medical Informatics Statistics and Documentation,
Auenbruggerplatz 2/V, 8036 Graz, Austria
andreas.holzinger@meduni-graz.at

Abstract. This paper presents initial research on a new learnability assessment methodology. We propose the use of electroencephalography (EEG) to further improve usability testing. We discovered whether and to what extend there is a correlation between brainwave patterns and the learnability of the software used. Our central hypothesis is that learnability can be assessed by analyzing the rise and fall of specific frequency bands in electroencephalographic recordings. In order to collect empirical evidence for our hypothesis, we conducted an experiment with N=32 participants. We developed a test environment comprising a low-cost EEG system and developed software for analysis and testing. Based on our findings, we consider our EEG-based learnability test applicable, either as a pre-test – in order to determine whether further testing is necessary – or as an augmenting method during standard usability testing. The users' emotions, registered on the EEG, can be applied as a baseline for detecting possible usability difficulties and employed in the development of a biological rapid-usability method for accessibility assessment.

Keywords: Learnability, Biological usability testing, EEG, evaluation method.

1 Introduction and Motivation for Research

Within our information overloaded world, the need for universal access to information is undisputed. In order to assure progress towards this objective, universal test methods have to be developed that can be applied to the information society *without the need for additional adaptations or specialised (re-)design* [1], [2]. *Learnability* denotes the ease with which users can employ a particular piece of software by observing and using it. The main question is: How long does it take a user to learn the handling of the software in order to efficiently achieve a particular goal? Learnability is especially relevant for designing and deploying complex software, mainly due to potential reductions in learning and training time.

C. Stephanidis (Ed.): Universal Access in HCI, Part III, HCII 2007, LNCS 4556, pp. 813–822, 2007.

2 Theoretical Background: Brainwaves, Arousal, Relaxation

Everything humans perceive is processed as electrochemical signals in the human brain [3]. These signals are rhythmic, consequently called *brainwaves* and can be measured on the scalp by using electroencephalography (EEG). Brainwaves are always present as different types of sinusoidal EEG activity, changing when the mental and physical state of the subject changes [4]. In order to get more detailed information on brainwaves, their spectrum has to be divided into particular frequency ranges. EEG allows the non invasive observation of electrical processes in the cerebral cortex, which is believed to be largely responsible for our individual thoughts, emotions and behavior. Cortical processes involve electrical signals. Any of the innumerable electrochemical discharges of the neurons in the cerebral cortex produces a small electromagnetic field with a frequency between 1 and 40 oscillations per second. The sum of these signals forms the aforementioned brainwaves, which can be recorded by EEG scanning equipment. Recording is obtained by placing electrodes on the scalp, thereby following a reliable and reproducible system of placement. Each electrode records electrical activity from potentials that contain approximately 30–500 million neurons [5]. The activity of a single neuron can not be interpreted as something intellectual or cognitive. Cognition, intellect, and awareness are bound to global activity states of the brain, so-called macrostates [6].

Similar to all periodic signals, brainwaves can be described by both frequency and amplitude. The frequency is the number of times a periodic event repeats within a second. Widely used EEG frequency bands include gamma (30-40 Hz), beta (12.5-28 Hz), alpha (8-12.5 Hz), theta (4-8 Hz), and delta (1-4 Hz). Frequency ranges of these bands overlap by at least 0.5 Hz. The amplitude represents the power of electrical impulses generated by the brain; its intensity is measured in microvolt (μV) [5].

As already mentioned, brainwave activity changes when the mental and physical state of a subject changes. When a person is aroused, excited, or alert, for example, the EEG pattern has a low amplitude and a fast frequency. When a person is calm and relaxing quietly, however, especially with her eyes closed, brainwaves with a high amplitude and a low frequency can be observed [7].

The ability to mobilize metabolic energy in order to meet environmental or internal demands on behavior is called an arousal response. Three types can be identified based on the duration of a response: alertness = short term, vigilance = long term, and arousal = longer term [8], [9]. Arousal has a direct impact on motivation [10], [11], [12], [13], [14]. Arousal responses of novice computer users, for example, are most likely associated with frustration and discomfort. Misunderstandings of the navigation, error messages, and crashes of the operating system may gradually change the duration and power of the arousal response. With ongoing training and help, however, users become more acquainted, thereby increasing their tolerance towards frustration. In this case, major developmental changes in the arousal response are the growth of processes that control and shape the response. Moreover, the range of internal and environmental stimuli that activate these processes is increased, which can be observed in the context of experienced users (e.g., they remain unaffected of certain types of system errors). EEG can be used for anticipating alertness, vigilance, and arousal. The type of electrical activity recorded can be associated with behavioral arousal; rhythmic fast activity (beta/gamma) appears to be involved in states of high

arousal (e.g., learning, problem solving), whereas rhythmic slow activity (alpha) can be associated with a relaxed, yet alert state of consciousness. Behavior as well as EEG patterns tend to gradually change over day, except for sleep phases, when abrupt transitions between slow-activity sleep and rapid eye movement (REM) occurs. An EEG recording of REM sleep is similar to those with high arousal but reflects episodes of dreaming. Alpha arousal is also called *relaxation.* It is the counter-response to stress and necessary for the body to recharge energy. When a person is in a relaxed physical state, alpha rhythms will be predominant in EEG recordings. Alpha rhythms are used in neurofeedback therapies for children with attention deficit disorder (ADD), whose lack of attention and focus leads to poor learning skills.

3 Arousal and Cognitive Performance

Yerkes and Dodson propose an inverted U-shaped function for modeling the relation between arousal and cognitive performance [15]. The so-called Yerkes–Dodson law consists of two parts (see figure 1). The first part states that cognitive performance increases with arousal up to a task-specific optimum, further decreasing with higher levels of arousal. The second part states that the optimum level of arousal for a given task is inversely related to task difficulty.

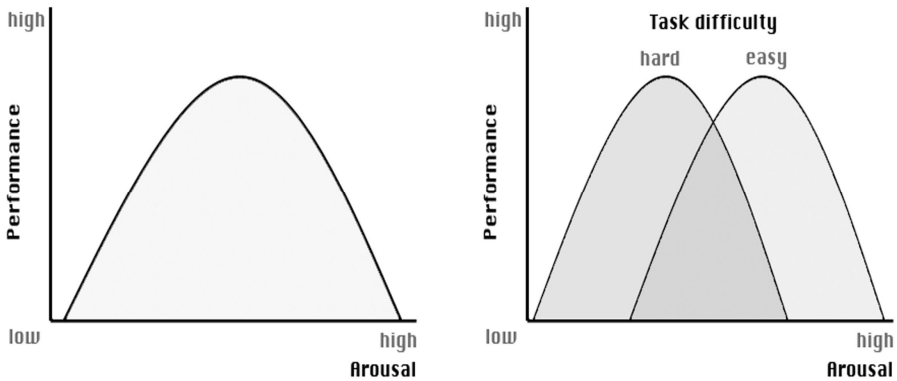

Fig. 1. The Yerkes–Dodson law [15], [16]

Optimal task performance is likely to be achieved in states of high alertness, rather than in states of high arousal. The upward part of the Yerkes–Dodson law can be interpreted as the energizing effect of arousal: increasing arousal improves alertness, which in turn improves performance. The downward part of the Yerkes–Dodson law can be interpreted as the negative effects (e.g. 'tunnel vision') of high arousal (or stress) on task performance. The second part of the Yerkes–Dodson law states that the optimum level of arousal decreases with increasing task difficulty. Cognitively demanding tasks may require a lower level of arousal for optimal performance to facilitate concentration, whereas tasks demanding prolonged efforts or persistence may be performed better with higher levels of arousal in order to increase motivation.

Cognitively demanding tasks can be regarded sensitive to disruption, for example by anxiety induced by aversive stimuli (e.g., a spider) and associated physiological effects, like rapid heart beat or sweating. Arousal and alertness conflict in these scenarios: Arousal can increase cognitive alertness but, as a side effect, cognitive performance suffers from sympathetic nervous system activation [13], [17]. Studies showed positive effects of optimal arousal on learning performance [18], [14].

4 Psychophysiologic Methods and Usability

Psychophysiologic measures are indicators for emotional and attentional responses. A response stimulus can be a startle tone, an emotionally loaded picture, video, or task presented to a subject. Measures of modern psychophysiology include brain activity, skin conductance level (SCL), heart rate (HR), muscle responses, and eye movement measures. Muter et al. [19] demonstrated that psychophysiologic measures are useful for usability studies. They reported about the use of HR and SCL in an experiment with a two-factor design: Each of their 25 subjects was presented a user-friendly and a user-hostile variant of a software tool for accomplishing banking transactions. In order to account for potential differences in task complexity, Muter et al. presented to their subjects memory tasks of varying difficulty in addition to the banking transaction (i.e., problem-solving) tasks. The authors found that the HR was higher during the interaction tasks than during memory tasks (see figure 2). Moreover, they found that HR was hardly related to task difficulty, whereas the SCL was significantly higher during the difficult (i.e. user-hostile) interaction task. Their HR finding can be interpreted as parasympathetic withdrawal, while their SCL finding may suggest that the user-hostile software leads to sympathetic excitation of the sort associated with a fight-flight-freeze response (FFF). Muter et al. conclude that the SCL may be regarded a good indicator for the overall usability of software.

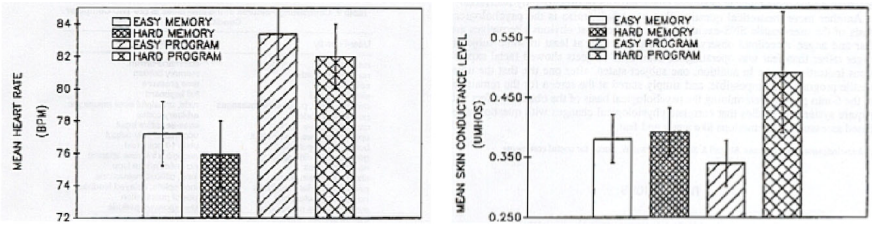

Fig. 2. This experiment from Muter et al. [19] suggests that *SCL (right side)* is a good indicator for usability, as opposed to *HR (left side)*

Both SCL and EEG can show the activation of the sympathetic nervous system. The fight-flight-freeze response (FFF) for example, can be identified in EEG recordings by a dominance of high beta and gamma brain waves. FFF rises cognitive alertness (arousal), as a side effect, however, it reduces the performance of certain higher cognitive functions including articulation, learning, decision making, and interpretation of paralinguistic signals.

From the point of view of evolutionary biology, FFF was key for mastering life-threatening situations. In today's software usability scenarios, however, the activation of FFF is counterproductive.

5 Hypothesis and Experimental Setting

Arousal is the ability to mobilize energy to meet a particular goal. Arousal may be affected by the look and feel of a software interface [20], [21]. Following the Yerkes-Dodson law, the level of arousal is correlated with cognitive performance when using information systems [16]. Specific types of brainwave activity recorded by EEG can be associated with levels of arousal [22], [23], [24]. Against this background, we formulated our working hypothesis: *Changes in the EEG recordings of a subject during a usability test can be used for assessing the learnability of the software used.*

Our experiment followed a two-phase design. In both phases, each subject (N=32) solved a memory game eight times. In the first phase, we used a conventional game with eight randomly presented pairs of pictures, four of them showing abstract and four realistic motives. There were eight different sets of pictures, each of them containing eight pictures. Subjects' task was to find all pairs of pictures. The memory game we used in the second phase used the same pairs of pictures, but the placement of abstract and realistic motives remained unchanged for the eight games in the second phase. We expected that as soon as our subjects discover this non-random placement strategy, they quickly focus their search accordingly, thereby reducing the number of uncovered non-matching pairs of pictures (i.e., error condition). The first phase of our experiment served as control condition. Throughout our experiment, we recorded subjects' brainwave activity with IBVA's EEG system (www.ibva.com). This relatively unobtrusive scanning equipment can be used in laboratory and field settings as well (see figure 3). Due to limitations on the number of electrodes used, IBVA's headband comprises merely 3 electrodes, so this system can record an EEG of the frontal lobe only. For our research, however, this limitation seems dispensable. All signals recorded by the IBVA system are sampled at a 512 Hz rate.

Fig. 3. The IBVA EEG system uses a 3-electrode headband that wirelessly communicates with the IBVA software running on a personal computer

After our experiment, we analyzed the recorded brainwaves according to our working hypothesis as follows: *We consider a software as learnable in case of alpha (8-12 Hz) wave dominance in subjects' EEG patterns. In case of high beta and gamma (25-40 Hz) dominance, however, we consider a software as hardly learnable.*

Following this, we expect that the recorded brainwaves of good learners show a significant alpha dominance, whereas the recordings of poor performers show a high beta/gamma (henceforth called betaH) dominance.

6 Results and Discussion

Following our aforementioned working hypothesis, we aimed at finding a correlation between subjects' individual learning progress and their alpha/betaH brainwave activity. In order to focus on those subjects with the highest and poorest individual learning progress, we selected N=10 subjects for each group, henceforth called 'top players' and 'poor players', respectively. For each game in the two test phases (henceforth called 'control scenario' and 'learning scenario'), we averaged the signal activity in the alpha and betaH frequency bands over all subjects of the particular group. Preceding to this, we averaged the signal activity of both brain hemispheres for each subject. Figure 4 shows the results of our analysis for the control scenario, where pairs of pictures are randomly selected from the card deck at the beginning of each game. On the left chart, we see the alpha signal activity for our learner groups, and on the right chart, the corresponding betaH signal activity is depicted.

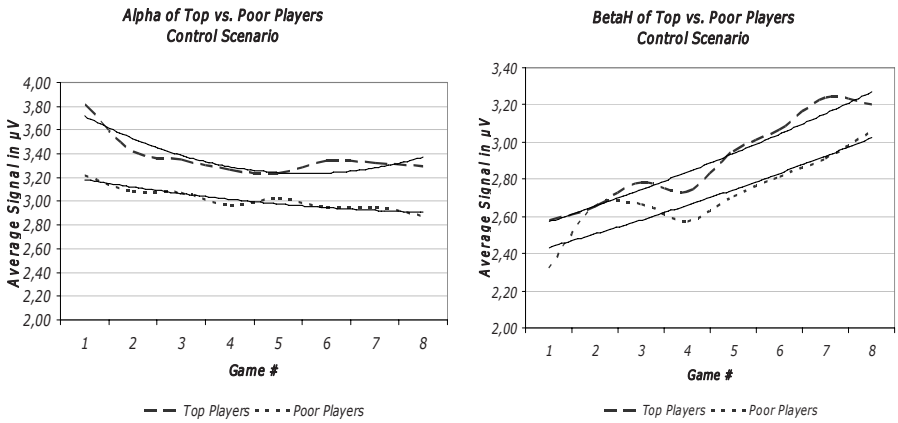

Fig. 4. Alpha and BetaH activity for top vs. poor players in the control scenario

Signal activity in the alpha and betaH frequency bands seems to be quite similar for both groups: apart from a difference in the level of signal activity between top and poor players, both charts show a decreasing (increasing) signal level over the eight games of our control scenario. The polynomial trend lines underline this observation.

The decreasing alpha signal activity over the eight games can be interpreted with the diminishing degree of relaxation for both top and poor learners, albeit on a different signal level.

The increasing betaH signal activity seems to reflect the increasing concentration efforts of both learner groups, again for top players on a higher level of signal activity. Due to our random selection strategy for pictures from the card desk, we

didn't expect to see meaningful learning progress, which, following Yerkes–Dodsons' law, might have reduced the mental efforts (and therefore the state of arousal) necessary for completing the random gaming task at hand.

Figure 5 depicts the results of our analysis for the learning scenario, where the placement of abstract and realistic picture motives remained unchanged for all eight games. Again, alpha signal activity is depicted on the left chart and corresponding betaH signal activity on the right chart.

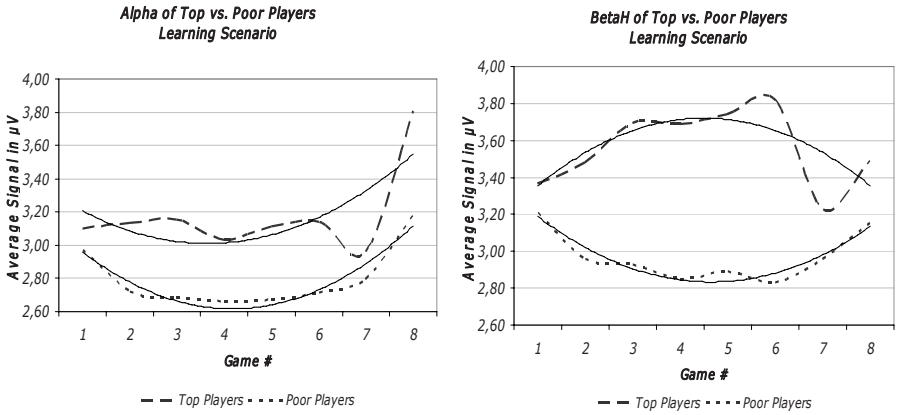

Fig. 5. Alpha and BetaH activity for top vs. poor players in the learning scenario

Now, average signal activity considerably differs from the one we presented so far. In the alpha band, the signal activity over the eight games now seems to form a U-shaped function for both groups, the polynomial trend line emphasizes this interpretation. And again, the signal level is higher for the top players, except for the seventh game, where their signal activity seems to plunge down. The betaH chart shows an even more differentiated recording for top vs. poor players. For the poor players, we see a U-shaped function which is quite similar to the one we have seen for their alpha. For the top players, however, we see an inverted U-shaped function, underlined by the polynomial trend line, with a maximum signal activity during the sixth, followed by a minimum activity during the seventh game.

Having the preceding games of the control scenario in mind, we suspect that the little variation in alpha signal activity up to and including the sixth game can be accounted to the more or less stable status of relaxation in both learner groups. For the last two games, however, we associate the significantly rising alpha with the fact that most users discovered the non-random placement strategy after the sixth game. This assumption is consistent with both users' number of errors when uncovering non-matching pairs of cards and their qualitative feedback during this phase of the game. The significant rise of alpha seems to be different in extent, though, with a steeper increase for the top players.

Regarding betaH, it seems that our poor learners don't show meaningful (or even reduced?) concentration efforts, whereas increasing efforts can be suspected for the top learners up to and including the sixth game (see above). So far, top learners'

investment is quite comparable to their former one during the control scenario. And as already mentioned above for alpha, top players' average betaH signal significantly changes during the last two games; this time inversely, showing reductions in average signal activity.

In order to investigate this a little further, Figure 6 depicts alpha and betaH average activity only for the top players in our learning scenario. Now, the simultaneous decrease for the seventh game, followed by the subsequent increase for the eight game in both frequency bands becomes more clearly visible.

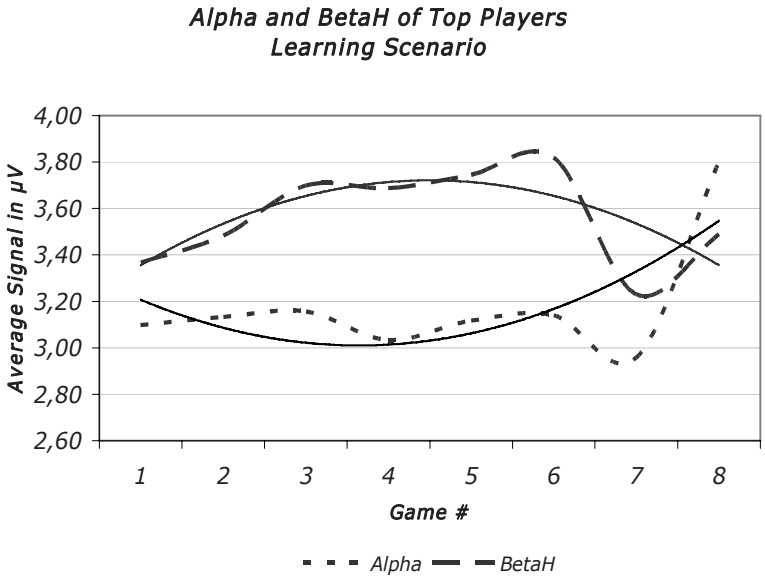

Fig. 6. Alpha and BetaH activity for top players in the learning scenario

The turning point after the sixth game, where all top players reportedly discovered the non-random placement of card pairs may be interpreted as a reduction of task complexity (cf. the Yerkes–Dodson law), which requires less mental efforts (and therefore lower states of arousal) for completing the eased gaming task. The raise of signal activity during a *flash of wit* has been reported by Machleidt et al. (1987). Following this, the subsequent alpha peak may be interpreted as a reward response.

7 Conclusion and Future Outlook

Our experiments showed some interesting, promising and significant differences in brainwave recordings of top performers versus weak performers. The correlations found, however, do not fully comply with our working hypothesis, especially regarding the dominance of alpha brainwaves in the EEG recordings of good learners.

Consequently, we must do further work based on our lessons learned. Limitations within our experimental setting include the (possibly) non-challenging amount of mental efforts necessary for completing our tasks, the small learning effect in the second phase of our experiments, potential noise stemming from a moderately controlled test environment, and the small sample size. Nevertheless, introducing rapid methods of biological usability testing can augment existing usability engineering methods [25] towards rapid biological testing [26]. In order to achieve this, further research is necessary, especially in real-life usability testing and on the systemic level of technology [27]. With the advent of new sensors and input devices, more psychophysiologic indicators of subtle responses will be available for adapting devices, interfaces, and interaction to heterogeneous user needs. Biological rapid usability testing may also offer a new set of methods and tools for software engineers, which can be easily integrated in the development process, at moderate costs in terms of efforts and time. This can further enhance to reach Universal Access in HCI, which aims at "*a conscious and systematic effort in order to advance a proactive approach towards interactive products and environments which should be accessible and usable by the broadest possible end-user population, anytime and from anywhere, without the need for additional adaptations or specialised (re-) design.*" [2], [28]

References

1. Stephanidis, C.: User Interfaces for All: Concepts, Methods and Tools. Lawrence Erlbaum, Mahwah (NJ) (2001)
2. Stephanidis, C., Salvendy, G., Akoumianakis, D., Arnold, A., Bevan, N., Dardailler, D., Emiliani, P.L., Iakovidis, I., Jenkins, P., Karshmer, A., Korn, P., Marcus, A., Murphy, H., Oppermann, C., Stary, C., Tamura, H., Tscheligi, M., Ueda, H., Weber, G., Ziegler, J.: Toward an Information Society for All: HCI challenges and R&D recommendations. International Journal of Human-Computer Interaction 11(1), 1–28 (1999)
3. Squire, L.R.: Encyclopedia of Learning and Memory. Macmillan, New York (1992)
4. Thompson, L.W., Obrist, W.D.: EEG Correlates of Verbal Learning + Overlearning. Electroencephalography and Clinical Neurophysiology 16(4), 332 (1964)
5. Nunez, P.L.: EEG. In: Ramachandran, V.S. (ed.) Encyclopedia of the Human Brain, pp. 169–179. Academic Press, San Diego (2002)
6. Roth, G., Schwegler, H., Stadler, M., Haynes, J.-D.: Die funktionale Rolle des bewußt Erlebten. Gestalt Theory 20, 186–213 (1998)
7. Kolb, B., Whishaw, I.: Introduction to Brain & Behavior. Worth Publishers, Cranbery (2001)
8. Duffy, E.: The psychological significance of the concept of arousal or activation. Psychol. Rev. 64, 265–275 (1957)
9. Marrocco, R.T., Field, B.A., Arousal.: In: Ramachandran, V. S (ed.) Encyclopedia of the human brain, Academic Press (2002)
10. Geen, R.G., Pigg, R.: Interpretation of Arousal and Its Effects on Motivation. Journal of Social Psychology 90, 115–123 (1973)
11. Brehm, J.W., Self, E.A.: The Intensity of Motivation. Annual Review of Psychology 40, 109–131 (1989)

12. Holzinger, A.: Effektivität von Multimedia - Motivation, Aufmerksamkeit und Arousal. GMW FORUM, Zeitschrift der Gesellschaft für Medien in der Wissenschaft,1(00) 10–13 (2000)
13. Holzinger, A.: Basiswissen Multimedia Band 2: Lernen. Kognitive Grundlagen multimedialer Informations Systeme Vogel, Würzburg (2000) http://www.basiswissen-multimedia.at
14. Holzinger, A., Pichler, A., Maurer, H.: Multi Media e-Learning Software TRIANGLE Case-Study: Experimental Results and Lessons Learned. Journal of Universal Science and Technology of Learning 0(0), 61–92 (2006)
15. Yerkes, R.M., Dodson, J.D.: The relation of strength of stimulus to the rapidity of habit formation. Journal of Comparative Neurology and Psychology 18, 459–482 (1908)
16. Holzinger, A.: Multimedia Basics, Learning. Cognitive Fundamentals of multimedial Information Systems. Laxmi, New Delhi vol. 2 (2002) http://www.basiswissen-multimedia.at
17. Brown, V.J., Bowman, E.M.: Alertness. In: Ramachandran, V.S. (ed.) Encyclopedia of the human brain, Academic Press, San Diego (2002)
18. Holzinger, A., Pichler, A., Almer, W., Maurer, H.: TRIANGLE: A Multi-Media test-bed for examining incidental learning, motivation and the Tamagotchi-Effect within a Game-Show like Computer Based Learning ModuleEducational Multimedia, Hypermedia and Telecommunication 2001. Association for the Advancement of Computing in Education, Charlottesville (VA) pp.766–771 (2001)
19. Muter, P., Furedy, J.J., Vincent, A., Pelcowitz, T.: User-Hostile Systems and Patterns of Psychophysiological Activity. Computers in Human Behavior 9(1), 105–111 (1993)
20. Norman, D.A.: Cognitive engineering. In: Norman, D., Draper, S. (eds.) User Centered System Design: New Perspectives on Human-Computer interaction, Lawrence Erlbaum Associates, Mahwah, NJ (1986)
21. Norman, D.A.: Emotional Design: Why we love or hate everyday things. Basic Books, New York (2004)
22. Gutjahr, L., Machleidt, W.: Influences of Psychic Variables on the EEG During a Conflict-Centered Interview. Electroencephalography and Clinical Neurophysiology 48(3), 25–26 (1980)
23. Hinrichs, H., Machleidt, W.: Basic Emotions Reflected in EEG-Coherences. International Journal of Psychophysiology 13(3), 225–232 (1992)
24. Bruggenwerth, G., Gutjahr, L., Kulka, T., Machleidt, W.: Music Induced Emotional EEG Reactions. Eeg-Emg-Zeitschrift Fur Elektroenzephalographie Elektromyographie Und Verwandte Gebiete, 25(2), 117–125 (1994)
25. Holzinger, A.: Usability Engineering for Software Developers. Communications of the ACM 48(1), 71–74 (2005)
26. Holzinger, A., Slany, W.: XP + UE -> XU Praktische Erfahrungen mit eXtreme Usability. Informatik Spektrum 29(2), 91–97 (2006)
27. Holzinger, A., Leitner, H.: Lessons from Real-Life Usability Engineering in Hospital: From Software Usability to Total Workplace Usability, In: Holzinger, A., Weidmann, K.-H (eds.) Empowering Software Quality: How can Usability Engineering reach these goals? Austrian Computer Society, pp. 153–160 (2005)
28. Stephanidis, C., Savidis, A.: Universal Access in the Information Society: Methods, Tools and Interaction Technologies. Universal Access in the Information Society 1(1), 40–55 (2001)