Monitoring of Cognitive Functioning by Measuring Reaction Times with Wearable Devices

Burcu Cinaz, Bert Arnrich, 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 cognitive tests. A drawback of existing RT tests is that they require the full attention of a test person which prohibits the measurement of cognitive functioning during daily routine tasks. In this contribution we present our first steps in designing and evaluating reaction time tests which can be operated throughout everyday life by means of wearable devices. In a feasibility study we induce changes in reaction times by applying cognitive load in 5 test subjects. We compare the obtained wearable reaction times with desktop-based reaction time tests. We show that relative changes in the mean duration and the variability of reaction times are similar for both desktop-based and wearable reaction time test. We conclude that wearable reaction time tests seems feasible to measure changes in reaction times and hence would allow the measurement of cognitive functioning throughout everyday life.**

Keywords-reaction time; response time; wearable; dementia; alzheimer; reaction time experiments

I. INTRODUCTION

Cognitive decline is commonly considered to be a consequence of typical aging processes. However, cognitive impairment may turn into Alzheimer's disease or dementia [3]. Early detection of cognitive decline in the elderly would provide the opportunity to start with the treatments early in the disease process.

For diagnosing cognitive decline there exist many different techniques such as neuropsychological tests, neuroradiological techniques, and laboratory testing [2,11]. Promising alternatives are computerized assessments of cognitive functioning which are also suited to early detect changes. In comparison to traditional neuropsychological assessment methods, computerized tests offer benefits such as cost reduction and time savings. An extensive review about computer-based cognitive tests for detecting cognitive decline can be found in [13].

A simple and sensitive cognitive test is the reaction time (RT) test which is defined as a measure of how rapidly information can be processed and a response to it can be activated [10]. In other words, it is the elapsed time between a stimulus and the individual's response to it. According to Jensen [8], RT tests are well suited for practical application in health care 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. The main drawback of 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. 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 contribution we present our first results in operating a reaction time test by hand movements. We induce changes in the duration and variability of reaction times by applying cognitive load to test subjects. In the following we propose and evaluate a setup to measure the response to a stimulus by recognizing certain hand movements of a subject with a 3D accelerometer. We compare the obtained wearable reaction times with desktop-based reaction time tests in two experimental conditions: (i) single-task in which the subject has to respond to a target stimulus, and (ii) dual-task in which the subject has to solve a cognitive task in parallel to the single-task.

In the following we first we present related work. Next, we describe our methods including experimental setup and data acquisition. Afterwards, we present the comparison of the wearable reaction time test with a desktop-based approach. Finally, we summarize our paper and provide an outlook for future work.

II. RELATED WORK

There exist mainly three kinds of reaction time tests [9]: simple, recognition and choice reaction. Simple reaction time 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 reaction time tests, subject has to respond to one stimuli type and ignore other stimuli types. This is sometimes called as "go/no-go" reaction time task. Recognition of a particular sound or symbol belongs to this category. Lastly, choice reaction time tests include multiple stimuli and multiple responses. The subject has to respond to each stimulus with a corresponding response e.g. pressing a 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 data can be found in [8,10,12].
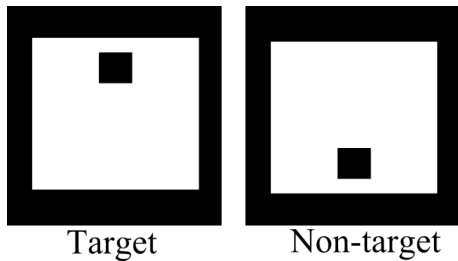
Figure 1: Target and non-target stimuli types of the TOVA test.

Increasing age and age-related diseases like cognitive impairment are important factors which influence length and variability of reaction times [9]. It has been known that with increasing age, reaction times become more variable and longer. Gorus et al. evaluated whether reaction times and performance variability are potential markers for the early detection of Alzheimer's disease. Persons with cognitive deterioration demonstrated more intra-individual performance variability and more slowing in their reaction times than cognitively healthy elderly. Thus, the authors suggest that intra-individual performance variability and RT are predictors for mild cognitive impairment (MCI) and Alzheimer's disease (AD) status [5]. Braverman et al. showed that the test of variables of attention (TOVA) is an accurate predictor of early attention complaints and memory impairments in a clinical setting [2]. Most of the studies have in common that reaction time tests are operated with a computerized test which require the full attention of the subject.

Ivorra et al. presented an approach to assess cognitive performance continuously throughout normal life activities [6]. The authors implement haptic stimuli to interrogate the subject and record the responses which are predefined hand movements detected by accelerometers. However, a comparison of the wearable implementation with desktop-based reaction time tests is missing.

III. METHODS

A. Experimental Setup

As outlined above, our goal was to compare reaction times obtained by a wearable sensor with a desktop-based approach in two experimental conditions (single-task and dual-task). For the desktop-based reaction time test and for generating the stimuli during the wearable 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 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). In the first half of the test, the occurrence frequency of targets are rare whereas for the second half they occur more frequently. Fig. 1 depicts the stimuli types of the TOVA test.

The collection of reaction times was performed using the following rules: (i) in the first half of the experiment, the participant responds to target stimuli by pressing the space bar, (ii) in the second half of the experiment, the participant
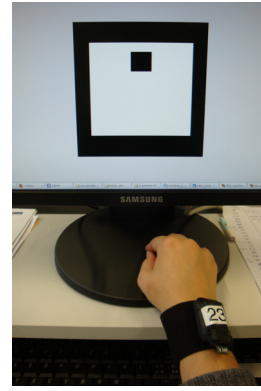

Figure 2: Experimental setup with one subject wearing acceleration sensor on her wrist.

responds to target stimuli by performing a wrist movement. We measured the wrist acceleration at 128 Hz with a three-axis MEMS acceleration sensor placed at the wrist of the person using a strap (see Fig. 2).

B. Experiment

Five healthy subjects (2 female, 3 male) participated in the experiment (mean age 26). The experiment has been conducted to assess and compare response times of the subjects from two different setups (responses performed as hand movements, responses performed on the keyboard). Each setup 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 1280 reaction times for each subject (2 setup x 2 conditions).

As cognitive task, a variant of the N-Back test, the so-called "Audio 2-Back", was employed [7] 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.

- *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.
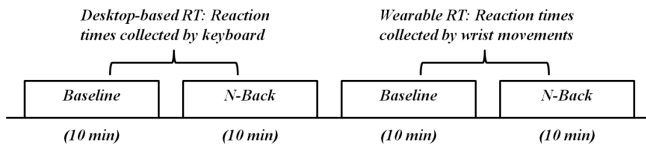
Figure 3: Experiment procedure including two conditions (baseline and N-back) for each setup.

- *Wearable RT (single-task):* The subject has to respond to each target stimulus by performing a wrist movement as quickly as possible and do nothing during 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 (dual-task).

In the following, we denote the single task of each setup as *"baseline"* and dual task as *"N-back"* condition. The experimental procedure can be seen in Fig. 3.

C. Data Acquisition

During desktop-based RT test, we collected reaction times of the subjects with the TOVA test. In order to compute reaction times during wearable RT test, we used a simple threshold approach: when the acceleration of the x-component exceeds a predetermined threshold, we assume that the user has reacted. The raw accelerometer data and the time when the visual stimuli (5 targets and 1 non-target) occurred are shown in Fig. 4. The violet points indicate the time point when the user was assumed to have reacted. This allowed us to compute the reaction times of the subject for all target stimuli.

IV. RESULTS

For the analysis, the mean reaction time and the standard deviation are considered as evaluation metrics. Mean reaction time is a measure of how fast the person responds whereas standard deviation measures the variability of a person's response times. Table 1 displays means and standard deviations of reaction times for all subjects in each condition. From the results in the table it can be observed that participants in both *N-Back* conditions were slower in their reaction times compared to the *baseline* conditions. In addition, it can be observed that the mean reaction times of the wearable approach are always higher compared to the desktop-based approach.
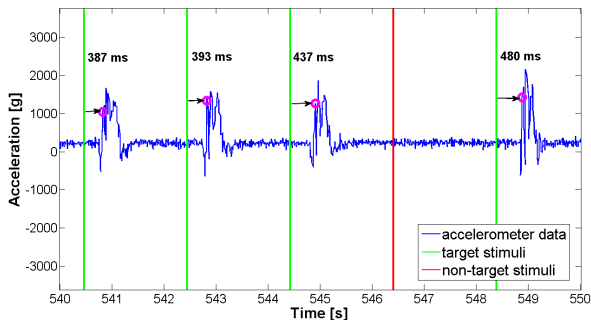

Figure 4: X-component (wrist-turn axis) of the acceleration data while reacting to four target stimuli. Based on a threshold approach, the time point when a subject has reacted was computed (violet circles).

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

Subjects	Desktop-based		Wearable	
	RT [ms] Baseline	RT [ms] N-Back	RT [ms] Baseline	RT [ms] N-Back
1	336 ± 53	487 ± 222	433 ± 56	634 ± 205
2	355 ± 62	505 ± 218	444 ± 66	628 ± 157
3	347 ± 67	432 ± 113	448 ± 57	534 ± 126
4	380 ± 75	507 ± 225	599 ± 108	659 ± 217
5	342 ± 61	397 ± 107	473 ± 90	513 ± 138
Mean ± Standard Deviation				

That might be explained that pressing a key can be operated faster than performing a hand movement. However, as outlined in the following, the relative changes between the experimental conditions are consistent. Since we are interested in measuring relative changes, the absolute differences between wearable and desktop-based approaches are not relevant.

In order to investigate the relative changes of reaction times due to the task complexity induced by the N-back task, we divided the mean and standard deviation of reaction times obtained during the *N-back* condition by the corresponding *baseline* reaction times. Fig. 5 depicts the relative changes of the mean reaction time values whereas Fig. 6 shows the relative changes of the standard deviations (SD) of reaction times for both desktop and wearable tests.
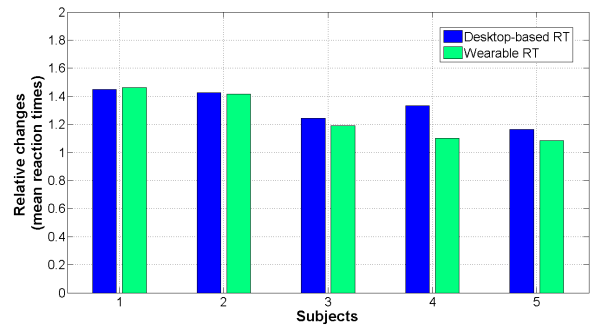

Figure 5: Ratio of mean reaction time between baseline and N-Back condition (mean(N-back) / mean(baseline)).
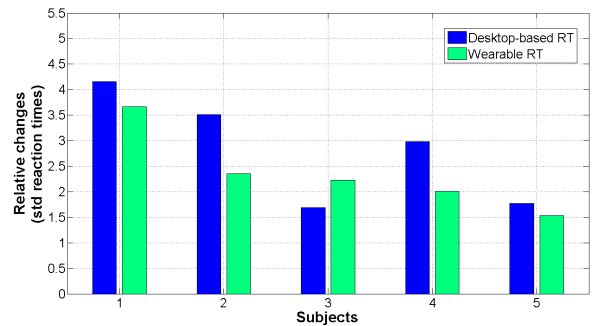

Figure 6: Ratio of standard deviation between baseline and N-Back condition (std(N-back) / std(baseline)).
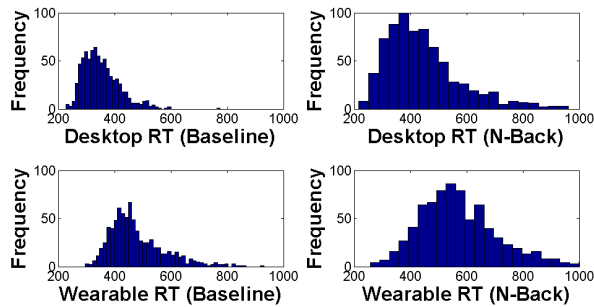
Figure 7: Histograms for baseline (left) and N-back (right) conditions for all subjects.

First, it can be observed that for both desktop and wearable RT test, the mean reaction time is increased during the *N-Back* condition since for all subjects the ratio exceeds 1. Second, it can be observed that the relative changes for desktop and wearable tests are similar for all subjects. The highest difference in the relative changes can be observed for subject 4. The relative changes of the SD of reaction times show that again for both desktop and wearable RT test, the SD is increased during the *N-Back* condition for all subjects. The increase in SD is higher during the desktop-based reaction time test for 4 subjects. However, in 4 out of the 5 subjects, the SD measured with the wearable RT test was more than 2 times higher in the *N-Back* condition.

The histograms in Fig. 7 illustrate the distributions of reaction times for both desktop-based and wearable RT test. It can be seen that the mean reaction times are shifted to the right for the wearable setup and the variability is increased during the *N-back* condition.

V. CONCLUSION AND FUTURE WORK

In this paper, we presented our experimental design and initial results in measuring reaction times of a person using a wearable sensor. In order to show to what extent a wearable sensor is convenient to measure reaction times, we designed an experiment in which we measured response times of five subjects from two different setups. In the first half of the experiment, the participants responded to each visual stimulus by pressing the space bar on the keyboard whereas in the second half of the experiment, they responded to each visual stimulus with a hand movement. In order to measure changes in the duration and variability of reaction times we induced additional cognitive load. The result showed that relative changes in the mean duration and the variability of reaction times were similar for both desktop-based and wearable

reaction time test. We conclude that wearable reaction time tests seem feasible to measure changes in reaction times and hence would allow the measurement of cognitive functioning throughout everyday life.

In future work we extend our setup by (i) recognizing more hand or body movements as response; (ii) implementing a wearable stimulus, e.g. vibration motor; and (iii) investigate long-term measurements of reaction times throughout daily life in elderly persons as cognitive performance indicator.

REFERENCES

[1] Pebl. http://pebl.sourceforge.net/.

[2] E. R. Braverman, A. L. Chen, T. J. Chen, J. D. Schoolfield, A. Notaro, D. Braverman, M. Kerner, S. H. Blum, V. Arcuri, M. Varshavskiy, U. Damle, B. W. Downs, R. L. Waite, M. Oscar-Berman, J. Giordano, and K. Blum. 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] Eric Braverman. *Cognitive Decline of Aging: Important Neuroendocrinological Predictors of Early Cognitive Decline in A Clinical Setting*. Weill Cornell Medical Center, PATH Medical, New York, NY USA, 2011.

[4] G. B. Forbes. Clinical utility of the Test of Variables of Attention (TOVA) in the diagnosis of attention-deficit/hyperactivity disorder. *J Clin Psychol*, 54:461–476, Jun 1998.

[5] E. Gorus, R. De Raedt, M. Lambert, J. C. Lemper, and T. Mets. Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. *J Geriatr Psychiatry Neurol*, 21:204–218, Sep 2008.

[6] A. Ivorra, C. Daniels, and B. Rubinsky. Minimally obtrusive wearable device for continuous interactive cognitive and neurological assessment. *Physiol Meas*, 29:543–554, May 2008.

[7] S. M. Jaeggi, M. Buschkuehl, J. Jonides, and W. J. Perrig. Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. U.S.A.*, 105:6829–6833, May 2008.

[8] Arthur R. Jensen. *Clocking the mind: Mental chronometry and individual differences*. Elsevier, 2006.

[9] Robert J. Kosinski. A literature review on reaction time, August 2009.

[10] R Duncan Luce. *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, 1986.

[11] R. Mistur, L. Mosconi, S. D. Santi, M. Guzman, Y. Li, W. Tsui, and M. J. de Leon. Current Challenges for the Early Detection of Alzheimer's Disease: Brain Imaging and CSF Studies. *J Clin Neurol*, 5:153–166, Dec 2009.

[12] Robert Whelan. Effective analysis of reaction time data. *The Psychological Record*, 58:475–482, 2008.

[13] Katherine Wild, Diane Howieson, Frank Webbe, Adriana Seelye, and Jeffrey Kaye. Status of computerized cognitive testing in aging: A systematic review. *Alzheimer's and Dementia*, 4(6):428 – 437, 2008.