

Human Performance Measurement In Interaction With Computer

Nebojša D. Đorđević¹ and Dejan Rančić²

Abstract – Regarding different human perceptual, cognitive and motor abilities can lead to universally usable interface development.

In order to evaluate user performance in interaction with interface, we developed software tool for testing sensomotor abilities of user in human-computer interaction. Test concept allows program-led testing of the intent-group and precisely quantifies user performance. Every experimental result is just a piece of a mosaic in the human performance in interaction with information systems based on computers.

In this study we obtained an efficient tool for graphical interpretation of the results, visual analyses of the tested groups averaged results, and easy creation of the user profiles.

Keywords – HCI, User interface, Cognitive models, HCI testing tool, User profile.

I. INTRODUCTION

Multidisciplinary nature of human-computer interaction requires contribution from different science disciplines; especially from computer science, cognitive psychology, social and organizational psychology, ergonomics and human factors, computer-aided design and engineering, artificial intelligence, linguistics, philosophy, sociology and anthropology. HCI can be defined as “a field of study related to design, evaluation and implementation of interactive computer systems used by humans, which also includes research of the main phenomena that surround it” [1].

Main goal of HCI is to improve interaction between the user and the computer in order to make computers more user friendly and designed systems more usable. The most important element in HCI is user interface. User articulates his requests to the system via dialogue with the interface. Interface is the point at which human-computer interaction occurs. Physical interaction with end user is provided using hardware (input and output devices) and software interaction interface elements.

II. COGNITIVE MODEL OF HCI

Cognitive modeling provides a description of user in interaction with the computer system; it provides a model of user's knowledge, understanding, intentions and mental processing. Description level differs from technique to technique and ranges from high-level goals and results

regarding thinking about a problem all the way to the level of motor activities of the user such as pressing a key on a keyboard or a mouse click. Research of these techniques is done by psychologists, as well as computer science specialists.

Classification of cognitive models is based on whether the focus is on the user and its task, or on transformation of the task into interaction language [1]:

- Hierarchical presentation of user's tasks and goals (GOMS model [2]);
- Linguistic and grammar levels;
- Models of physical level.

Models of the physical level relate to human motor skills and describe user's goals that are realizable in a short time period. An example is KLM model (*Keystroke-Level Model*) [3] used for determining user's performance with a given interface. In this mode, the task of accomplishing a goal is given in two stages:

- *Task acquisition*, during which user makes a mental picture of how to reach a given goal, and
- *Task performance* using the system.

Task acquisition closely connects KLM with GOMS level that gives an overview of the tasks for a given goal. By summing these time periods we get estimated time for completion of those tasks for a given goal. Precision of the KLM model depends on the experience of the designer, because he is required to make a realistic decision about the abilities of end user. Obviously, the development of high quality user interface is impossible without cognitive modeling and techniques.

Interaction models are descriptions of user inputs, application actions and obtained outputs. Interaction models are based on formalisms, which ensure their implementation within interface development tools.

In order to evaluate user performance as realistically as possible, we extend the interaction models (UAN, XUAN [4]). Extended model (XUAN/t – *Extended User Action Notification per Time*) treats equally the complexity of interactions, both from the system and from the user.

III. TESTING COGNITIVE CHARACTERISTICS

The classical methods of experimental psychology are under the constant development in order to cope with complicated cognitive tasks, specific to human interaction, on one side, and to computers on the other. The reliable and valid results of the interface performance rating can be achieved by observing the user efficiency through the repetitive assignment of similar tasks in similar environment conditions. The most important prerequisite to design an efficient interactive system is understanding cognitive and perceptual

¹Mr Nebojša D. Đorđević is with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Nis, Serbia, E-mail: djnebojsa@bankerinter.net

²Dr doc Dejan Rančić is with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Nis, Serbia, E-mail: ranca@elfak.ni.ac.yu

abilities of the user [5-7]. Modern computer systems are based on human ability to fast interpret affection of sense organs and respond with a sequence of complex actions. In the short time intervals, measured in milliseconds, users perceive changes on their screens and react adequately. A lot of working duties is tightly bound to perception, so designers should be aware of the boundaries of human perception [8]. The eyesight is especially important because the speed of human reaction depends on various visual stimuli, such as the time to accommodate to a very bright or very dim light, ability to recognize the appropriate part of a context, determine the speed or route of the moving point, etc. Visual sense reacts differently to different colors depending on spectral boundaries and color sensibility. The other senses, like these of hearing and touch, are also important. Ability tests represent measure tools for special HCI characteristics. These tests are standardized procedures for special HCI activities. These activities are measured for each user in order to compare to other users results achieved under similar conditions. There are several steps during user-computer dialogue, which we grouped into sensory, cognitive and motor activities. Test construction is based on recognition of activities in user-computer interaction, prominent user characteristics and the measurement method of individual production results. Based on the described model and psychometric concepts, we developed software CASE tool for evaluation of human cognitive characteristics in interaction with the computer.

IV. TESTING SENSORY ABILITIES

Cognitive processes, which represent response to specific stimulation, are represented using visual-information processing model [9].

Available information comes to special sensory register and remains in it about one second. Physical characteristics of the stimulation are determined at this level. After that, information is erased from the register (has been forgotten) or transferred into the short-time memory. At this level, some information has been lost, while the rest (along with information from long-time memory) has influence on user response. The goal of sensory ability tests (perception) is to determine reaction times of users to visual (TP1) and audio (TP2) stimuli. User's abilities in domains of seeing, hearing and kinesthetic senses are tested. The test lasts 20 seconds, during which time user is stimulated with series of stochastic visual and auditory stimuli. User's task is to react as quickly as possible by pressing a certain key (LIGHT-OFF, RINGER-OFF), confirming registration of the tested stimuli. System registers time lapse between giving the stimuli and user's response, as an evaluation parameter.

In order to acquire HCI ability information from different user groups, we performed special tests on group of 234 users. The group includes $n_1=116$ male and $n_2=118$ female users. We performed statistical analysis on obtained results in average reaction time on visual as well as audio stimuli in order to discover statistically significant difference between different user groups. For statistically significant difference estimation we used Student's t-test [11], which is based on

average reaction time difference between two independent user groups (with limitation that n_1+n_2 should be greater than 60). Hypothesis acceptance condition was that average reaction time difference between two independent user groups is significantly greater than standard average response time difference error. The standard average response time difference error for our test was 0.089419 sec. Obtained Student's t-value can be interpreted using Student's tables for limit t-values for chosen level of freedom n_1+n_2-2 ($=232$ in our case) and significant level ($p=0.01$, which means 99% of confidence).

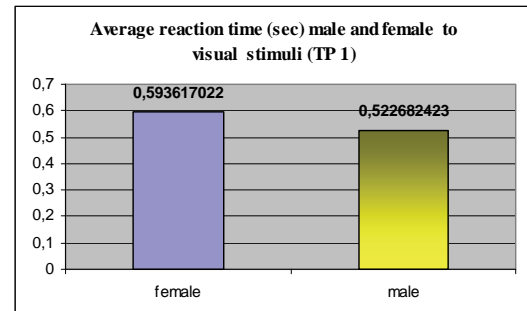


Fig.1. Average reaction time in sensory ability tests (visual stimuli)

In case of visual stimuli (TP1), obtained Student's t-value $t=0.79$ is less then limit value $t=2.58$, which means that there is no statistically significant difference between male and female users (Fig.1). The difference is effect of random variance, the samples belongs to the same basic set.

V. TESTING PSYCHOMOTOR ABILITIES

In order to articulate his demands, user utilizes certain interaction elements of user interface (hardware and software), which enable his physical interaction with the computer. In physical interaction with hardware device, user makes a voluntary activity, which is coordinated with visual senses (from the primary sensory zone) and kinesthetic senses (from the motor cortex). Kinesthetic senses provide muscle coordination and development of skills for performing different complex movements while working. The goal of psychomotor tests is to determine the precision in coordination, object manipulation, psychomotor orientation, reaction time, manipulation aptness and the ability of making visual-motor guesses. First group of tests (PM), so called "CLICK-A-FIELD", is aimed to probing psychomotor orientation, visual-motor guessing ability and coordinated manipulation of user-computer interaction tools, coordination of individual senses and body parts. Tests last 20 seconds, and user's task is to click a field (1×1 cm), which cyclically, using random coordinate generator, appears on the screen. During the test, the system on-line continually registers times related to certain events (PRESS-MOUSE-BUTTON, RELEASE-MOUSE-BUTTON) and connects them in database with the user and the test. After the event, RELEASE-MOUSE-BUTTON field is erased form the screen and it appears at a new randomly generated coordinates.

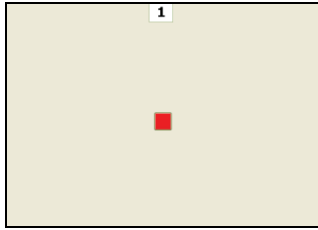


Fig.2. Tests PM2 - field is highlighted red on the interface

In order to determine the influence of different factors on user's psychomotor characteristics we developed four different tests. The goals of these tests are the same, however: PM1 field on the interface is darker shade of gray than the background; PM2 field is highlighted red on the interface (Fig.2); in PM3 test the field is 1×3 cm on the interface; in PM4 test after RELEASE-MOUSE-BUTTON event a beep sound is given in order to provide audio stimuli.

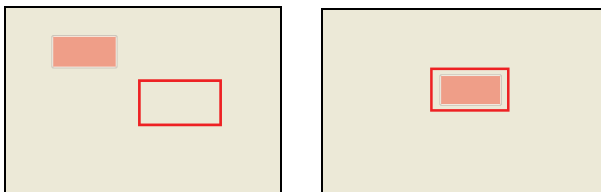


Fig.3. Test PM5 - "DRAG-ME" test

For determining precision and ability of fast, easy, correct and coordinated manipulation of visual objects with interaction technique of dragging objects on the screen, we developed PM5 test (called "DRAG-ME") (Fig.3). Test lasts 20 seconds, and user's task is to click on a red rectangular object on the screen and drag it into a rectangular window with red borders. After each attempt the object on the screen appears at a different randomly generated coordinate. System on-line registers successful attempts.

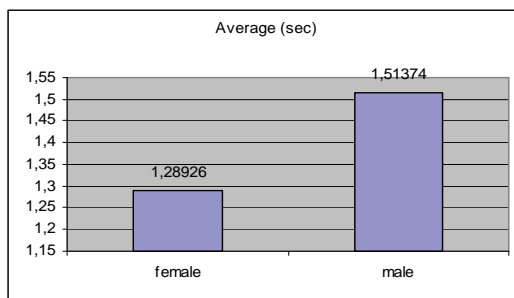


Fig.4. Average reaction time in psychomotor ability tests (psychomotor orientation)

We performed statistical analysis on obtained results in psychomotor tests. But, in case of psychomotor orientation tests (PM), average response time was 1.51374 sec for male and 1.28926 sec for female users. Obtained t-value $t=2.06$ is greater than limit value $t=1.96$, which means (with 95% confidence, $p<0,05$) that there is statistically significant difference between male and female users (Fig.4).

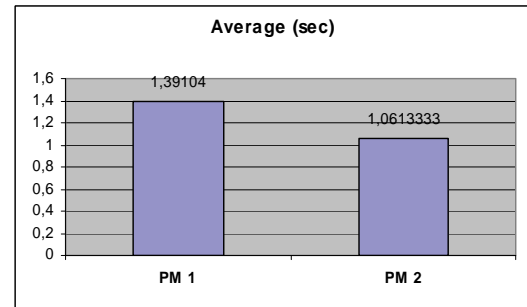


Fig.5. Average time in psychomotor tests with small and significant contrast difference in button color

Nevertheless, in case of psychomotor orientation tests with small (PM1) as well as significant (PM2) contrast difference in button color, average response time for entire testing population (both male and female users) was 1.39104 for PM1 and 1.061 sec for PM2. Obtained t-value $t=3.9567$ is greater than limit value $t=2.58$, which means (with 99% confidence) that there is statistically significant difference in response time between cases with small and significant contrast difference in button color (Fig.5).

VI. TESTING MEMORY ABILITIES

Memory is information-process structure composed from three components: sensory, short-time and long-time memory [10]. All memory components are necessary for successful information memorizing. Memory subsystem for sensory information deals with sensory representation of visual or audio event, which stimulates user sense during very short period. Short-time memory represents activity center in information processing system with limited capacity. In this zone, information comes from both sensory as well as long-time memory subsystem [10]. Information in long-time memory is persistent with potentially unlimited capacity. Crucial characteristic of long-time memory is that information, which is memorized, may differ from the original information because of the experience as well as other information influence.

The main goal of memory tests (TM1) is to investigate memory span through the ability of immediate reproduction of a series of elements after only one viewing of the series. This test is not time limited; it lasts until the first unsuccessful reproduction is made. User can see, in a certain time interval, series of randomly generated numerical signs of given length. Presentation time of the series is proportional to the length of series. User's task is to reproduce the entire series successfully. This step is repeated with each series one sign longer.

We also developed two more tests with the same scenario as TM1, with a certain difference: TM2 generated series are made of letter signs, and in TM3 the series are made with alphanumeric signs.

System registers the longest length of successfully reproduced series as a memory span parameter.

We performed statistical analysis on obtained results in memory tests. We used randomly generated numbers with average length of 7.322 numerical signs (for TM1 test) and randomly generated signs with average length of 5.88 letter signs (for TM2 test) (Fig.6). Obtained t-value $t=4.79$ is greater than limit value $t=2.58$, which means (with 99% confidence) that there is statistically significant difference in average length of repeated sequence for numbers and letter signs.

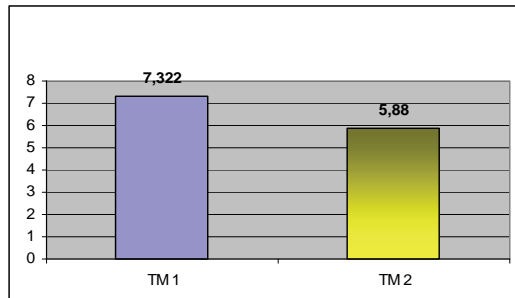


Fig.6. Average length randomly generated signs in memory tests

VII. CONCLUSIONS

Understanding physical, intellectual and personal differences between potential users defines the level of understanding and fulfilling user needs. Regarding different human perceptual, cognitive and motor abilities can lead to universally usable interface development. Taking into account different aspects of user profiles confronts us with the challenges of physical, cognitive, perceptual, personal and cultural differences between users. In order to evaluate user performance in interaction with interface, we extend the concepts of existing interaction models. Based on the described model and psychometric concepts we developed software tool for testing sensomotor abilities of user in human-computer interaction. Test concept allows program-led testing of the intent-group and precisely quantifies user performance.

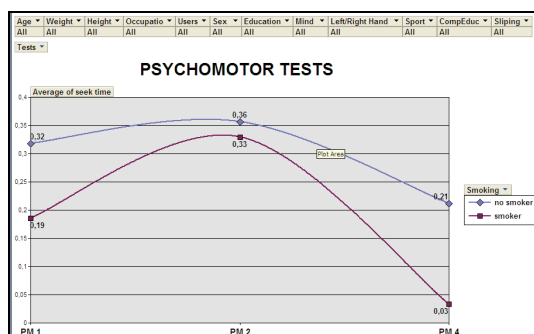


Fig. 7. Graphical interpretation of user profile for smokers and nonsmokers

In this study we obtained an efficient tool for making user profiles (Fig.7). The software tool enables graphical

interpretation of the results, more on the statistical capabilities (Fig.8), visual analyses of the tested groups averaged results, and easy creation of the user profiles.

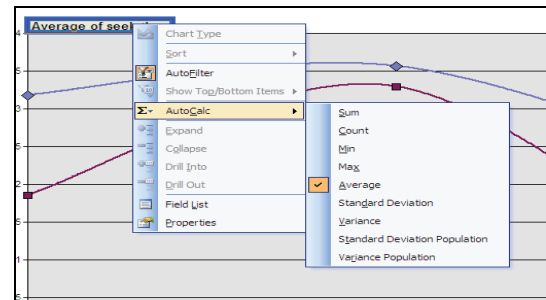


Fig. 8. Representation of the CASE Tool statistical capabilities

Differentiation of tested users is utilized to determine compatibility of individual interaction models with given intent-groups. Qualitative result analysis provides recommendations for design of individual interface parts, which are useful for the intent-group for which it is designed.

A future works can be based on extending a set of user characteristics which qualify perceptual and motoric performance, extending a set of tests using different interaction techniques.

REFERENCES

- [1] Dix, A., Finlay, J., Abowd, G., Beale, R., *Human-Computer Interaction*, 2nd ed. Prentice Hall Europe, 1998.
- [2] Kieras, D. E, Arbor, A., Towards a Practical GOMS Model Methodology for User Interface Design, In M. Helander: *Handbook of Human-Computer Interaction*, Elsevier Science Publishers B. V. (North Holland), 1988, pp. 135-202.
- [3] Card, S. K., Moran, T. P., Newell, A., The Keystroke-Level Model for user performance with interactive systems, *Communications of the ACM*, vol. 23, 1980, pp. 396-410.
- [4] Harrison, M. D., Duke, D. J., A review of formalisms for describing interactive behavior, *In the Software Engineering and Human-Computer Interaction – Notes in Computer Science*, vol. (896), Springer-Verlag, 1995, pp. 49-75.
- [5] Wickens, D., Hollands, G., *Engineering Psychology and Human Performance*, Prentice-Hall, Englewood Cliffs, NJ, 2004.
- [6] Ashcraft, M. H., *Cognition*, Third Edition, Prentice-Hall, Englewood Cliffs, NJ, 2001.
- [7] Goldstein, E. B., *Sensation and Perception*: 6th Edition, Wadsworth Publishing, Pacific Grove, CA, 2002.
- [8] Ware, C., *Information Visualization: Perception for Design*, 2nd edition, Morgan Kaufmann Publishers, San Francisco, CA, 2004.
- [9] Atkinson R. C., Shiffrin, R. M., Human Memory: a proposed system and its control processes, In: K. W. Spence and J. T. Spence (eds.): *The Psychology of Learning and Motivation*, vol. 2, New York, Academic Press, 1968, pp. 89-195.
- [10] Sperling, G. A.: A model for visual memory tasks, *Human Factors*, 5, 1963, pp. 19-31.
- [11] Spiegel, M. R., *Theory and Problems of Probability and Statistics*, New York: McGraw-Hill, 1992.