Realistic and adaptive cognitive training using virtual characters

D Sjölie

Department of Applied IT, University of Gothenburg,
412 96 Gothenburg, SWEDEN

daniel.sjolie@ait.gu.se

www.ait.gu.se

ABSTRACT

Computer-aided cognitive training has the potential to be an important tool in the fight against dementia and cognitive decline but many challenges remain. This paper presents an example of how realistic and adaptive training may address these challenges. Virtual characters were used as stimuli in a dual n-back working memory task in a realistic 3d-environment. Support for continuous adaptation was a priority, including adaption based on affective states such as arousal.

1. INTRODUCTION

An increasingly older population around the world makes the need for methods to combat dementia and age-related cognitive decline increasingly urgent. Computer-aided cognitive training has been championed as one such method with great potential (Klingberg, 2010; Li et al, 2008). Unfortunately, recent research has shown that it is very difficult to achieve general cognitive improvements with cognitive training, i.e., transfer effects (Owen et al, 2010). This paper provides an example of how solutions building on an understanding of human brain function in realistic environments can be used to meet this challenge. The importance of realistic interaction and the ecological validity of training are fundamental motivations for the use of virtual reality (VR) for rehabilitation and training (Pugnetti et al, 1995; Rizzo et al, 2001).

The difficulty to gain general improvements from cognitive training is commonly described in terms of a distinction between near transfer and far transfer. Training a particular task very often leads to improvements on similar tasks (near transfer). Transfer to tasks with no clear similarity to the trained task, however, (far transfer) has proved difficult (Dahlin et al, 2008; Li et al, 2008). Much recent work on cognitive training has been focused on trying to get general improvements related to, e.g., attention or working memory. Such general improvements should result in improved performance in everyday tasks such as remembering what to shop but they require far transfer. One response to this difficulty of achieving far transfer is to focus on near transfer and on the need to train the right thing, e.g., using interactive systems with high ecological validity. Reality-based human-computer interaction (HCI) in general and virtual reality (VR) in particular provides a foundation for ecologically valid computer applications by building on the user's skills and experiences from reality (Jacob et al, 2008; Rizzo et al, 2004).

2. PREVIOUS WORK

2.1 Cognitive training

One form of cognitive training that has attracted much attention is working memory (WM) training. Working memory capacity predicts performance in a wide range of cognitive tasks, and many neuropsychiatric conditions such as stroke or attention-deficit hyperactivity disorder (ADHD) coincide with impaired WM (Klingberg, 2010). Several studies have shown that performance on specific WM tasks such as 2-back (comparing the last number in a sequence to the one presented 2 steps before) does improve with training and that this effect does transfer to similar (near-transfer) tasks (Dahlin et al, 2008; Klingberg, 2010; Li et al, 2008; Owen et al, 2010). However, the magnitude and range of transfer, in particular the potential for far-transfer, remains disputed. Studies comparing transfer effects in old and young adults have presented seemingly conflicting results. A study by Dahlin et al. concluded that while transfer to untrained tasks is possible for both young and old the magnitude varies and it is often harder to demonstrate transfer in old adults (Dahlin et al, 2008). In other studies transfer effects in young and old have been compared without any reliable differences (Li et al, 2008). More generally, Owen et al. (2010) failed to show any general cognitive improvements for 11,430 participants training on cognitive tasks online for several weeks. Suggested reasons for such results include variations in the amount and intensity of the training (sometimes with significant individual variance) as well as differences in the overlap between trained tasks and the evaluated transfer task.

Figure 1. *This task was inspired by the dual N-back task used by Jaeggi et al. (2008). The stimuli to keep track of is who (which of the characters) is/was doing what (which animation) N steps back.*

One common working memory training task is *n-back*. In a basic implementation of the n-back task the subject may be presented with a series of numbers and asked to compare each new number to the one seen n steps before. E.g., with n=1 the question is if the new number is the same as the last, with n=2 if it is the same as the number before the last, etc. This requires the subject to constantly remember n previous numbers and to update this list each time a new number is presented. The numbers can be exchanged for any stimuli. In the *spatial n-back* version subjects need to remember *where* a stimulus was presented n steps back, and compare it to the location of new stimuli. If the stimulus itself is also varied this becomes a *dual n-back* task, where the subject must remember and compare both position and identity. This is a very demanding cognitive task, providing one way to increase the intensity of the training. Training on a dual spatial n-back test with sound as the second stimuli has been shown to improve measures of fluid (i.e., general) intelligence (Jaeggi et al, 2008).

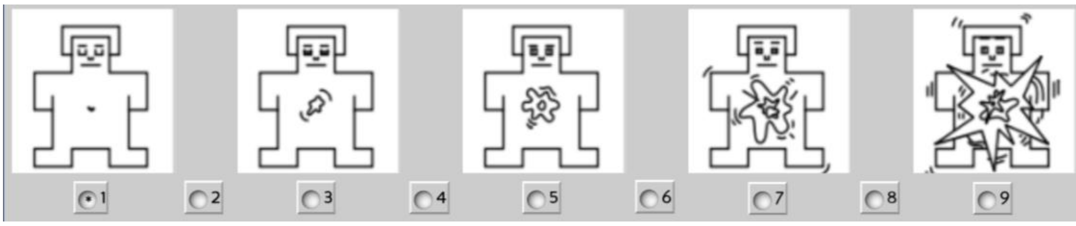

Figure 2. *The Self-Assessment Manikin (SAM) may be used to present a visual scale of affective dimensions to users to get quick responses. We used scales for arousal (shown here) and valence with 5 manikin images plus in-between scale steps.*

2.2 Presence and Synchronization for Virtual Rehabilitation

Presence has traditionally been described as the sense of "being there" in a virtual environment (Slater, 2002) but recent elaborations have a closer connection to cognition and the human brain. An emphasis on presence as hypothesis selection (Sanchez-Vives and Slater, 2005) connects to the importance of predictions and prediction errors in recent theories of brain function and presence as "the ability to act there" connects to the use of existing motor representations (Jäncke et al, 2009). The concept of *mental simulations* can be helpful in understanding why existing expectations and representations are so important and in getting a handle on how cognitive training can be designed with this in mind. Mental simulation includes unconscious and flexible reactivation of memories, employed to recognize the current context and to simulate, or predict, possible actions and expected results (Barsalou, 2008). The idea that predicting future events based on previous experience is a critical aspect of how the brain works has gathered increasing support in recent years. It is prominent in recent theories of cognition and brain function by Hawkins (2005), Friston (2010) and others. A recent paper by Clark (2013) provides a broad introduction.

Particularly interesting for presence and virtual rehabilitation is how this framework suggests that the brain essentially implements a running simulation of reality. This has prompted descriptions of presence as related to the synchronization between an external environment (real or virtual) and the subjective mental reality simulated

©2014 ICDVRAT; ISBN 978-0-7049-1546-6

by the brain (Sjölie, 2012). Related theories of brain function provide a basis for understanding how such synchronization develops and how we may design a virtual environment to facilitate synchronization of specific mental simulations, corresponding to cognitive skills. For example, the combination of familiar stimuli with a familiar context should provide an optimal foundation for internalization and adaptation of a specific task (i.e., learning).

3. REALISTIC COGNITIVE TRAINING

We have implemented a version of dual n-back task using a realistic 3d-environment with animated characters in order to increase the familiarity and realism of both the stimuli and the environmental context (fig. 1). The task is transformed to remembering which characters have made which movements over the last few steps. The moving character corresponds to the position and the different animations correspond to the second stimuli. Based on the reasoning presented above, learning to keep track of who did what just a minute ago, in a realistic 3d-environment, should have a greater chance of producing transfer to similar everyday situations than training using, for example, n-back with arbitrary images at different 2d positions.

Much effort was spent on making it possible to balance the difficulty of the training. A gamepad was used to provide a clear and simple interface with few buttons. Buttons on the front of the gamepad were used to answer the repeated n-back question of whether the stimuli were the same or not, with buttons on the right and left respectively for each question (what/who, see fig. 1). Normally, new stimuli is presented at regular time intervals but in order to reduce the risk of subjects giving up completely if they get behind we introduced an optional mode to wait for the subject to respond. This was coupled with feedback on late responses and a count of late responses. The feedback was very noticeable as the lighting in the entire virtual environment changed. Similar feedback was given in response to wrong answers. The duration of the regular time interval could also be changed to adapt the level of stress in the task. The application has been implemented using Panda3D, a full featured 3d game engine.

3.1 Adapting to Affective State

Synchronization and optimal training depends on a suitable level of prediction errors. According to activity theory (Kaptelinin and Nardi, 2012) any human activity is driven by a need to change something in the environment. Something important is not as we would prefer it to be, it does not match our “prediction” of the ideal world, and the mismatch arouses us to action. Based on such reasoning we attempted to measure the subject’s level of arousal throughout the training using Self-Assessment Manikin (SAM, fig. 2) questions at regular intervals.

In an attempt to provide automatic adaptation of the cognitive training we used a commercial EEG headset (Emotiv EPOC) to try to measure mental workload related to arousal. Data was collected over 13 trials with 4 subjects before data collection was aborted because of poor classification performance. The classification of the EEG data was not good enough to enable successful automatic adaptation for most users, but trends for selected subjects suggest that such adaptation may be possible given optimal conditions (fig. 3).

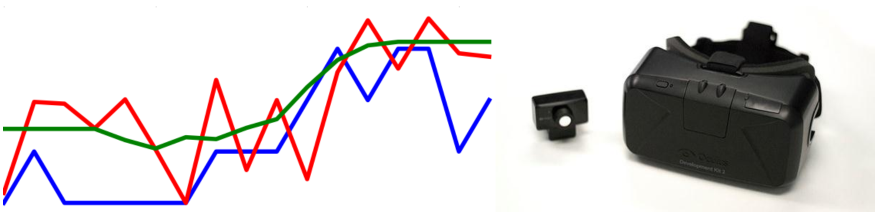

Figure 3. *Left: Example of classified arousal (red) and reported arousal (blue) based on EEG measurements. The Y-axis is arousal (on a 0.0-1.0 scale) and the X-axis is consecutive 1-minute task blocks over one trial. Notice the trend when classified arousal is smoothed (green). Right: The latest development version of the commercial and accordable Oculus Rift HMD*

4. FUTURE WORK

In order to further investigate the effect of the realistic and adaptive cognitive training presented here it needs to be compared to solutions with varying degrees of realism and adaptivity. Our system already implements more traditional n-back tasks such as images shown on different locations, in 2d- or 3d-space. This makes it possible to evaluate training with logically identical tasks while stimuli and environments vary in familiarity and realism. It is also possible to increase the realism of the environment and context by immersing the subject in the virtual environment using traditional VR technologies such as head-mounted displays (HMDs). The recent development of affordable HMDs (fig. 3) suggests that such systems may indeed become common in the near future. Support for the Oculus Rift is under development for the Panda3D engine used to build our application. Ideally, such different setups should be used to train a large number of subjects over a significant period of time, followed by an evaluation of the resulting transfer to different tasks, including everyday activities.

While the automatic adaptation using EEG failed initially the basic motivations for implementing such functions are valid and much remains investigate. Alternative psychophysiological measures such as galvanic skin response (GSR) or functional near-infrared spectroscopy (fNIRS) would be interesting to evaluate, and a combination of new headsets and improved training protocols may give EEG a new chance.

Acknowledgements: Thanks to Lars-Erik Janlert and Johan Eriksson for comments on the development of the prototype. Thanks to Kempe Foundations and Umeå University for funding.

5. REFERENCES

Barsalou, LW, (2008), Grounded Cognition. Annu. Rev. Psychol. 59, 617–645 doi:10.1146/annurev.psych.59.103006.093639

Clark, A, (2013), Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 36, 181–204, doi:10.1017/S0140525X12000477

Dahlin, E, Neely, AS, Larsson, A, Backman, L, Nyberg, L, (2008), Transfer of Learning After Updating Training Mediated by the Striatum, Science 320, 1510–1512. doi:10.1126/science.1155466

Friston, K, (2010), The free-energy principle: a unified brain theory?, Nat Rev Neurosci 11, 127–138. doi:10.1038/nrn2787

Hawkins, J, (2005), On Intelligence, Owl Books.

Jacob, RJK, Girouard, A, Hirshfield, LM, Horn, M.S, Shaer, O, Solovey, ET, Zigelbaum, J, (2008), Reality-based interaction: a framework for post-WIMP interfaces, in: Proc. CHI 2008. ACM Press.

Jaeggi, SM, Buschkuehl, M, Jonides, J, Perrig, WJ, (2008), Improving fluid intelligence with training on working memory, Proc. Natl. Acad. Sci. 105, 6829–6833. doi:10.1073/pnas.0801268105

Jäncke, L, Cheetham, M, Baumgartner, T, (2009), Virtual reality and the role of the prefrontal cortex in adults and children, Front. Neurosci. 3. doi:10.3389/neuro.01.006.2009

Kaptelinin, V, Nardi, B, (2012), Activity Theory in HCI: Fundamentals and Reflections. Synth. Lect. Hum.-Centered Inform, 5, 1–105. doi:10.2200/S00413ED1V01Y201203HCI013

Klingberg, T, (2010), Training and plasticity of working memory, Trends Cogn. Sci. 14, 317–324. doi:10.1016/j.tics.2010.05.002

Li, S-C, Schmiedek, F, Huxhold, O, Röcke, C, Smith, J, Lindenberger, U, (2008), Working Memory Plasticity in Old Age: Practice Gain, Transfer, and Maintenance, Psychol. Aging 23, 731–742. doi:10.1037/a0014343

Owen, AM, Hampshire, A, Grahn, JA, Stenton, R, Dajani, S, Burns, AS, Howard, RJ, Ballard, CG, (2010), Putting brain training to the test, Nature 465, 775–778. doi:10.1038/nature09042

Pugnetti, L, Mendozzi, L, Motta, A, Cattaneo, A, Barbieri, E, Brancotti, A, (1995), Evaluation and retraining of adults' cognitive impairments: Which role for virtual reality technology?, Comput. Biol. Med. 25, 213–227. doi:10.1016/0010-4825(94)00040-W

Rizzo, AA, Buckwalter, JG, McGee, JS, Bowerly, T, Zaag, C, van der, Neumann, U, Thiebaux, M, Kim, L, Pair, J, Chua, C, (2001), Virtual Environments for Assessing and Rehabilitating Cognitive/Functional Performance A Review of Projects at the USC Integrated Media Systems Center, Presence Teleoperators Virtual Environ. 10, 359–374. doi:10.1162/1054746011470226

Rizzo, AA, Schultheis, M, Kerns, KA, Mateer, C, (2004), Analysis of assets for virtual reality applications in neuropsychology, Neuropsychol. Rehabil. 14, 207–239.

Sanchez-Vives, MV, Slater, M, (2005), From presence to consciousness through virtual reality, Nat. Rev. Neurosci. 6, 332–9. doi:nrn1651

Sjölie, D, (2012), Presence and general principles of brain function, Interact. Comput. 24, 193 – 202. doi:10.1016/j.intcom.2012.04.004

Slater, M, (2002), Presence and the sixth sense, Presence Teleoperators Virtual Environ. 11, 435–439. doi:10.1162/105474602760204327