

# A Wearable User Interface for Measuring Reaction Time

Burcu Cinaz, Christian Vogt, Bert Arnrich, and Gerhard Tröster

ETH Zurich, Electronics Laboratory, Gloriastrasse 35, 8092 Zurich, Switzerland  
Burcu.Cinaz@ife.ee.ethz.ch

**Abstract.** Reaction time (RT) tests are known as simple and sensitive tests for detecting variation in cognitive efficiency. RT tests measure the elapsed time between a stimulus and the individual's response to it. A drawback of existing RT tests is that they require the full attention of a test person which prohibits the measurement of cognitive efficiency during daily routine tasks. In this contribution we present the design and the evaluation of a wearable RT test user interface which can be operated throughout everyday life. We designed a wearable watch-like device which combines the generation of a haptic stimulus and the recognition of subject's hand movement response. In order to show to what extent the wearable RT test is convenient to measure reaction times, we designed an experiment in which we measured the reaction times of ten subjects from two different setups. In the first half of the experiment, the participants performed a desktop-based RT test whereas in the second half of the experiment they performed the wearable RT test. In order to measure changes in the duration and variability of reaction times we induced additional cognitive load in both setups. We show that individual changes of reaction times occurred due to the cognitive load manipulation are similar for both desktop-based and wearable RT test. Additionally we investigate the subjective ratings of perceived workload. We conclude that the presented wearable RT test allows to measure changes in reaction times occurred due to additional cognitive load and hence would allow the assessment of changes in cognitive efficiency throughout everyday life.

**Keywords:** reaction time, wearable user interface, cognitive efficiency.

## 1 Introduction and Motivation

Reaction time (RT) tests measure how rapidly information can be processed and a response to it can be activated [14]. In other words, RT tests measure the elapsed time between a stimulus and the individual's response to it. According to Jensen [10], RT tests are well suited for cognitive assessment tests since in comparison to conventional psychometric techniques, RT tests offer a high sensitivity for detecting variation in cognitive efficiency and they can be virtually unlimited repeated. Several desktop-based RT tests have been developed in which users have to respond to visual stimuli by using keyboard, mouse or special buttons. An extensive review about computer-based cognitive tests can be found in [20]. There are several examples on applying RT tests to assess cognitive functioning: early detection of cognitive decline such as

dementia or Alzheimer's disease in elderly; determining the ability to manage complex activities such as driving, piloting or search and rescue; identifying of children with intellectual disabilities such as Attention Deficit Hyperactive Disorder (ADHD).

The main drawback of existing desktop-based RT tests is the requirement of the full attention of the subject, i.e. the subject has to interrupt his daily routine for several minutes in order to perform the task on the computer. This restriction prohibits the measurement of cognitive efficiency during daily routine tasks, e.g. to determine the ability to manage complex activities such as piloting. Our goal is to develop reaction time tests which can be operated throughout everyday life by means of wearable devices. An important step in the development is to ensure that wearable reaction time tests are suitable to measure changes in reaction times similar to desktop-based approaches.

In this paper, we present the design and the evaluation of a wearable RT test user interface. We designed a wearable watch-like device which combines the generation of haptic stimuli and the recognition of subject's responses. Haptic stimulus is generated by a vibration motor. The subject's responses to a stimulus are performed by a hand movement which is recognized with an inertial measurement unit (IMU). For the evaluation of the wearable interface, we conducted an experiment to investigate to what extent we can measure the user's reaction time with our interface compared to desktop-based tests.

## 2 Related Work

Three kinds of RT tests are commonly employed in literature [11]: simple, recognition and choice RT tests. Simple RT tests consist of one stimulus and one response. For instance the subject has to press a button as soon as the letter "X" appears at a pre-defined position or as soon as a light or sound appears. In recognition RT tests, the subject has to respond to a stimulus (target) and ignore other stimuli (non-target). This is sometimes called as "go/no-go" RT task. Recognition of a particular sound or symbol belongs to this category. Lastly, choice RT tests include multiple stimuli and multiple responses. The subject has to respond to each stimulus with a corresponding response, e.g. by pressing a certain key whenever a corresponding letter appears on the screen. A detailed series of recommendations on how to conduct experiments using reaction times and how to analyze the collected data can be found in [14], [10], [19].

Increasing age and age-related diseases like cognitive impairment are important factors which influence length and variability of reaction times [11]. It has been known that with increasing age, reaction times become more variable and longer. Gorus et al. showed that persons with cognitive deterioration demonstrated more intra-individual performance variability and more slowing in their reaction times than cognitively healthy elderly [5]. Braverman et al. showed in a clinical setting that the test of variables of attention (TOVA) is an accurate predictor of early attention complaints and memory impairments [2]. The effect of stress was investigated in an experiment which examines the cognitive performance under psychosocial stress [17]. The results showed that participants under stress were slower in their reaction times.

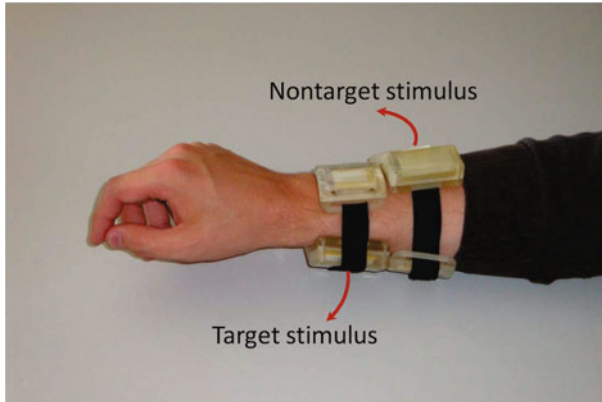
Another application area of RT tests is the Attention Deficit Hyperactive Disorder (ADHD) patients. Children with ADHD have often difficulties in focusing on tasks and one of the most consistent findings is increased moment-to-moment variability in reaction time [18].

Most of the studies have in common that RT tests are operated with a computerized test which requires the full attention of the subject. Since the user has to interrupt his current activity to perform the test, most of these techniques are not feasible to be used during normal life activities. There exist only a few studies which measure one's cognitive performance continuously during everyday activities. Lieberman et al. implemented visual stimuli (3 LEDs), auditory stimuli (a miniature speaker) and two push buttons on a wrist-worn device to assess vigilance [13]. Ivorra et al. implemented a haptic stimulus to interrogate the central nervous system in a minimally obtrusive way [8]. As the response, the detection of a wrist movement is defined. By doing so, they showed that a simple RT test can be continuously administered throughout the course of normal life activities. However, a comparison of the wearable implementation with desktop-based RT tests is missing.

### 3 Materials and Methods

#### 3.1 Design of the Wearable Reaction Time Test

The wearable user interface to measure reaction times consists of two main modules: the stimuli module to generate haptic stimuli and the inertial measurement unit (IMU) module for detecting wrist movements. According to the literature, the wrist is a recommended stimulus site for wearable tactile displays [3], [12], [15], [16]. Therefore we designed a wrist-mounted tactile display in order to deliver the stimulus information to the user. For generating vibro-tactile stimuli, we used a coreless mini DC vibration motor with a diameter of 6mm and a resonant frequency around 200 Hz (manufactured by Precision Microdrivers Ltd.). In order to maximize the vibration amplitude and to ensure a proper sense of the vibration, we placed the motor in a separate plastic enclosure resulted in WxLxH dimensions of 90x55x30mm which can be attached to the wrist of the user by using a strap. The stimuli module has its own battery supply. The vibration motor needs a continuous current of 83mA and a start current of 150mA. In a conservative calculation (continuous current of 150 mA, single stimulus duration 500ms, 160 stimuli in 12 minutes), a total of 400 mAh would be required to perform a continuous reaction time test over 24 hours. In order to guarantee a continuous operation during at least one day, we have selected a lithium ion battery with 650mAh. In addition, we have integrated an audio driver (MAX4410 by Maxim Inc.) in order to allow the generation of auditory stimuli through headphones. The IMU module consists of the ETH Orientation Sensor (ETHOS) which was developed in our laboratory [6]. The ETHOS includes a 3D accelerometer and gyroscope which allows to recognize the subject's gesture response.



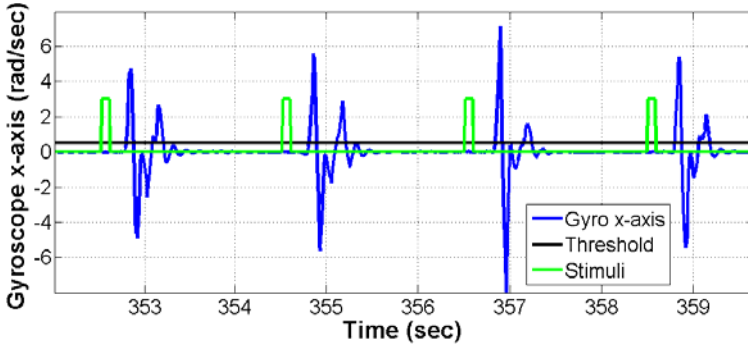
**Fig. 1.** Wearable implementation of a go/no-go RT test. The left module generates the target stimuli as vibrations on the lower side of the wrist, whereas the right module generates non-target stimuli on the upper side of the wrist. For both modules, the IMU is placed on the opposite side of the vibration motor to recognize the user's hand gesture response.

The accelerometer and gyroscope were sampled with a frequency of 128Hz. The detailed description of the ETHOS hardware platform can be found in [6]. We modified the firmware version of the ETHOS to control the vibro-tactile component. An implementation of a go/no-go task which is comprised of two wearable user interfaces to induce target and non-target stimuli can be seen in Fig. 1.

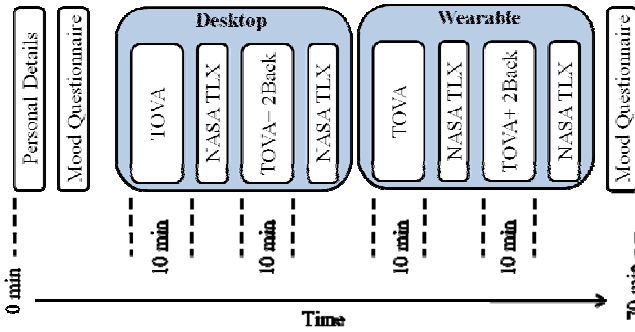
In order to automatically recognize a predefined hand gesture response to the haptic stimulus, we performed a preliminary experiment. Similar to the wearable RT test presented in [8], we have defined the response gesture as a fast rotation of the wrist. Three subjects performed a RT test on the wearable device during three different conditions. In the first condition, the subject was sitting on a chair while the arms were heading towards the floor. In the second condition, the arms were placed on the table. In the third condition, the subject was walking with a moderate speed (4km/h) on a treadmill. In each condition we recorded 3d acceleration and gyroscope data. Based on a visual inspection of the recorded data, we manually labeled each wrist response. For all conditions, it was clearly visible that the wrist-turn axis of the gyroscope ( $x$ -axis) was the most sensitive axis for detecting the fast rotation of the hand. In order to define a threshold for automatically detecting this hand gesture, we computed the correctly identified responses for different thresholds. With a threshold of 0.5rad/sec, we could correctly detect the occurrence of this hand gesture response in all conditions. The raw gyroscope data and the occurrence of haptic stimuli are exemplary shown in Fig. 2. According to the simple threshold approach mentioned above, we compute the time point when the user was assumed to have reacted.

### 3.2 Experiment: Comparison of Wearable and Desktop RT Tests

Ten healthy subjects (6 male, 4 female, average age 26.15 years) participated in our experiment. All participants were paid 30 Swiss Francs for participating in one session of approximately 70 minutes. The goal of this experiment was to evaluate our



**Fig. 2.** X-component (wrist-turn axis) of the gyroscope data while reacting to four stimuli with a fast rotation of the hand. Based on the threshold approach the time point when a subject has reacted was computed.



**Fig. 3.** Experimental procedure including two conditions (*baseline* as TOVA single task and *cognitive load* as TOVA+2Back dual-task) for each setup (desktop based RT and wearable RT)

wearable user interface by comparing the reaction times obtained by the wearable reaction time test with a desktop-based approach.

**Experimental Setup.** For the desktop-based reaction time test we used a free version of the TOVA test which is implemented with the psychology experiment building language (PEBL) [1]. The implementation of the test is based on the description in [4]. A white square appears briefly on the screen, with a black square within it. Participant must respond only to targets (the black square on top) and ignore the non-targets (the black square on the bottom). Each stimulus is presented for 100 ms at 2000 ms intervals. For the wearable RT test, we placed two RT modules on the dominant wrist of the user as shown in Fig. 1. The left module generates the haptic target stimuli on the lower side of the wrist (volar side), whereas the right module generates haptic non-target stimuli on the upper side of the wrist (dorsal side). Similar to the desktop-based RT test, each stimulus is generated for 100 ms at 2000 ms intervals. For recognizing the user’s wrist turn response, we only use the data collected with the IMU placed opposite to the left module.

**Experimental Procedure.** Each setup (wearable and desktop-based) consists of two experimental conditions: (i) single-task in which the subject has to respond to the target stimulus, and (ii) dual-task in which the subject has to solve a cognitive task in parallel to the single-task. Each condition lasts 10 minutes and contains 320 stimuli (160 targets and 160 non-targets). This leads to a total of 640 reaction times for each subject (160 targets x 2 setup x 2 conditions). As cognitive task we employed a variant of the N-Back test, the so-called “Audio 2-Back” [9] as explained in the following. The four phases used for each subject are:

- *Desktop-based RT (single-task):* The subject has to respond to each target stimulus by pressing the space bar on the keyboard and ignore non-target stimuli types. This is the typical variant of the test of variables of attention (TOVA).
- *Desktop-based RT with N-Back (dual-task):* In this condition a second task is added to the traditional desktop-based TOVA test. The subject has to solve an Audio 2-Back task which is presented to the user simultaneously with the TOVA test. Thereby a letter is presented to the subject via an audio message and the subject has to respond if the currently pronounced letter is the same as the one that was pronounced 2 positions back. The response to the Audio 2-Back was done by saying “match” whenever a sound match occurs. The investigator controls if the subject answers correctly and gives feedback continuously to the user about correct and false answers to keep him concentrated on both of the tasks.
- *Wearable RT (single-task):* The subject has to respond to each target stimulus generated on the wrist by performing a wrist movement and ignore the non-target stimuli types.
- *Wearable RT with N-Back (dual-task):* The subject has to respond to target stimuli with hand movements, and solve Audio 2-Back task simultaneously.

In the following, we denote the single task of each setup as “*baseline*” and dual task as “*cognitive load*” condition. Directly after each condition for both settings, each subject was asked to indicate his perceived workload by completing the multidimensional assessment tool NASA Task Load Index (TLX) [7]. The rating consists of the following six scales: mental demand, physical demand, temporal demand, own performance, effort and frustration. Based on the ratings, the total workload was computed as a weighted average. The experimental procedure can be seen in Fig. 3.

## 4 Results

### 4.1 Reaction Times

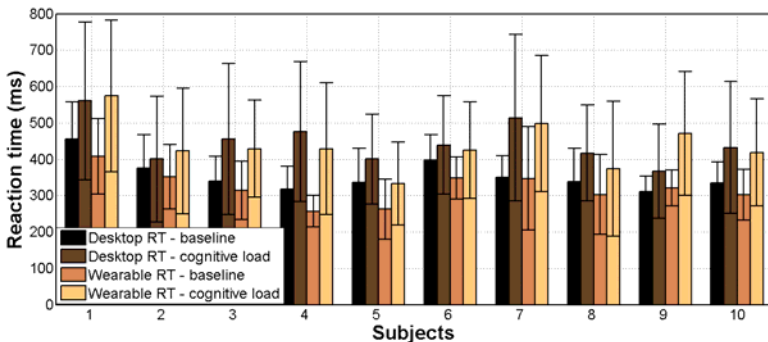
For the analysis, the mean reaction time and the standard deviation are considered as evaluation metrics. In Table 1 and Fig. 4 the means and standard deviations of the reaction times for all subjects in each condition are presented. First, it can be observed that for both desktop and wearable RT test, the mean reaction time is always increased during the *cognitive load* condition compared to the *baseline* condition. Mean reaction times of the desktop-based RT test are significantly correlated with the wearable RT test for the *baseline* condition ( $r = 0.8336$ ,  $P < 0.01$ ) as well as for the *cognitive load* condition ( $r = 0.7070$ ,  $P < 0.05$ ). Second, it can be observed that the

increase in mean reaction times from *baseline* to *cognitive load* conditions is similar within subjects for both desktop and wearable setting. The relative difference (mean RT during *cognitive load* minus mean RT during *baseline condition*) between desktop and wearable setting are significantly correlated ( $r = 0.7095$ ,  $P < 0.05$ ). Consistently, the variability of reaction times was always higher in the *cognitive load* condition compared to the *baseline* condition for both desktop and wearable setting. No significant correlations were observed for the standard deviation of reaction times. Besides, it can be observed that for most subjects the mean reaction time in the wearable setting is lower compared to the desktop-based approach during *baseline* (exception is subject 9). In the *cognitive load* condition the mean wearable reaction times are again lower for most subjects (exceptions are subjects 1, 2 and 9). This might be explained by the fact that the transduction of a visual stimulus takes generally longer than the perception of a haptic stimulus as known from literature [10].

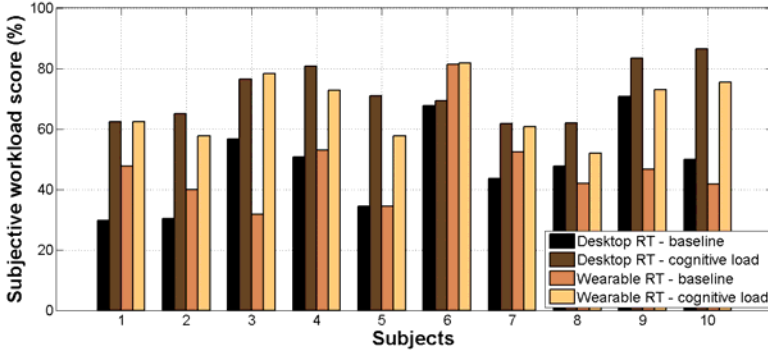
**Table 1.** Comparison of mean reaction times including standard deviation for the four experimental conditions

Subjects	Desktop-based Reaction Times		Wearable Reaction Times	
	Baseline[ms]	Cog. Load[ms]	Baseline[ms]	Cog. Load[ms]
1	455 (102)	560 (217)	408 (103)	574 (208)
2	375 (93)	400 (173)	351 (89)	423 (173)
3	339 (68)	455 (208)	314 (80)	429 (134)
4	317 (62)	475 (192)	257 (44)	429 (182)
5	336 (94)	401 (123)	263 (83)	333 (115)
6	398 (70)	439 (135)	348 (57)	424 (133)
7	350 (59)	513 (229)	347 (142)	498 (187)
8	338 (91)	417 (132)	303 (111)	373 (185)
9	311 (42)	367 (129)	321 (49)	471 (170)
10	334 (59)	432 (181)	302 (70)	418 (147)

*Mean (Standard Deviation)*



**Fig. 4.** Mean reaction times for each subject collected from two conditions in each setting. Error bars indicate standard deviation.



**Fig. 5.** Subjective workload scores of each subject obtained from the NASA task load index for each condition

## 4.2 Subjective Ratings

Fig. 5 shows the subjective NASA task load index for each subject. As intended from the experiment design, it can be observed that for both desktop and wearable RT test, the subjective ratings of the *cognitive load* condition are higher than the respective *baseline* condition. A comparison between both tests shows that 50% of the subjects perceived higher total workload during *baseline* condition when using the wearable device. This is due to the fact that 90% of all subjects rated the “physical demand” item of NASA-TLX with higher values for the wearable setting since additional physical demand was required for responding with the wrist movement. However, the comparison between both *cognitive load* conditions shows that 70% of the subjects perceived lower workload when using the wearable device. These results indicate that operating the wearable device results in lower perceived mental load when the user is engaged in a primary task which requires a certain amount of information processing.

## 5 Conclusion and Outlook

In this paper, we presented our experimental design and initial results in measuring reaction times of a person using a wearable RT test. In order to show to what extent a wearable interface is convenient to measure reaction times, we designed an experiment in which we measured response times of ten subjects from two different setups. In the first half of the experiment, the participants performed a desktop-based RT test whereas in the second half of the experiment they performed the wearable RT test. In order to measure changes in the duration and variability of reaction times we induced additional cognitive load in both setups. Besides the recording of reaction time data, subjective ratings of perceived workload were collected with the NASA-TLX. In a comparison of the obtained wearable reaction times with desktop-based reaction times, we showed that individual changes of reaction times due to the cognitive load are similar for both desktop-based and wearable RT test. According to the subjective ratings of the participants, we could show that all participants perceived the induced



cognitive load as intended from the experiment design. Furthermore, subjective ratings showed that operating the wearable RT test interface under cognitive load results in lower perceived mental load compared to desktop-based reaction time test. These results suggest that the wearable RT test is more appropriate when the user is engaged in a second task which requires a certain amount of information processing. Based on the achieved results, we conclude that wrist-mounted reaction time tests seem feasible to measure factors which influence length and variability of reaction times and would allow the measurement of variation in cognitive efficiency throughout everyday life where the individuals are engaged in multiple tasks.

In our future work, we will perform statistical comparisons of different wearable RT setups, e.g. generation of the target/non-target stimulus on the dominant/non-dominant hand. In addition we are going to conduct long-term measurements of reaction times throughout daily life as cognitive performance indicator. We are planning to measure reaction times in real time from employers which have to perform complex intellectual tasks like flight operators. The obtained reaction times would allow us to identify loss of cognitive efficiency and to reduce the risk of cognitive overload.

**Acknowledgements.** This research has been partly funded by the European FP7 project ProFiTex, grant agreement number CT-TP 228855-2.

## References

1. Pebl, <http://pebl.sourceforge.net/>
2. Braverman, E.R., Chen, A.L., Chen, T.J., Schoolfield, J.D., Notaro, A., Braverman, D., Kerner, M., Blum, S.H., Arcuri, V., Varshavskiy, M., Damle, U., Downs, B.W., Waite, R.L., Oscar-Berman, M., Giordano, J., Blum, K.: Test of variables of attention (TOVA) as a predictor of early attention complaints, an antecedent to dementia. *Neuropsychiatr. Dis. Treat.* 6, 681–690 (2010)
3. Chen, H.-Y., Santos, J., Graves, M., Kim, K., Tan, H.Z.: Tactor Localization at the Wrist. In: Ferre, M. (ed.) *EuroHaptics 2008*. LNCS, vol. 5024, pp. 209–218. Springer, Heidelberg (2008)
4. Forbes, G.B.: Clinical utility of the Test of Variables of Attention (TOVA) in the diagnosis of attention-deficit/hyperactivity disorder. *J. Clin. Psychol.* 54, 461–476 (1998)
5. Gorus, E., De Raedt, R., Lambert, M., Lemper, J.C., Mets, T.: Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. *J. Geriatr. Psychiatry Neurol.* 21, 204–218 (2008)
6. Harms, H., Amft, O., Winkler, R., Schumm, J., Kusserow, M., Tröster, G.: Ethos: Miniature orientation sensor for wearable human motion analysis. In: *Proceedings of IEEE Sensors Conference*. IEEE (2010)
7. Hart, S.G., Stavenland, L.E.: Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) *Human Mental Workload*, ch. 7, pp. 139–183. Elsevier (1988)
8. Ivorra, A., Daniels, C., Rubinsky, B.: Minimally obtrusive wearable device for continuous interactive cognitive and neurological assessment. *Physiol. Meas.* 29, 543–554 (2008)
9. Jaeggi, S.M., Buschkuhl, M., Jonides, J., Perrig, W.J.: Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. USA* 105, 6829–6833 (2008)

10. Jensen, A.R.: *Clocking the mind: Mental chronometry and individual differences*. Elsevier (2006)
11. Kosinski, R.J.: A literature review on reaction time (August 2009)
12. Lee, S.C., Starner, T.: BuzzWear: alert perception in wearable tactile displays on the wrist. In: *CHI 2010: Proceedings of the 28th International Conference on Human Factors in Computing Systems*, pp. 433–442. ACM, New York (2010)
13. Lieberman, H.R., Matthew Kramer, F., Montain, S.J., Niro, P.: Field assessment and enhancement of cognitive performance: Development of an ambulatory vigilance monitor. *Aviation, Space, and Environmental Medicine* 78 (2007)
14. Duncan Luce, R.: *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press (1986)
15. Matscheko, M., Ferscha, A., Riemer, A., Lehner, M.: Tactor placement in wrist worn wearables. In: *International Symposium on Wearable Computers (ISWC)*, pp. 1–8 (2010)
16. Oakley, I., Kim, Y., Lee, J., Ryu, J.: Determining the feasibility of forearm mounted vibrotactile displays. In: *Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS 2006* (2006)
17. Scholz, U., Marca, R.L., Nater, U.M., Aberle, I., Ehlert, U., Hornung, R., Martin, M., Kliegel, M.: Go no-go performance under psychosocial stress: Beneficial effects of implementation intentions. *Neurobiology of Learning and Memory* 91(1), 89–92 (2009)
18. Vaurio, R.G., Simmonds, D.J., Mostofsky, S.H.: Increased intra-individual reaction time variability in attention-deficit/hyperactivity disorder across response inhibition tasks with different cognitive demands. *Neuropsychologia* 47(12), 2389–2396 (2009)
19. Whelan, R.: Effective analysis of reaction time data. *The Psychological Record* 58, 475–482 (2008)
20. Wild, K., Howieson, D., Webbe, F., Seelye, A., Kaye, J.: Status of computerized cognitive testing in aging: A systematic review. *Alzheimer's and Dementia* 4(6), 428–437 (2008)