Volume 4, Issue 1, Spring 2011

JCR
Journal of CyberTherapy &Rehabilitation

Official Journal of the International Association of CyberPsychology, Training & Rehabilitation

www.vrphobia.eu

In This Issue

SPECIAL ISSUE!

Select Papers from the European Conference on Cognitive Ergonomics 2010, Including:

THE VIRTUAL REALITY MEDICAL INSTITUTE

INJURY CREATION SCIENCE

The Next Generation of Injury Simulation Today

Prosthetic tissue, wounds, and life saving skills training devices used in the training of medical professionals

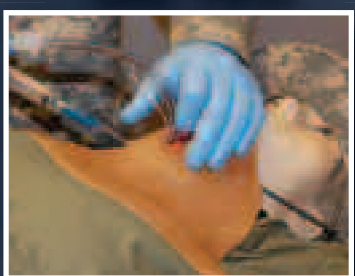
- Cricothyrotomy Skills Trainer
- Needle Decompression Skills Trainer
- Bleeding Wound Skills Trainer
- Amputation Skills Trainer
- Burn Wound Skills Trainer
- Odor Wound Skills Trainer

Merging latest special effects technology with medical and material sciences research to replace live tissue and training.

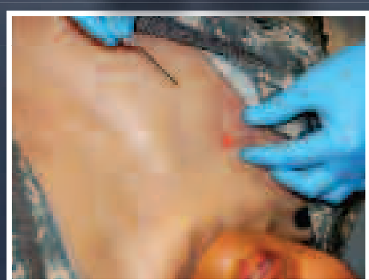
Physiologically based research and development program focused on providing enhanced training capabilities for medical professionals to include:

- Basic Life Support
- Patient Assessment
- Hemorrhage Control
- Fracture Management
- Shock Prevention & Treatment

Cricothyrotomy Skills Trainer

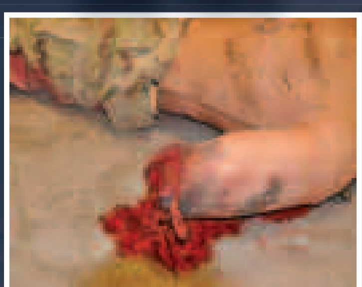


Needle Decompression Skills Trainer



Bleeding Wound Skills Trainer

Severe Amputation Skills Trainer



Simulated Burn Wound Package

Odor Simulation Wound Kit

Visually Realistic - Comfortable - Easy to Use
Durable - Reusable - Tactilely Realistic

VRMC
VIRTUAL REALITY MEDICAL CENTER

FOR MORE INFORMATION, CONTACT:
Mark D. Wiederhold, M.D., Ph.D. FACP
The Virtual Reality Medical Center
858.642.0267 mwiederhold@vrphobia.com
www.vrphobia.com

JOURNAL OF CYBERTHERAPY & REHABILITATION

Volume 4, Issue 1, Spring 2011

THE VIRTUAL REALITY MEDICAL INSTITUTE

iACToR
International Association of CyberPsychology, Training & Rehabilita

Brussels, Belgium – San Diego, California

JCR
Journal of
CyberTherapy
& Rehabilitation

Journal of CyberTherapy & Rehabilitation

Spring 2011
Volume 4, Issue 1

EDITORIAL

There is an emerging body of literature about the proliferation of social networking sites (SNS) and their effects on mental health. To date, much of it has focused on investigating the possible negative effects of SNS, such as Internet addiction. However, research also supports the benefits of SNS in mental health, addictions, stigmatized identities, trauma and violence recovery, and grief support. As clinicians and researchers, we are just beginning to harness the power of SNS to promote mental well-being.

Participation in SNS has increased dramatically over the past five years. A 2010 Pew report showed that 73% of online teens and 47% of online adults in the U.S. used SNS. Another survey conducted by Pew in April–May 2010 noted that Poland, Britain, and South Korea are close behind the U.S. in SNS usage, followed by France, Spain, Russia, and Brazil. Lower participation in other countries is due primarily to less-wired populations. Notable exceptions are Germany and Japan, where Internet usage is high but SNS usage is low.

The European Union has been investing in e-Health since 2004, when outgoing Public Health and Consumer Protection Commissioner David Byrne said, "We need a … Europe where people have easy access to clear and reliable information on how to be in good health and about diseases and treatment options." An outgrowth of the European Parliament hearing at which he testified was the creation of the ICT (information and communication technologies) for Health, enabling health service providers in different EU member states to work together to exploit these technologies. More recently, the First International E-Mental Health Summit in Amsterdam in 2009 organized by the Trimbos Institute in collaboration with the International Society for Research on Internet Interventions attracted 500 participants from more than 40 countries. In the U.S., the new healthcare reform law provides financial incentives for providers to use health information technology and electronic health records, and in March 2011 leaders in healthcare technology will share their innovations in San Diego and San Francisco, California for the Health 2.0 conference.

In one such innovation, a researcher used a GPS-enabled phone and a location-aware SNS to design a system to help trainees with cognitive impairment who felt lost to find a nearby caregiver. These individuals were enrolled in a supported employment program that provided them with a job coach to help them get to and from work for the first few weeks. The system was programmed to send text messages to the job coach and time and location alarms to help the trainee get to work on time. This type of SNS could enable parents, guardians, and caregivers to watch loved ones unobtrusively.

A recent study of 217 college-age participants in South Korea found that SNS network size was positively related to subjective well-being, and the results suggest that this is due to self-disclosure. In the SNS context, it is postulated that the positive association with well-being results from the self-disclosure "confession effect," the expectation of mutual self-disclosure, and the expectation of social support.

A case study report found that deploying the Three Good Things positive psychology exercise as a Facebook application was viable, with a 1% dropout rate, which is similar to or better than other online wellness applications. In the exercise, people post three good things that happened, along with the reasons they think they happened. People found that sharing with others and viewing other's posts were valuable, as long as they were able to choose which comments they made were public and which were private.

Specialized health SNS such as PatientsLikeMe and DailyStrength offer emotional support, social support, and patient empowerment; some also offer physician Q&A, quantified self-tracking, and clinical trials access. PatientsLikeMe includes support for mental disorders such as anxiety, bipolar affective disorder, depression, obsessive-compulsive disorder, and Posttraumatic Stress Disorder; DailyStrength provides support for an even broader array of mental health issues. In an online SNS, inhibitions may be lowered, anxiety may be lessened, and anonymity may be increased. This presents the ideal 24/7 support for treatment of people with disorders such as depression. Indeed, the Pew report showed that teens look online for health information about issues they find are embarrassing to talk about such as drugs, sex, and depression.

Of course, there are cautions. One study found that people with depression who used an online SNS spiraled down if they had friends who were moderately or severely depressed and had a negative opinion of the SNS. The researchers concluded that the SNS could be helpful if people take a break from it if their posts elicit these reactions.

A position paper on pervasive healthcare concludes that "[provided-designed systems and services] should include help for people to access peer-to-peer social support sharing and caring in order to encourage sustained engagement with self management to build positive healthy identities for themselves." Online health consumers are beginning to rely on "patient opinion leaders" for advice on chronic conditions such as mental disorders, and we need to be there with them. Of course, we must be mindful of issues such as privacy and data accuracy as we create tools to help SNS participants balance their needs to share information with their needs to manage self-presentation. Nonetheless, as clinicians and researchers, we should take advantage of SNS to extend the practice of evidence based medicine and mental health.

Brenda K. Wiederhold, Ph.D., MBA, BCIA
Editor-in-Chief,
Journal of CyberTherapy & Rehabilitation
Virtual Reality Medical Institute

Journal of CyberTherapy & Rehabilitation
Spring 2011, Volume 4, Issue 1

GUEST EDITORIAL
Cognitive Engineering in Mental Health Computing

Willem-Paul Brinkman, *Delft University of Technology, The Netherlands*

Computer applications in the support of mental healthcare and rehabilitation are becoming more widely used. They include technologies such as Virtual Reality (VR), electronic diaries, multimedia, brain computing and computer games. Research in this area is emerging, and focuses on a variety of issues such as clinical effectiveness of a computer supported intervention, usability of a system, human values that are affected by a technological intervention, actual use, acceptance, and accessibility of these systems. To classify and to understand the objectives of the work presented in this special issue the mental health computing research model is presented. This descriptive model identified four research categories: (1) technology elements, (2) technology intervention, (3) clinical effect, and (4) field use. Each category has its own focus, methods and set of researchers.

Keywords: Cognitive Engineering, Mental Health Computing, Research Model

Special Issue

Mental disorders make up a large amount of the global burden of disease, with an estimation of 14% attributed to neuropsychiatric disorders, and 28% of non-communicable diseases (Prince et al., 2007). Computer support in the diagnosis, care or rehabilitation of these disorders is increasing, motivated by the improvement of effectiveness or more efficient use of healthcare resources. Computer support in prevention programs to make individuals more resilient against mental stress and enhance their overall mental wellness is also receiving research attention. Considerable attention has been devoted to studying the efficacy of some of these systems, and some attention is now being given to the use of these technologies, for example, in a recent special issue in the Interacting with Computers journal (Doherty & Bickmore, 2010). As has been recognized in areas such as consumer electronics and office applications, the interplay between technology and humans is an important factor in determining the use and acceptance of technology. Likewise, as a result, these aspects also seem to be important in the domain of mental health computing, considering the use, acceptance and accessibility of computer support systems for clinical psychology and also the human values these systems effect. This special issue, therefore, focuses on cognitive engineering for technology in mental healthcare and rehabilitation, or more specifically, mental health computing. While cognitive engineering is an approach to analyze, model, design and evaluate interactive, complex systems, cognitive ergonomics is a relatively young branch of science, which focuses on the reciprocal influence between work and the human mind (Hollnagel, 1997), which often nowadays involves the use of computer technology. In this context, work could mean the activities of the therapist, for example, when administering treatment, but also include physical or mental efforts of a patient to accomplish a specific goal. Enhancing knowledge in this area seems essential, as computer support systems for mental health are becoming more complex as multiple actors are supported (e.g. patients, therapists, technical support, social support network, daily care takers), systems are distributed over time and place (e.g. a-synchronized, mobile, and remote care or treatment), and technologies are become more sophisticated (e.g. physiological sensors, artificial intelligent, multi modal interaction). Adding to the complexity of the situation is also the mixture of mental health researchers and technology oriented researchers that are active in this field. However, their involvement differs widely depending on the research focus, which often becomes visible in evaluation activities. For example, some talk about a two-stage approach for the collaborative process between these two types of researchers in the design of this type of technology (Coyle, Doherty, Matthews, & Sharry, 2007), one focusing on the design and development of technology, and the other on the clinical evaluation. Moving from a development view to a research view, research can be placed into four categories (Figure 1), each with their own focus, methods and involvement of researchers.

These categories are (1) technology element, focusing on the specific technology component and its effect on interaction with humans; (2) technology intervention, focusing on the design of a usable and acceptable computer support system; (3) clinical effect, focusing on clinical effectiveness of computerized intervention, and (4) field use, focusing on the use of the computerized intervention in the field. From an intervention perspective, these categories might appear to be successive stages. However, for the researchers involved, especially in the technology element category, a specific intervention might not yet be envisioned. In the opposite direction of the technological innovation is the information flow from the clinical field. This includes feedback about the effect of current technology intervention in the field; problems or limitations of non-technology supported treatment and care; and also new opportunities due to new theories or enhanced clinical knowledge. It is also important to acknowledge values. For example, patient autonomy, respect, and privacy are human values that drive demands and acceptance of technological innovations. With a focus on cognitive engineering, the contributions in the special issue particularly address research in the technology element and intervention categories. Whereas research in these two categories also includes evaluation, in the clinical effect category evaluation is clinical, often directly including the target group, e.g. the patients, with a strong involvement of mental health researchers. Before large scale randomized controlled trails (RCT) are conducted, such as a comparison of the treatment effect between treatment with and without technology intervention (Botella et al., 2007; Krijn et al., 2007), small pilot studies are often carried out, ensuring that researchers have a sufficient understanding about potential additional clinical factors, and ensuring that the technology is stable and usable. RCTs are time consuming, and technological innovations are often frozen once the RCT is underway. One such pilot study is presented by Hourani, Kizakevich, Hubal, Spira, Strange, Holiday, Bryant, and McLean in this issue. These authors compared the efficacy of a newly designed stress inoculation training (SIT) based intervention, predeployment SIT (PRESIT), with current best practices, in this case the U.S. Navy and Marines' combat and operational stress control programs. PRESIT includes a multimedia stressor environment (MSE) that allowed individuals to travel

← Identified problems, opportunities and values			
Technological innovations →			
Technology element	**Technology Intervention**	**Clinical effect**	**Field use**
Focus on effect of technology components that affect the interactions with a user Methods include, for example, lab studies, often with non-patients Strong involvement of technology-oriented researchers	Focus on establishing usable technological health intervention Methods include both design activities and, for example, usability studies, often with non-patients, but also with therapists Often multidisciplinary team	Focus on efficacy of treatment with technology intervention Methods include case studies and randomized controlled trials with patients Strong involvement of mental health researchers	Focus on daily practice on technology intervention Methods include field observations, or surveys among patients and/or therapists Involvement of clinicians

Figure 1. Mental health computing research model.

in a vehicle through an Iraqi village in a virtual scenario, encountering enemies or events such as explosions, but also cues individuals had to react on to test their reaction time. The results of the pilot study showed support for the MSE and PRESIT in general.

Research in the field use category can also examine clinical efficacy, however, this time in a practical, less controlled setting, for example, a study about observations made in nearly 500 virtual reality (VR) therapy sessions in a clinic (Wiederhold & Wiederhold, 1999). The use and usability problems can also be studied in the field, as was done, for example, during field observations on how therapists used a VR exposure therapy system in the treatment of fear of flying (Brinkman, Sandino, & van der Mast, 2009; Brinkman, van der Mast, Sandino, Gunawan, & Emmelkamp, 2010). As technological innovations for mental healthcare and rehabilitation will become more established, research in this category is likely to increase.

The majority of the work presented in the special issue is focused on the technology element and innovation categories (Figure 2). Research that falls in the technology element category often provides knowledge that can be applied across multiple technology interventions. A good example of this is the work of Busscher, de Vliegher, Ling and Brinkman presented in this issue. They present a study into physiological measurement and the evaluation on neutral VR worlds. Because of individual differences, the physiological baseline measurement is often done in a virtual environment without stressors. However, because of the technology novelty factor, individuals might already be aroused. Their study, however, showed that a neutral virtual world does not have to lead to an increase in arousal, a finding that is beneficial for VR interventions that use physiological measurements. The work presented by Brouwer, Neerincx, Kallen, van der Leer and ten Brinke also provides new insights which are relevant for interventions that apply a VR neuro-bio paradigm, for example, to treat stress-related disorders. They induced stress by simulating a bomb explosion in VR and by giving negative feedback to participants about their performance. Analyzing EEG, ECG and cortisol level, they found that stress was reflected in EEG mid-frontal alpha asymmetry, heart rate variability, and cortisol level. The last paper in the special issue that also can be placed in the technology element category is the work from Cherni, Kadri, Taruella, Joseph, Le Roy and Klinger. The technology element they studied was the screen size of a virtual system that delivered virtual information. They found that increasing the size of the screen improved their participants' perception of the visual information and had a positive effect on performance in the virtual task. Although this task is especially relevant as a part of cognitive rehabilitation intervention for patients with a brain injury, the finding also seems relevant for VR-based therapy in general, and for VR-based applications for non-patients such as virtual training environments. All three studies in this category used non-patients in their experiments, and in two studies this was combined with a group of patients.

As the work of Cherni et al. focuses on a specific intervention, it can also be placed in the technology intervention category. Likewise, the work of ter Heijden and

Technology element ⟷		Technology intervention	Clinical effect
Evaluating Neutral VR Worlds (Busscher, de Vliegher, Ling, & Brinkman, 2011) Stress Responses in VR (Brouwer, Neerincx, Kallen, van der Leer, & ten Brinke, 2011)	VRE with automated free speech (ter Heijden & Brinkman, 2011) Display size and cognitive rehab (Cherni et al., 2011)	ICT for reminiscing (Mulvenna et al., 2011) Combined use of VR and Music for Rehab (Trobia, Gaggioli, & Antonietti, 2011) The Military – 3MR system (Brinkman, Vermetten, van den Steen, & Neerincx, 2011)	Predeployment stress inoculation training (Hourani et al., 2011)

Figure 2. Contributions to the special issue.

Brinkman, which focused on the treatment of social phobia in VR, has a technology element that is beneficial for the intervention of other mental disorders. Still, their work can be seen as an exploration of various design solutions of a specific intervention – VR-exposure for social phobic patients. They compared three types of automatic free speech implementations that allowed individuals to have a free speech dialogue with a virtual character in a VR environment. They set this against a control condition in which a human operated the verbal response of the character. The study included a group of non-patients and two phobic patients. On several measures the human control condition did not outperform all the automatic conditions, suggesting that an automated semi-scripted conversation might create a similar experience as a manually controlled conversation. Trobia, Gaggioli, and Antonietti also studied the feasibility of a technology intervention – an integrated training approach that combined VR technology and music to support stroke patients in the performance of mental practice. Over a period of eight weeks two stroke patients received treatment with the system, and results showed an increase in motor scores and an improvement in reported activities of daily living. Stroke rehabilitation, therefore, seems to be supported by this intervention. The idea of combining multiple modalities can also be found back in the Military Multi-Modal Memory Restructuring System put forward by Brinkman, Vermetten, van den Steen and Neerincx. Their system was designed to treat combat veterans with posttraumatic stress disorder and focuses on the restructuring and relearning of past events. They report on a series of design and evaluation activities to establish a technology intervention that was usable and enhanced storytelling. As the focus was on establishing a complete design solution, they followed a situated cognitive engineering approach (Neerincx & Lindenberg, 2008) to study the various factors and establish requirements and specifications. The use of design methods seems specifically relevant for the technology intervention category. The focus of these methods, however, should not be limited to technology but should also consider interaction with the various actors and their values as emphasised by value sensitive design methods (Friedman, Kahn, & Borning, 2006). For example, Mulvenna, Doyle, Wright, Zheng, Topping, Boyle and Martin, with their empirical evaluation of card-based and device-based reminiscing not only looked at usability, but also looked at the attitude of older people towards their system. They concluded that results showed no specific barriers to the usage of the system for reminiscing activities.

People Involved

In August 2010 the workshop "Cognitive Engineering for Technology in Mental Health Care and Rehabilitation" was held in Delft, The Netherlands as part of the European Conference on Cognitive Ergonomics. The workshop was organized by Willem-Paul Brinkman, Gavin Doherty, Alessandra Gorini, Andrea Gaggioli, and Mark Neerincx. The special issue is one of the outcomes of that workshop.

Many people have been involved in this special issue and especially the help of the Managing Editor, Emily Butcher, was much appreciated. The following people have acted as reviewers in the special issue:

Mariano Alcañiz, Willem-Paul Brinkman, Yang Cai, Julian Dooley, Robert Hubal, Paul Kizakevich, James Spira Giuseppe Riva, Charles van der Mast, and Valentijn Visch.

References

Botella, C., Garcia-Palacios, A., Villa, H., Banos, R. M., Quero, S., Alcaniz, M., et al. (2007). Virtual reality exposure in the treatment of panic disorder and agoraphobia: A controlled study. *Clinical Psychology & Psychotherapy, 14*(3), 164-175.

Brinkman, W. P., Sandino, G., & van der Mast, C. (2009). Field Observations of Therapists Conducting Virtual Reality Exposure Treatment for the Fear of Flying. *Ecce 2009 - European Conference on Cognitive Ergonomics, 258*, 169-176.

Brinkman, W. P., van der Mast, C., Sandino, G., Gunawan, L. T., & Emmelkamp, P. M. G. (2010). The therapist user interface of a virtual reality exposure therapy system in the treatment of fear of flying. *Interacting with Computers, 22*(4), 299-310.

Brinkman, W. P., Vermetten, E., van den Steen, M., & Neerincx, M. A. (2011). Cognitive engineering of a military multi-modal memory restructuring system. *Journal of CyberTherapy & Rehabilitation, 4*(1), 83-99.

Brouwer, A. M., Neerincx, M. A., Kallen, V., van der Leer, L., & ten Brinke, M. (2011). EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress. *Journal of CyberTherapy & Rehabilitation, 4*(1), 27-40.

Busscher, B., de Vliegher, D., Ling, Y., & Brinkman, W. P. (2011). Psysiological measures and self-report to evaluate neutral virtual reality worlds. *Journal of CyberTherapy & Rehabilitation, 4*(1), 15-26.

Cherni, H., Kadri, A., Taruella, A., Joseph, P. A., Le Roy, C., & Klinger, E. (2011). Virtual information display for cognitive rehabilitation: choice of screen size. *Journal of CyberTherapy and Rehabilitation, 4*(1), 73-81.

Coyle, D., Doherty, G., Matthews, M., & Sharry, J. (2007). Computers in talk-based mental health interventions. *Interacting with Computers, 19*(4), 545-562.

Doherty, G., & Bickmore, T. (2010). Special Issue Supportive Interaction: Computer Interventions for Mental Health Introduction. *Interacting with Computers, 22*(4), 241-242.

Friedman, B., Kahn, P. H., & Borning, A. (2006). Value sensitive design and information systems. In P. Zhang & D. Galletta (Eds.), *Human-computer interaction and management information systems: foundations* (pp. 348-372). Armonk, NY: M.E. Sharpe.

Hollnagel, E. (1997). Cognitive ergonomics: it's all in the mind. *Ergonomics, 40*(10), 1170-1182.

Hourani, L. L., Kizakevich, P. N., Hubal, R., Spira, J., Strange, L. B., Holiday, D. B., et al. (2011). Predeployment stress inoculation training for primary prevention of combat-related stress disorders. *Journal of CyberTherapy & Rehabilitation, 4*(1), 101-116.

Krijn, M., Emmelkamp, P. M. G., Olafsson, R. P., Bouwman, M., Van Gerwen, L. J., Spinhoven, P., et al. (2007). Fear of flying treatment methods: Virtual reality exposure vs. cognitive behavioral therapy. *Aviation Space and Environmental Medicine, 78*(2), 121-128.

Mulvenna, M. D., Doyle, L., Wright, T., Zheng, H., Topping, P., Boyle, K., et al. (2011). Evaluation of card-based versus device-based reminiscing using photographic images. *Journal of CyberTherapy & Rehabilitation, 4*(1), 63-72.

Neerincx, M. A., & Lindenberg, J. (2008). Situated cognitive engineering for complex tasks environments. In J. M. C. Schraagen, L. Militello, T. Ormerod & R. Lipshitz (Eds.), *Naturalistic decision making and macrocognition* (pp. 373-390). Ashgate, UK: Aldershot.

Prince, M., Patel, V., Saxena, S., Maj, M., Maselko, J., Phillips, M. R., et al. (2007). Global mental health 1 - No health without mental health. *Lancet, 370*(9590), 859-877.

ter Heijden, N., & Brinkman, W. P. (2011). Design and evaluation of a virtual reality exposure therapy system with automatic free speech interaction. *Journal of CyberTherapy & Rehabilitation, 4*(1), 41-56.

Trobia, J., Gaggioli, A., & Antonietti, A. (2011). Combined use of music and virtual reality to support mental practice in stroke rehabilitation. *Journal of CyberTherapy & Rehabilitation, 4*(1), 57-61.

Wiederhold, B. K., & Wiederhold, M. D. (1999). Clinical observations during virtual reality therapy for specific phobias. *Cyberpsychol Behav, 2*(2), 161-168.

Journal of CyberTherapy & Rehabilitation
Spring 2011, Volume 4, Issue 1

PHYSIOLOGICAL MEASURES AND SELF-REPORT TO EVALUATE NEUTRAL VIRTUAL REALITY WORLDS

Bert Busscher[1,2], Daniel de Vliegher[3], Yun Ling[3] and Willem-Paul Brinkman[3]

Using virtual reality technology for exposure therapy to treat patients with anxiety disorders is attracting considerable research attention. The ability to monitor patient anxiety levels helps therapists to set appropriate anxiety arousing situations. Physiological measures have been put forward as objective indicators of anxiety levels. Because of individual variation, they need a baseline recording which is often conducted in neutral virtual worlds which do not include phobic stressors. Still, because of the novelty of the virtual worlds, reports in the literature suggest that individuals already show some level of arousal when placed in these worlds. This paper presents two studies which look at the effect two different neutral virtual worlds can have on individuals. Findings suggest that a neutral world does not have to result in an increased level of arousal.

Keywords: Virtual Reality, Exposure Therapy, Physiological Measurements, Neutral Worlds, Fear of Flying

Introduction

Virtual Reality Exposure Therapy (VRET) is receiving considerable research attention for treatment of patients suffering from anxiety disorders such as claustrophobia, fear of driving, acrophobia, spider phobia, social phobia, panic disorder with agoraphobia, Posttraumatic Stress Disorder (PTSD), and fear of flying. VRET is based on the idea of gradual exposure in vivo, considered the gold standard for treatment of phobias. Recent meta-studies (Gregg & Tarrier, 2007; Parsons & Rizzo, 2008; Powers & Emmelkamp, 2008) show that exposure in VR is as effective as exposure in vivo. An important element of the therapy is that the exposure is done gradually to more anxiety-arousing situations. Therapists are, therefore, continuously monitoring the anxiety level of a patient. This can be done using Subjective rating of Anxiety (SUD), behavioral observations or physiological measures. The latter has the advantage of being more objective and can be used directly by a computer to assist a therapist in a multi-patient VRET setting (Paping, Brinkman, & van der Mast, 2010). Physiological measures, however, need a baseline measurement because of individual variation. One often used procedure is to obtain a physiological baseline recording when the patient is placed in a neutral VR world, i.e. a VR world which should not include phobia-related stressors. Even if this world has no phobia-related stressor, it is not clear whether the experience of being placed in a Virtual Environment (VE) causes some level of anxiety. Some authors (Wiederhold & Wiederhold, 2005) have suggested that the majority of non-phobic individuals do get some level of arousal when placed in a VE. For example, Jang et al. (2002) report a study with non-phobic individuals and observed that participants were initially aroused in the VR exposure, but returned to a normal baseline after approximately seven minutes. In another study, Wiederhold et al. (1998) also report that non-phobics, when placed in a VE, initially show some level of anxiety. They argued that the VE is a new and novel stimulus and therefore causes this effect. Expanding on this line of reasoning, this paper explores whether the design of the neutral world can also contribute to this effect. Or in other words, would it be possible to design a truly neutral world? As reported in this paper, we were confronted with this question after the re-

Corresponding Author:
Bert Busscher MSc, Department of Clinical, Health and Neuropsychology, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands, VALK Foundation, Sandifortdreef 1a, 2333 ZZ Leiden, The Netherlands, Tel: +31 (0)71 5273733; Fax: +31 (0)71 5273796, E-mail: b.h.busscher@ umail.leidenuniv.nl

[1]Department of Clinical, Health and Neuropsychology, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands
[2]VALK Foundation, Sandifortdreef 1a, 2333 ZZ Leiden, The Netherlands
[3]Man-Machine interaction, Department of Mediamatics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, the Netherlands

sults of our first study suggested that both phobic and non-phobic participants showed higher heart rates during exposure in a neutral virtual world than in both the VE with phobic stressors and in the recovery phase after the VR exposure. Furthermore, both phobics and non-phobic individuals experienced moderate to severe nausea in the neutral VR condition. This called into question the neutrality of the neutral virtual world and led to our research into the creation of neutral virtual worlds.

The paper starts with briefly discussing key concepts such as VR systems, presence and problems experienced by patients. After this, the first study is presented in which both non-phobic and phobic individuals are placed in a neutral VR world, a virtual airplane, and a recovery phase. The second study starts with a discussion of the design of a new neutral VR world. This virtual world aims to be an almost identical representation of the room the individual is sitting in. Results are presented from data collected in four conditions – the real world room, the new neutral VR world, the virtual airplane, and recovery phase. The paper concludes by discussing the findings which suggests that it might be possible to design a truly neutral world. Also, no support was found for a possible transfer of habitation from the physical room to the neutral virtual room.

Background

The sense of being a part of the VE even when a person is physically situated in a totally different real world is considered a key element of VRET. This concept of presence is related to four components – technological devices; user-computer interactions; main task and the user (IJsselsteijn, de Ridder, Freeman, & Avons, 2000; Witmer, Jerome, & Singer, 2005). In the application field of VRET, the main technical devices used are a head mounted display (HMD) and a computer automatic environment (CAVE). The CAVE has a relatively higher immersion level with stereoscopic images on four to six sides around the user while the HMD has only one stereoscopic image in front of the user. In a study on the effects of VRET in patients with acrophobia using CAVE and HMD, it is reported that VRET was superior to no-treatment on anxiety, behavioral avoidance and attitudes towards heights. Although the therapy given in the CAVE resulted in higher level of presence than the therapy given through the HMD, no differences in effect were found between them and the results remained stable during the following six months (Krijn, Emmelkamp, Biemond, et al., 2004). This, therefore, seems to suggest that only a certain level of presence is needed for treatment to be effective. Even with devices such as a HMD or a CAVE, patients can still experience low levels of presence causing them to drop out of the treatment (Krijn, Emmelkamp, Olafsson, & Biemond, 2004). This underlines the idea that presence is also determined by individual factors (Ling, Nefs, Brinkman, Heynderickx, & Qu, 2010) such as vision ability, cognitive processing ability of the VE (T. W. Schubert, 2009), and personality (Wallach, 2010). It has also been suggested that imaginative power influences presence (Huang, Himle, & Alessi, 2000; Regenbrecht, Schubert, & Friedmann, 1998). Being able to visualize more vividly could intensify the experience of the VE. Cybersickness is another potential intervening factor. Cybersickness is a form of motion sickness that occurs as a result of exposure in a VE and can range from a slight headache to an emetic response (Stanney, Mourant, & Kennedy, 1998). Although physiological measurements can be used for determining anxiety during VRET, the side effects of cybersickness can also arouse physiological changes in people (Min, Chung, Min, & Sakamoto, 2004). Both cybersickness and presence therefore seem to be important factors that might explain, besides the initially suggested habituation, physiological effects in neutral virtual worlds.

Study 1

Method

The first study initially aimed to study physiological response of both phobic and non-phobic individuals in a VE with phobic stressors. Both groups were exposed to three conditions – a neutral virtual world, a virtual flight, and a recovery phase. Both the effects for two groups and conditions on physiological recordings and self-reported anxiety were analyzed.

VR System

The VRET Delft 2007 system is described in detail elsewhere (Aslan, 2007; Brinkman, van der Mast, Sandino, Gunawan, & Emmelkamp, 2010; Gunawan, Mast, Neerincx, Emmelkamp, & Krijn, 2004; Schuemie, 2003). Briefly, the HMD used was the stereoscopic Cybermind Visette Pro with a resolution of 640x480 per display and a 60 Hz refresh rate. An Ascension Flock of Birds was used as the tracking tool. Two personal computers (PC) were used in the system – the therapist's computer where the therapist controls the therapy session and a patient PC which gets input from the HMD and therapist computer. Both the neutral virtual world and the flight world were created with WorldUp R4 by Sense8.

Figure 1. Left, the neutral courtyard, right, virtual flight world.

Participants

Participants for Study 1 were aviophobics that applied for therapy at the VALK foundation, and non-paid volunteers without fear of flying who acted as a control sample. The VALK foundation is a mental health clinic that specializes in aviation-related anxiety. During the recruitment period 46 phobic clients who applied for treatment received written information regarding the VR study at their home address two weeks before their first visit. Out of this group, 40 phobics were willing to participate. One client was excluded because of the use of cardioactive medication (ß-blockers). This left 39 phobic clients (15 men) with an average age of 44.5 (SD = 12.4), who fulfilled the DSM-IV criteria for specific situational phobia furnishing usable data. In the same period 22 non-paid volunteers without fear of flying and an average age of 48.3 (SD = 11.4) successfully completed a part of the same protocol. Volunteers were recruited through the social network of the research institution's staff. One of them received a positive diagnosis for aviophobia during the intake and was excluded. Another control subject's questionnaire data was rendered unusable, her physiological data was included for analyses. The 21 non-phobic healthy subjects had flown at least several times; most of them had flown within 18 months of the experiment. None of the control subjects was ever treated for fear of flying. Before the start of the experiment, informed consent was obtained from all participants. The research protocol has been approved by the local medical ethics committee.

Measures

For the physiological recordings, the three target variables were Heart Rate (HR), Pre Ejection Period (PEP) and Respiratory Sinus Arrhythmia (RSA). PEP is considered a measure of (activating) sympathetic cardiac control (Sherwood, et al., 1990), whereas RSA is a measure of (calming) parasympathetic control (Berntson, et al., 1994). Scoring of these variables from thorax impedance and the ECG is described in detail elsewhere (Goedhart, Kupper, Willemsen, Boomsma, & de Geus, 2006; Goedhart, Van der Sluis, Houtveen, Willemsen, & De Geus, 2007). Briefly, from the ECG (sampling rate 1000 Hz) the HR was obtained from the time between two adjacent R waves. PEP was defined from the ECG and ICG as the time interval from the Q-wave onset, the onset of the electromechanical systole, to the B-point (from the ICG), which signals opening of the aortic valves (Sherwood, et al., 1990; Willemsen, DeGeus, Klaver, VanDoornen, & Carroll, 1996). RSA was obtained from the ECG and respiration signals by subtracting the shortest IBI during HR acceleration in the inspirational phase from the longest IBI during deceleration in the expirational phase (i.e. the peak-through method) (Grossman, 1990). When no phase-related acceleration or deceleration was found, the breath was assigned a RSA score of zero. Automatic scoring of PEP and RSA was checked by visual inspection of the impedance and respiratory signal from the entire recording. Our focus on cardiac parameters reflects two major considerations – measurements needed to be as non-invasive as possible and they needed to respond to changes in psychological state over a time scale of a few minutes. The PEP and RSA measures are uniquely qualified to meet both demands (Goedhart, et al., 2006; Willemsen, et al., 1996). Using a visual display of the output of an inbuilt vertical accelerometer, we identified artefact free periods in each condition that lasted at least five minutes each.

All questionnaires were administered in the Dutch language, including:

-The Subjective Units of Discomfort (SUD) scale was used to examine to what extent participants were feeling anxious at several moments. They had to indicate their perceived anxiety on a scale from 1 ("totally relaxed") to 10 ("extremely anxious") (Wolpe, 1973).

-The Visual Analogue Flight Anxiety Scale (VAFAS) was used to examine to what extent participants were anxious about flying. The one-tailed scale ranges from 0 ("no flight anxiety") to 10 ("terrified or extreme flight anxiety") (Van Gerwen, Spinhoven, Van Dyck, & Diekstra, 1999).

-The Igroup Presence Questionnaire (IPQ) was used to measure the feeling of being in the VE. Each of the 14 items is rated on a 7-point Likert-type scale, ranging from -3 ("totally disagree") to 3 ("totally agree"). The IPQ scale consists of three subscales: spatial presence, involvement and realness. The psychometric properties proved to be good to excellent (T. Schubert, Friedmann, & Regenbrecht, 2001). In the present study only the total score on the IPQ was used. The internal consistency in the present study was good, Cronbach's Alpha .95.

Procedure

All measurements took place at the VALK facility. Upon arrival participants were informed about the procedure. For the aviophobics it was emphasized that participating was voluntary and neither participation nor refusal to participate impacted on the quality of treatment. After informed consent was given, six electrodes were attached and connected to the Vrije Universiteit Ambulatory Monitoring System (VU AMS) which records the thorax impedance and the ECG in freely moving individuals (Goedhart, et al., 2006; Goedhart, et al., 2007; Houtveen, Groot, & de Geus, 2006; Riese, et al., 2003; Willemsen, et al., 1996).

Participants were then seated upright in a normal chair and partook in three experimental conditions, always in the same fixed order. Participants first received a 7-minute VR exposure in a neutral VE after which they were asked to fill out the IPQ. The neutral VE (Schuemie, 2003) consisted of a courtyard in which participants moved around under therapist control, i.e. locomotion is not controlled by the participants. The locomotion was standardized and automated. Participants completed two rounds along the outer perimeter of the courtyard (Figure 1, left). This condition was followed by a 7-minute VR flight simulation in a real airplane seat. Participants were seated upright and followed a standardized program consisting of taxi-out, take-off, a short cruise flight, descent, approach and landing. Subsequently, participants were given seven minutes of recovery time while seated in the airplane seat. Subjective units of distress (SUD) were measured at four discrete moments – before the start of the experiment, directly after both VR presentations and at the end of the recovery period. The VAFAS was administered before start of the experiment.

Results

Comparison of phobic and non-phobic control participants on sociodemographic characteristics and the VAFAS scale were performed with one-way ANOVA. Table 1 shows the main characteristics for the group of phobic participants and the control group.

An ANOVA was conducted on the SUD scores collected in three conditions (neutral VR world, virtual flight, and recovery) from the two participant groups (phobic, and control). Significant condition ($F(1.73,91.54) = 3.81$, $p = .031$) and group ($F(1,53) = 21.68$, $p < .001$) effects were found for self-reported distress. Phobics had higher levels throughout, while on average, participants reported less fear during recovery compared to the virtual flight. Follow-up analyses for both groups separately showed significant differences in reported anxiety between the recovery condition and both the neutral VR world ($t(1, 37) = 2.51$, $p = .017$) and the virtual flight ($t(1,33) = 3{,}09$, $p = .004$) for the flight phobics, while no significant differences between conditions were seen for the control group (Figure 2).

Table 1
Number of participants, gender, age, BMI and VAFAS score in study 1

	Phobics Mean (SD or %)	Non-phobics Mean (SD or %)
Number of participants		
Total	39	21
Men	15 (38%)	11 (52%)
Women	24 (62%)	10 (48%)
Age (years)	44.5 (12.4)	48.3 (11.4)
BMI	24.6 (3.8)	23.5 (2.4)
VAFAS	8.0 (1.4)*	0.6 (0.7)

* Phobics differ from non-phobics at $p < .001$

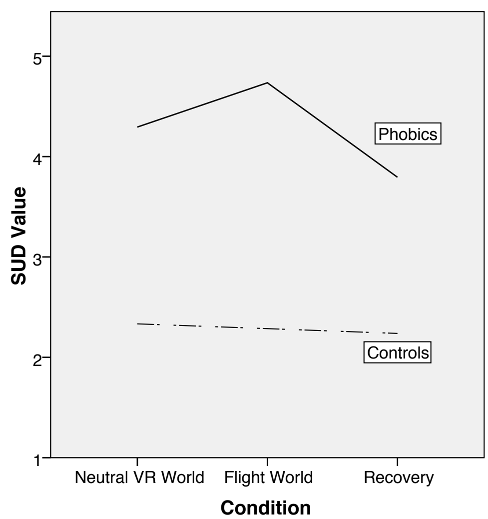


Figure 2. SUD scores for phobic and control participants.

An ANOVA with the same independent variable was also conducted on the physiological data. Of the physiological variables, RSA had to be log (ln) transformed to obtain normal distributions. For HR a significant main effect was found for condition ($F(1.78, 101.62) = 16.94$, $p < .001$). Both control participants and phobic participants had higher heart rates during the neutral VR world than in any other condition (Figure 3). In contrast to the main effect of condition for HR, no significant effects of condition were found in RSA and PEP data. In fact, there was no significant main or interaction effect in RSA at all. A significant main group effect did emerge in PEP data, phobic participants had significantly shorter PEP values than control subjects, indicating higher cardiac sympathetic control ($F(1,58) = 5.83$, $p = .019$).

A one-way ANOVA was conducted for the IPQ. Non-phobic subjects scored significantly higher on the total IPQ scale ($F(1,57)=10.42$, $p=0.002$) including its subscales Spatial Presence (SP: $F(1,57)=11.45$, $p < 0.05$), Involvement (INV: $F(1,57)=5.24$, $p < 0.05$), and Realism (Real: $F(1,57)=4.08$, $p < 0.05$). IPQ scores were relatively high compared with other studies[1]. No significant correlations between IPQ scores and SUD scores were found.

Almost all participants complained either during or directly after the neutral VE exposure about dizziness and nausea. This was corroborated by an elevated HR during this supposedly non-provocative neutral condition. No difference in anxiety between the neutral VR world and the flight world was reported by the flight phobics. All other measures did not differentiate between conditions. This led us to the conclusion that the neutral VR world probably was not truly neutral after all.

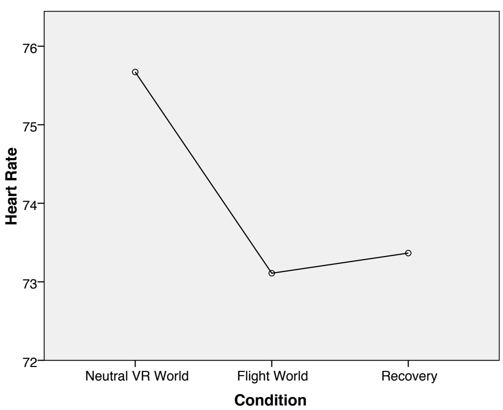


Figure 3. Average HR for phobic and non-phobic combined.

Study 2

Method

Novelty of a new environment or cybersickness might have caused the higher level of arousal in the neutral VR world in Study 1. This would suggest that arousal level could be reduced by removing the novelty and therapist controlled locomotion element from the neutral VR world. The aim of the second study was, therefore, to examine whether a new neutral world would still result in an elevated level of arousal. In addition, the study also set out to study the suggested novelty effect or possible transfer of habituation by changing the order in which physiological recording was collected (the actual room first, or virtual room first).

VR system and new neutral world

To make a possible transfer of habituation possible a virtual world was created which was a close replication of the actual room the individual was situated in, causing participants to see the same environment when they would put on or take off the HMD. Participants were seated in front of a television (Figure 4, right) showing a documentary about wildlife. In the new neutral VR world (Figure 4, left) participants were seated in front of the same television set showing the same documentary. Looking around with or without the HMD would give the same view of the room. The new neutral VR world ran on the same hardware as the VR flight and the old neutral VR world but used different software with the exception of the Windows XP operating system. The Vizard Virtual Reality Toolkit, Vizard 3.0 was used to create an executable that provided head tracking and the image for the HMD. The model of the room was created with Autodesk Maya 2008 and textures were edited with Adobe Photoshop CS2. The model consisted of the room in which a table, television set, room dividers and a metal rail were modeled in detail. There were 18 textures made with several different sizes ranging from 2048 x 2048 pixels (the wall closest to the patient) to 64 x 512 pixels (a table leg). The world was displayed with a resolution of 640 x 480 to match the resolution of the HMD used in VR flight. All textures were file textures. The textures were made out of photographs taken from the location where the patient would sit. Every visible face got its own unique texture. No dynamic lights or computer-generated shadows were used. Distortion was removed from the images and the color balance of several images was altered. Some objects were edited out of textures, like the table that was removed from the photograph that formed the texture for the wall behind the television set, which was facing the patient. Shadows belonging to

[1]For comparison data see www.igroup.org

objects that were removed from the scene were edited out like the radio and chair in front of the table. Some shadows had to be drawn in by hand like the shadow of the table on the part of the wall behind the table. A video could be displayed on the television set triggered by pressing a keyboard button. The video used was ripped from DVD and re-compressed with a resolution of 720 x 576 at 25 frames/second. The video format and codec used was VC-1, WMV3 (Windows) and the audio format and codec used was WMA2, 161 (Windows). The video was edited to a duration of 6 minutes and 29 seconds. A DVD with the exact same edited video was made so that the video could be played on the DVD player in the actual room.

Figure 4. Left, new neutral virtual world, right, picture of the actual room.

Participants

Forty-four subjects participated, comprised of 32 students who earned credits by participating and 12 non-paid volunteers recruited by means of word-to-mouth. All participants received an E-mail with information regarding the study before the start of the experiment. One participant was excluded because of the use of cardioactive medication. Another two participants were partly excluded from analyses because of equipment failure during physiological recordings; their questionnaire data was included for analysis. Average age of all 43 (16 men) participants was 25.6 (SD = 8.0), the youngest being 18 years old, the oldest 51.

Measures

In addition to the three physiological measurements and questionnaires used in the previous study the following two questionnaires were added:

-The Simulator Sickness Questionnaire (SSQ) was used to examine to what extent participants experienced symptoms associated with simulator sickness caused by the VR exposure. The SSQ consists of a checklist of 27 symptoms, each of which is rated in terms of degree of severity (none, slight, moderate, severe). It is normally administered twice, before and after a VR exposure (Kennedy, 1993). The instrument provides three subscales (Nausea, Oculomotor and Disorientation) and a composite Total Severity Score, which is used in the present study. The instrument's psychometric properties are good (Johnson, 2005). The internal consistency in the present study was good, Cronbach's Alpha .78.

-The Vividness of Visual Imagery Questionnaire (VVIQ) was used to examine the patients' ability to form mental pictures (Marks, 1973). The vividness of the image is rated along a 5-point Likert-type scale, from 1 ("perfectly clear and as vivid as normal vision") to 5 ("no image at all, you only 'know' that you are thinking of an object"). All items for images obtained are first answered with eyes open, secondly with eyes closed. Note that a low VVI score means vivid imagery and a high score means vague imagery. The psychometric properties proved to be good to excellent (Campos & Perez-Fabello, 2009). In the present study the average score on the VVIQ was used. The internal consistency in the present study was good, Cronbach's Alpha .95.

Procedure

Participants started with filling out the VAFAS and the SSQ (pre-exposure). After attachment of the electrodes of the VU AMS participants were seated upright in a normal seat. Participants randomly started either with the new neutral VR world, or the neutral real world. Participants were asked to complete the IPQ and SSQ-post-exposure directly after the neutral VR world. These two conditions were followed by a 7-minute VR flight simulation while seated in a real airplane chair. Participants were seated upright and followed a standardized program consisting of taxi-out, take-off, a short cruise flight, descent, approach and landing. Subsequently, participants were given seven minutes of recovery time while seated in the airplane seat. SUD scores were recorded at five discrete moments – before the start of the experiment, directly after both neutral worlds, after the virtual flight and at the end of the recovery period. Before the start of the experiment informed consent was obtained from all participants. The research protocol had been approved by the local medical ethics committee.

Results

As was done with Study 1, a series of ANOVAs was conducted to study the effect of two independent variables – condition (real world, new neutral VR world, virtual flight, recovery), and group (first real world then new neutral VR world, or first new neutral VR world and then real world).

Table 2
Number of participants, gender, age, BMI and VAFAS score in study 2

	Mean (SD or %)
Number of participants	
Total	43
Men	16 (37%)
Women	27 (63%)
Age (years)	25.6 (8.0)
BMI	22.6 (2.8)
VAFAS	0.8 (1.2)

A significant main effect was found for condition (F(2.46, 100.9) = 3.29, p =.032) in the SUD scores. Participants reported lower levels of anxiety during the real world than during any other condition. Follow-up analyses for both groups separately showed a significant difference in reported anxiety between the real world and both the new neutral VR world (t(1, 20) = -2.32, p =.031) and the virtual flight (t(1,20) = -2,35, p =.029) for the participants who saw the real world first (Figure 5). Interestingly, no significant differences between conditions were found when the new neutral VR world was presented first.

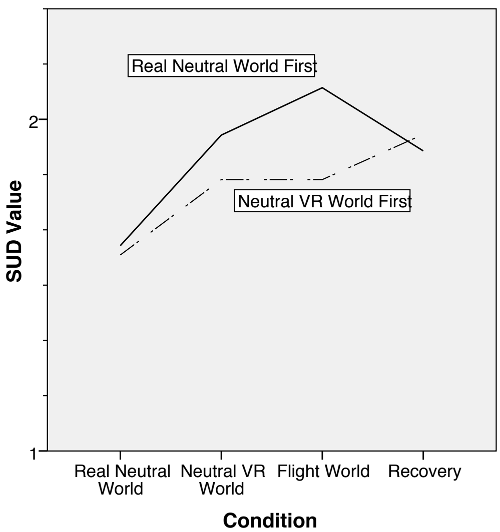


Figure 5. SUD scores for both groups.

The analysis of heart rate found a significant condition by group interaction (F(2.19, 85.32) = 5.48, p = .005), together with a main effect of condition (F(2.19, 85.32) = 10.12, p < .001). Follow-up tests revealed that HR during the virtual flight was significantly lower than HR in any other condition (all p <. 001), while the interaction with group was driven by an increase of HR during the real world condition for the participants who saw the real world first (Figure 6).

In contrast to the condition by group interaction for HR, no significant condition by group interactions were found for RSA and PEP. Significant main condition effects did emerge for overall RSA and PEP levels. Participants had significantly longer RSA values during virtual flight compared to all other conditions, indicating higher parasympathetic control during the virtual flight (F(2.72, 106.04) = 9.06, p < .001), and significant longer PEP values during virtual flight compared to the new neutral VR world and the recovery condition (F(3, 37) = 5.12, p =.003), indicating less cardiac sympathetic control during virtual flight.

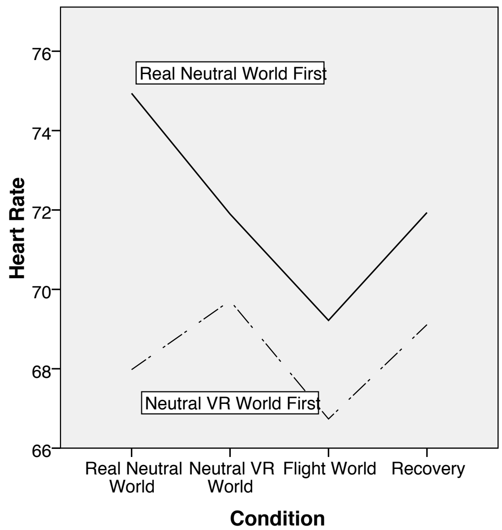


Figure 6. Average HR for both groups.

On average, participants had a significant decrease in SSQ from pre- to post-presentation measurement (t(1, 41)=2.65, p=.011). These changes from pre to post scores on the SSQ were significantly correlated to the SUD values from the Virtual Flight such that decreased simulator sickness was accompanied by a lower anxiety score during the flight condition (r = -.437, p = .003). SSQ-post scores were significantly correlated with SUD-Flight (r = .508, p = .001) and SUD-Recovery (r=.522, p<.001). Participants with lower post presentation simulator sickness scores reported lower anxiety during the virtual flight and the recovery condition, while participants with higher post presentation SSQ values reported more anxiety in both conditions.

No significant correlations between IPQ scores, SUD scores and VVIQ were found. A significant negative correlation was found between IPQ and SSQ-Post (r = -.325, p =.033) and a significant positive correlation was found between IPQ scores and SSQ pre-post (r=.391, p=.009). On average, participants with a higher presence score had a lower post-presentation simulator sickness score than participants with lower IPQ score, while participants with a higher presence score showed a stronger decrease in simulator sickness compared to participants with lower IPQ scores.

Cross Comparison

Cross comparison was performed on control participants from Study 1 with all participants from Study 2. Table 3 shows the main characteristics for both groups. ANOVAs were conducted with condition (Neutral VR world, Flight world, Recovery) and group (participants Study 1, participants Study 2) as independent variables. Age was added as a covariate.

Table 3
Number of participants, gender, age, BMI and VAFAS score in study 1 and 2

	Participants Exp.1 Mean (SD or %)	Participants Exp.2 Mean (SD or %)
Number of participants		
Total	21	43
Men	11 (52%)	16 (37%)
Women	10 (48%)	27 (63%)
Age (years)	48.3 (11.4)	25.6 (8.0)*
BMI	23.5 (2.4)	22.6 (2.8)
VAFAS	0.6 (0.7)	0.8 (1.2)

* Participants study 1 differ from Participants study 2 at p <.001

A significant group by condition interaction was found for self-reported distress (F(1.71, 104.1) = 6.76, p =.003), together with main effects of group (F(1,61) = 5.10, p = .028) and condition (F(1.71,104.1) = 9.19, p < .001). In Figure 7 it can be seen that the control participants in Study 1 had higher levels of distress throughout, while in contrast with the participants from Study 2, their reported distress was highest during the neutral VR condition.

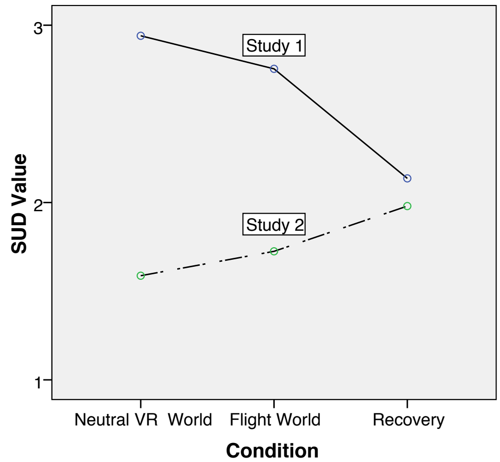


Figure 7. SUD scores for participants from both studies.

In parallel with self-reported distress, a significant group by condition interaction was found for HR (F(4.57,101.21) = 4.57, p = .017), with a main effect for condition, not for group (F(1.72,101.21) = 4.21, p = .023). Control participants in Study 1 had significantly higher HR during the neutral VR condition than during the virtual flight and recovery condition, while participants from Study 2 showed no significant difference in HR between the neutral and recovery condition.

Analogue with both subjective distress and HR, a significant group by condition interaction (F(2,118) = 4.52, p = .013) showed for the parasympathetic measure RSA, together with a main effect of condition (F(2,118) = 5.87, p = .004). Post-hoc analyses revealed that participants in the second study had nearly the same RSA values during the neutral VR world and the recovery condition, while RSA values during virtual flight were significantly longer than RSA values during the neutral VR world (t(1,40)= 4.55, p<.000) and recovery condition (t(1,40)= 2.50, p=.016). Differences between conditions did not reach significance for participants from the first study (Figure 8).

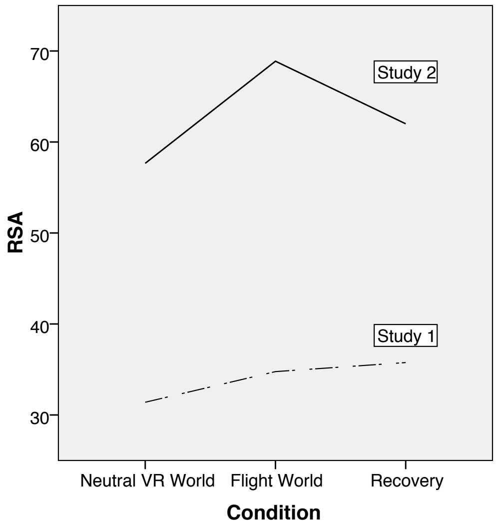


Figure 8. Average RSA for participants from both studies.

For the sympathetic measure PEP only a significant group by condition interaction was found ($F(1.57,92.68) = 3.57$, $p = .042$). Again, participants in the second study had no significant differences between the neutral VR world and the recovery period, with significantly longer PEP values during the virtual flight. Participants in Study 1 showed no significant difference in PEP between all three conditions.

A one-way ANOVA was conducted for the IPQ. No significant differences between both groups were found.

Conclusion and Discussion

In the second study all physiological measures differentiated between the flight condition on one side and the VR neutral world and recovery condition on the other side, while no physiological difference was apparent between the VR neutral world and recovery. Even self-reported distress showed no significant differences between conditions when the new VR neutral world was presented first. This seems to refute the idea that a virtual world by definition will generate arousal and anxiety (Wiederhold & Wiederhold, 2005). The second study found only an interaction effect between the condition and the groups in HR. Still, follow-up analyses only found a significant decrease, instead of an increase, in HR, when recordings were first collected in the actual room and then in the neutral VR room. This observation is therefore contradictory to the idea that novelty of VE would always cause arousal. Also, the follow-up analyses did not find a significant difference in HR of participants from the group in which the recording took place in the opposite order (first neutral VR world, second actual room). A significant decrease would have provided support for the hypothesis of transfer of habituation from one environment to another. The lack of interaction effects in the other physiological measures makes this hypothesis, again, less likely. Thus, this suggests that to obtain neutral physiological measurements the VE does not have to be a replication of the actual room the individual is situated in. Cross comparison of control participants of Study 1 and all participants from Study 2 strengthen our findings. Both subjective measures of anxiety, as well as all physiological measures of arousal, indicated equal anxiety and arousal in the neutral VR world and the recovery condition in the second study, while participants in Study 1 had elevated values for SUD and HR during the supposedly neutral VR condition when compared to the virtual flight and recovery condition. The study also found presence and cybersickness to be negatively related. Although only a certain level of presence is needed for treatment to be effective (Krijn, Emmelkamp, Biemond, et al., 2004), maximizing presence might reduce simulator sickness and thereby minimize drop out. No relationship was found between imaginative power, anxiety and presence. Our data, therefore, did not corroborate the idea that imaginative power influences presence (Huang, et al., 2000; Regenbrecht, et al., 1998). This might be caused by relatively higher IPQ scores compared to other studies. It is reported that imagination has an important effect on presence when VR is limited. However, if the VR is vivid enough, participants do not need to use their imagination to create a convincing virtual environment (Wallach, 2010). The average lower heart rate in the virtual flight condition compared to the neutral and recovery conditions could reflect different types of coping mechanisms. Our data are strongly reminiscent of similar data in dental phobics exposed to a stressful video showing surgical operations (Bosch, et al., 2001) as well as non-phobics exposed to neutral and flight-videos (Busscher, Van Gerwen, Spinhoven & De Geus, 2010). Exposure to phobic stimuli is a complex stressor in that it can invoke both fight-flight responses, characterized by increased sympathetic and reciprocal decreased parasympathetic activity, and a passive coping response (freeze), characterized by increased sympathetic activity paired with increased parasympathetic activity. A principal finding in Study 1 is that phobics were more anxious during the entire experiment than non-phobics, as expressed in significantly higher SUDs and sympathetic activation (PEP). This contributes to the validity of VR as a useful tool in exposure-based therapy.

References

Aslan, B. (2007). *A New VRET Framework Fear of Flying*. Unpublished Bachelor Thesis, Rotterdam University of Applied Sciences.

Berntson, G. G., Cacioppo, J. T., Binkley, P. F., Uchino, B. N., Quigley, K. S., & Fieldstone, A. (1994). Autonomic Cardiac Control .3. Psychological Stress and Cardiac Response in Autonomic Space as Revealed by Pharmacological Blockades. *Psychophysiology, 31*(6), 599-608.

Bosch, J. A., De Geus, E. J. C., Kelder, A., Veerman, E. C. I., Hoogstraten, J., Van Nieuw Amerongen, A. (2001). Differential effects of active versus passive coping on secretory immunity. *Psychophysiology 38*(5), 836-846.

Brinkman, W.-P., van der Mast, C., Sandino, G., Gunawan, L. T., & Emmelkamp, P. (2010). The therapist user interface of a virtual reality exposure therapy system in the treatment of fear of flying. *Interacting with Computers 22*(4), 299-310.

Busscher, B, Van Gerwen, L. J., Spinhoven, Ph., De Geus, E. J. C. (2010). Physiological reactivity to phobic stimuli in people with fear of flying. *Journal ofPsychosomatic Research 69*, 309-317.

Campos, A., & Perez-Fabello, M. J. (2009). Psychometric Quality of a Revised Version Vividness of Visual Imagery Questionnaire. *Perceptual and Motor Skills, 108*(3), 798-802.

Goedhart, A. D., Kupper, N., Willemsen, G., Boomsma, D. I., & de Geus, E. J. C. (2006). Temporal stability of ambulatory stroke volume and cardiac output measured by impedance cardiography. *Biological Psychology, 72*(1), 110-117.

Goedhart, A. D., Van der Sluis, S., Houtveen, J. H., Willemsen, G., & De Geus, E. J. C. (2007). Comparison of time and frequency domain measures of RSA in ambulatory recordings. *Psychophysiology, 44*(2), 203-215.

Gregg, L., & Tarrier, N. (2007). Virtual reality in mental health - A review of the literature. *Social Psychiatry and Psychiatric Epidemiology, 42*(5), 343-354.

Grossman, P., J van Beek, and C Wientjes. (1990). A comparison of three quantification methods for estimation of respiratory sinus arrhythmia. *Psychophysiology, 27*(6), 702-714.

Gunawan, L. T., Mast, C. v. d., Neerincx, M. A., Emmelkamp, P., & Krijn, M. (2004). Usability and Presemce in Virtual Reality Exposure Therapy for Fear of Flying. *Proceedings of the Euromedia 2004 Conference*, 125-132.

Houtveen, J. H., Groot, P. F. C., & de Geus, E. J. C. (2006). Validation of the thoracic impedance derived respiratory signal using multilevel analysis. *International Journal of Psychophysiology, 59*(2), 97-106.

Huang, M. P., Himle, J., & Alessi, N. E. (2000). Vivid visualization in the experience of phobia in virtual environments: Preliminary results. *Cyberpsychology & Behavior, 3*(3), 315-320.

IJsselsteijn, W. A., de Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: Concept, determinants and measurement. *Human Vision and Electronic Imaging V, 3959*, 520-529.

Jang, D. P., Kim, I. Y., Nam, S. W., Wiederhold, B. K., Wiederhold, M. D., & Kim, S. I. (2002). Analysis of physiological response to two virtual environments: Driving and flying simulation. *Cyberpsychology & Behavior, 5*(1), 11-18.

Johnson, D. M. (2005). *Introduction to and Review of Simulator Sickness Research:* U.S. Army Research Institute for the Behavioral and Social Sciences.

Kennedy, R. S., Lane, N. E. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology, 3*, 203-220.

Krijn, M., Emmelkamp, P. M. G., Biemond, R., de Ligny, C. D., Schuemie, M. J., & van der Mast, C. A. P. G. (2004). Treatment of acrophobia in virtual reality: The role of immersion and presence. *Behaviour Research and Therapy, 42*(2), 229-239.

Krijn, M., Emmelkamp, P. M. G., Olafsson, R. P., & Biemond, R. (2004). Virtual reality exposure therapy of anxiety disorders: A review. *Clinical Psychology Review, 24*(3), 259-281.

Ling, Y., Nefs, H. T., Brinkman, W.-P., Heynderickx, I., & Qu, C. (2010). The role of display technology and individual differences on presence. *Proceedings of ECCE2010*, 323-326.

Marks, D. F. (1973). Visual Imagery Differences in Recall of Pictures. *British Journal of Psychology, 64*(Feb), 17-24.

Min, B. C., Chung, S. C., Min, Y. K., & Sakamoto, K. (2004). Psychophysiological evaluation of simulator sickness evoked by a graphic simulator. *Applied Ergonomics, 35*(6), 549-556.

Paping, C., Brinkman, W.-P., & van der Mast, C. (2010). An explorative study into tele-delivered multi-patient virtual reality exposure therapy system. In K. Wiederhold (Ed.), *Coping with posttraumatic stress disorder in returning troops: Wounds of War II* (pp. 203-219): IOS press.

Parsons, T. D., & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 250-261.

Powers, M. B., & Emmelkamp, P. M. G. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders, 22*(3), 561-569.

Regenbrecht, H. T., Schubert, T. W., & Friedmann, F. (1998). Measuring the sense of presence and its relations to fear of heights in virtual environments. *International Journal of Human-Computer Interaction, 10*(3), 233-249.

Riese, H., Groot, P. F. C., Van Den Berg, M., Kupper, N. H. M., Magnee, E. H. B., Rohaan, E. J., et al. (2003). Large-scale ensemble averaging of ambulatory impedance cardiograms. *Behavior Research Methods Instruments & Computers, 35*(3), 467-477.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence-Teleoperators and Virtual Environments, 10*(3), 266-281.

Schubert, T. W. (2009). A New Conception of Spatial Presence: Once Again, with Feeling. *Communication Theory, 19*(2), 161-187.

Schuemie, M. J. (2003). *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy.* Delft, The Netherlands: Delft University of Technology.

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R., & Vandoornen, L. J. P. (1990). Methodological Guidelines for Impedance Cardiography. *Psychophysiology, 27*(1), 1-23.

Stanney, K. M., Mourant, R. R., & Kennedy, R. S. (1998). Human factors issues in virtual environments: A review of the literature. *Presence-Teleoperators and Virtual Environments, 7*(4), 327-351.

Van Gerwen, L. J., Spinhoven, P., Van Dyck, R., & Diekstra, R. F. W. (1999). Construction and psychometric characteristics of two self-report questionnaires for the assessment of fear of flying. *Psychological Assessment, 11*(2), 146-158.

Wallach, H. S., Safir, Marilyn P., Samana, Roy. (2010). Personality variables and presence. *Virtual Reality, 14*, 3-13.

Wiederhold, B. K., Davis, R., & Wiederhold, M. D. (1998). The effects of immersiveness on physiology. *Stud Health Technol Inform, 58*, 52-60.

Wiederhold, B. K., & Wiederhold, M. D. (2005). *Virtual reality therapy for anxiety disorders: Advances in evaluation and treatment.* Washington, DC: Amrican Psychological Association.

Willemsen, G. H. M., DeGeus, E. J. C., Klaver, C. H. A. M., VanDoornen, L. J. P., & Carroll, D. (1996). Ambulatory monitoring of the impedance cardiogram. *Psychophysiology, 33*(2), 184-193.

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The factor structure of the presence questionnaire. *Presence-Teleoperators and Virtual Environments, 14*(3), 298-312.

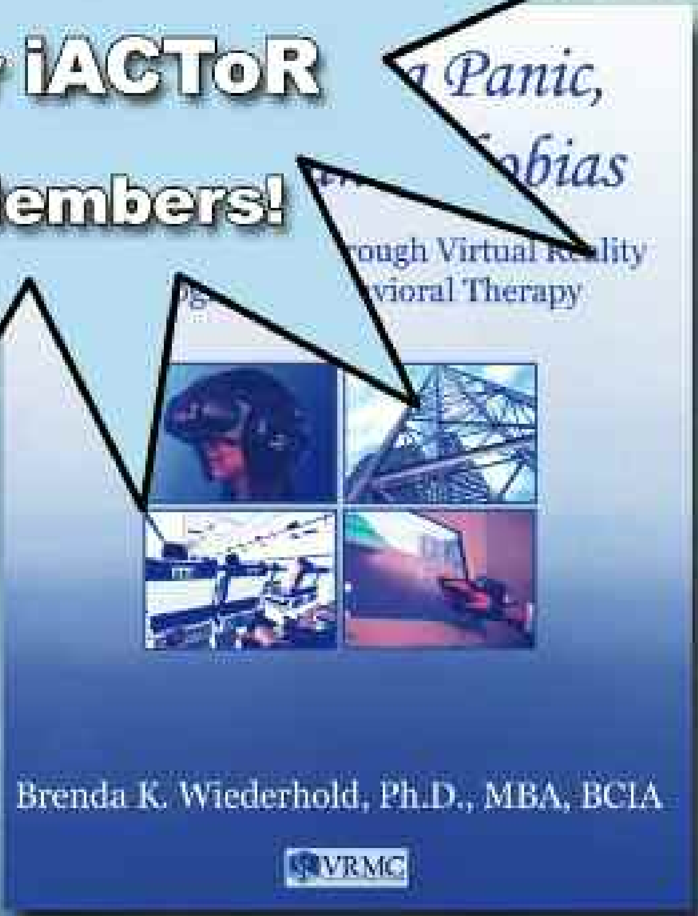
Wolpe, J. (1973). *The practice of behavior therapy*: New York: Pergamon Press.

20% OFF for iACToR Members!

Conquering Panic, Anxiety, & Phobias

Achieving Success Through Virtual Reality and Cognitive-Behavioral Therapy

By Dr. Brenda K. Wiederhold, PhD, MBA, BCIA



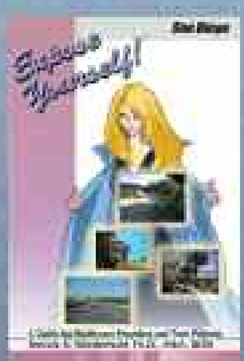
2003 - 132 PAGES - SOFTCOVER

$19.95

This book is written as a starting point toward helping the large portion of our population that suffers from anxiety disorders to overcome their fears and control their anxiety. It is a resource to enable those suffering from anxiety to take control of their lives and become an active participant in their own recovery.

This book is essentially divided into two parts: a discussion of anxiety and its physical and emotional effects on sufferers. While Virtual Reality Therapy is described, its use is not necessary in order to follow the suggestions in this book. The lessons and worksheets included can help in a variety of areas, not just anxiety, but anger, mild depression, and feelings of helplessness.

Also of Interest...



Expose Yourself! San Diego

A Guide for Healthcare Providers and their Patients
By Dr. Brenda K. Wiederhold, PhD, MBA, BCIA

$19.95

Virtual Reflections

By Dr. Brenda K. Wiederhold, PhD, MBA, BCIA

$12.95

For Ordering Options, Please Visit:

http://www.vrphobia.com/products.htm

or contact us at:

1-866-822-VRMC

frontoffice@vrphobia.com

Journal of CyberTherapy & Rehabilitation
Spring 2011, Volume 4, Issue 1

EEG ALPHA ASYMMETRY, HEART RATE VARIABILITY AND CORTISOL IN RESPONSE TO VIRTUAL REALITY INDUCED STRESS

Anne-Marie Brouwer[1], Mark A. Neerincx[1,2], Victor Kallen[1,3], Leslie van der Leer[1,4], and Michiel ten Brinke[1,4]

We propose to combine Virtual Reality (VR) and bio-neuro feedback to help treat stress-related disorders. As a first step in that direction, we here attempted to induce stress through VR and identify (neuro)physiological correlates. Nine participants performed a surveillance task in two different cities within VR while EEG, ECG and cortisol level were recorded over time. We aimed to induce stress by simulating a bomb explosion and providing negative feedback about the participant's performance. Associative stress was elicited by having participants return to the city where the bomb explosion occurred and they supposedly performed badly. (Associative) stress was reflected in EEG mid-frontal alpha asymmetry, heart variability and cortisol level. General stress levels as expressed by cortisol and mid-frontal alpha asymmetry correlated between participants. These results are promising for a successful implementation of a VR bio-neuro feedback system.

Keywords: Virtual Reality, Stress, PTSD, Physiological, EEG

Introduction

In diverse domains – such as defense, police force and first aid – there is a risk that professionals encounter traumatic events during their work. For some of these professionals, it is difficult to imagine, reminisce and discuss such events. For caregivers, it may be difficult to address the many differences in how professionals cope with traumatic events, show specific symptoms and handle confrontations. Desensitization through (gradual) exposure is a generally accepted component within diverse psychotherapy formats aimed at phobia and Posttraumatic Stress Disorder (PTSD: Shearer, 2007) and may also be useful to train professionals who are likely to encounter traumatic events. Virtual Reality (VR) could be a particularly suitable tool in exposure. It provides a flexible, controlled setting to systematically evoke personal experiences which may arguably be stronger than memorizing or imagining the experiences. Simultaneously obtaining information about an individual's (neuro)-physiological correlates of stress could be helpful in several different ways. Firstly, it can be used (online) to tailor the appropriate level of VR exposure. Secondly, it can be fed back to the individual online such that he or she can attempt to get the values within normal range, therewith hopefully speeding up recovery or boosting stress resilience (Repetto et al., 2009). Finally, (neuro)physiological information can provide objective data on patient's responses to stressors, which may be used in clinical decision making and in systematically monitoring treatment or training results over time.

The aim of this study is to make the first steps toward a VR neuro-bio paradigm to treat or train individuals who were or are likely to be exposed to stressful or traumatic events. We have two main questions. Firstly, can we elicit stress– direct and through memorized negative association – through a VR simulation of patrolling a city? Secondly, which of several (neuro)physiological variables correlate

Corresponding Author:
Anne-Marie Brouwer, TNO Human Factors, P.O. box 23, 3769 ZG Soesterberg, The Netherlands, E-mail: anne-marie.brouwer@tno.nl

[1]TNO Human Factors, P.O. box 23, 3769 ZG Soesterberg, The Netherlands
[2]Delft University of Technology, P.O. box 5, 2600 AA Delft, The Netherlands
[3]Leiden University, Institute of Psychology, Unit Developmental and Educational Psychology, P.O. Box 9500, 2300 RA Leiden, The Netherlands
[4]University College Utrecht, P.O. Box 80145, 3508 TC Utrecht, The Netherlands

with direct and associative stress under these circumstances? We investigate EEG (electroencephalography-frontal alpha asymmetry), ECG (electrocardiography-heart rate and its variability) and cortisol level in saliva. Identifying (neuro)physiological correlates of stress is the first step in using them online for neuro-bio feedback.

Virtual Reality Exposure

VR is increasingly used as an alternative to standard in vivo exposure to deliver graded exposure therapies (Coelho et al., 2009; Cukor et al., 2009; Emmelkamp et al., 2001). VR integrates real-time computer graphics, body tracking devices, visual displays and other sensory input devices to immerse patients in a computer-generated virtual environment. Graded VR exposure therapy (VRET) has been used clinically for treating a variety of anxiety disorders, including combat-related PTSD (Difede & Hoffman, 2002; Garcia-Palacios et al., 2006; Krijn et al., 2004), but the evidence base for the effectiveness of VRET in combat-related PTSD is currently based mainly on case studies (Ready et al., 2006; Rizzo et al., 2009; Wood et al., 2007). VR editors and libraries can accommodate the generation of personal worlds, so that the therapy can be tailored to the individual patient's experiences and needs. The VR behavior can be adapted by the therapist and/or the patient. Patient and therapist can set agreements on the specific adaptation mechanisms before the therapeutic sessions start. The integration of the VR usage into the overall therapeutic sessions and protocols, and therapists' trust in the system, are essential for acceptance and adequate utilization of these systems. Design of good user interface support can substantially improve therapist's performance and as such, therapy effectiveness (Gunawan et al., 2004).

(Neuro)physiological Correlates of Stress

Figure 1 schematically represents the relation between different stress relevant neurophysiological structures. Stress causes an increase in heart rate through the activated sympathetic system, a decrease in heart variability through the suppression of the parasympathetic system, and an increased release of cortisol through the Hypothalamus – Pituitary gland – Adrenal gland (HPA) – axis (Kallen et al., submitted; Kirschbaum et al., 1993; De Kloet, 2003; Benarroch, 1993; Brownley et al., 2000). Especially, Heart Rate Variability has been identified as a strong parameter of sympathetic-parasympathetic balance and as such, as a prominent marker of effective (physiological) stress management (Friedman & Thayer, 1998; Friedman, 2007). Additionally, within the neurophysiological dynamics shown in Figure 1, lowered heart rate variability, either before or during a challenge, has been identified as a potential predictor for the cortisol response following a stressor (Kallen et al., 2010). This association is particularly strong in anxious individuals (Kallen et al., submitted). The timing of the (expected) physiological responses is typically different. The involvement of multiple biochemical mechanisms within the HPA-axis cause a delay – only approximately 15-20 minutes after exposure to a perceived stressor can significant increases in cortisol concentrations be found, whereas the Autonomic Nervous System (controlling heart rate and its variability) responds immediately.

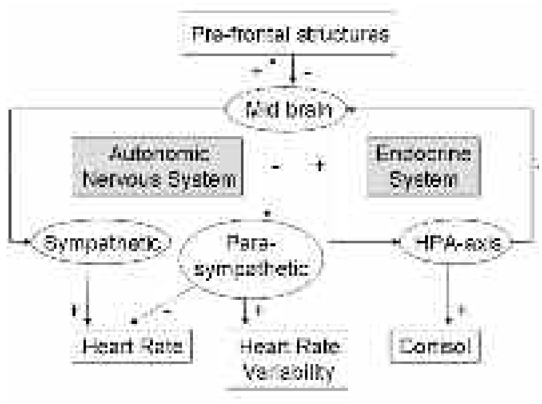


Figure 1. Stress and the interactions between neurophysiological systems.

Several studies report either higher or lower right (relative to left) frontal brain activity to be correlated with stress, in EEG (Lewis et al., 2007; Verona et al., 2009; Crost, Pauls & Wacker, 2008), fMRI (Masten et al., 2009; Wang et al., 2007) and NIRS (Tanida et al., 2007). The different directions of asymmetry may depend on the individual's coping strategy (Harmon-Jones et al., 2010) or personality (Crost, Pauls & Wacker, 2008). In EEG, frontal asymmetry of brain activity can be formalized by taking the (absolute) difference of alpha power (8-13 Hz) as measured by electrodes left and right. The exact underlying relationship between stress and frontal asymmetry is still unclear (Lewis et al., 2007).

Neuro-bio Feedback

With neurofeedback, individuals receive direct (e.g. auditory) feedback about their own brain signals with the goal of learning to modulate them. The idea is that neurofeedback helps to change certain characteristics of the brain signals, as identified by research on neural correlates of behavior or disorders, which again leads to changes in this behavior or to the recovery of the disorder. Individuals are indeed able to modulate certain types of brain signals with help of feedback (Birbaumer et al., 1992; Birbaumer, 2006; Lubar & Shouse, 1976; Weiskopf et al., 2003; Heinrich et al., 2007) and it can help to reduce certain disorders or increase (cognitive) performance (Tan et al., 2009; Keizer et al., 2010ab; Vernon et al., 2003; Egner & Gruzelier, 2001, 2004). Although more empirical evidence is needed, it has been argued that neurofeedback reduces symptoms of PTSD (Peniston & Kulkolsky, 1991; Peniston et al. 1993; Hammond, 2005; 2006). Instead of brain signals, other physiological signals can be used for feedback as well. This is called biofeedback. Frequently used variables are changes in skin conductance and heart rate or heart rate variability (Cukor et al., 2009; Repetto et al., 2009). Ultimately, we are interested in the applicability of both types, possibly combined, and use the term neuro-bio feedback.

Usually, neuro-bio feedback refers to feedback that is given to the measured individual him or herself. However, with neuro-bio feedback we also keep the possibility in mind that the feedback, or online information, is given to the therapist who could then decide whether and how to inform the individual. In the context of VR-mediated exposure therapy, physiological information could directly be fed into the executive VR-software, which in turn optimizes the presented VR scenario to fit the patient‘s needs (Popovic et al., 2006).

Neuro-bio feedback fits very well into a growing interest of clinical psychologists, psychiatrists and medics in self-management or self-regulation. Self-management has the potential to improve therapeutic outcomes and reduce the workload on therapists.

Study Overview

Our ultimate aim is to arrive at an integrated neuro-bio VR system to prevent and treat stress-related disorders. As a first step, we here try to elicit stress, direct and associative, through (elementary) VR and find corresponding (neuro)physiological correlates. We instructed healthy participants to pay attention during first-person virtual squad car patrols through two cities, and report situations they deemed in need of the civil force's attention. Of the two city scenarios, both presented alternately two times, one city (the "stress city") was accompanied by stress-inducing manipulations during the first presentation. These consisted of negative verbal feedback about the participant's performance and premature discontinuation of the scenario, marked by a sudden bomb explosion in the virtual world, due to "aberrant performance." The other city (the "non-stress city") was completed both times without trouble. With this design we aimed to manipulate direct stress (being highest during the first visit to the stress city – the only condition where negative events were actually presented) and associative stress (being higher during the second visit to the stress city than during the second visit to the non-stress city). We investigated whether direct stress and associative stress were reflected in EEG (frontal alpha asymmetry), ECG (heart rate and heart rate variability) and cortisol levels obtained using saliva samples.

Methods

Participants

Nine participants (six men, three women; age range: 21-37) were recruited from the TNO research participant database. All participants were right-handed, had normal or corrected-to-normal vision and hearing, did not have traumatic experiences with bomb explosions or related events, did not have a diagnosis of a DSM-IV axis I disorder(s), and did not have endocrine abnormalities (Kirschbaum et al., 1999; Lewis et al., 2007). Female participants did not use the oral contraceptive pill (Kirschbaum et al., 1999). Participants received a monetary reward to make up for their travel and time.

Experimental Setup

The experiment was conducted in a sound-attenuated, electrically shielded Faraday chamber. During the experiment, participants sat in a comfortable armchair approximately 70 cm in front of a 19" Dell 1907ftp LCD monitor, with a desk light illuminating the room from behind the participants. For the experimental task, participants used a number keypad.

Vibrations of the TNO tactile vest were employed to enhance the experience of a bomb explosion (see Stimuli). The vest was worn over their clothes and lined with five rows of 12 equally-spaced tactors spanning the whole trunk circumference. The tactors were custom built. They consisted of plastic cases with a contact area of 1x2 cm, containing 160 Hz vibrating motors (TNO, The Netherlands, model JHJ-3).

Headphones were used to provide sound during scenarios, including the explosion, and for feedback from the experimenter. A webcam enabled the experimenters to see the participants. The experimenters used a microphone to provide (negative) feedback during the first visit to the stress city in the second scenario.

Ag-AgCl electrodes of the EEG equipment were placed according to the standard international 10-20 system, using the mid frontal (F3/F4), lateral frontal (F7/F8), and mid parietal (P3/P4) channels referenced to linked mastoids (M1/M2; Crost et al., 2008; Verona et al., 2009; Lewis et al., 2007). For the present analysis, only data from the frontal electrodes were used. All electrode impedances were below 5 kΩ. EEG data was amplified using g.USBamp (g.tec, AT), with a bandpass filter of .1-100Hz. Analog signals were digitized at 256 Hz using a 24-bit A/D converter.

The reference Ag-AgCl electrode of the ECG was placed on the manubrium of the sternum; the ECG channel electrode was placed in the left, fifth intercostal space (i.e., between ribs 5 and 6); the ECG ground electrode was placed 5-8 cm below the ECG channel electrode.

Stimuli

The visual and auditory scenarios were created using Virtual Battle Space 2 (VBS2; Bohemia Interactive Australia) and recorded with Fraps (Beepa). Three preprogrammed VBS2 cities were chosen for the scenarios. Figure 2 shows screenshots of the five scenarios. The baseline scenario, during which no task had to be performed, was a tour through a European looking city. One of the task-cities was a fictional city on an island with a Mediterranean appearance (Mediterranean city). The other city represented Bacau, East Timor and had a tropical background (tropical city). We counterbalanced between participants which city was associated with stress. In Figure 2, the Mediterranean city is the stress city and the tropical city is the non-stress city. In both cities, ambiguous events were created with animated virtual people, animals, vehicles and objects so that participants could perform their task of reporting suspicious situations. Examples of these events were two women in burqas who walked past the police car or a running woman who was followed by three men.

For each participant, the first visit to the stress city ended with a bomb explosion. As the assignment of which city was the stress city was counterbalanced across participants, scenarios with and without a bomb explosion were programmed for both task cities. The scenarios including a bomb explosion that lasted four and a half minutes, whereas the full scenarios lasted eight minutes total. The

Figure 2. Example screenshots of the scenarios. Here, the Mediterranean city functions as the stress city and the tropical city is the non-stress city.

explosion consisted of a birds-eye view of the police officer walking or driving past an object that exploded on the screen, accompanied by a loud explosion and screaming sounds presented at maximally 100 dB. To enhance the effect, all 60 tactors of the tactile vest were activated.

Task and Procedure

Table 1 provides a chronological overview of the experimental events.

As a cover story, participants were told that the rationale behind the experiment was to evaluate a recruit assessment tool for use in military and civil forces. It was conveyed that earlier research had investigated behavioral and cognitive aspects of the program, but that this time, research focused on physiological effects. The earlier studies had supposedly employed a tactile vest to improve the experience of the virtual environment. Participants were told that the vest was not used for this purpose in this study, but that they did have to wear it because it could influence body temperature, one of this study's measures.

Table 1
Overview of the experimental procedure

Informed consent Salivary sample 1 Instruction Placing of electrodes
Baseline Block - no task (8 minutes) Eyes closed Salivary sample 2 Ratings subjective state
Experimental block 1 - first visit stress city (4.5 minutes, with negative feedback after 1.5 and 3 minutes; 'premature' end with bomb explosion) Eyes closed Negative feedback by experimenter in person Ratings subjective state Salivary samples 3, 4 and 5 at 5, 10 and 15 minutes after the bomb explosion.
Experimental block 2 - first visit non-stress city (8 minutes) Eyes closed Positive feedback by experimenter in person Salivary sample 6 Ratings subjective state
Experimental block 3 - second visit stress city (8 minutes) Eyes closed Positive feedback by experimenter in person Salivary sample 7 Ratings subjective state
Experimental block 4 - second visit non-stress city (8 minutes) Eyes closed Positive feedback by experimenter in person Salivary sample 8 Ratings subjective state
BIS-BAS scale STAI Debriefing

Participants were asked to sit as still as possible during the recordings to prevent artifacts contaminating the data. Participants were told that they would see VR simulations of tours through cities from a first-person perspective of a patroller driving a police car. They did not have to steer as the simulation was already recorded. In the first (baseline) block, they did not need to do anything but watch. In the subsequent four experimental blocks participants were instructed that they had to report objects or events which were in need of the civil force's attention to the police station by using the buttons 1, 2, and 3 of a number keypad. The 1 had to be clicked for events on the left side, the 2 for events in the middle, and the 3 for events on the right side of the screen. Double clicking indicated urgent events, while single clicks had to be used to report less urgent events. Participants were told that in the experimental blocks, two city scenarios had to be performed twice.

The scenarios were said to consist of several segments, which would follow each other as long as the performance was above a certain threshold. If the performance dropped below this threshold level, the scenario would end prematurely. Participants were told that they had to stay above the performance threshold, because the data would not be comparable to that of earlier research if they did not. In addition, they were reassuringly told it was unlikely that their performance would drop below threshold level. The experimenters mentioned that global feedback would be provided during the first experimental block to give participants an indication of their performance. In reality, clicking behavior was not recorded and there was no performance; negative and positive feedback was predetermined and given regardless of the participant's responses.

Participants were told that the first city scenario (presented in the first and third experimental block) had been rated as difficult by participants in earlier research, while the second city scenario (presented in the second and the fourth experimental block) had been rated as easy. During the first experimental block participants were told through the intercom that they performed poorly. There was no feedback during the other experimental blocks. The first experimental block ended prematurely, allegedly due to poor performance. After each block, participants closed their eyes for about two minutes until one of the experimenters entered the room. The experimenter provided negative feedback after the eyes closed interval of the first experimental block and positive feedback after the closed eyes intervals of the other blocks. After each closed eyes interval, participants provided a salivary sample. Then they rated their subjective state (see Questionnaires).

After the experiment, participants filled out two personality questionnaires (the BIS-BAS and the STAI – see Questionnaires). Then participants were told the true purpose of the experiment. They were told that their task performance was neither good nor bad because there was no actual task. Participants' wellbeing after the experiment was secured by the experimenters, and participants were given the opportunity to ask any remaining questions.

We chose to always present the second visit to the stress city before the second visit to the non-stress city because in this way, participants never visited a city twice in a row.

Questionnaires

On the subjective state questionnaire, participants indicated to what extent they felt nauseous, physical discomfort, mental pressure, focused and content with their performance by placing crosses along lines going from "not at all" to "a lot" (ratings subjective state).

Participants also filled out two personality questionnaires. We used the Dutch version of the trait subscale of the state-trait anxiety inventory (STAI; Lewis et al., 2007). The 20 statements (e.g., "I feel at ease") were rated on a 1 ("almost never") to 4 ("almost always") Likert-scale. The STAI-score calculated from these responses indicates general, chronic anxiety, where higher scores (>40) indicate high trait-anxiety and lower scores (<33) indicate low trait-anxiety. Finally, participants filled out the Dutch version of the behavioral inhibition system/behavioral activation system (BIS-BAS) scale (Carver & White, 1994; Verona et al., 2009). This scale consists of 20 items, each rated on a 1 ("strongly agree") to 4 ("strongly disagree") Likert-scale. The responses lead to five scores: one BIS-score, three subscale-BAS-scores and one total BAS-score. The BIS-score (based on seven items; e.g., "I feel pretty worried or upset when I think or know somebody is angry at me") indicates how strongly participants experience negative affect or behavioral inhibition in the face of aversive stimuli or threats. The total BAS-score (13 items from BAS-subscales) indicates how strongly participants generally experience positive affect or behavioral approach in the face of rewards or goals (Sutton & Davidson, 1997). The higher the scores on the scales, the more sensitive the participants were to the respective underlying concept.

Variables

Stress

Our independent variable was stress. We aimed to manipulate both direct and associative stress.

We attempted to induce direct stress in the first experimental block through several means. We told participants that they would enter a relatively difficult scenario, but that it was important to achieve a certain level of performance because otherwise the data would be useless. Subsequently, they were told through the intercom twice that they were performing badly. A bad performance was underlined by the premature end of the scenario. The sudden and intense bomb explosion was also meant to induce a stress experience. Finally, the experimenter came into the room to tell the participant that performance had been poor. Since there was no bomb explosion, premature ending of the scenario and only positive feedback about performance during the other experimental blocks, we assume that the level of direct stress in those blocks was low compared to the first.

We hoped to induce associative stress by presenting the scenario in which the participant allegedly performed poorly (the stress city) a second time. We assume that associative stress would be higher during the second visit of the stress city compared to the second visit to the non-stress city.

EEG

EEG was used to measure frontal asymmetry, the main candidate for a neural correlate of stress.

For the analyses of the EEG data artifacts were not rejected. This was intentionally not done, as the ultimate goal of the present study was to investigate the combined use of VR therapy and neurofeedback, a situation in which artifacts would arise as well. We investigated intervals with eyes open (i.e. during the scenario) and intervals with eyes closed (immediately after the scenario). Eyes-closed data probably included less eye movement and blink artifacts compared to eyes-open data.

Total alpha power (8-13 Hz, μV2) was calculated for each participant, block, electrode and eyes open/closed condition from the power spectra derived with a fast Fourier transform (FFT). Alpha power was natural-log-transformed to normalize the data (Allen et al., 2004). For the eyes open, we calculated alpha power over an interval of three minutes, starting after the first minute of each scenario because the EEG signals were still stabilizing during this period. For the eyes closed, we used an interval of one minute, starting after the first 30 seconds because participants sometimes forgot to close their eyes immediately.

Asymmetry scores were calculated by subtracting the alpha power recorded at the left scalp location from the alpha power at the corresponding right scalp location for each participant, block, electrode and eyes open/closed condition. The direction of the asymmetry has been reported to depend on withdrawal (avoidance) or approach strategies to cope with stress, with more left alpha power for withdrawal and more right alpha power for approach (Verona et al., 2009). Since coping strategy likely varies between participants, we used the absolute value of the asymmetry scores for further analysis.

ECG

Intervals between successive heart beats and variability were recorded using ECG as another possible correlate of stress.

ECG was analyzed over the same time intervals as EEG. The median RRI (the interval between subsequent R waves in the ECG: the interval between subsequent heartbeats) was determined for each participant, block, electrode and eyes open condition. Heart variability was computed as the root mean squared successive difference (RMSSD: Goedhart et al., 2007) between the RRIs[1].

Cortisol

Cortisol level was the third potential physiological measure of stress.

At the start of the experiment and after each scenario (see Table 1), salivettes were used to obtain the salivary samples, which were analyzed for cortisol levels. In total, eight salivary samples were obtained to cover the expected cortisol response curve, with cortisol peak levels generally occurring 10-30 minutes after cessation of the stressor (Kirschbaum et al., 1993). Intervals of 15 minutes between stress and non-stress salivary samples have been used in previous studies (e.g., Montgomery et al., 2006). For this study, we used the sample taken right after the baseline block as a baseline (sample 2 in Table 1). Peak cortisol level was expected at salivary sample 5 (15 minutes after the bomb explosion) or at salivary sample 6 (about 25 minutes after). Since cortisol response curves strongly differ between individuals, we also defined the baseline as the lowest cortisol concentration before the first experimental block, and the peak as the highest cortisol concentration afterwards, separately for each individual participant.

Statistical Analyses

Measures that reflect direct stress are expected to show a significantly high stress level during the first experimental block (first visit stress city) as compared to the other three experimental blocks. Measures that reflect associative stress are expected to show a significantly higher stress level during the third experimental block (second visit stress city) as compared to the fourth (second visit non-stress city). To test these expectations, we performed a repeated measures ANOVA once for the eyes open and once for the eyes closed for each physiological variable that was recorded over the whole experiment (frontal alpha asymmetry, RRI and RMSSD) with experimental block as independent variable. For frontal alpha asymmetry this was done for each of the two measurement locations: mid frontal (F3/F4) and lateral frontal (F7/F8). Significant effects were further investigated using Fisher LSD post-hoc tests. In addition, we performed paired t-tests to specifically compare the third and the fourth block. This was also done for each variable that was recorded over the whole experiment, separately for eyes open and eyes closed and separately for the two measurement locations for EEG. We used paired t-tests to compare the baseline to the peak cortisol levels.

Results

Questionnaires

The subjective state questionnaires indicated that subjective stress did not vary across blocks. The only notable variation seemed to be that participants evaluated their performance as worse in the first experimental block (when they received negative feedback about their performance) than in the other blocks.

The STAI-scores ranged from 25 to 47 (median=32, SD=7.0), with five low-anxious participants (i.e., STAI-scores <33) and two high-anxious participants (i.e., STAI-scores >40). BIS-scores ranged from 12 to 17 (median=14, SD=1.6), which is quite low considering the possible range of 7 to 28. The total BAS-scores were also somewhat low (range=18-33, median=23, SD=5.2). In sum, the participants tended to be rather low-anxious and not very sensitive for negative stimuli or rewards.

[1]For the eyes open, we also computed heart variability as the power in the high frequency range (0.15-0.5 Hz) of the RRI over time using Welch's method. This measure of variability correlated very well with RMSSD as indicated by significant linear correlations for each block (all p-values <0.01, R2s ranging from 0.89 to 0.97 with a mean of 0.93). We thus decided to only use RMSSD, because in order to compute variability as the power in the high frequency range, at least 100 seconds of data are needed. This would mean that for eyes closed, we would have to use another time window as the one used for EEG.

EEG

Figure 3 shows the asymmetry for the mid frontal (A) and lateral frontal (B) locations averaged over participants and blocks for the eyes closed intervals. The repeated measures ANOVAs indicated a close to significant effect of experimental block for the mid frontal location ($F(3,24)=2.74$, $p=0.07$) and no effect lateral-frontally ($F(3,24)=1.60$, $p=0.22$). The paired t-test showed that asymmetry was higher for the second visit of the stress city than for the second visit of the non-stress city ($t_8=2.62$, $p=0.03$) midfrontally. The same pattern was close to significance for the lateral frontal location ($t_8=2.19$, $p=0.06$).

For the eyes open intervals, the ANOVAs and paired t-tests did not show any effect of experimental block on asymmetry scores for either the mid frontal location or the lateral frontal location (all p-values>0.30).

ECG

Figure 4 shows the median RRI (A) and the RMSSD (B) averaged over participants and blocks for the eyes open intervals. RRI (i.e., the inverse of heart rate) was significantly affected by experimental block ($F(3,24)=16.89$, $p<0.01$). Post-hoc tests indicated that the median RRI was higher (i.e., heart rate lower) in blocks with the second visits compared to the blocks with the first visits (all p-

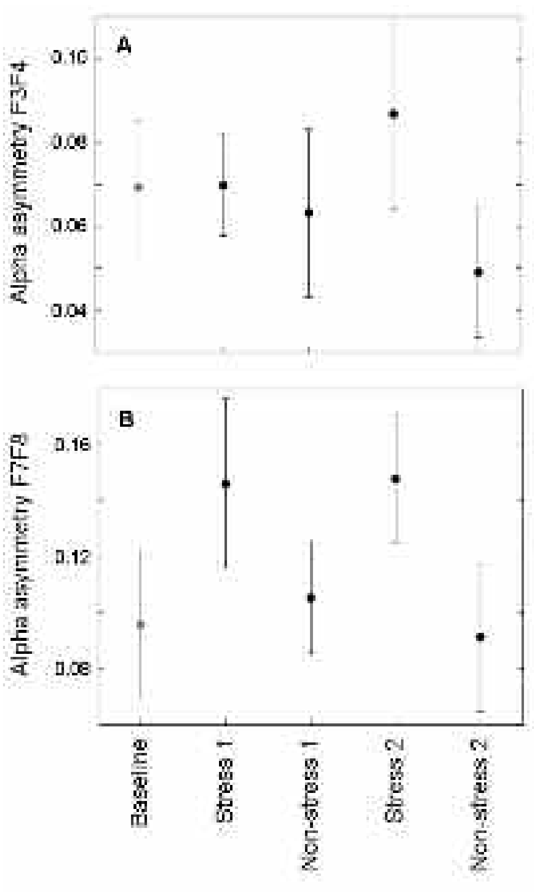


Figure 3. Alpha asymmetry for the mid frontal (A) and lateral frontal (B) locations averaged over participants and blocks for the eyes closed intervals. Error bars represent standard errors of the mean.

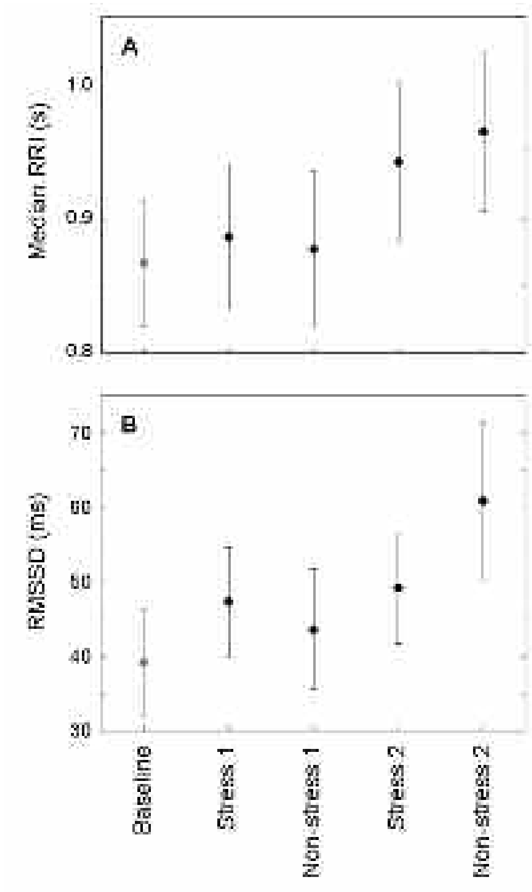


Figure 4. Median RRI (inverse of heart rate, A) and the RMSSD (heart variability, B) averaged over participants and blocks for the eyes open intervals. Error bars represent standard errors of the mean.

values <0.01). Consistent with this, the paired t-test did not indicate a difference between the third and the fourth experimental block (t8=1.91, p=0.09). RMSSD (ie., heart rate variability) was significantly affected by experimental block as well (F(3,24)=6.46, p<0.01). Post-hoc tests indicated that RMSSD was higher during the second visit to the non-stress city than in all other experimental blocks (all p-values <0.01). In line with this, the paired t-test showed that RMSSD was higher in the fourth than in the third block (t8=2.88, p=0.02).

Cortisol

Average cortisol values for each sample are presented in Figure 5. The paired t-test showed a significant difference between the lowest cortisol levels before and the highest cortisol levels after the stressor (t8=3.92, p<0.01). The average percentage increase was 59% (standard deviation of 46). Cortisol level was not significantly different from baseline as defined by sample 2 15 minutes after the bomb explosion (t8=0.93, p=0.38) but it was altered after 25 minutes (t8=2.32, p<0.05, with an average increase of 30% with a standard deviation of 29).

Correlations (neuro) physiological Stress Correlates

The results as presented above show the hypothesized effect of stress manipulation on EEG alpha asymmetry for the mid frontal locations (eyes closed), heart rate variability (eyes open) and cortisol level. As a post-hoc analysis, we checked whether these measures correlate to each other. Since all of these measures reflect stress at different time points, we transformed them into one value for each participant that, as we assume, reflects the general stress level over the experiment. For EEG we computed for each participant the average F3F4 asymmetry across blocks (eyes closed). For heart rate variability we computed for each participant the average RMSSD across blocks (eyes open). For cortisol, we computed for each participant the area under the curve (AUC) of cortisol level over time (Preussner et al., 2003). There was a strong positive linear correlation between cortisol AUC and EEG asymmetry (R2=0.84, p<0.01, Figure 6). The other measures did not correlate (cortisol AUC and heart rate variability: R2=-0.01, p=0.97; EEG asymmetry and heart rate variability: R2=-0.09, p=0.81).

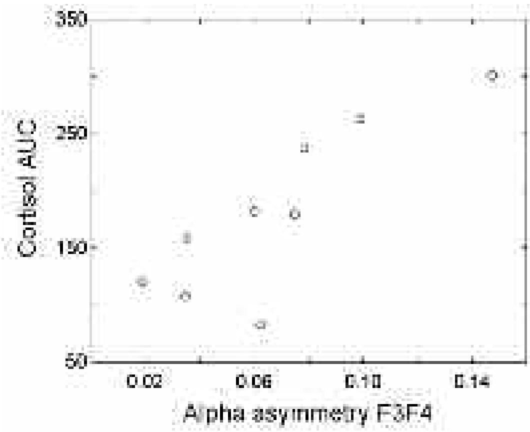


Figure 6. Cortisol AUC (Area Under the Curve) and mid frontal alpha asymmetry averaged over blocks: each data point represents one participant.

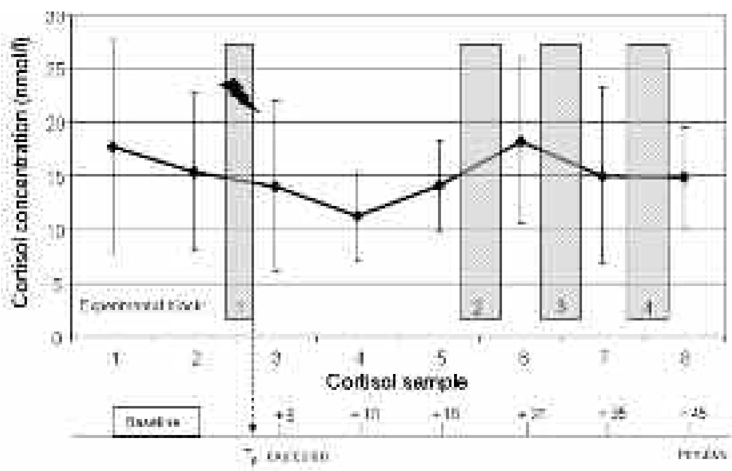


Figure 5. Average cortisol level for each sample.

Discussion

Most of our physiological variables responded (under some conditions) consistent to the hypothesized level of direct and associative stress. Consistent with associative stress, alpha asymmetry levels mid frontally (closed eyes) indicated more stress for the second visit to the stress city compared to the second visit to the non-stress city. At the lateral frontal location this was close to significant. Heart variability (eyes open) showed a pattern that was consistent with associative stress as well: heart variability was lower for the second visit to the stress city compared to the second visit to the non-stress city. Cortisol levels correlated with direct stress, peaking after the bomb explosion. The only physiological variable that was affected by experimental block number in a manner inconsistent with both direct and associative stress was RRI (eyes open). Heart rate was lower for both second visits compared to the first visits.

It is encouraging that we found physiological correlates of stress, even though we only recorded from a small sample of participants that included men and women in a rather large age range. Crost et al. (2008) noted that including only men increases the power of studies like these by eliminating the influences of menstrual cycle phase and the use of oral contraceptives on salivary cortisol levels. Furthermore, men and women may differ with respect to the susceptibility to a stress-inducing experimental procedure (Matud, 2004; McLean & Anderson, 2009). Earlier frontal asymmetry research mostly used undergraduates as participants (e.g., Crost et al., 2008; Lewis et al., 2007; Verona et al., 2009), whereas the present study's sample consisted of participants between 21-37 years old.

Overall, our results indicate that we elicited both direct and associative stress using VR. This is even despite our participants being generally low-anxious and insensitive to positive or negative rewards (as indicated by the STAI and BIS-BAS questionnaires). Together with the fact that the virtual environment and task could have been more realistic and engaging, e.g. by having the participants steer through the environment themselves, this may explain why subjective reported stress did not vary over experimental blocks. By measuring stress even though the participants do not report it, our study suggests an added value of physiological correlates over subjective measures.

Different physiological measures work under different circumstances. We here found that alpha asymmetry correlated with associative stress when the eyes were closed, but not when they were open. For heart variability, it was the other way around. This is probably caused by EEG (and alpha in particular) being strongly affected by visual input, eye movements and blinks, whereas heart variability is not. If EEG is (online) corrected for artifacts, it may work as a correlate for eyes open intervals as well. An advantage of the eyes open interval in this study is that it was longer than the eyes closed interval, probably resulting in more reliable data. In addition, eyes open intervals coincided with a period involving more stress than eyes closed intervals (at least for block 2-4), the latter taking place after the scenario had ended and the performance was known. Different physiological measures also differ in time resolution and delay. Cortisol levels in saliva respond with a relatively long delay and are limited in time resolution. Heart variability, especially as determined by the power in the high frequency range of the RRI, requires quite long windows of analysis. However, variables like these can still be of added value, e.g. if cortisol is a suitable variable to measure direct stress and high frequency heart variability is particularly sensitive to stress. Online neuro-bio feedback brings its own requirements. Cortisol levels cannot be used for online feedback and closing the eyes would not be convenient for most procedures as well. In future experiments we plan to record skin conductance as well.

Our finding that overall cortisol AUC and EEG asymmetry are positively correlated is worth following up on. While in a stress study, Lewis et al. (2007) report a lack of within participant correlation between cortisol and right frontal asymmetry, Schmidt et al. (1999) and Tops et al. (2005) suggest that individuals with relatively larger right frontal brain activity also have higher cortisol levels.

In conclusion, this study showed that more absolute frontal alpha power asymmetry and a lower heart variability are triggered by returning to a virtual environment in which stress had been induced during the first visit, compared to a virtual environment in which stress was not induced. Considering the ultimate goal of the study, finding a basis for the use of VR and neuro-bio information in the treatment of PTSD, it is encouraging that effects were found for associated stress. VRET would usually not include the actual stressor, but merely the environment in which the stressor occurred. Our findings affirm a link between frontal alpha asymmetry, heart variability and cortisol on the one hand and stress on the other hand. Although the significance was only found in eyes-closed data for alpha symmetry and in eyes-open data for heart variability, it is reasonable to expect more successful findings in future research since this study's sample was rather small and heterogeneous; no on-line filtering of EEG data was implemented, short windows for analysis were used when the eyes were closed, and only mild stress was induced. A more comprehensive approach, perhaps even involving the use of a head-mounted display, as well as patients with an anxiety disorder, might give more conclusive results concerning the feasibility of a simultaneous application of VRET and neuro-bio feedback. With respect to neurofeedback, a sequential neurofeedback treatment can be suggested as an option – patients would first be immersed in VR and then close their eyes and try to establish a more symmetrical distribution of alpha power.

Acknowledgements

The authors gratefully acknowledge the support of the BrainGain Smart Mix Programme of the Netherlands Ministry of Economic Affairs and the Netherlands Ministry of Education, Culture and Science. Additional support was provided by the "Brain & Cognition – Social Innovation in Health Care, Education and Safety" programme (project "Resilience and vulnerability after stress").

References

Allen J.J.B., Coan, J.A., & Nazarian, M. (2004). Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. *Biological Psychology, 67*, 183-218.

Beauchaine, Th.P., Gatzke-Kopp, L., & Mead, H.K. (2007). Polyvagal theory and developmental psychopathology: emotion dysregulation and conduct problems from preschool to adolescence. *Biological Psychology, 74*, 174-184.

Benarroch, E.E. (1993). The central autonomic network: functional organization, dysfunction, and perspective. *Mayo Clinical Proceedings, 68*, 10, 988-1001.

Birbaumer, N. (2006). Brain-computer-interface research: Coming of age. *Clinical Neurophysiology, 117* (3), 479-483.

Birbaumer, N., Elbert, T., Rockstroh, B., Daum, I., Wolf, P., & Canavan, A. (1992). *Clinical-psychological treatment of epileptic seizures: A controlled study.* In A. Ehlers, I. Florin, W. Fiegenbaum, & J. Margraf (Eds.), Perspectives and promises of clinical psychology (81–96), New York: Plenum Press.

Brownley, K.A., Hurwitz, B.E., & Schneiderman, N. (2000). *Cardiovascular psychophysiology.* In J.T. Cacioppo, L.G. Tassinary, & G.G. Berntson (Eds.), Handbook of Psychophysiology, 2nd ed. pp. 342-367. Cambridge (UK): Cambridge University press.

Carver, C.S., & White, T.L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319-333.

Coelho, C.M., Waters, A.M., Hine, T.J., & Wallis, G. (2009).The use of virtual reality in acrophobia research and treatment. *Journal of Anxiety disorders, 23*, 563–574.

Crost, N.W., Pauls, C.A., & Wacker, J. (2008). Defensiveness and anxiety predict frontal EEG asymmetry only specific situational contexts. *Biological Psychology, 7*8, 43-52.

Cukor, J., Spitalnick, J., Difede, J., Rizzo, A., & Rothbaum, B.O. (2009). Emerging treatments for PTSD. *Clinical Psychology Review, 29*, 715–726.

De Kloet, E.R. (2003). Hormones, brain, and stress. *Endocrine Regulations, 37*, 51-68.

Difede, J. & Hoffman, H.G. (2002). Virtual reality exposure therapy for world trade center post-traumatic stress disorder: a case report. *Cyber Psychology & Behavior, 5*(6), 529-535.

Egner, T., & Gruzelier, J.H. (2001). Learned self-regulation of EEG frequency components affects attention and event-related brain potentials in humans. *NeuroReport, 12*, 4155-4159.

Egner, T., & Gruzelier, J.H. (2004). EEG Biofeedback of low beta band components: Frequency-specific effects on variables of attention and event-related brain potentials. *Clinical Neurophysiology, 115*, 131-139.

Emmelkamp, P.M.G., Bruynzeel, M., Drost, L., & van der Mast, C.A.P.G. (2001). Virtual Reality Treatment in Acrophobia: A Comparison with Exposure in Vivo. *CyberPsychology & Behavior, 4*(6), 335-341.

Friedman, B.H. (2007). An autonomic flexibility-neurovisceral integration model of anxiety and cardiac vagal tone. *Biological Psychology, 74*, 185-199.

Friedman, B.H., & Thayer, J.F. (1998). Anxiety and autonomic flexibility: a cardiovascular approach. *Biological Psychology, 49*, 303 – 323.

Garcia-Palacios, A., Botella, C., Hoffman, H., Banos, R.M., Osma, J., Guillen, V. & Perpina, C. (2006). *Treatment of mental disorders with virtual reality.* In: Roy, M.J. (Ed.). Novel approaches to the diagnosis and treatment of posttraumatic stress disorder (pp. 170-182). Amsterdam, Netherlands: IOS Press.

Goedhart, A.D., Van der Sluis, S., Houtveen, J.H., Willemsen, G., & De Geus, E.J.C. (2007). Comparison of time and frequency domains of RSA in ambulatory recordings. *Psychophysiology, 44*, 203-215.

Gunawan, L.T., van der Mast, C., Neerincx, M.A., Emmelkamp, P., & Krijn, M. (2004). *Usability of therapist's user interface in virtual reality exposure therapy for fear of flying.* In J. Schreurs & R. Moreau (eds.), Proceedings of the Euromedia 2004 Conference, April 19-21 2004, Hasselt, Belgium, pp. 125-132.

Hammond, D.C. (2005). Neurofeedback with anxiety and affective disorders. *Child and Adolescent Psychiatric Clinics of North America, 14* (1), 105-123.

Hammond, D.C. (2006). What is neurofeedback? *Journal of Neurotherapy, 10* (4), 25-36.

Harmon-Jones, E., Gable, P. A., & Peterson, C. K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biological Psychology, 84* (3), 451-462.

Heinrich, H., Gevensleben, H., & Strehl, U. (2007). Annotation: Neurofeedback - train your brain to train behaviour. *Journal of Child Psychology and Psychiatry 48*(1), 3-16.

Kallen, V.L., Stam, J.V., van Pelt, J., & Westenberg, P.M. (2010). Associations between ANS and HPA-axis responsiveness to stress in adolescence. *International Journal of Psychophysiology, 77*, 284.

Kallen, V.L., Stam, J.V., van Pelt, J., & Westenberg, P.M. (submitted). Towards a neuro-peripheral model of stress rescilience: associations between ANS and HPAaxis responsiveness to stress in adolescence.

Keizer, A.W., Verment, R.S., & Hommel, B. (2010a) Enhancing cognitive control through neurofeedback: A role of gamma-band activity in managing episodic retrieval. *NeuroImage, 49* (4), 3404-3413.

Keizer, A.W., Verschoor, M., Verment, R.S., & Hommel, B. (2010b). The effect of gamma enhancing neurofeedback on the control of feature bindings and intelligence measures. *International Journal of Psychophysiology, 75* (1), 25-32.

Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C., & Hellhamer, D. H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic Medicine, 61*, 154-162.

Kirschbaum, C., Pirke, K.M., & Hellhammer, D.H. (1993). The 'Trier Social Stress Test' – a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*, 76-81.

Krijn, M., Emmelkamp, P.M.G., Olafsson, R.P., & Biemond R., (2004). Virtual reality exposure therapy of anxiety disorders: a review. *Clinical Psychology Review, 24*(3), 259 – 281.

Lewis, R.S., Weekes, N.Y., & Wang, T.H. (2007). The effect of a naturalistic stressor on frontal EEG asymmetry, stress, and health. *Biological Psychology, 75* (3), 239-247.

Lubar, J.F., & Shouse, M.N. (1976). EEG and behavioral changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR): A preliminary report. *Biofeedback and Self Regulation, 1*, 293–306.

Masten, C.L., Eisenberger, N.I., Borofsky, L.A., Pfeifer, J.H., McNealy, K., Mazziotta, J.C., & Dapretto, M. (2009). Neural correlates of social exclusion during adolescence: Understanding the distress of peer rejection. Social Cognitive and Affective Neuroscience, 4 (2), 143-157.

Matud, M. P. (2004). Gender differences in stress and coping styles. *Personality and Individual Differences, 37*, 1401-1415.

McLean, C. P., & Anderson, E. R. (2009). Brave men and timid women? A review of the gender differences in fear and anxiety. *Clinical Psychology Review, 29*, 496-505.

Montgomery, A. J., Mehta, M. A., & Grasby, P. M. (2006). Is psychological stress in man associated with increased striatal dopamine levels?: A [11C]raclopride PET study. *Synapse, 60*, 124-131.

Peniston, E.G., & Kulkosky, P.J. (1991). Alpha-theta brainwave neuro-feedback therapy for Vietnam veterans with combat-related post-traumatic stress disorder. *Medical Psychotherapy, 4*, 47-60.

Peniston, E.G., Marrinan, D.A., Deming, W.A., & Kulkosky, P.J. (1993). EEG alpha-theta brainwave synchronization in Vietnam theater veterans with combat-related post-traumatic stress disorder and alcohol abuse. *Advances in Medical Psychotherapy, 6*, 37-50.

Popovic, S., Slamic, M., & Cosic, K. (2006). *Scenario self-adaptation in virtual reality exposure therapy for posttraumatic stress disorder.* In M. J. Roy (Ed.): Novel approaches to the diagnosis and treatment of posttraumatic stress disorder. Amsterdam: IOS Press.

Preussner, J.C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D.H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology, 28*, 916-931.

Ready, D.J., Pollack, S., Rothbaum, B.O. & Alarcon, R.D. (2006). Virtual reality exposure for veterans with posttraumatic stress disorder. *Journal of aggression, maltreatment & trauma, 12*(1-2), 199-220.

Repetto, C., Gorini, A., Vigna, C., Algeri, D., Pallavicini, F., & Riva, G. (2009). The use of biofeedback in clinical virtual reality: the INTREPID project. *Journal of Visualized Experiments, 12*, 1554.

Rizzo, A., Graap, K., Mclay, R.N., Perlma, K., Rothbaum, B.O., Reger, G., Parsons, T., Difede, J., Pair, J. et al. (2007). *Virtual Iraq: initial case reports from a VR exposure therapy application for combat-related post traumatic stress disorder.* Proceedings of the International Workshop on Virtual Rehabilitation, (pp. 124-130). Washington, USA: IEEE Press.

Sapolsky, R.M., Romero, L.M., & Munck, A.U. (2000). How do glucocorticoids influence stress response? Integrating permissive, suppressive, stimulatory, and preparative actions. *Endocrine Reviews, 21*, 55-89.

Schmidt, L.A., Fox, N.A., Goldberg, M.C., Smith, C.C., & Schuklin, J. (1999). Effects of acute prednisone administration on memory, attention and emotion in healthy human adults. *Psychoneuroendocrinology 24*, 461–483.

Shearer, S.L. (2007). Recent advances in the understanding and treatment of anxiety disorders. *Primary Care, 34*, 475-504.

Sutton, K.S., & Davidson, R.J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. *Psychological Science, 8*(3), 204-210.

Tan, G., Thornby, J., Hammond, D.C., Strehl, U., Canady, B., Arnemann, K., & Kaiser, D.A. (2009). Meta-analysis of EEG biofeedback in treating epilepsy. *Clinical EEG and Neuroscience, 40* (3), 173-179.

Tanida, M., Katsuyama, M., & Sakatani, K. (2007). Relation between mental stress-induced prefrontal cortex activity and skin conditions: A near-infrared spectroscopy study. *Brain Research, 1184* (1), 210-216.

Tops, M., Wijers, A.A., van Staveren, A.S.J., Bruin, K.J., Den Boer, J.A., Meijman, T.F., & Korf, J.

(2005). Acute cortisol administration modulates EEG alpha asymmetry in volunteers: relevance to depression. Biological Psychology 69, 181–193.

Vernon, D., Egner, T., Cooper, N., Compton, T., Neilands, C., & Sheri, A. (2003). The effect of training distinct neurofeedback protocols on aspects of cognitive performance. *International Journal of Psychophysiology, 47*, 75-85.

Verona, E., Sadeh, N., & Curtin, J.J. (2009). Stress-Induced Asymmetric Frontal Brain Activity and Aggression Risk. *Journal of Abnormal Psychology, 118* (1), 131-145.

Wang, J., Korczykowski, M.d , Rao, H.d , Fan, Y., Pluta, J.d , Gur, R.C., McEwen, B.S., & Detre, J.A. (2007). Gender difference in neural response to psychological stress. *Social Cognitive and Affective Neuroscience,* 2 (3), 227-239.

Weiskopf, N., Veit, R., Erb, M., Mathiak, K., Grodd, W., Goebel, R., & Birbaumer, N. (2003). Physiological self-regulation of regional brain activity using realtime functional magnetic resonance imaging (fMRI): Methodology and exemplary data. *NeuroImage, 19*, 577–586.

Wood, D.P., Murphy, J., Center, K., McLay, R., Reeves, D., Pyne, J., Shilling & Wiederhold, B.K. (2007). Combatrelated post-traumatic stress disorder: a case report using virtual reality exposure therapy with physiological monitoring. *CyberPsychology & Behavior, 10*(2), 309-315.

DESIGN AND EVALUATION OF A VIRTUAL REALITY EXPOSURE THERAPY SYSTEM WITH AUTOMATIC FREE SPEECH INTERACTION

Niels ter Heijden[1] and Willem-Paul Brinkman[1]

Research on Virtual Reality Exposure Therapy (VRET) to treat social phobia is not new. Still, few studies focus on creating an elaborate conversation between the patient and characters in a virtual environment. This study focuses on techniques to run a semi-scripted conversation between virtual characters and a patient considering both manual and automatic speech response. Techniques evaluated are a speech detector and a speech recognizer. They were compared to a human control condition. We analyzed the flow and interaction individuals (N = 24) experienced and did a Turing-like test. A case study with two phobic patients was also conducted. Both the patients and therapist, and their interaction with the system, were observed. The study showed that the different automatic techniques had their (dis)advantages, but often did not show any significant difference with the human control condition. A VRET system with semi-scripted conversations might therefore be suitable for the treatment of patients with social phobia. Using automatic speech response techniques might reduce the system workload demand placed upon therapists, allowing them to devote more attention towards monitoring the patient.

Keywords: Social Phobia, Virtual Reality Exposure Therapy, Natural Speech, Public Speaking, Dialogue

Introduction

People that suffer from social phobia fear social situations in which they believe embarrassment may occur. This often leads to avoidance behavior. They fear scrutiny and negative evaluation by others. Making a phone call, asking for assistance in a shop, or speaking in public, are all situations they might dread. Social phobia is one of the most common types of anxiety disorders, estimated to affect 13.3% of the U.S. population in their lifetime (Kessler, et al., 1994). The disorder is associated with depression, substance abuse (e.g. alcoholism, drug abuse), restricted socialization, and poor employment and education performance (Katzelnick, et al., 2001; Kessler, 2003). The disorder leads to intensive use of health services in the western world (Wiederhold & Wiederhold, 2005).

Exposure in vivo is the gold standard for the treatment of phobias. However, for social phobia this treatment might be difficult to arrange (e.g. arranging an audience), and for the therapist difficult to control (e.g. a patience or a hostile audience). Exposure in Virtual Reality (VR) has therefore been suggested as an alternative with some initial encouraging results (Klinger, et al., 2005; Robillard, Bouchard, Dumoulin, Guitard, & Klinger, 2010). Most VR research focuses on one specific social situation i.e. speaking in front of a small group of virtual characters, also called avatars (Anderson, Rothbaum, & Hodges, 2003; Klinger, et al., 2005; Pertaub, Slater, & Barker, 2001; Slater, Pertaub, & Steed, 1999). Still, in the development of these settings the main focus is often on the body posture of the avatars (Anderson, et al., 2003; Herbelin, 2005; Klinger, et al., 2004; Slater, et al., 1999) and less on oral communication between the patient and the avatar. In work (Grillon, Riquier, Herbelin, & Thalmann, 2006) that does report on oral communication, implementations are often relatively limited in their flexibility to support free natural dialogue. This has motivated the following study into a public speaking scenario with virtual

Corresponding Author:
Willem-Paul Brinkman, Department of Mediamatics, Man-Machine Interaction, Delft University of Technology, Mekelweg 4, 2628 CD Delft, the Netherlands, Tel: +31 (0)15 2783534, Fax: +31 (0)15 2787141, E-mail:w.p.brinkman@tudelft.nl

[1]Delft University of Technology, The Netherlands

avatars that can ask patients questions and respond to their reply with follow up questions using automatic free speech interaction.

Presently, conversational software agents are used in several areas such as automated phone reservation systems (McTear, O'Neill, Hanna, & Liu, 2005), e-retail (McBreen & Jack, 2001) and real estate agents (Cassell, et al., 1999). Most of these dialogues are goal-oriented or follow instructions from a user by scanning for expected keywords. They are not developed with the aim to carry on a casual conversation. Still, attempts have been made to create such an agent. The earliest versions might be chatbots (Hutchens & Alder, 1999; Wallence, 2009; Weizenba, 1966). These type of conversational agents have their own contest ("The Loebner Prize in Artificial Intelligence," 1991) where they compete in a Turing test setting with the aim of developing a conversational agent of which the dialogue cannot be distinguished from a human dialogue. Thus far, none have passed the test successfully and chatbot technology might therefore not be ready to simulate a conversation in a VR treatment setting. Most work seem to focus on Artificial Intelligence Markup Language AIML (Wallace, 2001) or variants-type bots (Galvao, Barros, Neves, & Ramalho, 2004) , although other techniques have also been suggested, such as utilizing sentence grammatical analysis (Araki & Doshita, 1995; Li, Zhang, & Levinson, 2000).

Besides the dialogue reasoning component, the oral user input itself is also problematic. The perfect speech recognizer that converts human speech into language that can be understood by a computer does not yet exist (Jurafsky & Martin, 2009). Therefore, free speech seems too ambitious to realize with existing technology (Jurafsky & Martin, 2009; Martin, Botella, García-Palacios, & Osma, 2007). Instead, semi-automatic alternatives seem more feasible. For example, therapists could control most of the avatars' behavior by simply listening to the patient and selecting appropriate responses. In this way, avatars might have a realistic (oral) behavior; the drawback, however, might be the extensive workload forced upon the therapists. An alternative approach, which removes much of this workload, is to have patients read one of the patient-responses out loud from a short list of possible responses shown on the screen (Brinkman, van der Mast, & de Vliegher, 2008). This method can be implemented fairly successfully with speech recognition. But with a list of sentences on the screen it is unclear how this affects the level of presence and the associated anxiety response. Instead, an intermediating solution, using automatic keyword detection from the users' free speech (McTear, 2002), is explored here. This technique was combined with semi-scripted dialogues that were controlled by a computer algorithm deciding which response the avatar should give to the patient. The focus of the research was to examine what level of technology sophistication would yield the best results while keeping the amount of work needed to create the dialogues in mind. Three automatic speech response techniques and a manual control condition were empirically compared in a controlled experiment and in a case study with two social phobia patients. The results suggest that a VR system with semi-scripted conversations might be a suitable exposure environment for social situations that have a conversational component. Using automatic response techniques are likely to reduce therapist workload, while giving individuals a similar level of presence.

System Design

An experimental Virtual Reality exposure therapy (VRET) system was created to provide a controlled environment for exposing patients to a public speaking situation. Patients were asked to talk for a few minutes about a specific subject in front of a small virtual audience after which virtual avatars asked the patients questions about the subject, and responded to the answers of patients in follow-up questions.

System Set Up

Figure 1 shows the set up of the system used in the study. As the presence of the therapist might have an unwanted effect on the patient, a remote set up was developed in which the patient and therapist were separated into two different rooms. The system was distributed over two computers; one handled the visualization of the virtual world including the avatars and all the audio input by the patient, while a second computer handled the interaction with the therapist and the reasoning logic of the dialogue. For the net communication and the therapist interface the Remote Delft Virtual Reality Exposure Therapy (RD-VRET) framework was used. This framework handled the network communication between the two computers and the different software components, and also provided the basic graphic user interface components for the therapist user interface. The therapist could monitor the patient on a screen with speakers by a live video feed from a camera at the back of the patient room and an audio link with the patient room. The therapist could also open an audio channel to talk to the patient.

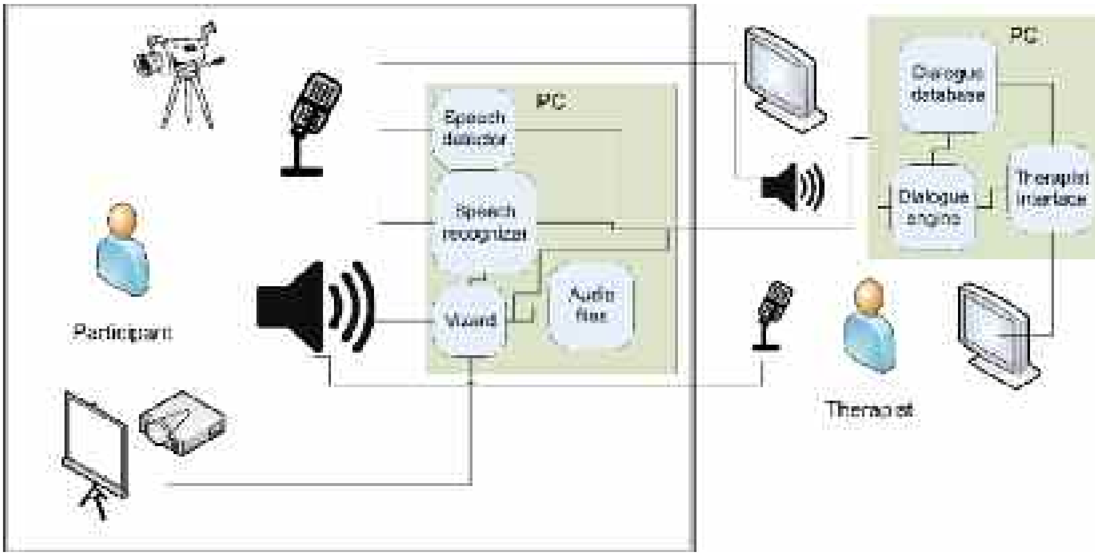


Figure 1. System set up.

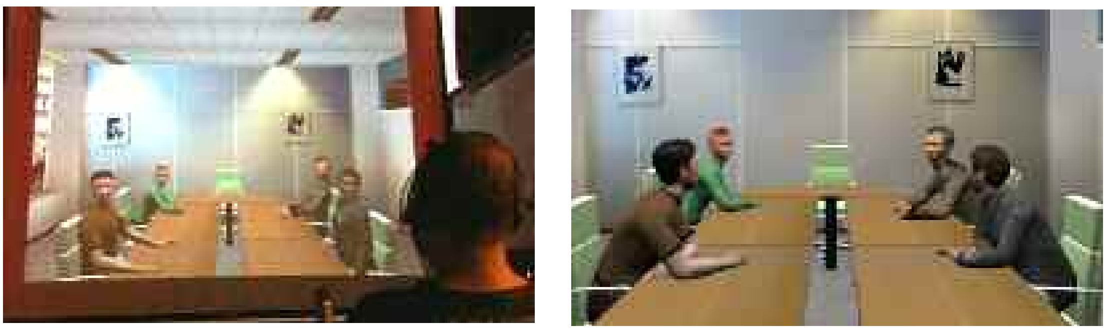


Figure 2. Virtual reality room. Right, an avatar is talking.
Left, avatars look at the patient indicating they are listening to what the patient is saying.

The software package Vizard was used for the visualization of the virtual room and the avatars. The avatars came from a commercial avatar package especially developed for Vizard. Extra animations for the avatars were modelled using key-framing. To support natural turn taking between the avatars and the patient, a software-routine was implemented ensuring that the avatars looked at the avatar that was talking or at the patient that was talking. Patients sat behind a table with a microphone and wore a head set. They sat two meters away from the screen on which a 3.5 by 2.5 meters virtual room was projected with a screen resolution of 1280 × 1024 pixels.

To study different ways of interpreting and responding to patients, four different speech response conditions were implemented. The first condition did not use any speech recognition but solely used the amount of time the patient was talking to base its response on. The second condition did use the speech recognizer but only checked on a predefined finite list of key-words that would be the same for all dialogues. The third condition added the ability to check for certain specific key-words appropriate for the specific point in the dialogue. Finally, the last condition was a control condition, where a therapist selected the responses instead of a computer algorithm.

Dialogue Development

The system was designed to support an alternating computer-patient dialogue, in which the computer (i.e. the avatars) first asked the patient something, at which point the patient replied, to which the computer replied on its turn again, etc. As patients could speak freely, the computer should be able to react on various patient responses. Figure 2 shows an example of a part of such a dialogue. Each computer reply was linked with several potential patient replies. Although in theory there could be an infinitive number of different user replies, potential patient replies were grouped, and for each of these groups an appropriate computer reply was written. To control for the potential exponential growth of a dialogue tree, the depth of the tree was limited to a maximum of five turns after the computer leaf nodes were again brought together in a new opening question node by the computer, who always took the lead in the dialogue. In this way the final dialogue followed a repeating pattern of a widening tree merging back to a single node and widening again from there. A special editor tool, Editor3 (ter Heijden, Qu, Wiggers, & Brinkman, 2010), was developed to support the development of the dialogues. The dialogues were saved to Sqlite database files that could be read later on by the dialogue engine (Figure 1).

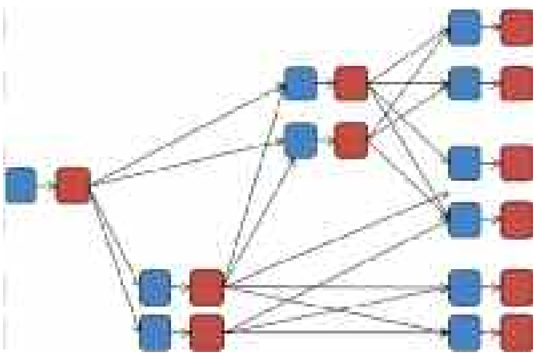


Figure 3. Example of dialogue structure in which the computer reply was linked with various patient responses.

For the study four dialogues were created on the following subjects: democracy, France, dogs, and penguins. For a Dutch target group these subjects were considered general enough that most people would have some knowledge and opinions on them, and allowed the avatars to pose some knowledge and opinion questions about it. Questions about recent events were avoided to make the dialogue more pertinent in future studies. After the initial draft, the dialogues were evaluated and rewritten in iterative cycles. This was done by implementing the dialogue in a chatbot and asking users to chat with it. The results of chatbot sessions (ter Heijden, et al., 2010) were used to evaluate the dialogues and to extract typical user replies, including keywords.

Avatar Response Categories

Besides the questions avatars could pose, special consideration was given to the patient reply. The grouping of these replies was based on two computational parameters – the length of the speech and the detection of specific keywords in the patient's reply by the speech recognizer. Potential patient responses were grouped into seven categories. The first category was the short response category. This was the most basic category. In the speech detection condition this was implemented by detecting whether a patient had given a reply that took less than one second, while in the two conditions with the speech recognizer an additional check was implemented to include only sentences containing less than six words. Possible avatar responses for this category were "Could you tell us something more?" or simple "Why?" The main aim of the avatars' replies was to engage patients in lengthier responses, and thereby overcome potential avoidance behavior.

The second category was the Yes/No category. This category was used if the speech recognizer found the word "Yes" or "No" in the patient response. Possible avatar responses for this category were, for example, "Why is that?" or more context-related such as "What characteristics of birds do they have then?" on a patient's "yes" reply to the question whether they consider a penguin a bird. Likewise, a "no" response would result in a reply such as 'Why not?' The only additional action for a "no" response was a check for double negative, in which case that response was ignored for this category.

The third category was the positive/negative category. Patient responses which included mainly positive words such as "of course" or "rightly" were considered positive replies. Responses which included mainly negative words such as "not" and "nothing" were considered negative replies. Replies to patients' responses in this category included related questions that were based on the implications of acceptance or rejection of assumptions in the previous questions. For example, a patient's negative response on the question "What is your favourite penguin?" would lead to the avatar asking "What is your favorite polar animal?"

The fourth avatar response category was the "don't know" category. This included patients' responses such as "I don't know" or "I don't have an opinion on that." Possible replies in this category were "Could you really not think of anything?" or "That is okay."

The fifth category was the keywords category. This category included responses in which a specific keyword was detected. Compared to the other categories, developing replies for responses in this category took more developmental effort. First, a list of keywords had to be identified that patients would often use in their response. This was done based on a word frequency analysis on the replies obtained in chatbot evaluations. Second, for each keyword a specific avatar reply was written. In the other categories the avatar reply was less dependent on the question leading up to the patient response, and these replies could therefore be used at multiple occasions in the dialogue. This was not the case for patient responses in the keyword category. Here, unique replies were written for each specific preceding avatar question. For example, when the patient response to the question "What kind of penguin do you know?" included the word emperor penguin, the avatar would continue the dialogue with the follow up question "What is the difference between an emperor penguin and a kings penguin?" The expected added value for this category was that patients might feel that the avatars were really listening and responding to their replies, making the social stressors more authentic.

The sixth avatar response category was the general category. These were avatar responses that did not relate to what the patient might have said. This included opening questions, but also follow-up question on aspects patients might not have addressed yet, making them still seem natural in dialogue. For example, a follow-up question for the question "What is the greatest threat to the penguin species?" was "The penguin is not a protected species. Do you think they should be?" Both questions have synergy with each other, while it is unlikely that the patients might have already addressed the issue of penguins being unprotected species in their initial reply. The seventh and last category is the end category. These are the avatar responses to round up a question line with a remark such as "But we are straying too far away from the main topic, so back to the penguins," or shorter ones like "okay" or "If you say so." Although these avatar responses could provoke a reply from the patient, the avatar did not directly respond to that comment. Instead, another avatar started with a new question line.

Table 1 shows that not all avatar response categories were implemented in each speech response condition. No speech recognition technology was used in the speech detector condition. In this condition the computer only measured when a patient talked or stopped talking, and the amount of time a patient talked. On the other hand, in the (limited) speech recognition conditions, the speech recognizing software Nuance NaturallySpeaking was used. Whereas the limited speech recognition condition only included the more content generic avatar response category, the speech recognition condition also included the keyword category. These two conditions were included to study the effect of including the keyword category since the development of this category is more labor intensive. In the human control condition the therapist selected an avatar reply from all the avatar response categories. For example, across the four dialogues, this meant the therapist could select from an average of 4.35 avatar responses after the avatars had posed their opening question of a new question line.

Table 1
Implementation of avatar response categories in the four speech response conditions

Avatar response category	**Speech detector**	**Limited speech recognition**	**Speech recognition**	**Human Control**
Short	X	X	X	X
Yes/No		X	X	X
Positive / Negative		X	X	X
Don't know		X	X	X
Keywords			X	X
General	X	X	X	X
End	X	X	X	X

Method of the Experiment

The experiment was set up as a within-subject design. All participants were exposed to the four speech response conditions – speech detector, limited speech recognition, speech recognition, and human control. To avoid possible learning effects about a subject, a participant talked about another subject in each condition. Furthermore, the order of condition and the assignment of the subjects to conditions were balanced. This resulted in 24 sequence orders. Each participant was randomly assigned exclusively to one of these sequence orders.

Measures

Before the exposure, participants were asked to complete the Personal Report of Confidence as a Public Speaker (PRCS) questionnaire (Paul, 1966) to measure the participant's fear of public speaking. Furthermore, they were asked to complete a questionnaire to collect basic information such as gender, age, etc. and the level of proficiency and knowledge concerning computers, 3-D techniques, and VR. After the each condition, the participant's sense of presence was measured with a modified version of the Igroup Presence Questionnaire (IPQ) (Schubert, Friedmann, & Regenbrecht, 2001) including only the general sense of being there question and the experienced realism questions to measure their subjective experience of realism in the virtual environment. In addition, participants completed the Dialogue Experience Questionnaire (DEQ) (see appendix), to measure their experience of the dialogue and the avatars. After the four dialogues, participants were asked to complete the full IPQ questionnaire considering all four dialogues, and to complete three Turing test type questions. Here, they were asked to rate the dialogues and place them in order of likeliness that they were controlled by a computer or a human. As a check participants were also asked to order the dialogues on the smoothness of the conversation. Besides subjective data, behavioral data was also collected about the number of times participants were interrupted by the avatars when they had not yet completed their answer. This also included situations where they had just paused for a moment and continued talking.

Procedure

At the start of the experiment participants received a short introduction about the overall aim of the study, and had to sign a consent form. Participants were, however, not informed about the different speech response conditions. After signing the form, they completed the PRCS and basic information questionnaire. Once this was complete, the speech recognizer was trained. The main part of the

Table 2
Cronbach's α results from reliability analysis on factors items of the DEQ

Factors	**Speech detector**	**Limited speech recognition**	**Speech recognition**	**Human control**	**Mean**
Flow					
dialogue speed	0.83	0.81	0.79	0.82	0.81
interruption	0.91	0.87	0.86	0.95	0.90
correctness locally	0.86	0.61	0.88	0.76	0.78
correctness globally	0.76	0.81	0.74	0.82	0.78
Interaction					
involvement	0.80	0.73	0.88	0.87	0.82
discussion satisfaction	0.84	0.82	0.75	0.83	0.81
reality	0.79	0.77	0.83	0.76	0.78

experiment consisted of four sessions with the virtual audience, talking about four different subjects. To help them during the initial three minute presentation about the subject, they were given a sheet with some general pointers to talk about, which, however, did not overlap with the question set of the avatars. The participants were also instructed not to pose questions to the avatars. The presentation phase lasted between 1.5-3 minutes, after which avatars would start the question and answer (Q&A) phase. This consisted of 8-10 main questions with 1-5 follow-up questions each (ter Heijden, et al., 2010). After this, participants filled out IPQ and DEQ. Once completed, they received the presentation sheet to prepare themselves for the next session. Between the second and third session, participants were allowed to take a short break to drink something and walk around. After the four sessions, participants completed the Turing test type questions and were interviewed about the overall experience. The entire experiment took between 1.5-2 hours. Afterwards, the participants received a small gift in the form of a chocolate bar as thanks for their participation.

Participants

Participants were recruited from the researchers own social network. The age of the 24 participants (seven female) ranged from 17-66 years (M = 32.6, SD = 13.4). A higher education was being pursued or had been completed by 14 participants, while 10 participants had no education, or had completed or were following an education below this level. Of the 24 participants 21 had seen 3-D stereoscopic images, but only nine had ever used a VR system before. Participants had an average score of 9.8 (SD = 6.5) on the PRCS questionnaire, with three participants scoring 16, 25, and 26, falling into the category of subjects Paul (1966) had included in his anxiety treatment study.

Results

Presence

The overall IPQ results were compared with the online IPQ data set[1] in a MANOVA which used the source as independent variable and the general presence and three factors as dependent variables. No clear significant (F(4,632) = 2.13, p. = 0.076) difference was found between the online IPQ data set and IPQ rating from this experiment. However, univariate analysis found a significant effect (F(1,635) = 7.79, p. = 0.005) on the experienced realism factor. Participants in this experiment (M = 2.5, SD = 1.13) rated the experienced realism higher than ratings in the online IPQ data set (M = 2.0, SD = 0.83). After each dialogue session participants had also filled out a modified version of the IPQ questionnaire that only included general presence and experienced realism factor questions. A MANOVA with repeated measure taking the speech response condition as within-subject variable and the two IPQ measures as dependent measures found no significant effect (F(6,18) = 1.7, p. = 0.184).

Dialogue Experience

The results of the DEQ were analyzed for their internal consistency. Items that had low correlations with other factor items were removed. The mean Cronbach's α over the four conditions ranged from 0.78 to 0.90, above the 0.7 threshold (Table 2). The participants mean score on the factor items were therefore used for future analysis.

To examine the effect of the conditions on DEQ, a MANOVA with repeated measure was conducted. The speech response condition was taken as a within-subject variable and the four flow factors as dependent measures. The results revealed a signification overall effect (F(12,12) = 5.63, p. = 0.003) for the condition on the flow factors. Univariate analyses on the individual factors, only found a significant effect for speech response condition on dialogue speed (F(3,69) = 14.74, p. <0.001), and the correctness locally (F(3,69) = 7.09, p. < 0.001) factor. Figure 4 shows the mean factor scores. Pair wise comparisons with Bonferroni correction showed that for the dialogue speed rating the human control condition (M = 4.3, SD =1.36) was significantly higher than the limited speech recognition (M = 3.1, SD = 1.14, p. = 0.001) and the speech recognition (M = 2.9, SD = 1.22, p. < 0.001) condition. Likewise, the speech detector (M = 3.9, SD = 1.25) was also rated significantly higher than the limited speed recognition (p. =0.016) and speech recognition conditions (p. = 0.003). It seems, therefore, that the delay in the avatars response when using the speech recognition was noticeable for participants. For the factor correctness locally participants rated the human control condition (M = 4.8, SD = 1.04) higher than the speech detector (M = 4.3, SD = 1.16, p. = 0.007) and the limited speech recognition (M = 3.8, SD = 0.90, p. < 0.001) condition.

A similar analysis was done on the three DEQ interaction factors. This analysis found a significant (F(9,15) = 5.36, p. = 0.002) overall effect for the conditions. Univariate analysis revealed a significant effect in the involvement

[1]Fwww.igroup.org/pq/ipq/data.sav downloaded on 12 Nov 2010

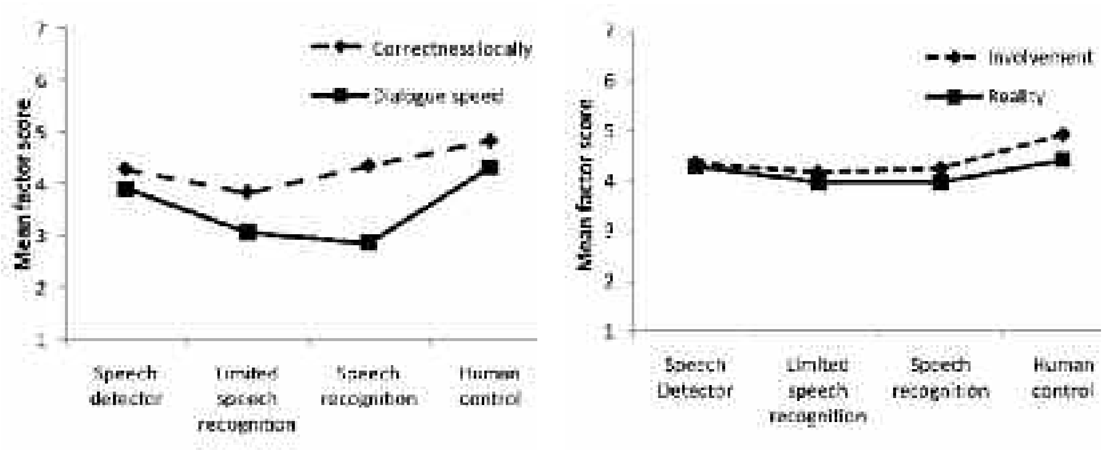


Figure 4. Mean factor on DEQ factor on which significant effects were found for the speech response conditions.

factor (Greenhouse-Geisser correction: $F(2.2, 50.7) = 5.65$, $p. = 0.005$) and in the reality factor ($F(3, 69) = 3.56$, $p. = 0.019$). Pair wise comparisons with Bonferroni correction for the involvement factors showed that the participants had rated the human control ($M = 4.9$, $SD = 1.19$) significantly higher than any of the other conditions (speech detector: $M = 4.4$, $SD = 1.21$, $p. = 0.046$: limited speech recognition: $M = 4.2$, $SD = 1.12$, $p. < 0.001$; speech recognition $M = 4.3$, $SD = 1.40$, $p. = 0.002$). For the reality factor human control ($M = 4.4$, $SD = 1.12$) was only rated significantly higher than the limited speech recognition condition ($M = 4.0$, $SD = 1.15$, $p. = 0.05$).

Turing Test

During the experiment participants were not informed about the different speech response conditions. Because of the speech recognition training, it seems likely they might have expected some computer control, if not all. After completing all dialogue sessions they were asked to rate their sessions on the likeliness that they were computer or human-controlled. Figure 5 shows the mean ratings. An ANOVA found a significant effect ($F(3, 69) = 3.35$, $p. = 0.024$) for the speech response conditions. Pair wise comparison with Bonferroni correction showed that the human control condition ($M = 4.4$, $SD = 1.53$) was

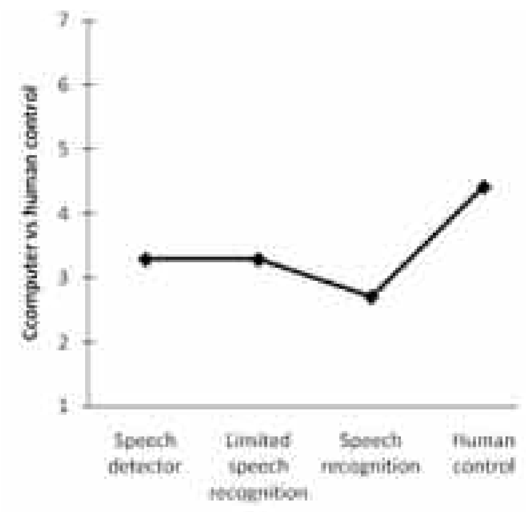


Figure 5. Mean rating on likeliness speech response was controlled by a computer or a human.

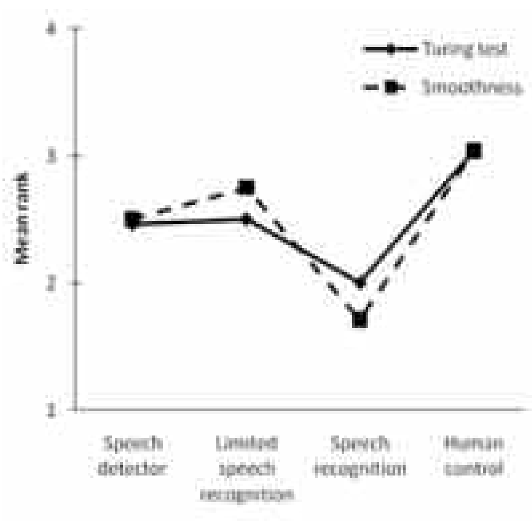


Figure 6. Mean ranking speech response condition on likeliness they were controlled by a computer or a human, and the smoothness of the dialogue.

rated as significantly ($p. = 0.015$) more likely to be controlled by a human than the speech recognition condition ($M = 2.7$, $SD = 1.85$). Interestingly, the rating for human control condition did not significantly ($t(23) = 1.33$, $p. = 0.195$) deviate from the middle of the scale, even though participants were informed one of the four sessions was human controlled before they rated this question. A Friedman test on the ordering task of the sessions on the likeliness of computer or human control also revealed a significant effect ($\chi^2(3) = 7.85$, $p. = 0.049$), as was also the case for the smoothness ordering task ($\chi^2(3) = 14.15$, $p. = 0.003$). As can be seen in Figure 6, and was confirmed by Wilcoxon signed ranks pair wise comparison tests with Bonferroni correction, participants ranked the human control condition significantly higher than the speech recognition condition (Turing test: $p. = 0.037$; Smoothness: $p. = 0.037$).

Avatar Interruptions

Figure 7 shows the number of times an avatar had interrupted the participants when they were still speaking. Interestingly, this was not only a problem for the automatic response conditions as it had also occurred in the human control condition. A Friedman test revealed a significant effect ($\chi^2 3) =10.02$, $p. = 0.018$) for speech response condition. Wilcoxon signed ranks pair wise comparison tests with Bonferroni correction, found only that significantly ($p. = 0.034$) more interruptions were made in the speech detector ($Mdn = 2$) condition than in the human control condition ($Mdn = 0.5$).

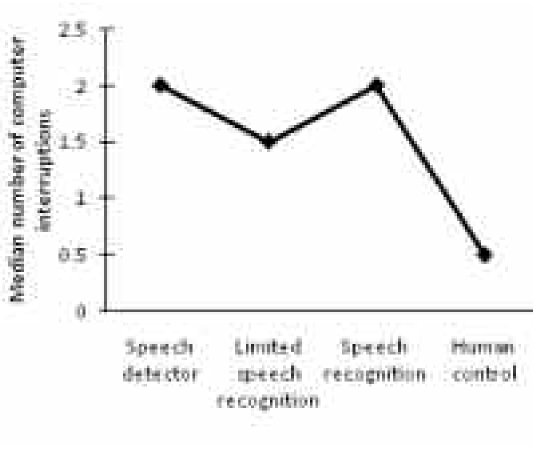


Figure 7. Median of the number of computer interrupts while participants had not finished their answer.

Discussion

Although several significant differences were found, it is more interesting to note that on several places no significant difference was found, especially with the human control condition. A sample size of 24 for a One-sample *t*-test would give an 80% change of finding a significant effect ($d = 0.6$) with a size somewhere halfway between what Cohen (1992) classified as a large ($d = 0.8$) and a medium ($d = 0.5$) size effect. Therefore, the absence of a significant difference could be interpreted that it is unlikely that a difference with a large effect would exist. Only on one measure did the human control condition outperform all the automatic speech response conditions – the DEQ interaction involvement factor. Clearly, participants felt that avatars were listening to them more when they were controlled by a human. However, this did not seem to have had any effect on their overall feeling of presence as no difference was found between any of the speech response conditions. Likewise, in the Turing test, the participants seemed to be unable to make a clear difference between the human control condition and some of the automatic speech response conditions. A drawback of the current implementation of the speech recognizer was the noticeable response delay. If this were to be improved, it would be beneficial since the human control condition did not noticeably outperform the speech recognizer in correctly responding to a participant's answer or holding a conversation that could take place in real life. Ideally, the automated system should combine the speed of the speech detector with the information the speech recognizer gives. Partial sentence information might even be enough for the limited speech recognizer because much of the information concerning attitude can already be found in the first few words of the sentence.

Case Study

As the experiment was conducted with non-patients, a case study was conducted to see how actual patients suffering from social phobia would behave and perceive the system. As this was a preliminary study, any potential treatment effect was not examined. Besides the patient, the case study also provided an opportunity to study how an actual therapist would use and perceive such a system – a user side of VRET systems only explored in a few reports (Brinkman, Sandino, & Van der Mast, 2009; Brinkman, Van der Mast, Sandino, Gunawan, & Emmelkamp, 2010; Wrzesien, Burkhardt, Alcañiz Raya, Botella, & Bretón López, 2010).

Method

One of the university student psychologists and two patients, referred to here as client 1 and client 2, volunteered to participate in the case study. Client 1 was a 43-year-old male MSc student that had not finished his final thesis project, as he avoided his thesis presentation to his supervisors. He wanted to participate because of his personal interest in the technology. Furthermore, it was an opportunity for him to enter the building of his faculty, something he had avoided for some time. Client 1 had a PRCS score of 25, and on the dichotomous version of the Tellegen Absorption Scale (TAS) he scored 11. Client 2 was a 25-year-old male BSc student that had anxiety of presenting in public, believing that he would shake, stutter and speak incoherently when talking before a group of people. He was currently attending sessions from another psychologist. He had a PRCS score of 16, and a TAS score of 16. The therapist was a female psychologist that worked at the university and treated students from the university. The set up of the system and room was similar to the experiment. Because of time constraints, the patients were not exposed to the limited speech response condition. The patients were invited for two sessions. The first session was used to introduce the patients to the system, and trained the speech recognizer on their voice. They also practiced with the system using a few questions from the France dialogue, using a human response condition. In the second session, the patients were exposed to the democracy, dogs, and penguins dialogues. They started with the speech detection condition, followed by the human control condition, took a short break, and finished with the speech recognizer condition. After each dialogue session they completed DEQ, and after the last session they completed IPQ. This was followed by an interview on how they had experienced the sessions. After this, the clients were thanked and for their participation received a small gift that had a value of about €10.

Patient Results

Whereas Client 2's presence score on IPQ did not differ from

Table 3
DEQ results case study clients and experiment participants

	Client 1 (experiment M)			Client 2 (experiment M)		
Factors	**Democracy – Speech detection**	**Dogs – Human control**	**Penguins – Speech recognizer**	**Penguins – Speech detection**	**Democracy – Human control**	**Dogs – Speech recognizer**
Flow						
dialogue speed	2.8 (3.3)	3.0 (5.4)*	3.0 (3.3)	2.0 (3.7)*	4.3 (4.1)	5.5 (2.4) ¥
interruption	3.4 (3.2)	6.0 (5.7)	4.2 (4.5)	5.6 (4.6)	5.8 (4.6)	3.4 (4.4)
correctness locally	3.8 (3.8)	3.3 (5.3)¥	3.8 (4.9)	2.8 (4.7)*	4.8 (5.0)	5.5 (4.0)*
correctness globally	4.0 (4.6)	3.8 (5.5)*	4.3 (4.5)	4.5 (4.2)	5.0 (5.1)	6.8 (4.8)*
Interaction						
involvement	4.4 (3.7)	4.2 (5.3)	4.0 (5.1)	4.2 (4.7)	4.8 (5.5)	5.6 (3.5)*
discussion satisfaction	3.4 (4.3)	3.6 (5.9)¥	4.4 (5.1)	5.6 (5.5)	6.2 (4.8)	5.0 (4.1)
reality	4.0 (3.4)	5.0 (5.2)	4.2 (5.0)	4.6 (4.1)	4.8 (4.6)	5.6 (3.2)¥

* Sign < 0.05; ¥ sign < 0.01

Table 4
Session time and individual response time of clients and experiment participants in same dialogue – speech response condition

		Session time in seconds		Individual response time in seconds	
Dialogue	**Speech response**	**Experiment M(SD)**	**Client**	**Experiment M(SD)**	**Client M**
Client 1					
Democracy	Speech detection	423(66)	361	6.7(2.5)	4.4
Dogs	Human control	403(89)	385	8.6(4.4)	7.5
Penguins	Speech recognizer	587(49)	560	7.6(2.0)	4.7*
Client 2					
Democracy	Speech detection	389(71)	490*	4.8(2.5)	9.2¥
Dogs	Human control	690(292)	1191 ¥	15(11.7)	26.2
Penguins	Speech recognizer	427(65)	553 ¥	8.8(2.9)	13.0*

* Sign < 0.05; ¥ sign < 0.01

the participants from the experiment, Client 1 rated the overall presence the maximal score of 6, which was significantly ($t(23) = -8.62$, $p. < 0.001$) higher than the participants average rating ($M = 3.5$, $SD = 1.44$). However, on IPQ sub-scales he scored significantly lower (spatial presence 0.4: $M = 3.4$, $SD = 1.14$, $t(23) = 12.87$, $p. < 0.001$; involvement 1.5: M = 3.3, SD = 1.15, t(23) = 7.57, p. < 0.001; experienced realism 1.75: $M = 2.5$, $SD = 1.13$, $t(23) = 3.29$, $p. = 0.003$). Client 2 scored the experienced realism as a value, which was also significantly ($t(23) = 2.21$, $p. = 0.037$) lower than the mean score obtained in the experiment.

The results of DEQ showed (Table 3) that Client 1 was rating the flow factors and the interaction factor discussion satisfaction significantly lower than participants had done after their dogs dialogue session with the human control speech response condition. This was the first time with a patient that the therapist controlled the system, which might partly have caused this. Client 2 gave a lower score to the flow factors than participants in the penguins dialogue with the speech detection condition. However, he gave a much higher rating for almost all DEQ factors for the dogs dialogue with the speech recognizer. Overall, it seems, therefore, that both clients did not have a greater tendency than the participants in the experiment in experiencing the dialogue with human control as more positively, nor was the dialogue with the speech recognizer received more negatively. The data for the speed detector condition seems, however, less conclusive.

Table 4 gives an overview of the session times and the mean length of the answers. Client 1 and Client 2 seem to have used different avoidance strategies. Where Client 1 gave relatively shorter answers, Client 2 gave relatively longer answers compared to participants in the experiment that received the same dialogue – speech response condition. With longer answers Client 2 might have tried to control the conversation. His session times were also significant longer. Client 1, on the other hand, avoided watching the screen, especially in the presentation phase. The other client also often watched the presentation sheet instead of the screen. Both could have been signs of avoidance behavior. Interestingly, the participants in the experiment had also commented on the intense stare of the avatars when it was the participant's turn to talk. The avatars eye gaze might therefore be a phobic stressor, an element also controlled in other VRET environment for the treatment of social phobia (Grillon, et al., 2006).

During the human control conditions, clients were also asked to rate their anxiety on the Subjective Unit of Discomfort (SUD) scale (Wolpe, 1958). Although the scores are relatively low, the SUD score of both clients went up in the Q&A phase (Figure 8). Both clients were also unable to point out the dialogue that was controlled by the therapist in the Turing test. Both rated the speech recognition (score 6) as most likely to be human controlled and the speech detection (score 2) condition as most likely to be computer controlled, with the human control (score 4) being rated at the middle of the scale. In the session with the speech recognizer, Client 2 was, for the first time, interrupted by the computer in the middle of his sentence. He suspected that the therapist had done this on purpose, which might therefore have influenced his rating of the Turing test. Both clients were enthusiastic about the program after the second session and they saw its uses for therapy. Client 2 felt more confident about the way he presented afterwards.

Therapist Results

Observing the therapist, it was apparent that she was more able to monitor the patient in the automatic response conditions than in the human control condition. In the latter, much of her attention focused on selecting avatar responses. Instead of selecting the most fitting response, she often tried to select a response that would provoke the client. In the automatic speech response condition, the computer interrupted the clients in the middle of the sentence at points. This was very interesting for her. She saw it as a potential phobic stressor, and suggested to include this computer behavior on purpose. In addition, she also saw this as a strategy to respond to patients that talk for a long time as a form of avoidance behavior. She also wanted to have more control over the length of the presentation phase, being able to cut off the phase and move to the Q&A phase. She also preferred the order of phases not to be fixed. For some patients, the Q&A phase could be a kind of introduction to the subject, with the presentation phase being the more anxiety-provoking part, especially for patients with fear of presenting to a group. Also, letting the patient stand instead of sit behind a desk while presenting without a bullet point sheet, she considered might be more anxiety provoking. Having more control over the type of avatars and the number attending the presentation she thought would also be an effective way to control the exposure. The intense eye gaze of the avatars might have triggered avoidance behaviors, such as looking away by the clients – an observation also made by others when studying eye tracking behavior of social phobic patients (Grillon, Riquier, & Thalman, 2007). Therefore, the therapist saw this as a future option to altar the ability of the avatar to respond to this kind of behavior. Overall, the therapist was enthusiastic about the system and saw a future of using such a system for homework assignments.

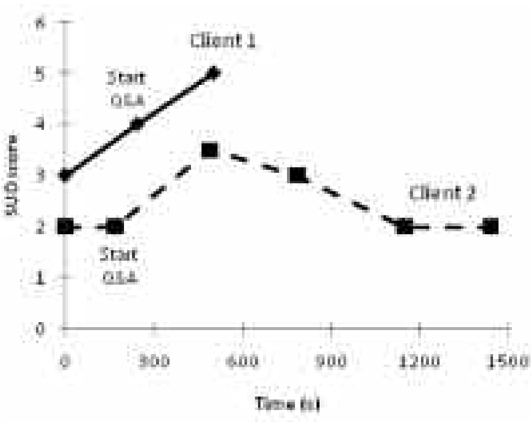


Figure 8. SUD score in the human control speech response condition.

Conclusion and Discussion

Manual speech response did not seem to outperform all automatic speech response techniques both for non-patients and phobic patients on factors such as presence, dialogue flow, discussion satisfaction, dialogue reality, and avatar interruption. The exception, however, was the ability to create the feeling that avatars are really listening. Here, manual control still seems superior. Still, both non-patients and phobic-patients were, on average, unable to distinguish manual control from each of the other types of automatic control. The benefit of automatic control was observed clearly on the therapist side, reducing system workload demands placed upon therapists, thereby allowing the therapist to devote more attention towards monitoring the patient.

The study also had a number of limitations that should be considered. First, to limit the range of potential patients' responses, the dialogues were designed with the computer taking the lead. However, some social phobic patients might also need exposure to situations where they have to take the lead. Second, the results of the speech recognizer condition might depend on the quality of the speech recognizer software used in this study. Other packages might give other results. Third, the social setting only focused on public speaking, whereas social phobic patients might also fear other social situations. Exploring the techniques in other virtual scenes therefore seems interesting, especially whether they provide room for extensive exposure to conversations. Finally, treatment response was not explored. The case study only focused on two therapy sessions. For an actual treatment more sessions seem to be needed, also raising the question of whether multiple dialogues would be required or even multiple social scenes. The study also points out a number of potential social phobic stressors, such as: (1) avatar eye gaze directed towards or away from the patient; (2) responding on patient avoidance behavior when looking away from the avatars; and (3) interrupting the patients when they are talking too long as a form of avoidance behavior. When varied, they might allow control of anxiety-provoking elements in the exposure. Combining these with automatic speech response techniques and automatic anxiety measure instruments might reduce therapist workload demands even further, maybe even to a point where direct continuously-monitored sessions with a therapist might not always be needed. This would open up options such as homework assignments, or a therapist monitoring multiple VRET sessions simultaneously. In the last case, an intelligent software agent in the VRET platform might support the therapist to cope with situations were multiple patients require attention simultaneously (Paping, Brinkman, & van der Mast, 2010). This study also demonstrates that the therapist and patient might not have to be in the same room, as during the sessions they were in separate rooms communicating over a computer network. Still, for any of these technological solutions, it will remain essential to study them empirically, demonstrating they provide at least a similar level of exposure with a reduced therapist workload. The results collected in this study seem to suggest this for automatic speech response techniques used in a VRET system to treat patient suffering from public speaking anxiety.

References

Anderson, P., Rothbaum, B. O., & Hodges, L. E. (2003). Virtual Reality Exposure in the Treatment of Social Anxiety. *Cognitive and Behavioral Practice, 10*(3), 240-247.

Araki, M., & Doshita, S. (1995). Cooperative Spoken Dialogue Model Using Bayesian Network and Event Hierarchy. *Ieice Transactions on Information and Systems, E78d*(6), 629-635.

Brinkman, W.-P., Sandino, G., & Van der Mast, C. A. P. G. (2009). *Field Observations of Therapists Conducting Virtual Reality Exposure Treatment for the Fear of Flying.* Paper presented at the ECCE2009, (pp.169 - 176).

Brinkman, W.-P., van der Mast, C. A. P. G., & de Vliegher, D. (2008). *Virtual Reality Exposure Therapy for Social Phobia: a Pilot Study in Evoking Fear in a Virtual World.* Paper presented at the HCI2008 workshop - HCI for technology enhanced learning, (pp. 85 - 88).

Brinkman, W.-P., Van der Mast, C. A. P. G., Sandino, G., Gunawan, L. T., & Emmelkamp, P. (2010). The therapist user interface of a virtual reality exposure therapy system in the treatment of fear of flying. *Interacting with Computers, 22*(4), 299-310.

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjalmsson, H., et al. (1999). *Embodiment in Conversational Interfaces:* Rea. Paper presented at the CHI'99.

Cohen, J. (1992). A power primer, *Psychological Bulletin,112*(1), 155-159.

Galvao, A. M., Barros, F. A., Neves, A. M. M., & Ramalho, G. L. (2004). Adding Personality to Chatterbots Using the Persona-AIML Architecture. *LNAI 3315*, 963-973.

Grillon, H., Riquier, F., Herbelin, B., & Thalmann, D. (2006). *Use of Virtual Reality as Therapeutic Tool for Behavioural Exposure in the Ambit of Social Anxiety Disorder Treatment.* Paper presented at the Proceedings of the 6th International Conference on Disability, Virtual Reality and Associated Technology.

Grillon, H., Riquier, F., & Thalman, D. (2007). Eye-tracking as diagnosis and assessment tool for social phobia. *Virtual rehabilitation,* 138 - 145.

Herbelin, B. (2005). *Virtual Reality Exposure Therapy for Social Phobia.* Ecole Polytechnique Federale de Lausanne.

Hutchens, J. L., & Alder, M. D. (1999). *Introducing MegaHAL.* Workshop on Human Computer Conversation.

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing : an introduction to natural language processing, computational linguistics, and speech recognition* (2 ed.): Upper Saddle River, N.J.: Pearson Prentice Hall.

Katzelnick, D. J., Kobak, K. A., DeLeire, T., Henk, H. J., Greist, J. H., Davidson, J. R. T., et al. (2001). Impact of generalized social anxiety disorder in managed care. [Article]. *American Journal of Psychiatry, 158*(12), 1999-2007.

Kessler, R. C. (2003). The impairments caused by social phobia in the general population: implications for intervention. [Proceedings Paper]. *Acta Psychiatrica Scandinavica, 10*8, 19-27.

Kessler, R. C., McGonagle, K. A., Zhao, S. Y., Nelson, C. B., Hughes, M., Eshleman, S., et al. (1994). Lifetime and 12-Month Prevalence of DSM-III-R Psychiatric-Disorders in the United-States - Results from the National-Comorbidity-Survey. [Article]. *Archives of General Psychiatry, 51*(1), 8-19.

Klinger, E., Bouchard, S., Legeron, P., Roy, S., Lauer, F., Chemin, I., et al. (2005). Virtual Reality Therapy Versus Cognitive Behavior Therapy for Social Phobia: A Preliminary Controlled Study. *CyberPsychology & Behavior, 8*(1), 76-88.

Klinger, E., Légeron, P., Roy, S., Chemin, I., Lauer, F., & Nugues, P. (2004). Virtual Reality Exposure in the Treatment of Social Phobia. *Studies in Health Technology and Informatics, 99*, 91-119.

Li, Y., Zhang, T., & Levinson, S. E. (2000). *Word concept model for intelligent dialogue agents.* Text, Speech and Dialogue, Proceedings, 1902, 445-449.

The Loebner Prize in Artificial Intelligence (1991). Retrieved 19-11-2008, 2008, from http://www.loebner.net/Prizef/loebner-prize.html

Martin, H. V., Botella, C., García-Palacios, A., & Osma, J. (2007). Virtual Reality Exposure in the Treatment of Panic Disorder With Agoraphobia: A Case Study. Association for Behavioral and *Cognitive Therapies, 14*(1), 58-69.

McBreen, H. M., & Jack, M. A. (2001). *Evaluating humanoid synthetic agents in e-retail applications.* [Article]. Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 31(5), 394-405.

McTear, M. F. (2002). Spoken Dialogue Technology: Enabling the Conversational User Interface. *ACM Computing Surveys, 34*(1), 90-169.

McTear, M. F., O'Neill, I., Hanna, P., & Liu, X. (2005). Handling Errors and Determining Confirmation Strategies -An Object-based Approach. *Speech Communication, 45*(3), 249-269.

Paping, C., Brinkman, W.-P., & van der Mast, C. (2010). *An explorative study into tele-delivered multi-patient virtual reality exposure therapy system.* In K. Wiederhold (Ed.), Coping with posttraumatic stress disorder in returning troops: Wounds of War II (pp. 203-219): IOS press.

Paul, G. L. (1966). Insight VS. *Desensitization in Psychotherapy.* Stangord, CA: Standfort University Press.

Pertaub, D. P., Slater, M., & Barker, C. (2001). *An Experiment on Fear of Public Speaking in Virtual Reality.* Paper presented at the Conference on Medicine Meets Virtual Reality 2001.

Robillard, G., Bouchard, S., Dumoulin, S., Guitard, T., & Klinger, E. (2010). *Using virtual humans to alleviate social anxiety: preliminary report from a comparative outcome study.* In B. Wiederhold, G. Riva & S. I. Kim (Eds.), Annual review of cybertherapy and telemedicine 2010 (pp. 57-60). Amsterdam, The Netherlands: IOS Press.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: factor analytic insights. *Presence, 10*(3), 266-281.

Slater, M., Pertaub, D. P., & Steed, A. (1999). *Public Speaking in Virtual Reality: Facing an Audi-*

ence of Avatars. [Editorial Material]. IEEE Computer Graphics and Applications, 19(2), 6-9.

ter Heijden, N., Qu, C., Wiggers, P., & Brinkman, W. P. (2010). *Developing a dialogue editor to script interaction between virtual characters and social phobic patients*. Paper presented at the ECCE2010 Workshop on cognitive engineering for technology in mental health care and rehabilitation, (pp. 111-123).

Wallace, R. (2001). Artificial Intelligence Markup Language Retrieved 12-12-08, 2008, from http://www.alicebot.org/TR/2001/WD-aiml/

Wallence, R. S. (2009). *The anatomy of A.L.I.C.E.* In R. Epstein, G. Roberts & G. Beber (Eds.), Parsing the turing test (pp. 181-210): Springer.

Weizenba, J. (1966). ELIZA - A Computer Program for Study of Natural Language Communication between Man and Machine. [Article]. *Communications of the Acm, 9*(1), 36-&.

Wiederhold, B. K., & Wiederhold, M. D. (2005). *Virtual reality therapy for anxiety disorders : advances in evaluation and treatment* (1 ed.): DC: American Psychological Association.

Wolpe, J. (1958). *Psychotherapy by Reciprocal Inhibition.* California: Stanford University Press.

Wrzesien, M., Burkhardt, J. M., Alcañiz Raya, M., Botella, C., & Bretón López, J. M. (2010). *Analysis of distributed-collaborative activity during augmented reality exposure therapy for cockroach phobia.* In B. K. Wiederhold, G. Riva & S. I. Kim (Eds.), Annual Review of Cybertherapy and Telemedicine 2010 - Advanced Technologies in Behavioral, Social and Neurosciences (pp. 134-139). Amsterdam, The Netherlands: IOS Press.

APPENDIX

Table - *Items of the Dialogue Experience Questionnaire included in final analysis*

No	Statement (original Dutch text) ¥
	Flow: dialogue speed
1*	The discussion partners needed a long time to think (De gesprekpartners moesten lang na denken)
2*	The discussion partners often went quiet (De gesprekpartners lieten vaak stiltes vallen)
3*	On occasions I had to wait long for a reaction of the discussion partners (Ik moest soms lang wachten op een reactie van de gesprekpartners)
4*	The conversation did not run smoothly (De conversatie verliep stroef)
	Flow: interruption
1	I was always able to finish (Ik kon altijd volledig uitpraten)
2*	On occasions I was unable to tell everything that I would like to have told (Soms kon ik niet alles vertellen wat ik wilde vertellen)
3*	On occasions the discussion partners talked before their turn (De gesprekpartners praten soms voor hun beurt)
4*	On occasions, the discussion partners started talking while I was talking (De gesprekpartners praten soms door me heen)
5	I got enough time from the discussion partners to explain everything calmly (Ik kreeg genoeg tijd van de gesprekspartners om alles rustig te vertellen)
	Flow: correctness locally
1	The discussion partners addressed my answers (De gesprekpartners gingen in op mijn antwoord)
2	The discussion partner responded to my answers (De gesprekpartners reageerde op mijn antwoorden)
3	I got the feeling that the discussion partners understood my answers (Ik had het gevoel dat de gesprekpartners mijn antwoorden begrepen)
4	The questions had a logical order (De vragen volgde een logisch vervolg)
	Flow: correctness globally
1*	The discussion partners rambled (De gesprekpartners sprongen van de hak op de tak)
2*	On occasion, the discussion partners asked things I had already answered. (Soms vroegen de gesprekpartners dingen die ik al beantwoord had)
3*	On occasion, I had to repeat myself (Ik moest me zelf soms herhalen)
4*	I did not always understood why things were asked (Ik snapte niet altijd waarom iets gevraagd werd)
	Interaction: involvement
1	The discussion partners did listen to my answers (De gesprekpartners luisterde naar mijn antwoord)
2*	The discussion partners acted detached (De gesprekpartners reageerden afstandelijk)
3	The discussion partners were interested in my answers (De gesprekpartners waren geïnteresseerd in mijn antwoorden)
4	I got a feeling that I was listened to (Ik kreeg het gevoel dat er naar me werd geluisterd)
5*	On occasion, it felt like the discussion partners were not interested in me (Soms voelde het of de gesprekpartners niet geïnteresseerd in mij waren)
	Interaction: discussion satisfaction
1	The discussion was pleasant (Het gesprek verliep prettig)
2	I was left with a good feeling about the discussion (Ik heb een goed gevoel aan het gesprek overgehouden)
3	I have experienced the question round as pleasant (Ik heb de vragen ronde als prettig ervaren)
4*	The questions of the discussion partners made me nervous (Ik werd nerveus van de vragen die de gesprekpartners stelden)
5*	I did not feel comfortable during the question round (Ik voelde me niet op mijn gemak tijdens de vragen ronde)
	Interaction: reality
1	I got the feeling that this type of conversation could happen in real life (Ik heb het gevoel dat dit soort gesprek ook in het echt kan voorkomen)
2	The discussion partners seemed natural (De gesprekpartners kwamen natuurlijk over)
3	I can imagine that this could happen to me in real life (Ik kan me voorstellen dat ik dit in het echt ook mee zou kunnen maken)
4*	I had to adjust myself to the discussion partners (Ik moest mij zelf aanpassen aan de gesprekpartners)
5	The discussion partners seemed realistic (De gesprekpartners kwamen realistisch over)

¥Rated on 7-point Likert scale ranging from Strongly disagree to Strongly agree (helemaal oneens – helemaal eens); * Score revered

Journal of CyberTherapy & Rehabilitation
Spring 2011, Volume 4, Issue 1

COMBINED USE OF MUSIC AND VIRTUAL REALITY TO SUPPORT MENTAL PRACTICE IN STROKE REHABILITATION

Jonathan Trobia[1], Andrea Gaggioli[1] and Alessandro Antonietti[1]

Mental practice consists of rehearsing a movement with the goal of improving performance. Recent clinical studies suggest that mental practice can be an effective way to facilitate motor recovery after stroke. Though healthy subjects can easily learn to visualize a movement mentally, brain-injured individuals may perceive this task as difficult and overwhelming. We report progress of a research project, which has investigated the feasibility of combining music and Virtual Reality (VR) to support stroke patients in performing mental practice. We tested this approach in two chronic stroke individuals. After four weeks of treatment, both patients showed improved motor function and reported reduced feelings of anxiety. The results of this pilot study are encouraging and deserve further research.

Keywords: Stroke, Virtual reality, Music, Mental practice, Neurorehabilitation

Introduction

Recent years have seen a growing interest towards the application of motor imagery-based training, or “mental practice,” in stroke rehabilitation. According to this approach, patients mentally rehearse a movement with the goal of improving motor performance (Dickstein and Deutsch, 2007; Gaggioli et al, 2009). Neurophysiological studies have shown that prolonged mental practice induces plastic changes in the brain, which are similar to those resulting from physical training. For example, Pascual-Leone and colleagues (1995) used transcranial magnetic stimulation to test patterns of functional reorganization of the brain after mental or physical training of a motor skill. Participants practiced a one-handed piano exercise over a period of five days. Results showed that the size of the contra-lateral cortical output map for the long finger flexor and extensor muscles increased progressively each day and that the increase was equivalent in both physical and mental training.

Sharma and colleagues (2006) conceptualized motor imagery as a "backdoor" to accessing the motor system after a stroke because "it is not dependent on residual functions yet still incorporates voluntary drive" (p. 1942). Brain imaging studies indicated that motor imagery involves a complex distributed neural circuit, which includes the activation of primary motor cortex (M1), supplementary motor area, dorsal and ventral lateral pre-motor cortices, superior and inferior parietal lobules, pre-frontal areas, inferior frontal gyrus, superior temporal gyrus, primary sensory cortex, secondary sensory area, insular cortex, anterior cingulate cortex, basal ganglia and cerebellum (Decety, 1996). However, the use of mental practice in stroke rehabilitation is problematic. Patients often report difficulties in performing mental simulation. Also, neuropsychological evidence suggests that after stroke motor imagery is not symmetrical and that motor imagery vividness is higher when imagining movements on the unaffected than on the affected side. Finally, it is not a simple task to instruct patients to imagine movements using a first-person perspective (kinesthetic imagery), an approach that is believed to be effective to train fine motor skills (Jackson et al, 2001).

In order to support stroke patients in performing mental practice, we developed an integrated training approach, which combines the use of VR technology and mental

Corresponding Author:
Jonathan Trobia, Department of Psychology, Catholic University of Milan, Italy, E-mail: jonathan.trobia@gmail.com

[1]Department of Psychology, Catholic University of Milan, Italy

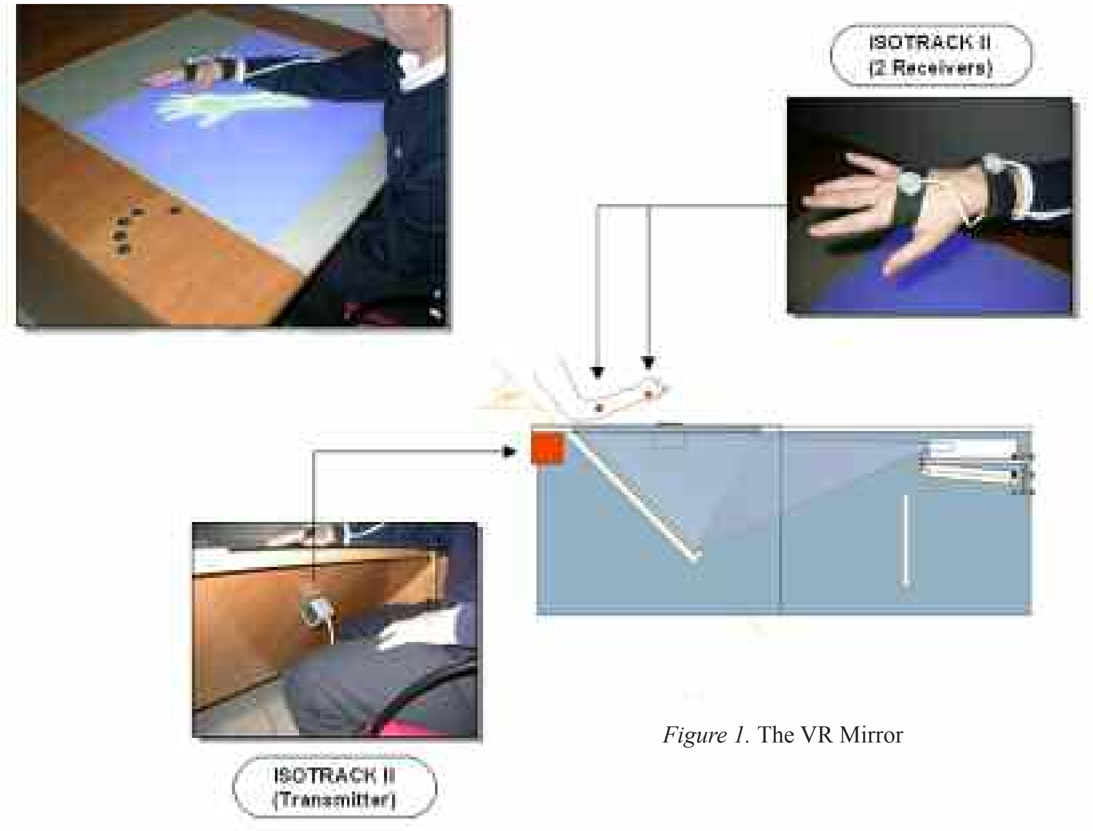


Figure 1. The VR Mirror

practice (Gaggioli, et al, 2006; Gaggioli et al, 2009). In this approach, a VR system called "VR Mirror" displays previously recorded and mirrored movements of the nonparetic arm to the patient. The observed movement is used to support mental rehearsals of the desired movement, and to promote the movement of the impaired limb by following the mirror image. The VR Mirror consists of a table with a back projected horizontal screen, a projector, a mirror and sensors for movement tracking (see Figure 1).

In the present study, we further extended this approach by adding the use of music. We argue that music can support mental practice in two ways – by providing the sequential and rhythmical structures required to perform the movements to be imagined and trained, and by leading patients to relax and experience positive emotional states during the VR exercises. From a neuropsychological perspective, listening to music is a complex experience recruiting different neural circuits (Janata et al, 2003). In fact, listening to music involves the simultaneous activation of several mental codes and the coordinated activation of three cognitive formats – motor, iconic, and verbal (Antonietti, 2009). As concerns the motor level, several studies have highlighted the link between music and body reactions for rehabilitation purposes. For example, Antic et al. (2008) investigated the effects of music in a sample of patients with acute ischemic stroke. They found that listening to music for 30 minutes caused an increase in the mean blood flow velocity in the middle cerebral artery in 80% of participants. In a computer-assisted dynamic posturography study, Carrick and colleagues (2007) found that music induced positive changes in stability scores in individuals with balance abnormalities, indicating that music can be a way to prevent fall and/or vertigo and to rehabilitate persons showing postural disorders. Synchronization between sounds and gestures induced by music has also been successfully used to teach brain-damaged patients to perform the appropriate movements required to dress autonomously (Gervin, 1991).

The association of music to computerized exercises for motor rehabilitation has been investigated by previous studies. For example, Schneider and colleagues (2007) developed a music-supported training program devised to induce an auditory-sensorimotor representation of movements in stroke patients. The equipment included electronic drum pads (to train gross motor skills) and a MIDI-piano (to train fine motor skills). Stroke patients used such tools to reproduce a musical pattern with the impaired arm. Compared to control subjects, patients who underwent music training showed significant improvement after treatment with respect to speed, precision, and smoothness of movements, as well as improvement of motor control in everyday activities.
Both Hummelsheim (1999) and Aldridge (2001) suggested that music can promote motor recovery by stimulating the activation of specific motor cerebral areas. However, besides the specific role that music plays in motor rehabilitation, the positive psychological effects of music should not be overlooked. For example, Särkämö and colleagues (2008) showed that listening to favorite songs daily improved cognitive recovery and mood after mild cerebral artery stroke. Specifically, music treatment determined an increase of positive emotions and a decrease of feelings of confusion and anxiety, which resulted in an enforcement of patients' motivation.

Starting from these premises, the goal of the present study was to investigate the feasibility of combining music and VR to support stroke patients in performing mental practice. The integrated protocol was tested in a pilot study involving two chronic stroke patients with upper limb impairment.

Method

Participants

Two hemiplegic cerebrovascular accident (CVA) patients participated in the study. Patient TG (male, age 68, right handed, 24 months post-CVA), had suffered a left-hemisphere lesion resulting in a chronic hemiplegia of the right arm. Patient LS (male, age 39, right handed, 24 months post-CVA) had suffered a right-hemisphere lesion resulting in a chronic hemiplegia of the left arm.

Materials and Procedure

Neuropsychological assessment included the evaluation of communication and cognitive skills, memory, attention, visuo-spatial, and executive functions according to standardized assessment procedures. Imagery ability was assessed through the Vividness of Movement Imagery Questionnaire (VMIQ) (Isaac et al, 1986). The Montreal Battery Test of Evaluation of Amusia (MBEA) was used to exclude music disorders[1] (Peretz et al, 2003).

The selection of the music pieces that accompanied patients' movements was based on the following criteria:

- High degree of congruence between the speed, duration, and rhythm of the music piece and the corresponding temporal parameters of the target movement exercise

- Association with positive emotional states, using the emotive clues that major tonality suggests (Sloboda, 1985; Gabrielsson and Lindström, 2001; Zatorre, 2003)

- Patients' familiarity with the music piece: Previous research found that well-known music is more effective in enhancing neural activation and mental association than unfamiliar music pieces (Demorest and Morrison, 2003; Peretz and Zatorre, 2005). Patients' attitudes regarding the musical pieces were determined by an interview about their own musical education and preferences.

The treatment protocol consisted in three weekly hospital sessions for four consecutive weeks. In each training session, patients observed the virtual movements displayed on the VR Mirror (Gaggioli et al. 2009) while listening to music pieces, which were previously selected by subjects in accordance with their music preferences. Then, patients were requested to imagine the movement observed on the VR mirror with the accompanying music tracks. In parallel to hospital treatment, patients performed mental practice exercises at home with the support of a DVD, storing audio/video clips of the rehabilitation exercises. Home exercises were executed on a daily basis.

Results and Discussion

Primary pre-treatment and post-treatment measures included the Action Research Arm Test (ARAT) (Lyle, 1981) and Fugl-Meyer Upper Extremity Assessment Scale (Fugl-Meyer, Jaasko and Leyman, 1975). The ARAT is an outcome measure designed specifically for use with patients with strokes. The test is divided into four categories (grasp, grip, pinch, and gross movement), with each item graded on a 4-point scale (0=can perform no part of the test, 1=performs test partially, 2=completes test but takes abnormally long time or has

[1]A lite version is available online, at http://delosis.com/listening/home.html.

Table 1
Functional Improvements After Treatement

Patient	Scale	Baseline	4-weeks	Follow-up (12 weeks)	Improvement
TG	FMA-UE	31	42	43	17%
	ARAT	32	40	43	13%
LS	FMA-UE	21	14	16	3%
	ARAT	1	4	4	5%

great difficulty, 3=performs test normally) and a total possible score of 60. The Fugl-Meyer Scale is a measure of impairment in studies measuring functional recovery in patients with strokes. In this study we used the upper-extremity motor component of the Fugl-Meyer Scale, which consists of 66 points. In addition, patients kept a diary in which they recorded the amount of mental practice exercises performed at home and their personal impressions about the treatment. Pre-post and follow-up scores are reported below. After four weeks of treatment, we observed a sensible improvement in motor function for patient TG and a moderate, albeit notable, improvement in patient LS. Both patients maintained the improvements on the follow-up session.

Conclusions

The main objective of this pilot study was to evaluate the feasibility of using music and VR to support mental practice in stroke rehabilitation. This strategy was tested in two post-stroke patients with chronic motor impairment of the upper limb. After eight weeks of treatment, both patients showed increased motor scores and reported an improvement in ADL. Though the limited number of patients and the absence of a control condition do not allow us to draw any conclusion about the efficacy of this intervention, these preliminary observations suggest that that music and VR can be successfully integrated into mental practice protocols in stroke rehabilitation.

References

Antic S., Galinovich I., Lovrendic-Huzjan A., Vukovic V., Jurasic M.J. and Demarin V. (2008), Music as an auditory stimulus in stroke patients, *Collegium Anthropologicum 32*, 19-23.

Antonietti A. (2009), *Why is music effective in rehabilitation?*, in Gaggioli A., Keshner E., Riva G. and Weiss P. L. (Eds.) (2009), Advanced technologies in neurorehabilitation: Emerging applications in evaluations and treatment, Amsterdam: IOS Publisher.

Aldridge D. (Ed.) (2005), *Music therapy and neurological rehabilitation.* Philadelphia, PA: Jessica Kilngsley Publishers.

Carrick F.R., Oggero E. and Pagnacco G. (2007), Posturographic changes associated with music listening, *Journal of Alternative and Complementary Medicine 13*, 519-526.

Decety J. (1996), Do imagined and executed actions share the same neural substrate?, *Cognitive Brain Research 3*, 87-93.

Demorest S.M. and Morrison S.J. (2003), *Exploring the influences of cultural familiar and expertise on neurological responses to music,* in Avanzini G. (Ed.), Neuroscience and music, Annals of N.Y. Academy of Science, 999.

Dickstein R. and Deutsch J.E. (2007), Motor imagery in physical therapist practice, *Physical Therapy87*, 942-953.

Fugl-Meyer A., Jaasko L. and Leyman I. (1975), The post-stroke hemiplegic patient, I: A method for evaluation of physical performance, *Scandinavian Journal of Rehabilitation Medicine 7*, 13-31.

Gabrielsson A. and Lindström S. (2001), *The influence of musical structure on emotional expression,* in Juslin P.N. e Sloboda J.A. (Eds) (2001), Music and emotion, Oxford University Press, New York, pp.223-248.

Gaggioli A., Meneghini A., Morganti F., Alcaninz M. and Riva G. (2006) A Strategy for Computer-Assisted Mental Practice in Stroke Rehabilitation, *Neurorehabilitation and Neural Repair 20* (4), 503-507.

Gaggioli A., Keshner E., Riva G. and Weiss P. L. (Eds.) (2009), *Advanced technologies in neurorehabilitation: Emerging applications in evaluations and treatment,* Amsterdam: IOS Publisher.

Gervin A.P. (1991), Music therapy compensatory technique utilizing song lyrics during dressing to promote independence in the patient with a brain injury, *Music Therapy Perspectives 9*, 87-90.

Hummelsheim H. (1999), Rationales for improving motor functions, *Current Opinion in Neurology, 12*, 697-701.

Isaac A., Marks D. and D. Russell (1986), An instrument for assessing imagery of movement: The vividness of movement imagery questionnaire (VMIQ), *Journal of Mental Imagery 10*, 23-30.

Jackson P.L., Lafleur M.F., Malouin F., Richards C. and Doyon J. (2001), Potential role of mental practice using motor imagery in neurologic rehabilitation, *Archives of Physical Medicial Rehabilitation 82*, 1133-1141.

Janata P., Tillmann B. and Bharucha J.J. (2002), *Listening to polyphonic music recruits domain general attention and working memory circuits,* Department of Psychological and Brain Sciences, Darmouth College.

Lyle R.A. (1981), Performance test for assessment of upper limb function in physical rehabilitation treatment and research, International Journal of Rehabilitation Research 4, 483-492.

Pascual-Leone A., Nguyet, L.G. Cohen D., Brasil-Neto J.P., Cammarota A., and Hallett M. (1995), Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills, *Journal of Neurophysiology 74*, 1037-1045.

Peretz I. and Coltheart M. (2003), Modularity of music processing, *Nature Neurosciences, 2*, 688-691.

Peretz I., Champod A.S. and Hyde K.L. (2003), Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. Annals of the N.Y. Academy of Sciences 999.

Peretz I. and Zatorre R. J. (2005), Brain organization for music processing, *Annual Review Psychology 56*, 89-114.

Särkämö T., Tervaniemi M., Laitinen S., Forsblom A., Soinlia S., Mikkonen M., Autti T., Silvennoinen H. M., Erkkilä J., Laine M., Peretz I. and Hietanen M. (2008), Music listening enhances cognitive recovery and mood after mild cerebral artery stroke, *Brain 131*, 866-876.

Schneider S., Schönle P.W., Altenmüller E., and Münte T.F. (2007), Using musical instruments to improve motor skill recovery following a stroke, *Journal of Neurology 254*, 1339-1346.

Sharma N., Pomeroy V.M., and Baron J.C. (2006), Motor imagery: a backdoor to the motor system after stroke?, *Stroke 37*, 1941-1952.

Sloboda J.A.(1985), *The musical mind. The cognitive psychology of music,* Oxford: Oxford University Press.

Ward N.S. (2005), *Plasticity and the funtional reorganization of the human brain,* Wellcome Department of Imaging Neurosciences, Institute of Neurology, University College, London.

Zatorre R. (2003), *Music and the brain,* in Avanzini G. (Ed.), Neuroscience and music, Annals of N.Y. Academy of Science, 999,

Journal of CyberTherapy & Rehabilitation
Spring 2011, Volume 4, Issue 1

EVALUATION OF CARD-BASED VERSUS DEVICE-BASED REMINISCING USING PHOTOGRAPHIC IMAGES

Maurice D. Mulvenna[1], Laura Doyle[1], Terence Wright[1], Huiru Zheng[1], Pamela Topping[1], Karl Boyle[1] and Suzanne Martin[1]

Reminiscence activity is commonplace as an activity that is widely recognized as beneficial to people with dementia. It can offer alleviation of the burden of care for those who look after people with this disease. In reminiscence activities, people may use items including photographs representing their lives, in some form. Reminiscence systems are the use of technology to support reminiscence work. This paper describes a study that carried out an evaluation of card-based versus device-based reminiscing using photographic images. The outcome of the study demonstrated no difference between traditional and device-based reminiscing of photographic images, indicating no barriers to the use of systems for reminiscing activities.

Keywords: Reminiscence Systems, Reminiscing, Assistive Technology, Reminiscence Therapy, Reminiscence Work

INTRODUCTION

The response at policy level to the demographic ageing or the “greying” of our populations in developed economies is a move towards self-management where possible, as well as supporting “ageing in place” where the older person can receive care at home, in the supportive circle of family and friends. However, many of these people may not have opportunities for social contact, even at home, and as such, potentially face a degree of social isolation. The paper outlines work in support of reminiscing, which is recognized as an activity that provides benefit to, for example, people with dementia. The activity also is of benefit as a therapeutic intervention to this group and is recognized as beneficial also to the wider, older population. As we age, we gather a large number of life experiences, many of them signifying important life stages – for example, as our family grows, as we impact on the world, and as the world impacts on us. An old photo, of sentimental value, can mean everything to a person, becoming imbued with tremendous significance and often-talismanic importance. These artefacts, whether a location, person or event, or indeed a photo of such an artefact, become the stuff of reminiscing, fuelling what is viewed as a therapeutic process, that, when managed, offers benefits (Koretsky, 2001; Sandoz, 1996) but can reinforce feelings of isolation and depression when unmanaged. As people age, they also increasingly face old age alone, especially in developed economies, as the demography of the post-war (1939-45) period is realized in societies today. The baby boomers of the post-war period are now at or past retirement age, and this increase in numbers of older people results in a significant strain on social and health services. It is projected that by 2025 over 70% of UK households will comprise of people living alone, where a majority will be elderly people. This large body of people, each of whom may have gathered many sets of memories as photographs, has no real facility to use material for reminiscing or share these and to enjoy the therapeutic benefit arising from sharing.

Reminiscing includes a range of activities and traditional tools aimed at stimulating thoughts, feelings and memories of times gone by. For example, these could

Corresponding Author:
Maurice Mulvenna, TRAIL Living Lab, School of Computing and Mathematics, Faculty of Computing and Engineering, University of Ulster, Shore Road, Newtownabbey BT37 0QB, UK, Tel: +44 2890368602, E-mail: md.mulvenna@ulster.ac.uk

[1]University of Ulster, Shore Road, Newtownabbey BT37 0QB, UK

be recalling significant cultural issues, events, old friendships or places. Reminiscing can help older people to improve their own health and wellbeing. In scientific literature, the impact of reminiscing therapy as an intervention has been demonstrated for a range of populations, but primarily for people with dementia.

In this paper, the issue of reminiscing is explored. As a social phenomenon, reminiscing offers social benefits and brings communities together while discussing personal and historical issues of relevance to the participants. However, as, for example, older people may lead increasingly isolated lives, there is a risk that they are excluded from the opportunities to engage socially in activities and acts such as reminiscing, without the effective employment of technology to support the acts. The key question being asked is should the act of reminiscing be supported using information and communications technology (ICT)?

Reminiscence is defined and explained in the paper, in particular from a perspective of the work carried out supporting people with dementia and their carers. Reminiscence systems are then defined and explored and in this section, particular attention is given to commercially available systems including social networking-based systems. The following section examines older people as users of ICT technology and services based upon this technology. The particular scope of studies in the realm of reminiscence systems is then outlined and the comparative evaluation study is then described, encompassing method, study and results. The key question of reminiscence being supported by technology is examined in the study, where photographic images are used by the study group for reminiscing with some participants viewing them as traditional photographs mounted on cards, while others view the images using a touch screen iPad device. Another aspect of the research was to examine if different types of images were treated differently in the study, whether on a device or mounted on card. The research question being examined was whether "personal" images belonging to the participant would be viewed for a longer time than general images, such as "generic" images of the participants' home area or images reflecting some "shared experience" such as a beach holiday. The paper concludes by examining the results of the study.

Reminiscence

Reminiscence has been defined as a "process of thinking or telling others about one's past experiences" (Cappeliez et al. 2005). As a therapeutic intervention, reminiscence is "using the recall of past events, feelings and thoughts to facilitate pleasure, quality of life, or adaptation to present circumstances" (Dochterman and Bulechek, 2003). Reminiscence therapy is generally based one of two theoretical perspectives – Erikson's theory of psychosocial development or Atchley's continuity theory. Erikson's theory emphasises "achievement of ego-integrity and a sense that one's life has meaning and significance, which can be fostered through reminiscence" (Bohlemeijer et al. 2005; Wang, 2005). Continuity theory, on the other hand, "emphasises adjustment by clients to maintain continuity in internal (self concept) and external (social behaviour and social interaction) structures" (Atchley, 1989). Reminiscence is seen as an attempt to maintain continuity between past and present through recall of "familiar knowledge, skills and strategies" (Lin et al. 2003). Reminiscing includes activities and the use of traditional prompts aimed at stimulating feelings and memories; e.g., the use of multi-sensory triggers to stimulate recall (Gibson, 2004). The majority of research in reminiscence systems has been carried out to assist people with dementia and related illnesses (Astell et al. 2008; Sarne-Fleischmann and Tractinsky, 2008). The impact of reminiscing therapy as an intervention has been examined; e.g., Wang (2005) demonstrated how it was valuable and beneficial to people with dementia although Woods et al. (2005) found inconclusive evidence of the efficacy of reminiscence therapy for dementia in a Cochrane Review. However, it has been shown that reminiscence in general, especially life review, are potentially effective methods for the enhancement of psychological well-being in older adults (Bohlemeijer et al. 2007) and the therapeutic potential of place-based reminiscence has been proposed as an avenue in enhancing the quality of life for older people in long-term care facilities (Chaudbury, 2003), sometimes using remote reminiscing facilities (Kuwahara et al. 2006).

Reminiscence Systems

The easiest way in which to describe reminiscing and how ICT supports reminiscing activities is to review some of the systems from research literature as well as systems in use. Reminiscence systems have been defined as the use of technology to support reminiscence work (Mulvenna et al. 2009). The use of multimedia in reminiscence systems was arguably the first stage in the growth of research in reminiscence systems supported by ICT, and there are a significant number of research projects and publications highlighting such work (Newell et al. 2002). It is natural, perhaps, that reminiscing work, which uses visual and hearing senses (as well as other senses) could be enriched with multimedia material encompassing photographs, videos, audio recordings, and music, as well as historical material from newspapers, for example. The multimedia paradigm also lends itself to extending the concept of memory books, used in traditional reminiscence work, where a caregiver or family member compiles a personal scrapbook with images and pictorial mementos of a person's life. Using multimedia, the reminiscence systems can animate the material, thus making it more attractive and attention holding than a paper-based scrapbook. However, since the process of creating a memory book is itself a process rich in reminiscing opportunities, care must be taken not to replace this type of work with a more mundane and less user-centered multimedia authoring process. One such example of an application that has been designed to support reminiscing is MemoryMiner[1]. This is a multimedia application where photographs and other visual material can be added to a repository, annotated and then published to the Internet where further annotations can be made to the material.

In order to make reminiscence systems as accessible as possible for reminiscing by people (and their caregivers, if relevant), the interface of such systems must be as user friendly as possible. This is particularly important where the user is not normally a computer user and/or for when the system has to provide cognitive support, for example, to people with dementia. Touch screen devices are becoming increasingly used in reminiscence systems for people with dementia as the primary mode of interaction, obviating the need for mouse and keyboard combinations. One example of such a device is the CIRCA system (Astell et al. 2008). Touch screen devices, used as a direct input device, have been highlighted as requiring little or no training for users (Pak et al. 2002) and offer a "supportive interaction environment" for people with dementia and their caregivers (Astell et al. 2010a).

Another trend is the increasing use of the Internet and core "platform" services from companies such as Facebook[2] and Google. These services are popular with users and one of the most popular uses is the sharing and annotation of personal photographs and videos, especially amongst younger users. This user-generated content is becoming a significant component of all Internet content, and it can be argued that this content, on these platforms, will become the destination that people will go to in the future when they want to reminisce. One such system used for reminiscing and recall of past events was used to engage people in reminiscing activities using e-mails linking to social media content. The research found that users valued the system and that prompts with images interestingly drew more responses but less thoughtful ones than textual prompts (Cosley et al. 2009).

In terms of social media, existing social networking sites such as Facebook can support reminiscing interactions using specialized applications or "apps." There are also emerging social network sites that cater specifically to older people. These include Sagazone[3] and Finerday[4] in the UK. It is unclear if they support specific reminiscing activities, but a part of their attraction is in bringing older people together in a social network, and supporting their interactions. While not explicitly supporting reminiscing, the MyHeritage[5] site is a social networking application that is strongly orientated to linking family members in support of sharing experiences. In 2009, it held over 550 million individual profiles, over 14 million family trees and 81 million photographs.

Older People as Users

The number of older people is increasing rapidly in coming years, and reminiscing activities offer a social benefit to this group of older people as they age. Technology continues to be adopted and utilized across many areas of society and there is evidence of continued uptake of ICT by older people to assist them in their daily life activities. Therefore, it is valid to take a position that the act of reminiscing can be supported by ICT, and that using ICT to support and extend the capabilities of reminiscing activities offers a wide social benefit.

However, older people's attitude to ICT is different than the attitude of the wider population and is widely divergent from the attitude of the technologically savvy younger generations. In a 2006 research report on the attitude of older people to TV, radio, the Internet and mobile/portable devices, the majority of older people were classified as reluctant participants, that is, "abstainers" or "resistors," where around 74% of over people aged over 65 were either resistors (32% c/f 21% of UK adults) or abstainers (42% c/f 11% of UK adults)

[1]http://www.memoryminer.com/
[2]http://www.facebook.com/apps
[3]http://www.sagazone.co.uk/
[4]http://www.finerday.com/
[5]http://www.myheritage.com/

(Thickett, 2006). Of course, while the majority of older people may be characterized as resistors or abstainers of technology, older people are also embracing some technologies such as digital cameras. Using digital devices with their computers can eliminate many of the issues with traditional photograph storage and use, can make indexing and finding particular images much simpler, and can facilitate sharing of such media much more easy to achieve. As well as having differing attitudes towards technology and technology use compared to the wider population, older people may also have issues with usability and accessibility arising from the impact of age on their senses, primarily on hearing (hearing loss), eyesight (decline in amplitude for ocular accommodation) and cognitive abilities (decline in power of recall, for example), but also on physical dexterity, where, for example, grip strength declines with age. The number of people suffering from chronic diseases increases with age and many of these diseases impact on the ability of a person to use a particular modality of interaction with technology; for example, a person with Parkinson's disease, which causes tremors in hand movement or a slowing of physical movement in general, may not be able to interact well with a touch screen interface. There is also a significant difference between the youngest old (55+) and the oldest old (80+) across these three areas of attitude towards technology, age impairment of senses and cognitive and physical ability, and impairment due to the impact of chronic disease.

If the reminiscing act is supported by technology, then how can it support this group of people, each of whom are old but all have very different needs, requirements and attitudes due to their particular situation? What design decisions would need to be taken for a reminiscence system in order to accommodate such different needs? For example, in the CIRCA system (Astell et al. 2008), designed for people with dementia, the system has been designed to be "failure-free," meaning that there is no obligation for the user to complete any task or to navigate to a particular screen. CIRCA was designed to provide an "intuitive, expansive and fail safe reminiscence experience, utilizing contemporary computer touch screen technology and interactive media design to assist people with dementia and their caregivers and relatives in prompting conversation in one on one or group situations" (Gowans et al. 2009).

The Potential Support Ratio (PSR) is the ratio of the number of 15-64 year olds who could support one person 65+ (retired). In the UK in 1950 the PSR was 12:1, in 2000 9:1 and it is projected to be 4:1 by 2050 and 2:1 in the developed world (United Nations, 2002). There is also, then, an impending crisis in terms of the burden facing formal and informal (families, neighbors and friends) caregivers. The acts of reminiscing, supported using ICT, must also then support the caregiver of the older person, if necessary, and alleviate their care-giving stress or burden. This is particularly important as the number of people with dementia, estimated at 35.6 million people worldwide, is forecast to be 65.7 million in 2030 rising to 115.38 million in 2050 (Prince and Jackson, 2010).

In summary, older people as users are a complex, heterogeneric group, with different needs and characteristics, where physical, cognitive and sense impairment is typically accelerating in old age, where propensity to suffering from chronic disease is increasing, and critically, where the number of caregivers available to support older people is expected to decrease.

Scope of Studies for Reminiscence Systems

We believe that there are three main modalities of use for reminiscence systems. Firstly, the use of a reminiscence system by an individual; secondly, more than one person (may be a person and their caregiver, for example) sharing reminiscences in the same physical space; and thirdly, shared reminiscing where people are physically remote from each other but inter-connected by the Internet. In order to explore how the reminiscing act may be supported by ICT in these three modalities, several experiments have been designed. The methods for these are set out below, and broad topics include:

- Examine act and activities of reminiscing, in particular, our understanding of reminiscence;

- Explore the affordances of existing artefacts used in reminiscence, and the way such artefacts are used in conversation;

- Carry out ethnographic work with older people;

- Establish reminiscence needs identification for older people;

- Develop a user-friendly visual computer system designed to be used by older people for purposes of reminiscing;

• Develop evaluation frameworks using WHO QOL and/or PIADS (Demers et al. 2002) for system in use, as appropriate;

• Assess value of different types of multimedia information (photographs, film, generic media, person-specific media, or "shared experience" media); and

• Evaluate ease of organization and change of media by older people, people with dementia and caregivers.

Research has shown that the function of photographic-based images as memory aids, or as stimuli for reminiscing, can be placed in the context of the narratives that can be constructed around the image (Wright, 2010). Personal memory that relates to cultural memory should be considered, such as how the "objects of memory" can be drawn upon and how they are recounted.

In this paper, the first modality of use is examined in our study in the following section, which compares how older people reminisce using ICT devices, as well as traditional card-based images.

Study and Methods

The study was designed to explore how older people use technology for reminiscing. Research, discussed earlier in this paper, outlined how technologies are being used for reminiscing activities, in particular for people with dementia. In our small study, we wanted to see if there was any significant difference in how people reminisced with and without technology to support the reminiscing process. In order to test this question, participants in the study were randomly allocated to reminisce using either a device (an iPad) or more traditional images, on cards. Therefore, the overall aim of the study was to examine the attitudes of older people in using a device to reminiscence as opposed to a card-based approach. The study measured the impact of card- versus device-based reminiscing using the amount of time spent with each image. The small scope of the study made it difficult to evaluate more complex measures of impact.

Another aim of the study was to see if the different types of images made a difference in reminiscing. These three types of images were personal images belonging to the participants, generic images of life events and shared experience images. For example, a generic image may be an image of musical instruments, a shared experience image may be a photograph of people at a dance hall in 1950's Ireland, while an example of a personal image in this context may be a photograph taken at a dancehall, involving the actual participant. The rationale for examining if there was a difference in how people treated different types of images was founded on recent research, suggesting that "personal items, which have a certain story and set of information attached to them, may limit people's reminiscing to recollection of this information, making it more of a memory test than enjoyable social activity. By contrast, generic items, which have no single story attached, may spark off different recollections in different people, thereby encouraging the sharing of stories and social reminiscing" (Astell et al. 2010b). This recent research related to people with dementia and we wanted to see if there were any indications that older people generally would treat different types of images differently. The shared experience type of image was, in effect, a kind of generic image, designed to highlight people working or playing together.

The inclusion criterion for the study was older people over 55 years of age and normally under 85 years of age in 2010. This included people born between 1925 (for 85-year-olds) and 1955 (for 55-year-olds). Participants were recruited from two locations (Newtownabbey and Newry) in Northern Ireland to reflect a mix of urban and rural dwellers. There were 19 participants in total, with an average age of 71. The oldest participant was 85 and the youngest was 62. Seven of the participants were male.

The participants were informed that the study would help to develop the understanding of computer technology in the area of reminiscence and life story work.

The study divided the participants into two groups, the control group using photographs mounted on cards and those using a device. The study control group used traditional photographs, mounted on card (card-based), while the device group used Apple® iPad® devices with touch-screen navigation of photographs (device-based). The allocation of devices and photo cards to participants was randomized within each location.

In the study, each participant was presented with fifteen photographic images in random sequence, either on the device or mounted on card. The fifteen photographs were drawn equally from personal, shared experience and generic photographs, that relate to when the participant was 18-40 years old, therefore, photographs from

1943 (age 18 for people born in 1925) to 1995 (age 40 for people born in 1955) were used.

In order to prepare for the study, there was one initial group meeting at each location. At this meeting each participant met with the fieldwork supervisor who explained the study to all those participating. Informed consent was obtained from the participants and all relevant questions, queries and concerns were answered. At this initial meeting participants also got the opportunity to become familiar with the hand held device. The device was passed from one participant to another, supervised and supported by the researchers pointing out features of the device, in particular the use of the application designed for the study, displaying sample photographs from Northern Ireland. Care was taken to ensure that each participant then used the device on their own browsing through the photographs.

The fieldwork supervisor asked the participants to consider five images from their personal photograph record that are of personal significance to their lives. The specific wording was to "consider five images that are personally important within your own life, which trigger memories for you." This initial meeting was also an opportunity for interaction with the participants to help guide the selection of the generic and shared experience images.

At a second drop-in meeting at each location, the five personal images selected by the participants were scanned and immediately returned to the participants. The fieldwork supervisor also carried out a pre-study survey with each participant to gauge what images would be appropriate for use as generic and shared experience images, and from this, culturally relevant generic and shared images were selected to avoid presenting an image with which the participant had no interest. At this meeting, the fieldwork supervisor also finalized the schedule of individual meetings.

In the study itself, each participant sat with the facilitator and the participant held the cards or device and decided when to move to the next photograph. Forty-five minutes was set aside for study, but in the event no participant exceeded this time period. The facilitator, who remained the same for all 19 engagements, encouraged each participant to lead their session and invited him or her to tell their story.

After the study, a post-study questionnaire was used to gain feedback from participants on the study.

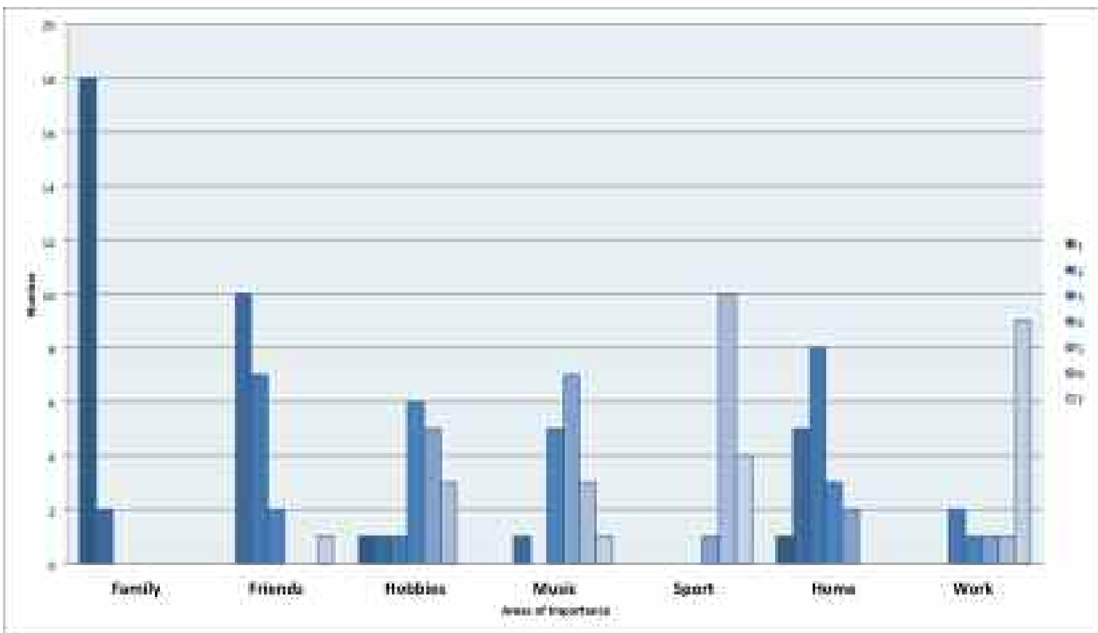


Figure 1. Pre-study Questionnaire on what is important in life.

Results

The pre-study survey asked the participants to explicitly rate what was important in their lives. In Figure 1, each of the following areas were rated from 1-7 on a Likert scale by the participants where "1" equalled important and "7" equalled un-important – family, friends, hobbies, music, sport, home, and work. The responses were then summed. The figure shows that, not surprisingly, the majority of people rated family as very important, with friends, then home ranking highly. This information was used to guide image selection.

A summary of the results of the average time (in seconds) spent by the participants on the image types, divided by locations (A and B) and whether they used a device or card is presented in Table 1.

Table 1
Time spent viewing photographic images (seconds)

	A	B	A	B
	Card		Device	
Generic	115	128	83	93
Shared	127	130	91	82
Personal	114	137	87	112

The data shows that, regardless of location, participants, on average, spent longer viewing photographs that were mounted on card than those on devices, for each of the different photograph types. In terms of differences between the types of photographs, while people spent slightly longer on average looking at personal images, no statistically relevant difference in the durations of viewing between any of the types of photographs (personal, generic, or shared experience) was found.

The results of the post-participation survey provided interesting views on perceptions of card versus device-based photograph reminiscing.

Table 3 shows that almost 80% of the participants found photographs to be emotionally moving.

When asked about their attitude towards card-based and device-based reminiscing, there was a broad range of comments, supportive of both cards and device.

Table 2
Attitudes to reminiscing with photographs

	n	%
Frequently reminiscence	10	53
Occasionally reminiscence	3	16
Not much reminiscence	6	32

Table 3
Attitudes to photographs

	n	%
Yes, I find photos moving	15	79
No, I don't find photos moving	4	21

Participants liked that the device was "Easy to use, tidier than paper," "Images don't get destroyed by handling," "Kept all images together," "Ease of use, image quality," "Ease of use," "Exciting," and commented "It is immediate," and "Very handy and quick." They disliked "Holding it," and commented "Very cold," "Left arm weak after break," and "Sensitive screen initially."

In terms of traditional images, the participants said they liked that "Cards can be handled around friends," "Could hold it in any hand," "Photographs are handy," "Handle images and look at full image closer," "It is more personal," "Being able to hold them," and "Not great with new technology." They disliked that "Traditional images are almost obsolete, everything now is CD," "Would like to try device also."

Discussion

This paper has described reminiscence, and how ICT can support the activities of reminiscing using reminiscence systems. It has provided a review of related work in the area of reminiscence systems, including commercial systems. The different types of technology supporting reminiscence work are described from the use of multimedia and touch screen technology to the growing importance of user-generated content and Internet-connected systems for reminiscing.

No significant issues in using devices for reminiscing were identified, and the study has shown that the participants had positive expectations of using either cards or devices for reminiscing. Participants perhaps enjoyed the novelty of using a device, even though the study sought to offset this intrinsic haptic reward (the fact that some things feel good when we touch them) by providing familiarization sessions for the participants to use the device.

Participants took longer on average interacting with the photographs mounted on card than for device-based photographs. However, perhaps the most important outcome of this study was that the participants did not reject the device-based reminiscing; the results from the surveys indicate that they enjoyed using the device.

When designing the study, it was felt that the participants would relate more strongly to images of their children, for instance, than to generic or shared experience images that did not have a personal family member included. However, a result that was unexpected was that there was no difference in how the participants viewed the three types of images (personal, generic and shared experience). There was also no evidence that participants spent longer viewing and discussing images that were not personal as suggested by Astell and colleagues (2010b).

The benefits in developing support for reminiscence using technology are that such systems offer advantages in ease of use, sharing of photographs and other media, as well as alleviating the burden of care with informal or formal caregivers, if appropriate. People may use devices such as the iPad® to communicate, play games, and browse images and videos from their own chair in their homes. The risks in adopting such technologies are that the essence and richness inherent in such a human activity as reminiscing are lost in translation, and that older people are expected to use technology as a proxy for interaction with other people including family and friends for social visits; and carergivers, for medication and home care support.

Acknowledgements
We acknowledge the participation of Joan Cosgrove and Wilma Neilly of the Newtownabbey Senior Citizens Forum, Colette Ruddy of the Confederation of Community Groups in Newry & Mourne and Denise McBride of the Newry & Mourne Senior Citizens Consortium. We also gratefully acknowledge the participation from the members of the Newtownabbey Senior Citizens Forum, Newry & Mourne Senior Citizens Consortium and Volunteer Now Newry. We also recognise the support of the University of Ulster's Research Impact Award to the Art of Memory project, and the guidance from the Reminiscence Network of Northern Ireland (RNNI), in particular from Prof Faith Gibson and Alexey Janes of RNNI.

References

Astell, A. J. Alm, N., Gowans, G., Ellis, M. P., Dye, R., & Campbell, J. (2008). *CIRCA: A communication prosthesis for dementia*. A. Mihailidas, L. Normie, H. Kautz & J. Boger (Eds). Technology and Aging. IOS Press.

Astell, A. J., Ellis, M. P., Bernardi, L., Alm, N., Dye, R., Gowans, G., & Campbell, J. (2010a). Using a touch screen computer to support relationships between people with dementia and caregivers. *Interacting with Computers, 22*, 267-275.

Astell, A. J., Ellis, M. P., Bernardi, L., Alm, N., Dye, R., Gowans, G., & Campbell, J. (2010b). Stimulating people with dementia to reminisce using personal and generic photographs, *Int. J. of Computers in Healthcare , 1*(2), pp.103-105.

Atchley, R.C., 1989. *A continuity theory of normal ageing.* The Gerontologist, 29, pp. 183-190.

Bohlmeijer, E., Valenkamp, M., Westerhof, G., Smith, F. and Cuipers, P., 2005, Creative reminiscence as an early intervention for depression: results of a pilot programme. *Ageing and Mental Health, 9*, pp. 302-304.

Bohlmeijer, E., Roemer, M., Cuijpers, P., Smit, F., (2007), The effects of reminiscence on psychological well-being in older adults: a meta-analysis. *Aging and Mental Health, 11*(3) 291-300.

Cappeliez, P., O'Rourke, N and Chadbury, H., 2005. Functions in reminiscence and mental health in later life. *Ageing and Mental Health, 9*, pp. 295-301.

Chaudhury, H., (2003), Quality of life and place-therapy, *Journal of Housing for the Elderly, 17*,1/2, pp85-103.

Cosley, D., Schwanda, V., Peesapati, S.T., Schultz, J., Baxter, J., (2009) *Experiences with a Publicly Deployed Tool for Reminiscing,* In: Mulvenna, M.D., Astell, A.J., Zheng, H., Wright, T., (Eds.) Proceedings of First International Workshop on Reminiscence Systems, CEUR Workshop Proceedings, pp. 31-36.

Dochterman, J.M and Bulechek, G.M., 2003, eds. *Nursing Interventions Classification: Iowa Intervention Project.* 4th edn. St Louis: Mosby.

Demers, L., Monette, M., Descent, M., Jutai, J., Wolfson, C., (2002), The Psychosocial Impact of Assistive Devices Scale (PIADS): Translation and preliminary psychometric evaluation of a Canadian-French version, *Quality of Life Research, 11*, 583-592.

Gibson, F., (2004), *The Past in the Present: Using reminiscence in health and social care*. Baltimore: Health Professions Press.

Gowans, G., Campbell, J., Astell, A., Ellis, M., Norman, A and Dye, R., 2009. *Designing CIRCA (Computer Interactive Reminiscence and Conversation Aid). A multimedia conversation aid for reminiscence intervention in dementia care environments.* Dundee: University of Dundee - School of Design.

Koretsky, P. (2001). Using photography in a therapeutic setting with seniors. Afterimage: *The Journal of Media Arts and Cultural Criticism, 29*:3 (Nov/Dec), 8.

Kuwahara, N., Abe, S., Yasuda, K., Kuwabara, K., (2006) *Networked reminiscence therapy for individuals with dementia by using photo and video sharing*, Assets '06: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility, pp.125-132, ACM Press, New York, NY, USA.

Lin, Y.C., Dai, Y.T and Hwang, S.L., 2003. The effect of reminiscence therapy for the treatment of depression in rural-dwelling older adults. *Issues in Mental Health Nursing, 23*, pp. 279-290.

Mulvenna, M.D., Astell, A.J., Zheng, H., Wright, T., *Reminiscence Systems,* In: Mulvenna, M.D., Astell, A.J., Zheng, H., Wright, T., (Eds.), Proceedings of First International Workshop on Reminiscence Systems, Cambridge, UK, September, 2009, pp. 2-4.

Newell, A. F. Carmichael, A. Gregor, P. Alm, N., *Information technology for cognitive support,* pp.464-481, In: Jacko, J.A., Sears, A., (eds.), The Human-computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications (Human Factors & Ergonomics), Lawrence Erlbaum Associates Inc, 2002.

Pak, R., McLaughlin, A.C., Lin, C., Rogers, W.A. & Fisk, A.D. 2002, "*An Age-Related Comparison of a Touchscreen and a Novel Input Device"*, Annual Meeting Proceedings Human Factors and Ergonomics Society, pp. 189-192.

Prince, M., AND Jackson, J. 2010. Alzheimer's Disease International World Alzheimer Report 2009. London.

Sandoz, C.J. (1996). Photographs as a tool in memory preservation for patients with Alzheimer's disease. *Clinical Gerontologist, 17*, 69-71.

Sarne-Fleischmann, V., Tractinsky, N., (2008) Development and evaluation of a personalised multimedia system for reminiscence therapy in Alzheimer's patients, *International Journal of Social and Humanistic Computing, 1* (1), pp. 81-96.

Thickett, J., 2006 Connecting Older People: Consumer Engagement with Digital Services, London, Ofcom.

United Nations 2002. World Population Ageing: 1950-2050. Available at http://www.un.org/esa/population/publications/worldageing19502050/index.htm Accessed on 01.08.07

Wang J.J., (2007), Group reminiscence therapy for cognitive and affective function of demented elderly in Taiwan, *International Journal of Geriatric Psychiatry, Vol. 22* (12), pp.1235-1240.

Wang, J.J., 2005. The effects of reminiscence on depressive symptoms and mood status of older institutionalised adults in Taiwan. *International Journal of Geriatric Psychiatry, 20*, pp. 57-62.

Woods, B., Spector, A., Jones, C., Orrell, M., Davies, S., (2005), Reminiscence therapy for dementia, *Cochrane Database of Systematic Reviews, Issue 1*, Art. No.: CD0011

Wright, T., (2010) Drawn from Memory: reminiscing, narrative and the visual image, *Int. J. Computers in Healthcare, Vol. 1*, No. 2

Spring 2011, Volume 4, Issue 1

VIRTUAL INFORMATION DISPLAY FOR COGNITIVE REHABILITATION: CHOICE OF SCREEN SIZE

Heni Cherni[1], Abdelmajid Kadri[1], Annabelle Taruella[2], Pierre-Alain Joseph[2], Catherine Le Roy[3], and Evelyne Klinger[1]

In VR-based cognitive rehabilitation, there is a particular interest in the subject's performance in the virtual tasks in which patients are engaged. This performance is a function of many factors, among which includes the characteristics of visual information delivered by the virtual system. This study was designed to examine the impact of the size of the display screen on the performance in a virtual task in the Virtual Action Planning Supermarket (VAP-S) among patients with brain injury and healthy control subjects. We designed two configurations – Config L with a large screen and Config S with a small screen. Results show that participants of both groups made a fewer number of incorrect actions in the virtual task in the large screen condition. We conclude that, in our study, increasing the size of the screen improved subjects' perception of the visual information and consequently, the performance of the task.

Keywords: Cognitive rehabilitation, Virtual Reality, VAP-S, Field of view, Performance

This work was conducted at Arts et Métiers ParisTech-LAMPA, in the Handicap & Innovative Technologies Entity, Laval, France.

Introduction

After stroke or traumatic brain injury, cognitive rehabilitation aims to promote the recovery of autonomy using training in Activities of Daily Living (ADL). Due to the frequent lack of efficient resources to deliver the necessary interventions for patients' rehabilitation, therapists are interested in functional virtual environments (VE) (Rizzo, Schultheis, Kerns and Mateer, 2004) that afford the simulation of instrumental ADL, e.g. shopping in a virtual supermarket (Klinger, Chemin, Lebreton and Marié, 2004). The information related to the task has to be delivered in an appropriate way by the virtual system in order to allow the patient to perceive it realistically, and to result in the best conclusions and thus, generate successful reactions. Given the necessary choice of an adequate device to display visual information, we are interested in the impact of the screen size on information perception and on the performance in completing the virtual task. Performance can be defined as a set of recorded data that documents the way in which the subject performs the task.

This study is a continuation of previous works that aimed to compare visual interfaces for better treatment of cognitive dysfunction with virtual reality systems. The use of a large screen is often accompanied by the intuition that such display affords a better presentation of the information and a stronger feeling of immersion. However, there is a lack of research that empirically demonstrates how users benefit from an increase in the screen size (Ni, Bowman and Chen, 2006a). The physical field of view (PFOV) has to be considered. PFOV corresponds to the space seen by both patient's eyes, not moving and fixed in front of them (Zanglonhi, Avital and Prigent, 2000). It depends on the screen size and the distance between the user and the screen.

Corresponding Author:
Evelyne Klinger, Arts et Métiers ParisTech – LAMPA, 4 rue de l'Ermitage 53000 Laval France, Tel:: +33 (0) 243 567 994, E-mail : evelyne.klinger@ensam.eu

[1]Arts et Métiers ParisTech, Laval, France
[2]Hopital Pellegrin, Bordeaux, France
[3]Kerpape Rehabilitation Center, Ploemeur, France

State of Art and Research Aims

Miscellaneous works studied this question with healthy people, mainly in spatial memory tasks, adopting two approaches using two different sizes of screens. The first approach consists in maintaining a constant PFOV by adjusting the viewing distance to screens with different sizes. This configuration induces the same FOV regardless of the size of the screens. The result is that objects are seen with the same visual angle, and as a result have exactly the same size on the retina independent of their real size on the screens (Rodieck, 2003). In this approach, the size of the information perceived by the user is the same in the various types of display.

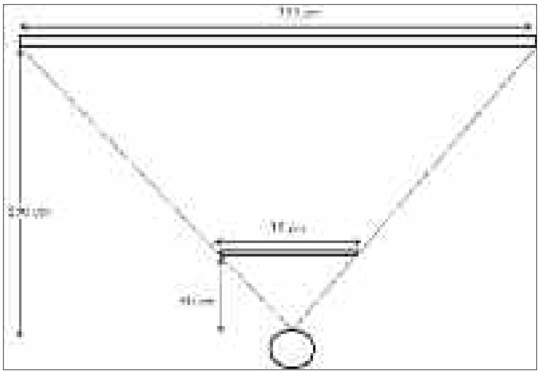


Figure 1. Two sizes of screens and constant PFOV (variable distance).

In their study, Patrick et al. studied 48 participants in a spatial recognition task in a VE for which they had to complete a cognitive task. With a PFOV fixed at 42°, two configurations were implemented, depending on the size of the screens. In the first one, users were placed at a distance of 0.69 m from a screen of 0.53 m wide. In the second configuration, users were placed at a distance of 2.66 m from a projection screen of 3.35 m wide. The results showed that participants had a poor comprehension of distances on the small screen which might be due to the large variation of the PFOV because of user's head movements during the experiment (Patrick, Cosgrove, Slavkovic, Rode, Verrati and Chiselko, 2000). In their study, Tan et al. designed two experimental conditions in which users were placed in front of 0.36 and 1.93 m wide screens of at a distance of 0.64 m from the small screen and 3.45 m from the large screen. A PFOV of 52° was adopted. Participants were involved in two different tasks: – a spatial orientation task in a 3-D complex VE and a reading comprehension task. Results showed that in the case of egocentric strategies (i.e. the user considers his point of view as if he is in the environment), the large screen resulted in the improvement of user's performance in spatial orientation tasks. But no difference was found between the two configurations in the reading comprehension tasks (Tan, Gergle, Scupelli and Pausch, 2006).

The second approach consists in maintaining the same viewing distance to two screens with different sizes in order to increase the PFOV and as as a result, the size of the objects on the retina. In their study, Ni et al. placed the user at about 0.61 m from two 0.50 m and 1.20 m wide screens, inducing a PFOV of 48° with the small screen and 90° with the large screen. Results showed that increasing the screen size, and thus the PFOV, improves the efficiency in spatial navigation tasks (Ni et al., 2006a).

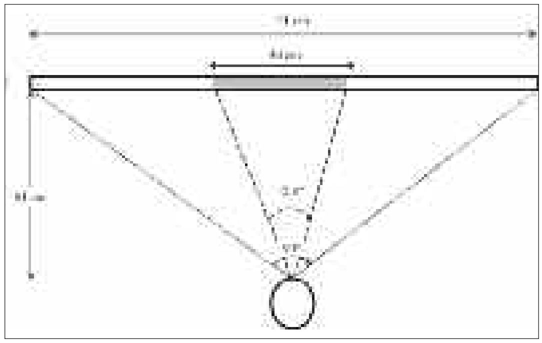


Figure 2. Two sizes of screens and two PFOV (constant distance).

Through these related studies, we understood the importance of the choice of the display device in virtual tasks. In order to design adapted VR-based training conditions for cognitive rehabilitation, we wish to study this question in the context of patients with brain injury. The main aim of this work is to study the impact of the size of the screen on the performance in a shopping task in the VAP-S (Virtual Action Planning Supermarket) among control subjects and patients with brain injury.

Method

Participant

Fifteen patients after AVC, hemiplegics in the chronic phase, (5 males and 10 females, mean age= 53±12) and 26 healthy control subjects (15 males and 11 females, mean age= 32±13) participated in the experiments.

Instrumentation

The VAP-S (Klinger, 2006; Klinger et al, 2004) was de-

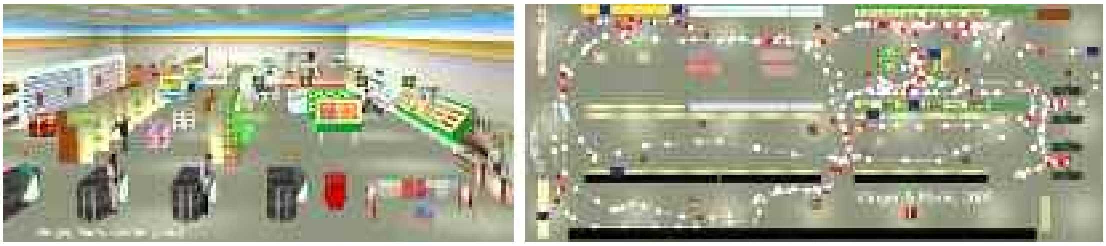


Figure 3. The Virtual Action Planning Supermarket (Klinger, 2006; Klinger et al, 2004).

signed to assess and train the ability of patients to plan and execute the task of purchasing items on a shopping list. Operating the VAP-S includes executing a series of actions, described as a task, and allows an analysis of the strategic choices made by clients and thus their capacity to plan, such as the "test of shopping list" (Martin, 1972). The VAP-S simulates a fully functional medium-size supermarket with multiple aisles displaying most of the items that can be found in a real supermarket. There are also four cashier check-out counters, a reception desk and a shopping cart. Some obstacles, like stacks of bottles or cartons, may hinder the advancement of the shopper along the aisles. In addition, virtual humans are included in the supermarket such as a fishmonger, a butcher, check-out cashier and some costumers. While sitting in front of a PC monitor, the participant enters the supermarket situated behind a cart as if he is pushing it, and moves around freely pressing the keyboard arrows. He is able to select items by pressing the left mouse button. The test of the task is to purchase seven items from a clearly defined list of products, to then proceed to the cashier's desk and to pay for them. Many variables can be calculated from the recorded data such as the total distance traversed by the patient, the total task time, the number of correct actions (e.g., selecting the correct product) and the number of incorrect actions (e.g., selecting a wrong product or selecting the same product several times). Performance in the task is calculated by the set of recorded data which include distance, time, correct actions and incorrect actions.

Studies showed the efficiency of the VAP-S can be used as a tool of cognitive evaluation for various populations such as patients suffering from Parkinson disease (Klinger, Chemin, Lebreton and Marié, 2006), mild cognitive impairment (Werner, Rabinowitz, Klinger, Korczyn and Josman, 2009), and schizophrenia (Josman, Schenirderman, Klinger and Shevil, 2009).

Experimental Conditions

Our objective was to design two experimental conditions, one with a large screen (Config L) and the other one with a small screen (Config S). In order to define the two configurations, we referred to the knowledge on human FOV (Zanglonhi et al, 2000) as well as on the VR-based studies that explored the relation between size of screen and the performance in VEs (Ni et al, 2006a; Ni, Schmidt, Staadt, Livingston, Ball and May, 2006b; Patrick et al, 2000; Tan et al, 2006; Tyndiuk, Thomas, Lespinet-Najib and Schlick, 2005). We chose a pragmatic approach considering the means available in the Bordeaux CHU and the Kerpape Reeducation Center. Therefore, we used a video-projection on a classic wall screen. We found a limitation in increasing the PFOV because the distance between the user and the screen had to still be sufficient in order to preserve an easy way for the patient to explore the space. In fact, the frequent head movements may be difficult for some patients. For the comfort of the patients we chose a 0.4 m wide screen for the small screen resulting in a PFOV of 45°. The large screen was 2 m wide resulting in a PFOV of 70°.

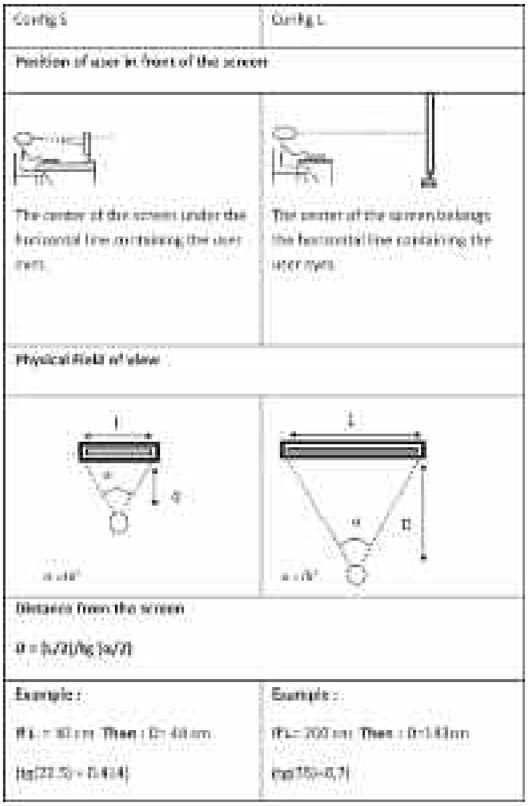
Procedure

All subjects carried out a familiarization session in which they had to buy three items evenly distributed throughout the VAP-S. Then they were engaged in an assessment session in which they had to purchase seven items from a clearly defined list of products without any time constraint. Each subject carried out the assessment session twice within a one week period, once in "config S" and once in "config L" ("S then L" or "L then S"). The eight variables used to describe the performance are the total distance traversed by the patient (DP), the total task time (T), the number of purchased items (NbA), the number of correct actions (BA), the number of incorrect actions (MA), the number of pauses (NbP), the duration of pauses (TP) and the time used to pay (Tp).

Choice of Statistic Tests

Descriptive data analysis (means, standard deviations,

Table 1
Experimental conditions : Config S (small screen= 0.4 m, PFOV= 45°)and Config L (large screen= 2 m, PFOV= 70°)



ranges)was used to describe the population and the main variables representing performance. For each group (patients and controls), we compared the performances in the two conditions (Config S and Config L) using a T test for paired samples (Kinnear and Gray, 2005). In each group, we also compared performance according to gender using the Mann-Whitney test as a non-parametric test since the number of participants in the experiment was less than 30.

Results

During the experiments, three out of 15 patients were unable to finish the purchasing task in the VAP-S, so we did not take these participants into account in the data analysis. Twelve patients (8 F, 4 M, mean age = 50±12) and 26 controls (11 F, 15 M, mean age = 32±13) performed the task in the two conditions – Config S and Config L.

Table 2 and Table 3 present a comparison between the performances in both conditions for patients and controls. The results do not show significant differences between Config L and Config S except for the number of incorrect actions (Controls: MA = 2.8±3.5 in Config L and MA = 7.7±4.5 in Config S, Patients: MA = 16±16 in Config L and MA = 26±14 in Config S). That means that patients with brain injury, as well as healthy controls, make a higher number of incorrect actions in the small screen configuration than in the large screen configuration. These results were main-

Table 2
Performance comparison between Config L and Config S in patients with brain injury

Patients	Config S		Config L		P*
	Patients N= 12 (8 F, 4 M)	Patients Extended value	Patients N= 12 (8 F, 4 M)	Patients Extended value	
Age	50 ±12	[23 ; 67]	50 ±12	[23 ; 67]	-
NbA	7	[7 ; 7]	7	[7 ; 7]	-
BA	11,4 ±1,4	[7 ; 12]	11,7 ±0,8	[9 ; 12]	0,104
MA	26 ±14	[9 ; 52]	16 ±16	[5 ; 59]	0,004
DP (m)	336,4 ±175 ,4	[131 ; 629]	324,4 ±120,6	[189 ; 573]	0,796
T (min)	19,1 ±9,6	[7,3 ; 33]	17,5 ±6,8	[9 ; 33,9]	0,534
NbP	44,9 ±22	[14 ; 78]	37,1 ±19	[12 ; 79]	0,222
TP (min)	10,6 ±6,6	[3 ; 22,9]	9,3 ±5,6	[3,7 ; 25,5]	0,409

* Bilateral signification: $p<0.05 \Rightarrow$ significant result

tained when we compared the performance according to gender (see Table 4, Table 5, Table 6 and Table 7).

For each group of participants, we also compared males and females in both conditions (see Table 8 and Table 9). Although these comparisons do not show significant differences in performance between males and females in both groups, male patients present an advantage in performance in both configurations. More data should be collected to confirm the results; the difference may be due to the fewer number of male patients compared to the number of female patients.

Table 3
Performance comparison between Config L and Config S in control healthy subjects

Controls	Config S		Config L		P*
	Controls N= 26 (11 F, 15 M)	Controls Extended value	Controls N= 26 (11 F, 15 M)	Controls Extended value	
Age	32 ±13	[22 ; 64]	32 ±13	[22 ; 64]	-
NbA	7	[7 ; 7]	7	[7 ; 7]	-
BA	12	[12 ; 12]	12	[12 ; 12]	-
MA	7,7 ±4,5	[2 ; 19]	2,8 ±3,5	[0 ; 17]	0,000
DP (m)	182 ±41,9	[129,2; 295,5]	183 ±51,4	[128 ; 347,2]	0,838
T (min)	6 ±1,8	[3,8 ; 10,3]	6,2 ±2,4	[4 ; 11,8]	0,417
NbP	13,3 ±5,1	[5 ; 23]	13,8 ±6,5	[4 ; 30]	0,632
TP (min)	2,2 ±0,9	[1 ; 5]	2,4 ±1,3	[1 ; 6]	0,181

* Bilateral signification: $p<0.05 \Rightarrow$ significant result

Table 4
Performance comparison between Config L and Config S in control healthy subjects

Controls females	Config S		Config L		P*
	Controls females N= 11	Extended value	Controls females N= 11	Extended value	
Age	33 ±12	[22 ; 59]	33 ± 12	[22 ; 59]	-
NbA	7	[7 ; 7]	7	[7 ; 7]	-
BA	12	[12 ; 12]	12	[12 ; 12]	-
MA	8,2 ± 4,8	[3 ; 19]	2,5 ± 2,3	[0 ; 8]	0,001
DP (m)	169,8 ± 38,5	[129,2 ; 251]	172,5 ± 38	[128 ; 235,7]	0,734
T (min)	5,9 ± 1,8	[3,8 ; 10,3]	6,3 ± 2,4	[4 ; 11,1]	0,299
NbP	14 ± 5	[5 ; 22]	14,6 ± 7,6	[8 ; 30]	0,763
TP (min)	2,2 ± 0,8	[1 ; 4]	2,6 ± 1,4	[1 ; 6]	0,375

* Bilateral signification: $p<0.05 \Rightarrow$ significant result

Table 5

Performance comparison between Config L and Config S in Controls males group

Controls males	Config S		Config L		P*
	Controls males N= 15	Extended value	Controls males N= 15	Extended value	
Age	31 ± 13	[22 ; 64]	31 ± 13	[22 ; 64]	-
NbA	7	[7 ; 7]	7	[7 ; 7]	-
BA	12	[12 ; 12]	12	[12 ; 12]	-
MA	7,2 ± 4,3	[2 ; 19]	2,9 ± 4,2	[0 ; 17]	0,000
DP (m)	190,9 ± 43,2	[140,9 ; 295,5]	190,7 ± 59,5	[129,5 ; 347,2]	0,980
T (min)	6,1 ± 1,8	[4,3 ; 10,2]	6,1 ± 2,4	[4,2 ; 11,8]	0,886
NbP	12,9 ± 5,2	[6 ; 23]	13,3 ± 5,9	[4 ; 27]	0,705
TP (min)	2,2 ± 0 ,9	[1 ; 5]	2,4 ± 1,2	[1 ; 5]	0,317

* Bilateral signification: $p<0.05 \Rightarrow$ significant result

Table 6

Performance comparison between Config L and Config S in patients' females group

Patients females	Config S		Config L		P*
	Patients females N= 8	Extended value	Patients females N= 8	Extended value	
Age	50 ± 15	[23 ; 67]	50 ± 15	[23 ; 67]	-
NbA	7	[7 ; 7]	7	[7 ; 7]	-
BA	11,1 ± 1,7	[7 ; 12]	11,6 ± 1	[9 ; 12]	0,104
MA	26 ± 12,2	[14 ; 51]	16 ± 18,2	[5 ; 59]	0,015
DP (m)	386,9 ± 172,7	[154 ; 629]	355,4 ± 128,8	[219 ; 573]	0,542
T (min)	21,6 ± 8,9	[7,8 ; 33]	19,2 ± 7,9	[9 ; 33,9]	0,362
NbP	49,5 ± 18,7	[17 ; 78]	40,3 ± 22,4	[12 ; 79]	0,156
TP (min)	12,2 ± 6,5	[3 ; 22,9]	10,3 ± 6,6	[3,7 ; 25,5]	0,260

* Bilateral signification: $p<0.05 \Rightarrow$ significant result

Table 7
Performance comparison between Config L and Config S in patients' males group

Patients males	Config S		Config L		P*
	Patients males N= 4	Extended value	Patients males N= 4	Extended value	
Age	52 ± 1	[51 ; 53]	52 ± 1	[51 ; 53]	-
NbA	7	[7 ; 7]	7	[7 ; 7]	-
BA	12	[12 ; 12]	12	[12 ; 12]	-
MA	25 ± 19,4	[9 ; 52]	17,2 ± 13 ,5	[5 ; 30]	0,226
DP (m)	235,5 ± 151,4	[131 ; 457]	262,5 ± 83,6	[189 ; 364]	0,810
T (min)	14 ± 10	[7,3 ; 29]	14,3 ± 1,6	[12,3 ; 16,2]	0,968
NbP	35,7 ± 28,1	[14 ; 77]	30,7 ± 8,4	[24 ;43]	0,767
TP (min)	7,4 ± 6,4	[12,3 ; 16,2]	7,3 ± 1,9	[5,3 ; 9,7]	0,985

* Bilateral signification: p<0.05 ⇒ significant result

Table 8
Performance comparison between males and females in controls' group in Config L and Config S

Controles (15M, 11F)	Config S			Config L		
	Males	Females	P*	Males	Females	P*
Age	31 ± 13	33 ± 12	0.631	31 ± 13	33 ± 12	0.145
NbA	7	7	-	7	7	-
BA	12	12	-	12	12	-
MA	7,2 ± 4,3	8,2 ± 4,8	0,620	2,9 ± 4,2	2,5 ± 2,3	0,751
DP (m)	190,9 ± 43,2	169,8 ± 38,5	0,126	190,7 ± 59,5	172,5 ± 38	0,484
T (min)	6,1 ± 1,8	5,9 ± 1,8	0,678	6,1 ± 2,4	6,3 ± 2,4	0,687
NbP	12,9 ± 5,2	14 ± 5	0,450	13,3 ± 5,9	14,6 ± 7,6	0,855
TP (min)	2,2 ± 0 ,9	2,2 ± 0,8	0,500	2,4 ± 1,2	2,6 ± 1,4	0,938

* Bilateral signification: p<0.05 ⇒ significant result

Table 9
Performance comparison between males and females in patients' group in Config L and Config S

Patients (4M, 8F)	Config S			Config L		
	Males	Females	P*	Males	Females	P*
Age	52 ± 1	50 ± 15	0,02	52 ± 1	50 ± 15	0,02
NbA	7	7	-	7	7	-
BA	12	11,1±1,7	0,280	12	11,6±1	0,480
MA	25 ± 19,4	26 ± 12,2	0,610	17,2 ± 13 ,5	16 ± 18,2	0,864
DP (m)	235,5 ± 151,4	386,9 ± 172,7	0,174	262,5 ± 83,6	355,4 ± 128,8	0,089
T (min)	14 ± 10	21,6 ± 8,9	0,174	14,3 ± 1,6	19,2 ± 7,9	0,234
NbP	35,7 ± 28,1	49,5 ± 18,7	0,234	30,7 ± 8,4	40,3 ± 22,4	0,497
TP (min)	7,4 ± 6,4	12,2 ± 6,5	0,234	7,3 ± 1,9	10,3 ± 6,6	0,396

* Bilateral signification: $p<0.05 \Rightarrow$ significant result

Discussion

In the VAP-S, we used eight variables to describe the participant performance in the purchasing task. For each group, the results of the statistical analyses highlighted a significant decrease in the number of incorrect actions (MA) in the large screen condition. We can infer that virtual objects in the VAP-S are more visible when they are presented on a large screen. In both groups, for males as well as for females, we can conclude that the perception of the visual information delivered by the virtual system is better in the large screen condition. This leads participants to improve their behavior and interaction during the task, and consequently, to improve their performance. Furthermore, the stimuli, which are source of information, can also be a source of confusion and slow down participants if they are delivered in a large quantity. So, the improvement of performance is also explained by a fewer number of stimuli seen simultaneously when the size of the screen is increased. In addition, we expected participants to stop more frequently in the large screen condition in order to move their head to explore the large FOV. In fact, the number of stops was larger in the small screen condition, even if the difference was not significant. This can also be explained by the better perception of information in the large screen condition, inducing less necessity of frequent stops. Inter-subjects comparison showed significant differences between healthy controls and patients with brain injury in all measured variables. These results were not presented in this paper because of the large difference in ages between both groups of subjects.

Conclusion

Thanks to this original work carried out with healthy controls and patients with brain injury, we showed that using a large screen in a virtual task can improve participants' performance by reducing the number of incorrect actions. We also found that subjects have a better perception of the visual information delivered by a virtual system in the large screen condition, leading to a positive impact on their performance in the task. According to our results, we recommend the use of large screens in VR-based cognitive rehabilitation of patients with brain injury. We also recommend assuring the participant is presented with a comfortable physical field of view (e.g., 70°) in order to avoid frequent head movements which may be difficult for some patients with brain injury. Then, we believe that the outcome of this work can be used in VR-based therapy, as well as in VR-based applications for healthy people such as virtual environments for training and virtual simulators like driving simulators.

Although our results indicate that a large screen can improve subjects' performance in a virtual task, additional research is clearly needed to document its usefulness and delineate the limitations of increasing the size of the screen. To extend this work, we would like to explore in more detail the comparisons between the small screen and the large screen with different software field of view (SFOV), which corresponds to the visual angle in the virtual environment. This parameter is often ether confused with the PFOV or not taken in account in most of the studies.

Finally, the outcome of this work will be taken into account in experiments in current VR-based therapy projects we are carrying out in close collaboration with clinical and industrial partners.

Acknowledgements
We thank Laval authorities (Laval Agglomération and Conseil Général de la Mayenne) for their financial support.

References

Josman, N., Schenirderman, A.E., Klinger, E. and Shevil, E., Using virtual reality to evaluate executive functioning among persons with schizophrenia: a validity study, *Schizophr Res, 115*, pp.270-7, 2-3, (2009).

Kinnear, P. and Gray, C., *SPSS facile appliqué à la psychologie et aux sciences sociales,* Paris, (2005)

Klinger, E. (2006), Apports de la réalité virtuelle à la prise en charge des troubles cognitifs et comportementaux, In *Informatique*, Edition228. Paris: ENST.

Klinger, E., Chemin, I., Lebreton, S. and Marié, R.M., A Virtual Supermarket to Assess Cognitive Planning, *Cyberpsychol Behav, 7*, pp.292-293, 3, (2004).

Klinger, E., Chemin, I., Lebreton, S. and Marié, R.M., Virtual Action Planning in Parkinson's Disease: a control study, *Cyberpsychol Behav, 9,* pp.342-347, 3, (2006).

Martin, R., *Test des commissions (2nde édition)*, Editest, Bruxelles, (1972)

Ni, T., Bowman, D.A. and Chen, J., Increased display size and resolution improve task performance in Information-Rich Virtual Environments, In *Proceedings of Graphics Interface 2006*, Quebec, Canada, pp.139-146. (2006a)

Ni, T., Schmidt, G.S., Staadt, O.G., Livingston, M.A., Ball, R. and May, R., A Survey of Large High-Resolution Display Technologies, Techniques, and Applications, *Proceeding IEEE VR Conf*, (2006b).

Patrick, E., Cosgrove, D., Slavkovic, A., Rode, J., Verrati, T. and Chiselko, G., Using a large projection screen as an alternative to head-mounted displays for virtual environments, In CHI '00: *Proceedings of the SIGCHI conference on Human factors in computing systems* ACM, New York, pp.478--485, (2000).

Rizzo, A.A., Schultheis, M.T., Kerns, K.A. and Mateer, C., Analysis of assets for virtual reality applications in neuropsychology, Neuropsychol Rehab, 14, pp.207-239, (2004).

Rodieck, R.-W., *la vision* (2003)

Tan, D., Gergle, D., Scupelli, P. and Pausch, R., Physically Large Displays Improve Performance on Spatial Tasks, *ACM Transactions on Computer-Human Interaction* (ToCHI), 13, pp.71-99, 1, (2006).

Tyndiuk, F., Thomas, G., Lespinet-Najib, V. and Schlick, C. (2005), Cognitive comparison of 3D interaction in front of large vs. small displays, In *Proceedings of the ACM symposium on Virtual reality software and technology*, Edition. Monterey, CA, USA: ACM.

Werner, P., Rabinowitz, S., Klinger, E., Korczyn, A.S. and Josman, N., The use of the virtual action planning supermarket for the diagnosis of mild cognitive impairment, *Dementia and Geriatric Cognitive Disorders*, 27, pp.301-309, (2009).

Zanglonhi, X., Avital, L. and Prigent, N., Champ visuel et expertise, champ visuel et aptitude à la conduite, *Réflexions phtalmologiques*, 5, pp.32-36, (2000).

COGNITIVE ENGINEERING OF A MILITARY MULTI-MODAL MEMORY RESTRUCTURING SYSTEM

Willem-Paul Brinkman[1], Eric Vermetten[2, 3], Matthew van den Steen[1], and Mark A. Neerincx[1, 4]

Several methods have been proposed to treat combat veterans with Posttraumatic Stress Disorder (PTSD). Still, a recent review reports on high drop-out and non response rates. This has motivated work in the design of a software application to support and increase the appeal of traditional face-to-face trauma-focused psychotherapy. The research followed a situated cognitive engineering approach, which included a domain analysis, scenarios and claims analysis based on experts reviews (N=10), user evaluations, and a case study. This resulted in the identification of nine core application functions: (1) provide a flexible way of storytelling, (2) provide a structured way of storytelling, (3) prevent losing track of changed and added events, (4) ensure patient trust, (5) ensure usage for therapists with different backgrounds, (6) ensuring awareness of treatment, (7) provide a personal approach, (8) prevent unexpected exposure to emotional material, (9) and ensure an appealing and motivating approach throughout the therapy. These functions formed the basis for the design of a military multi-modal memory restructuring (Military - 3MR) system, which focuses on restructuring and relearning of past events. The system allows the patient and therapist to visualize past events using personal photos, narrative text, online geographical maps, webcam snapshots, and patient created 3-D virtual worlds. Results of the usability evaluation (N=18) suggest key design features, such as the time line, content management, and the 3-D world editor, meet an acceptable perceived usability level. Results of a storytelling experiment (N=18) between telling an autobiographical story with or without the Military-3MR system found that with the system, time referencing and event description were more precise, and a smaller time period in the story was covered. In the case study, the veteran suffering from combat-related PTSD was pleased with the system and felt encouraged talking about past events.

Keywords: PTSD, Trauma-focused Psychotherapy, Memory, Multimedia, Cognitive Engineering

Introduction

Warriors that served in combat or peacekeeping operations have often witnessed or experienced traumatic events, such as attacks, serious personal injuries, and the death of comrades or civilians, including children. On their return from deployment these warriors can suffer from Posttraumatic Stress Disorder (PTSD). For example, a survey (Hoge, et al., 2004) among members of four U.S. combat infantry units found that between 12-20% of them reported PTSD symptoms a few months after their duty in Iraq or Afghanistan. Similarly, in another survey (Milliken, Auchterlonie, & Hoge, 2007) among U.S. Iraq war veterans 17-25% exhibited PTSD symptoms three to six months after their return. Several methods have been proposed to treat patients with PTSD ranging from pharmacological approaches to Cognitive-Behavioral Treatments (CBT), with exposure therapy currently considered the first-line treatment for PTSD (Cukor, Spitalnick, Difede, Rizzo, & Rothbaum, 2009). Still, a recent review (Schottenbauer, Glass, Arnkoff,

Corresponding Author:
Willem-Paul Brinkman, Department of Mediamatics, Man-Machine Interaction, Delft University of Technology, Mekelweg 4, 2628 CD Delft, the Netherlands, Tel: +31 (0)15 2783534, Fax: +31 (0)15 2787141, E-mail: w.p.brinkman@tudelft.nl

[1]Delft University of Technology, Mekelweg 4, 2628 CD Delft, the Netherlands
[2]Military Mental Health Research, Central Military Hospital, Utrecht, The Netherlands
[3]Rudolf Magnus Institute of Neuroscience, University Medical Center, Utrecht, The Netherlands
[4]TNO, Soesterberg, The Netherlands

Tendick, & Gray, 2008) of 55 studies on empirically supported treatments for PTSD reported that drop-out rates and non response to the treatment are often high. While dropout rates ranged widely from 0-54%, a 50% non response rate was not uncommon to find. This has motivated the work reported here in the design of a military multi-modal memory restructuring (Military - 3MR) system, a software application to support and increase the appeal of traditional face-to-face therapy. The application especially focuses on helping the patient to restructure and relearn memories relating to events of their past deployment.

One of the symptoms of PTSD as indicated by DSM-IV-TR is the inability to recall an important aspect of the trauma. For example, Clark and Beck (2010) discussed the case of Edward, a 42-year-old man with 20 years of distinguished service in Canadian infantry who, after several deployments, was diagnosed with PTSD. Some of this time included traumatic events, especially during the experience of his 6-month United Nation peacekeeping Rwanda tour in 1994. One of these traumatic events that Edward experienced was the apparent murder of a 5-year-old orphaned girl and her friends by soldiers of the Rwandan Patriotic Army (RPA). Edward's beliefs were based on his last visit to the orphanage, when the children were no longer there. Instead, a RPA soldier was there, smiling and gesturing by sliding his hand across this throat. However, when they studied his recollection in depth, they noticed inconsistencies regarding other information such as: (1) there was no indication from the caretaking nuns that the children had been taken away and murdered, and (2) the incident would have taken place after the genocide had stopped when many children returned to their village. Reports on perceptual and memory distortion of high stress events, sometimes referred to as critical incident amnesia (Grossman & Siddle, 2001), are not uncommon. For example, a study of police officers involved in shooting incidents (Artwohl, 2002) reports the case of an officer that saw a suspect suddenly point a gun at his partner. As the officer shot the suspect, he saw his partner go down in a spray of blood. When the officer ran over to help his partner, he found him standing unharmed. In fact, the suspect never fired a shot. Critical incident debriefings have been suggested (Grossman & Christensen, 2008) as a vital method to reconstruct the events, learn from it, and help those involved to come to terms with it. As everyone involved is brought together, potential memory loss, memory distortion, irrational guilt, etc., can be addressed. The absence of such a debriefing can have a large impact on someone's life. For example, Grossman and Christensen discuss the case of Tim, a Vietnam veteran who, for over 20 years, never told anyone about an event in which he thought he had abandoned his man at an observation point when it was overran by North Vietnamese soldiers. Only after telling his story with pain and tears in his eyes to a group of West Point psychology students, Grossman explained to him that with his actions he had actually saved every man in his company, by ordering his man at the observation point to open fire and afterwards running back to the company's defensive position on a hilltop. Tim had never realized this. An important element of cognitive therapy for PTSD is therefore directed toward evaluating and restructuring the traumatic memory (Clark & Beck, 2010).

As is pointed out in the "collaborative cycles for the design and evaluation of technology for mental health interventions" (Coyle, Doherty, Matthews, & Sharry, 2007), research into the design of technology of talk-based mental health intervention often follows a two stage approach. Where stage one focuses on design and development of the technology, the second stage focuses on clinical evaluation. The work reported here should be seen as part of this first stage. In this stage, because of ethical constraints, access to patients is often limited, and the focus of this stage is to evaluate and establish confidence about the usability of the technology for the target end users group. Furthermore, to design a software application to support therapies for combat-related PTSD, the situated cognitive engineering (Neerincx & Lindenberg, 2008) approach was applied. Instead of working from a pre-set of requirements, this approach recognizes the complexity and situational dependence of a working domain, and the need for exploration of technological innovations to establish a set of testable claims that can drive a design solution. The work went through several iterations where the requirements baseline was continuously refined as new insights were acquired through prototype evaluations and reviews with therapists. The first iteration of this approach was to do a thorough domain analysis. This was done in multidisciplinary meetings with cognitive engineers and a psychiatrist experienced in treating veterans suffering from combat-related PTSD, which eventually led to the establishment of an inventory of human factor knowledge, operational demands, and envisioned technology. This knowledge was used to create several scenarios and prototypes. PTSD experts reviewed these scenarios and discussed various possibilities and limitations. In addition, cognitive engineering experts reviewed the prototypes on their usability. In a following iteration, the system was empirically tested with users on its usability and its ability to support storytelling. The actual use of the system was eventually studied in a small case study with a veteran suffering from a combat-related PTSD. The paper gives an account of all these iterations, and concludes with summarizing the contributions made by this study and discussing directions for future work.

Domain Analysis

Designing a useful mental health care system requires a multidisciplinary approach (Coyle, et al., 2007). Besides studying the literature, an important element of the domain analysis, therefore, were several follow-up brainstorming sessions between cognitive engineers and a military psychiatrist. An important patient characteristic which was found in this analysis and which would later play an important role in the design and implementation of the system was the availability of personal material. Many of the patients who were treated at the medical center kept personal photos, pictures and various documents related to their past dispatches at their homes. Although this material may not directly be related to a problematic stressor, it could still contain information necessary for the treatment or it could be used for reappraising the past dispatch as a whole.

Operational Demands

Several operational demands were identified as relevant for the design of the system. First, the patient's memories play an important role in addressing fears and stressors. A system should therefore support the therapist and patient in organizing various memory elements. Patients exposed to their memories of past events will remember more facts and details of the traumatic stressors (Foa & Kozak, 1986). Biographic material could be useful in accessing these memories. However, patients might not be in possession of this material; therefore, a system should also offer other facilities to support recollection. Second, there seem to be remarkable differences in the selected treatments between therapists and between patients. Some treatments only focus on imagery-based treatment, while others focus on cognitive reappraisal or restructuring. While a single system might not support all these different types, a system that would have features that could support various treatment methods would be more attractive for more therapists than those that only support one single method. Third, the setting of a session can vary from individual homework or individual face-to-face sessions with a therapist, to a group session with other patients. Next, patients in a group session do not necessarily have to have participated in the same deployment. Managing these different experiences in a group can be difficult. Currently, a flap-over was used in the therapy, but not so often. After a session the drawings would be lost and rarely mentioned again. Fourth, patients that better understand their treatment are less likely to terminate their treatment (Clark & Beck, 2010). Therefore, psycho-educational features in the system could make the progress in a treatment program more insightful.

Human Factors Knowledge

Although research on mental health systems often focuses on the patient issues such as efficacy and acceptability, the therapist's point of view also requires considerable attention addressing human factor issues as workload and usability (Brinkman, van der Mast, Sandino, Gunawan, & Emmelkamp, 2010) and the interaction between the patient, the therapist and the system (Wrzesien, Burkhardt, Alcañiz Raya, Botella, & Bretón López, 2010). As part of the domain analysis a People, Activity, Context, and Technology (PACT) analysis (Benyon, Turner, & Turner, 2005) was conducted that identified the following three main human factors: trust, emotion, and cognitive task load. As traumatic events might involve moral judgement or have legal consequences for those involved, treatment has to be set in an environment patients can trust. Because of the perception of stigma still associated with receiving mental health services among soldiers (Hoge, et al., 2004), patient privacy is an important issue, especially when considering using autobiographical material of the patients. If data is stored, patients should be aware that data is handled correctly. A technology solution should not endanger the patient-therapist relationship as this is a key factor of therapy (Rizzo & Kim, 2005). Trust also involves the reliability and confidence with which therapists use a system. As was reported (Brinkman, et al., 2010) in the case with a system to treat people with a fear of flying, therapists were less confident using a system if they fear that it would crash during a session and shatter patients' trust in their ability to control a situation.

Emotion was also identified as an important human factor as the autobiographical material could trigger memories of traumatic events. As these can be very upsetting for patients, unintended exposure should be avoided. Additionally, in a group session, some patients do not like to be stared at during exposure and others might feel un-comfortable looking at the person in a distressing state. Finally, additional cognitive tasks the system placed upon the therapist and patient should be limited. High task load can potentially hinder treatment; it would divert the therapist's attention away from the patient and the treatment. For the patients, it would divert their attention away from their memory exposure and restructuring activities.

Envisioned Technology

Considering the work and the success of others on virtual reality (VR) exposure therapy systems, such as virtual Vietnam environment (Rothbaum, 2006), Virtual Iraq environment (Pair, et al., 2006), Bus world (Josman, et al., 2006), World Trade Center environment (Difede, Cukor, Jayasinghe, & Hoffman, 2006) and the clinical results (Wood, Wiederhold, & Spira, 2010), the project initially set out to develop a virtual environment tailored for the Dutch combat and peacekeeping operations. However, when exploring various 3-D development environments in the first multidisciplinary meeting, it was realized that the very act of creating such a world by a patient might already be part of the treatment, as it might be a visual way of storytelling. For the treatment of PTSD the latter has been regarded as a helpful strategy to modify memories of past traumatic experiences, for example, in expressive writing (Clark & Beck, 2010), in creative therapies (Johnson, Lahad, & Gray, 2009) or Narrative Exposure Therapy (NET) (Bichescu, Neuner, Schauer, & Elbert, 2007). In NET the generally fragmented, gap-filled reports of the events are transformed into a coherent narrative. This can cause the habituation to emotional responses to reminders of these traumatic events. The memories are included in the patient's biography in a narrative form, facilitating the reorganization of the autobiographical memory and attribution of meaning. Therefore, the first idea was to solely create a 3-D virtual world editor, allowing the patients themselves to select 3-D objects, such as houses, tanks and other vehicles, and place them on an empty template. To support the group process, a projector would display the computer screen on the wall. This way a patient could explain to other group members what the situation looked like, what they experienced, and in which order specific events had occurred. In following multidisciplinary meetings where the first scenarios were discussed the idea evolved into an even more personalized support environment, allowing patients to use multi-media modalities by using, besides the 3-D world, their own pictures, and consequently, the need to structure media elements on a timeline to support storytelling of past experiences. The 3-D world editor initially remained the main storytelling feature.

Scenarios and Claims

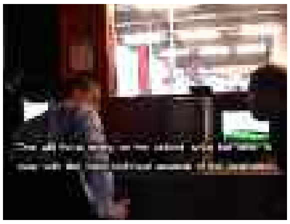
Scenario-based design (Carroll, 2000) emphasizes exploring the use of a system before it has actually been developed and this method has been successfully used in establishing design requirements for other mental health systems (Paping, Brinkman, & van der Mast, 2010). Therefore, a set of user scenarios was created to describe and discuss possible situations in which the envisioned application was presented. The gathered knowledge was later used to establish a preliminary requirements baseline. By creating these scenarios several assumptions had to be made explicit. The assumptions were linked to possible effects on the involved actors and were therefore important to analyze. A claims analysis was done with experts to reflect on possible effects as either advantages or disadvantages, and to establish a more elaborate understanding of the claims and assumptions without already verifying them empirically.

Use Scenarios

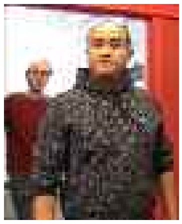
The scenarios described three possible situations in a therapeutic setting in which the system could be used: (1) the general use of the system showing all the major features of the application, (2) the use of the 3-D editor as a feature of the system, and (3) modifying or extending data related to an event discussed in a previous session. To explain these scenarios better, they were transformed into three separate movie clips showing actors using a preliminary low-fidelity prototype of the envisioned application.

Scenario A: Start of a Session and General Use of the Application

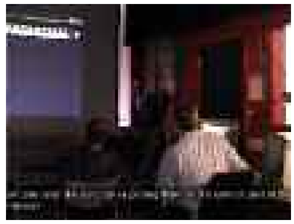
Scenario A showed general uses of the application and introduced the idea of personalization, a timeline, and a progress bar. This scenario started with two therapists preparing the group session (Figure 1a). The therapists had different roles during the session; one led the session and interacted with the patient, while the other operated the system. The session started with focusing on one patient and his experience (Figure 1c). The therapist asked the patient to place the pictures he had brought on the timeline and explain what could be seen on the pictures. Two main claims examined in this scenario were that (1) adding personal photographs and text was sufficient enough for a patient to restructure and relearn past deployment events; and (2) the usage of the timeline would help the patient to organize various memory elements.



(a) Two therapists: one focuses solely on the patient, the other on the technical aspects of the application



(b) Group session starts and the patients enter the room



(c) “Can you label the pictures by putting them on the timeline?”

(d) “All right, what do you still remember of that specific day?”

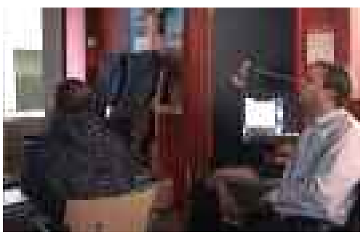
Figure 1. **Stills of the scenario describing the general use of the application.**

(a) “Let’s now go to the fifth of October. I read there were some firefights that day. Can you show me what happened exactly and what the area looked like?”

(b) Patient Guntur starts adding buildings and other structures, while telling what he had seen that day

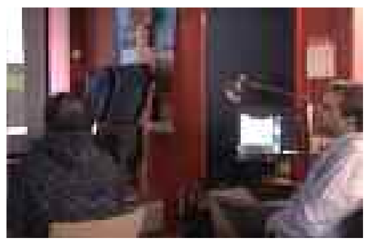
(c) “And you were standing here. Right Ingmar?”

Figure 2. **Stills of the scenario describing the 3-D editor.**



(a) “It has been a week since we last met, let us continue by going to the 20th of October”

(b) “Uhmm, yeah, we can continue but I’d like to go back to the day we edited last time...”



(c) “I was worried and actually couldn’t sleep last week because I didn’t tell you all the details.”

Figure 3. **Stills of the scenario describing the modification of a previously discussed day.**

Scenario B: Using the 3-D Editor

Scenario B described how a patient used a simple version of the Armed Assault 1 editor to recreate an event in the 3-D environment (Figure 2.b). The scenario showed that both group members and the therapist could interrupt the patient when recreating the event, asking him questions or giving their view on his story and the situation (Figure 2.c). The main claims examined here were that a patient could create such a world, and that editing work can be paused, and a patient can talk, or be calmed down by a therapist.

Scenario C: Returning to a previously edited day

The last scenario showed that a patient could continue to work on their story in follow up sessions (Figure 3). It also showed that patients could go back to what they discussed in previous sessions and add additional details. In the scenario the patient reveals that before the fire fight he had not told his colleague, who was shot, that he had seen someone behind the structure. In the previous session the patient had said that they did not know where the shooting was coming from.

Use Scenarios Evaluation

The three scenarios were used in individual expert reviews with 10 experts, which included therapists working with soldiers, war veterans, and civilians with war traumas, as well academics specialized in the treatment of traumatic memory, PTSD and grief. Six were non-Dutch experts, coming from Croatia, the UK, the US, and Poland, attending the NATO Advanced Research Workshop entitled "Wounds of War: Coping with Posttraumatic Stress in Returning Troops" that was held in October 2009 in Austria. The experts used several approaches such psycho-dynamics, imaginary exposure, cognitive behavior therapy and VR exposure. Some worked often in group settings, while others treated patients in one-on-one sessions. During the reviews, each expert was shown the video of the scenario, and asked to comment on the underlying claims to start the discussion.

The overall feedback from the experts was positive and the general idea underlying the envisioned approach was supported. The acquired comments were mainly related to features which could improve the system, including additional options to facilitate memory content, such as adding maps and photos of drawings. Other suggestions, which also resulted in refinements to the requirements baseline, were (1) the possibilities to personalize the application for an individual patient, (2) the use of keywords to tag or summarize a specific day and (3) methods to manage and store data and session information. Although not presented in the scenario, the latter suggestion came up multiple times during the reviews. Some mention that soldiers might not seek treatment because of the fear of stigmatization and the potential consequences. Not all experts supported the use of the 3-D editor, especially experts with a background in psychodynamic and group therapy without exposure. Comments were also made on the set up of the room. In the scenarios the therapist was placed slightly behind the patients. Instead, some suggested the therapist should be positioned to see the face of the patients and other non-verbal signals. Next, as only the patient that is working with the system needs a table, other group members only need a chair, which should be situated in the shape of a circle or the letter "U," creating the feeling of a "safe zone."

Design

The results of the domain analysis and scenario-based evaluation led to a list of nine core functions of the Military-3MR system, which formed the requirement baseline for the design and implementation of the system. The design and the evaluation process, reported later, also contributed to the final list shown in Table 1. An iterative approach was followed for the design of the system. Several prototypes were made and reviewed in the meetings of the multidisciplinary team. The storyboard, shown in Figure 4, gives an overview of specific features of the final system and how they are used. The actual system was implemented using two development packages: (1) Adobe Integrated Runtime (AIR) for the main application, and (2) Vizard for the 3-D world editor feature.

Since ensuring patient trust was identified as a core function, the system supported three types of installations set up: (1) full USB-stick installation, (2) full therapist computer installation, and (3) photo separation installation. In the first installation the software application is installed on a USB stick (or another external memory device), and a folder is created on the USB stick where photos can saved prior to a session. During the session, the application is run from the USB stick and internal files such as text elements and snapshots of maps are also stored on the USB stick. After the session, the USB stick can again be taken out of the computer. The USB stick could be locked up in a safe place or given to the patient. If the therapist thinks it is appropriate, the patient can also be given a homework assignment involving working on the timeline at home in preparation of the next session. The full therapist computer installation, on the other hand, ensures that the application is only used in the session with the therapist as the application and autobiographical material is stored on the computer of the therapist. A folder has to be created on the therapist's computer to copy the photos brought by the patient. An alternative would be to install everything on the patients' laptop on which their autobiographical material is stored. In this way, patients are in complete control of the access to their material which might enhance their trust. The last installation type is a combination of the previous two. The application is installed on the therapist's computer, but the photo material is kept in a folder on a USB stick belonging to the patient. Only when the USB stick is plugged into the therapist's computer will the photos be accessible for the application. This installation does not allow for the patient to complete unsupervised work on their timeline, but gives the patient assurance that they are still in control of their own photographic material.

Timeline

Patients' intrusive recollection of traumatic events is, among other things, caused by their fragmented and

Table 1
Core functions of the Military-3MR system and their underlying claims

No.	Core Function
1	*Provide a flexible way of storytelling*. Restructuring and relearning about a past deployment events can have a positive effect in traditional treatment. Only the patient has access to the various memory elements related to the events and stressors at that time. More efficient and effective results may be achieved by allowing the patient to facilitate memory content in a flexible and motivating way.
2	*Provide a structured way of storytelling*. Deployments usually cover multiple months. Within these months various events occur. Next to giving the patient the tools to flexibly add and edit memory elements, it might also be important to do this in a more structured environment.
3	*Prevent losing track of changed and added events*. Currently a physical folder is kept of every patient, containing notes, drawings and other deployment related information. These files can get rather large and it is difficult to keep track of the continuous changes which may occur. Especially if a therapist works with multiple different patients every week.
4	*Ensure patient trust*. Deployment related stories can contain very sensitive information. This data should therefore be stored securely. Trust might be another problem if the storage of private data is not properly implemented. The look and feel of the system can also be important when establishing trust.
5	*Ensure usage for therapists with different backgrounds*. Therapists use several methods to treat patients suffering from a combat-related PTSD. If the system does not force the patients to follow one specific procedure, while providing several tools to facilitate memory, the system might be interesting for a variety of different therapists and methods.
6	*Ensuring awareness of treatment*. Besides the restructuring and relearning element, a psycho-educational element can be introduced to give the patient more insights about the treatment in general. Details of sessions, including information about recent changes, future goals and accomplished goals may help patients learn and cope with their disorder.
7	*Provide a personal approach*. The patient should feel comfortable using the application. Personal preferences should be respected, making the application look and feel like a helpful companion throughout the treatment.
8	*Prevent unexpected exposure to emotional material*. During the first review with an actual patient, one of the suggestions was to implement an option to 'tag' photographs that are linked to very emotional situations. For example, a patient may not like the idea of being greeted with a very emotional memory when starting the application or navigating through the timeline.
9	*Ensure appealing and motivating approach throughout the therapy*. One of the problems with current treatment is the high drop-out rates. Keeping the patient motivated and aware of the progress may prevent the patient to quit the therapy too early.

poorly elaborated memory structure of these events (Clark & Beck, 2010). Providing a structured way of storytelling was therefore identified as another core function. A key element in the Military-3MR system to support this is the timeline which is a chronological visual representation of the events in the patient's life. This idea is not entirely new. For example, the life-chart method (Osuch, et al., 2001) provides a graphical presentation of events in patients' lives and their treatment interventions with an aim of visualizing the longitudinal course of the illness. In the Military-3MR system the purpose of the timeline, however, is (1) to help patients to structure their memory, (2) to enhance their awareness of the relatively small time scale of the traumatic event compared to their entire life, and (3) to enhance their awareness that the traumatic events that dominate their current life all happened in the past. The overall timeline is presented at the top of the screen (Figure 5), and is referred to as the lifeline of the patient. It shows which year the patient was born, the times of deployment, and current and future events. When a deployment is selected (Figure 4 step 5), the indicator moves back to the year of the deployment. As the years are counted

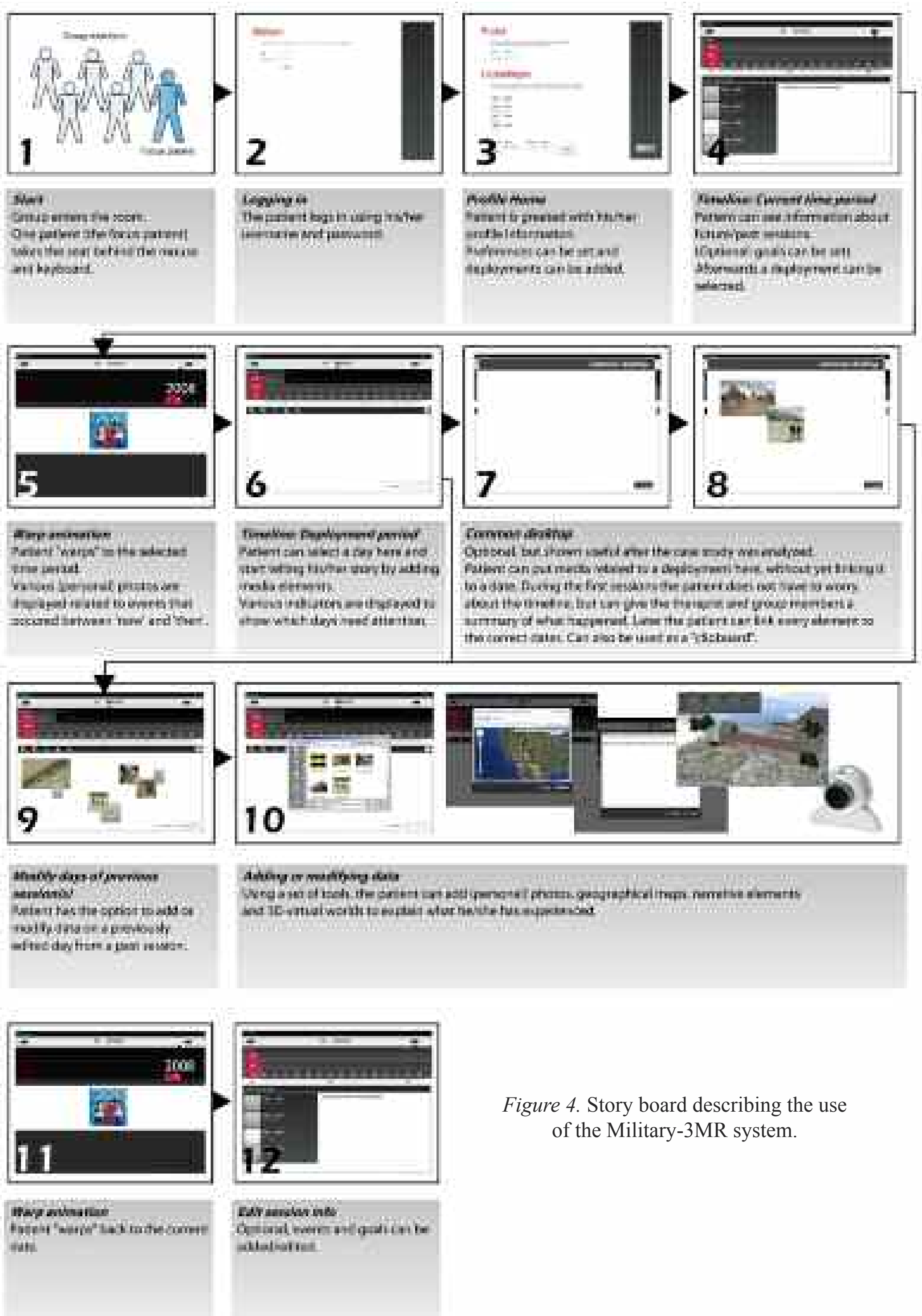


Figure 4. Story board describing the use of the Military-3MR system.

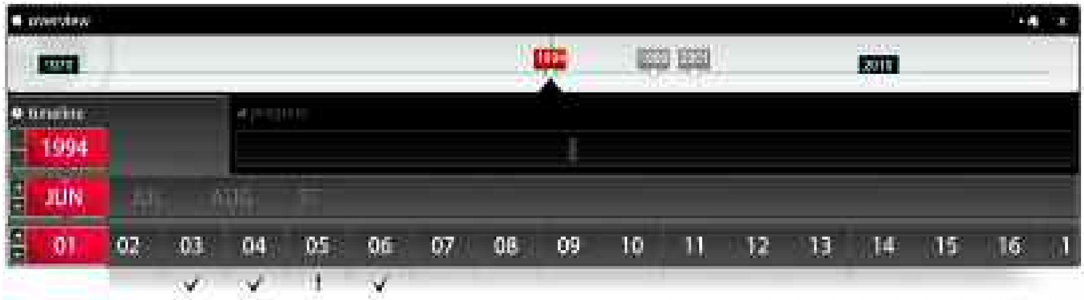


Figure 5. Timeline.

back, photos are shown of a memorable event for each year. To emphasize further that the deployment occurred in the past, the photos in the warp animation start in full color, but become grayer as they date further in the past. The timeline slider (Figure 5) allows patients to select a specific date. The symbols below the days indicate whether a discussion of that day has been completed or is still in progress. To avoid the patient only focusing on past events, the application also supports the creation of events in the present or the future. Here, the patient can be encouraged to talk about the present and set goals for the future.

Photo, Text, Maps, and Webcam

The system provides five different features to facilitate recall of memories. The patient has the option to add (1) images or photographs, (2) geographic maps using Google Maps, (3) story text, (4) webcam snapshots and (5) virtual representations of a situation using the 3-D editor. When the patient adds an image, a webcam snapshot or a map, the system creates an interactive thumbnail and puts this thumbnail on the content panel. Clicking on these thumbnails gives the patient the option to display the picture on the full size screen, to remove it, and to hide or unhide it. The hiding option avoids unwanted exposure of emotional material. After the case study, a common desktop function was also added. This is a clipboard on which photos and text can be placed temporarily, and copied to the content panel of specific days. With non-recent deployments, patients might not own digital photos, but only printed photos of their deployment. In these cases, the patient can digitize these photos by taking a snapshot of them with the webcam. The webcam can also be used to take pictures of drawings made in the session, or documents or objects brought in by the patient. It could also be used to make pictures of recreated situations with physical objects. For example, Playmobil has been used to successfully recreate accident and to study emergency responses (Gunawan, Ooms, Neerincx, Brinkman, & Alers, 2009).

3-D World Editor

The 3-D world editor allowed the patient to create a simple 3-D world by selecting objects such as buildings, trees and vehicles from a menu. Besides placing and rotating the objects, the patient can also change the camera position using the mouse and some keyboard keys. The patients can therefore look at the world from different angles (Figure 6) to help them reflect on the event from multiple perspectives.

Figure 6. 3-D world editor.

Evaluation

Besides the expert reviews of the system design and its underlying claims as discussed previously, another core activity of the situated cognitive engineering approach is the evaluation of the system by the users. In this study, the evaluation of the system focused on three issues: (1) the usability of the system, (2) the effect of the system on storytelling, and (3) the practical use of the system in the field. All these issues were empirically evaluated. While only the last issue was evaluated with an actual veteran suffering from combat-related PTSD, the other two issues were evaluated with participants that did not suffer from combat-related PTSD.

Usability Analysis

The usability of the system was evaluated using both an analytical method, i.e. heuristic evaluation, which involves designers to critically analyze the design, and an empirical method, i.e. a usability test, which involves testing the system with users. This approach allowed potential generic usability problems to be identified and gave direct input on how potential users would use and perceive the usability of the system.

Heuristic Evaluation

Several prototypes and evaluations were conducted before the final Military-3MR system was developed. Once the first set of core functions and requirements were defined, a first high-fidelity prototype was created. The ten principles for user interface design (Jakob & Rolf, 1990), also known as the ten heuristics, were used to obtain feedback on this high-fidelity prototype. Six MSc students, trained in the field of cognitive engineering, were all asked to individually complete a form containing several questions related to the ten heuristics. Later, a part of this group participated in a follow-up group evaluation that aimed at identifying possible usability issues present in the high-fidelity prototype. The obtained feedback from both evaluations mainly concerned: (1) the state of the system, which was not clear, (2) system behavior, as it did not follow common Graphical User Interface standards, and (3) clarity of the icons and buttons.

A formative usability evaluation approach was chosen to address these issues and to continuously acquire more feedback of the rapidly generated prototypes in a period of four weeks. During this phase various elements of each prototype were inspected by the cognitive engineering team, which included the six students and two lecturers. During this time multidisciplinary meetings were also held to improve the requirements baseline, and a small interview was conducted with a patient suffering from combat-related PTSD to review the prototype. The reviews and the small interview resulted in two of the core functions (Table 1): providing a personal approach and preventing unexpected exposure. Eventually, the evaluation reached a point where no usability issues were reported anymore. This last prototype, referred to as the final Military-3MR system, was now considered suitable to be used in a usability test with users and a case study with a veteran suffering from combat-related PTSD.

Usability Test

The evaluations of the previous prototypes led to several refinements regarding both the usability and the functionality of the system. With the final prototype finished, a user-based usability test was conducted in which 18 individuals (12 males, six females) participated, none suffering from a combat-related PTSD. The age varied between 21 and 59 (M = 36.2, SD = 15.1) years old. The participants had never worked with the final prototype before and none were given instructions beforehand. To evaluate the perceived usability of timeline navigation, content manager, and the 3-D editor, a component-based usability questionnaire (Brinkman, Haakma, & Bouwhuis, 2009) was used, as component-specific measures have been reported (Brinkman, Haakma, & Bouwhuis, 2008) to be, on average, statistically more powerful than overall usability measures. The questionnaire asked participants to rate the usability of each of these interaction components on six statements. Before participants were asked to complete the questionnaire they were asked to complete a number of tasks: select a specific deployment, navigate through the timeline to pick a date and eventually add, modify and delete content elements. In addition, the participants were asked to create a virtual world based on a screenshot provided to them. Participants were not provided with any additional instructions or help material.

The individual component measures, each based on six 7-point Likert scale statements, received an acceptable level of reliability with Cronbach's alpha ranging from 0.73 to 0.81 (Table 2). The average of the scores on these six items was therefore taken as a usability measure for the component. One-Sample t-tests showed that all three measures deviated significantly from the norm value of 5.29 (Brinkman, et al., 2009). As shown in Table 2, all means were above the norm value, suggesting that the perceived usability of these components were more comparable with the perceived usability of easy-to-use components in the norm set than the perceived usability of difficult to use components in the norm set. The questionnaires also allowed participants to add additional comments and suggestions. One reoccurring issue was that of the “maps” icon; a large number of participants thought that the “maps” icon was actually the icon to open up an Internet browser. Because there were not many other options, everyone still managed to add the map, but some found this a bit confusing. A suggestion made by multiple participants was to add the ability to drag and drop objects in the 3-D world editor. No other usability issue was found. All participants stated that they were very pleased with how the application worked, and most of them also mentioned that the user interface looked appealing.

Table 2
Result reliability analysis and One-Sample t-test with test value 5.29 on component-based usability questionnaire data

Component	***α***	***M (SD)***	***t***	***df***	***p.***
Timeline	0.81	6.37 (0.43)	10.58	17	<0.001
Content manager	0.70	6.08 (0.45)	7.37	17	<0.001
3D world editor	0.73	5.62 (0.57)	2.44	17	0.026

Storytelling Analysis

After participants in the usability test filled out the component-based usability questionnaire they were asked to tell two different autobiographic stories in six minutes each. One story they were asked to tell using the Military-3MR system, the other story they were asked to tell without the system. To control for possible learning or fatigue effects, the order of these two conditions was counterbalanced. The events, for example, a holiday or business trip, should have taken place at least three years in the past. The order between the time of the events was also counterbalanced, i.e. half the participants started with describing their more recent event, while the other half started with describing their less recent event. While speaking, they could use photos they had brought with them specifically for the experiment. At the start of the session, the experimenter selected only a small set of photos they could use to reduce the chance that prior to the experiment participants could have mentally rehearsed a specific story line.

During the storytelling, voice recordings were made of both stories, resulting in a total of 36 different audio clips. Each sentence was coded for five non-mutually exclusive categories (Table 3) reference to (1) time, (2) an event or (3) a location. All references in each story were counted so they could be used for a statistical analysis. Only 24 recordings were used in this analysis, as six participants used the text feature to write about their stories without saying what they were typing. These voice recordings did not contain any relevant information and could therefore not be used. For each story the number of references for each category was counted, and the difference between the stories told with or without the system was analyzed by Wilcoxon Signed-ranks tests. As Table 3 shows, participants made a significantly higher number of precise date references and were more specific concerning the time frame when they used the system. The system, therefore, seems to have encouraged participants to place their story in a specific time. This was also supported by the analysis of the number of months covered in a story. With the system ($Mdn = 0.13$) this was significantly ($Z = -2.43$, $p. = 0.015$) shorter than without the system ($Mdn = 0.5$). Once they picked a starting date for their story, they seemed to continue telling more about this day by adding events. Furthermore, they seemed to talk about days that were close together. This was not the case when participants did not use the system. In some cases these stories even covered several months. Table 3 also shows that participants made a higher number of detailed event descriptions and a lower number of more general event descriptions when using the system. This seems to suggest that the system encouraged participants to be more specific in describing the events in their story. Although the online geographical maps functionality was expected to also have an effect on the number of references made to specific locations, no significant difference was found. This, however, might have been an artefact of the story topic e.g. a holiday or business trip, where the location and the geographical setting might be less relevant than for events in a combat situation.

Another observation was the individual difference between participants in how they told their story with the system. Although it might be caused by the six-minute

Table 3
Results Wilcoxon signed-rank test on sentence analysis for story told with or without the system

No	Category	*Median frequency*		Z	p.
		With system	Without System		
1a	*Precise date reference*; e.g. "12th of March"	2.5	0.0	-2.96	0.003
1b	*Less precise time frame*; e.g. "the summer holidays" and "my second vacation that year"	1.5	4.0	-2.77	0.006
2a	*Detailed event description*; e.g. "buying a cola at the supermarket", "picking up the phone", and "getting in the car"	4.0	3.0	-2.56	0.011
2b	*General event description*; e.g. "going on a business trip", "leaving town", and "taking care of the pets".	1.5	3.0	-2.46	0.014
3	*Reference to location*; e.g. "the bank in the same street as the hotel we resided", "...I saw someone in my kitchen...", "...so we moved my furniture to the Weimarstraat..." , "Right before the small town which name I've forgotten..."	4.0	3.5	-1.51	0.130

time constraint, some participants only used the text feature to write about past events and what they saw, even when they brought a photo book with them containing pictures which may have supported the storytelling. However, others were very dependent on the images and photographs. Two participants did not even mention any events that were not found in the pictures.

After the storytelling session, participants were also asked to rate statements on a 7-point Likert scale about their attitude towards the system's ability to support storytelling. Table 4 shows the results of One-sample *t*-tests comparing the score with the middle point of the scale, e.g. four. Overall, the participants' attitude seems to lean towards the positive side of the scale, in favor of the system to support storytelling. Interestingly, the participants did not seem to agree with the statement that the system encouraged them to add media. This could be interpreted to mean that they felt in control of their storytelling and did not feel dominated by the system or under pressure to add content. Participants seemed to find that they were also significantly ($t(17) = 4.91$, $p. < 0.001$) more able to discuss everything they wanted to when using system than compared to not using the system.

Case Study with a Veteran

The usability analysis and storytelling analysis was conducted with people not suffering from a combat-related PTSD. However, actual experience with a system by a person actually suffering from this disorder might be different. Therefore, a small case study was conducted to gather insight on how a patient would interact and experience the system. The patient that participated in the case study was a Dutch veteran that had experienced stressful events during his deployments in the Former Yugoslavia: 1994 - Dutchbat I stationed Srebrenica; 1999- SFOR 6 operating in Bosnia; and 2001 - Task Force Harvest operating in Macedonia. The patient was already receiving outpatient PTSD treatment at the time of the case study. In January 2010 two sessions were observed that took place in the office of a psychiatrist in Central Military Hospital, in Utrecht, The Netherlands. Prior to the session, the psychiatrist informed the patient about ideas underlying the design of the system, the purpose of the case study, and the presence of an observer during the sessions. Figure 7 shows the office where the session took place and the set-up of the system, which included a projector, a laptop on which the application ran, and a cell phone with a Bluetooth which facilitated the Internet connection needed to access the online geographical maps. The patient was located behind the laptop and the psychiatrist was sitting to the right of him, both facing the wall on which the application was projected.

Table 4
Result One-Sample t-tests, with test value 4, on attitude questionnaire towards storytelling support

No	Statement	*M(SD)*	*t*	*df*	*p.*
1	I prefer a timeline, such as the one in the application, to tell my story	5.50 (1.30)	4.92	17	<0.001
2	I found the story telling with the application more enjoyable	5.28 (1.27)	4.25	17	0.001
3	I was able to tell everything I wanted (story telling **without** the application)	3.78 (1.59)	-0.59	17	0.562
4	I was able to tell everything I wanted (story telling **with** the application)	5.50 (1.25)	5.10	17	<0.001
5	I think I put more details in the story told with the application	6.00 (0.67)	12.37	17	<0.001
6	By using the application, I was encouraged to add media such as text, maps or pictures to explain things better	4.39 (1.79)	0.92	17	0.369
7	Comparing the two stories, I found that more memories came back with the application	5.94 (0.80)	10.28	17	<0.001

The first session took about an hour, and started with a few minutes introducing the system and allowing the patient to explore the system. Afterwards, the patient started with creating his profile including his deployments. Once this was done the psychiatrist decided to work on the first 1994 deployment. Although the patient understood the idea of the warp animation, he suggested the option to switch this feature off since it would be annoying to see it during each session. After the timeline was set at the starting day of the first deployment, the patient loaded all the images that he had brought on a USB stick on to this first day, with the intention to sort the images later to occupy specific days. In the reflection discussion after the session, both the patient and psychiatrist agreed on the need for a common desktop function to temporarily store images, to copy them from here to specific days. After the images were uploaded onto the screen, the psychiatrist asked the patient to discuss these photos in detail and suggested that he add text elements to them. As the psychiatrist entered the text, the patient was able to read back his comments on the screen and added details. Besides the photos, the webcam feature was also used to take a snapshot of a situation drawing made in a previous session with the patient, and again text was added to this snapshot. Often, the patient also used the online geographical map function to look up specific places and buildings he remembered from his deployment. Again, the psychiatrist encouraged the patient to explain his recollections in detail. The patient talked about his experiences and also how things had changed since then. Access to the online geographical map took a relatively long time because of the limited Internet connection. Interestingly, the download waiting time was used to let the patient tell more about the event that had happened.

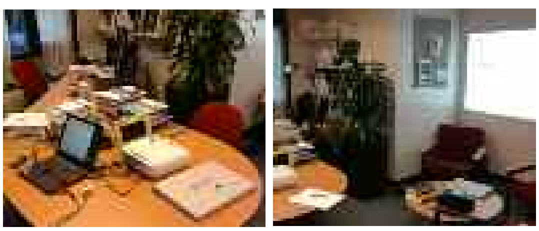


Figure 7. Office and set up of the Military-3MR system.

A week later, the second session was carried out, which lasted around 40 minutes. It started with a short reflection about the first session. The patient pointed out that talking and presenting his memories in this visual way was very appealing, and he considered it useful for him. He felt that this application helped him to become aware of events that had happened, and he felt in control of his own treatment, working on his own timeline. The patient also suggested that the year indicators of each deployment on the timeline should be easy to click on to allow moving between events that occurred at different points in time. This was not implemented on purpose, to focus the discussion at one deployment at a time, and to avoid memories from different deployments from becoming mixed. Table 5 shows a summary of his comments. After this reflection, the remaining time of the second session was devoted to discussing a document the patient had brought with him. This document was related to the events discussed in the first session, and the patient thought that this would help with specifying events on the timeline discussed in the previous session. The main interaction in the second session was directly between the patient and the psychiatrist without the use of the system.

Table 5
Summary of patient feedback on the use of Military-3MR system

No	Feedback
1	The application makes you aware of the time-related events
2	It is positive that you working on your own deployment, you are more in control to add story elements yourself
3	Normally everything is done mainly verbal, now treatment is done with visual aids, which more appealing
4	Might also be good for group session, but also for individual session
5	Going back to these past events is painful, but it is still important to do so

Although the 3-D editor was not used during the sessions, both the psychiatrist and patient were hopeful that this feature would contribute to elaborating on complex, stressful events. Later on, this was also confirmed in a review by a US soldier who had just returned from Iraq. This soldier was a behavioral health specialist stationed in the US Tripler Army Medical Center in Hawaii. After reviewing the Military-4MR system he wrote down that with the help of the 3-D editor soldiers would be able to draw a clearer image for the therapist. He thought that this would be beneficial for both therapists and patients since therapists would no longer need their own imagination to try and recreate what patients were trying to explain. With this clearer understanding, therapists would be able to help patients more effectively. Interestingly, his comments also related to the need for personalization of these worlds, as he wrote down that the 3-D editor currently provided too few objects to recreate events in Iraq, missing objects such as Humvees (high mobility mutipurpose wheeled vehicles) and Mine Resistant Ambush Protected Vehicles (MRAPs), and the style of buildings typically found in Iraq. Overall, he considered the 4MR-system a potential helpful tool for soldiers that have been avoiding material that reminds them of their deployments. The system could also help to identify factors that upset a specific patient if this was unclear beforehand.

Conclusion and Discussion

Several conclusions can be drawn from the presented study. Firstly, after several design and review cycles, the Military-3MR system now seems to have established an acceptable usability level without major usability problems, and with key design features such as timeline, content management, and the 3-D editor being received as easy-to-use. Second, the system seems to facilitate storytelling as time referencing and event description become more precise, and smaller time periods are covered. Third, the case study with the veteran showed that he enjoyed working with the system, giving him a sense of control and a visually appealing way of talking and reflecting on past experiences. Besides these points, the nine core functions identified are put forward as scientific contributions. A practical limitation of the current system is the use of the Internet for online geographical maps, or the use of personal files, such as digital photos. Due to security risks involved, military environments often limit the use of the Internet and do not allow USB sticks to be plugged into military computers. At the moment, a solution could be to use a standalone computer that is not connected to the military network, but instead connected to a public Internet provider using a mobile phone or USB Internet dongle. Another limitation is the limited involvement of patients in this study. No general statement can yet be made about the response of a treatment using the Military-3MR system. An extensive randomized controlled trail would be needed for this, with a validated treatment protocol.

In the case study, no use was made of the 3-D editor feature. Probably, the effort involved in creating a 3-D world suggests that this feature should mainly be used to recreate specific traumatic and stressful events to reflect on the patient's recollection of these events. Future work might also look at the possibility of combining the Military-3MR system and a Virtual Reality Exposure Therapy (VRET) system, whereby a VRET system could be used to expose patients to generic 3-D worlds, but also to the highly personalized 3-D worlds created by patients in the Military-3MR system. Exploring the use of generic material might also include using generic picture databases of combat and peacekeeping operations, or using historical event databases such as online newspapers databases or Wikipedia. This might help the patient to remember personal events and place them in the context of major news events of that day. Instead of relating to a specific date, generic material could also include music that was popular at the time of the deployment or iconic video material. For example, historic video material of Saddam Hussein, as a US army clinical psychologist at Tripler Army Medical Center wrote down in his review.

Although the Military-3MR system was designed for the treatment of patients with combat-related PTSD, the identified core functions and possibly parts of the Military-3MR system might also be considered relevant for the treatment of non combat-related PTSD patients suffering from traumatic events that stretch over a longer period of time, for example, civilians in war situations, child abuse, or rescue workers that respond to natural disasters. In these events and in a military context the 3MR system might help patients to restructure and relearn their traumatic events as part of their treatment for PTSD using autobiographic multi-modal memory cues.

Acknowledgement
We would like to thank the veterans, soldiers, and PTSD experts that helped review the system, especially Major Melba Stetz of the US Tripler Army Medical Center.

References

Artwohl, A. (2002). Perceptual and memory distortions in officer involved shootings. *FBI Law enforcement bullentin, 71*(10), 18-24.

Benyon, D., Turner, P., & Turner, S. (2005). *Designing interactive systems : people, activities, contexts, technologies.* Harlow, England ; New York: Addison-Wesley.

Bichescu, D., Neuner, F., Schauer, M., & Elbert, T. (2007). Narrative exposure therapy for political imprisonment-related chronic posttraumatic stress disorder and depression. *Behaviour Research and Therapy, 45*(9), 2212-2220.

Brinkman, W. P., Haakma, R., & Bouwhuis, D. G. (2008). Component-specific usability testing. *Ieee Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 38*(5), 1143-1155.

Brinkman, W. P., Haakma, R., & Bouwhuis, D. G. (2009). The theoretical foundation and validity of a component-based usability questionnaire. *Behaviour & Information Technology, 28*(2), 121-137.

Brinkman, W. P., van der Mast, C., Sandino, G., Gunawan, L. T., & Emmelkamp, P. M. G. (2010). The therapist user interface of a virtual reality exposure therapy system in the treatment of fear of flying. *Interacting with Computers, 22*(4), 299-310.

Carroll, J. M. (2000). *Making use : scenario-based design of human-computer interactions.* Cambridge, Mass.; London: MIT Press.

Clark, D. A., & Beck, A. T. (2010). *Cognitive therapy of anxiety disorders: science and practice.* New York ; London: Guilford.

Coyle, D., Doherty, G., Matthews, M., & Sharry, J. (2007). Computers in talk-based mental health interventions. *Interacting with Computers, 19*(4), 545-562.

Cukor, J., Spitalnick, J., Difede, J., Rizzo, A., & Rothbaum, B. O. (2009). Emerging treatments for PTSD. *Clinical Psychology Review, 29*(8), 715-726.

Difede, J., Cukor, J., Jayasinghe, N., & Hoffman, H. (2006). Developing a virtual reality treatment protocol for posttramatic stress disorder following the world trade center attack. In M. J. Roy (Ed.), *Novel approaches to the diagnosis and treatment of posttraumatic stress disorder* (pp. 219-234). Amsterdam, The Netherlands: IOS Press.

Foa, E. B., & Kozak, M. J. (1986). Emotional Processing of Fear - Exposure to Corrective Information. *Psychological Bulletin, 99*(1), 20-35.

Grossman, D., & Christensen, W. (2008). *On combat: The psychology and physiology of deadly conflict in war and in peace* (3rd ed.): Warrior science publication.

Grossman, D., & Siddle, B. K. (2001). Critical incident amnesia: the psysiological basis and the implications of memory loss during extreme survival stress situations. *The firearms instructor: the official journal of the international association of law enforcement firearms instructors*(31).

Gunawan, L. T., Ooms, A. H. J., Neerincx, M., Brinkman, W.-P., & Alers, H. (2009). Collaborative situational mapping during emergency response. *Proceedings of ECCE2009* (pp. 85-91).

Hoge, C. W., Castro, C. A., Messer, S. C., McGurk, D., Cotting, D. I., & Koffman, R. L. (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *New England Journal of Medicine, 351*(1), 13-22.

Jakob, N., & Rolf, M. (1990). *Heuristic evaluation of user interfaces.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people.

Johnson, D. R., Lahad, M., & Gray, A. (2009). Creative therapies for adults. In E. B. Foa, T. M. Keane, M. J. Friedman & J. A. Cohen (Eds.), *Effective treatment for PTSD: practice guidelines from the international society for traumatic stress studies* (2nd ed., pp. 479 - 490). New York: Guilford.

Josman, N., Somer, E., Reisberg, A., Weiss, P. T., Garcia-Palacios, A., & offman, H. (2006). BusWorld: Designing a Virtual Environment for Post-Traumatic Stress Disorder in Israel: A Protocol. *CyberPsychology & Behaviour, 9*(2), 241-244.

Milliken, C. S., Auchterlonie, J. L., & Hoge, C. W. (2007). Longitudinal assessment of mental health problems among active and reserve component soldiers returning from the Iraq war. *Jama-Journal of the American Medical Association, 298*(18), 2141-2148.

Neerincx, M. A., & Lindenberg, J. (2008). Situated cognitive engineering for complex tasks environments. In J. M. C. Schraagen, L. Militello, T. Ormerod & R. Lipshitz (Eds.), *Naturalistic decision making and macrocognition* (pp. 373-390). Aldershot, UK: Ashgate.

Osuch, E. A., Brotman, M. A., Podell, D., Geraci, M., Touzeau, P. L., Leverich, G. S., et al. (2001). Prospective and retrospective life-charting in posttraumatic stress disorder (The PTSD-LCM): A pilot study. *Journal of Traumatic Stress, 14*(1), 229-239.

Pair, J., Allen, B., Dautricourt, M., Treskunov, A., Liewer, M., Graap, K., et al. (2006). Virtual Reality Exposure Therapy Application for Iraq War Post Traumatic Stress Disorder. *Virtual Reality Conference* (pp. 67-72).

Paping, C., Brinkman, W.-P., & van der Mast, C. (2010). An explorative study into a tele-delivered multi-patient virtual reality exposure therapy system. In B. K. Wiederhold (Ed.), *Coping with post-traumatic stress disorder in returning troops: Wounds of War II* (pp. 203-219). Amsterdam, The Netherlands: IOS Press.

Rizzo, A., & Kim, G. J. (2005). A SWOT analysis of the field of virtual reality rehabilitation and therapy. *Presence-Teleoperators and Virtual Environments, 14*(2), 119-146.

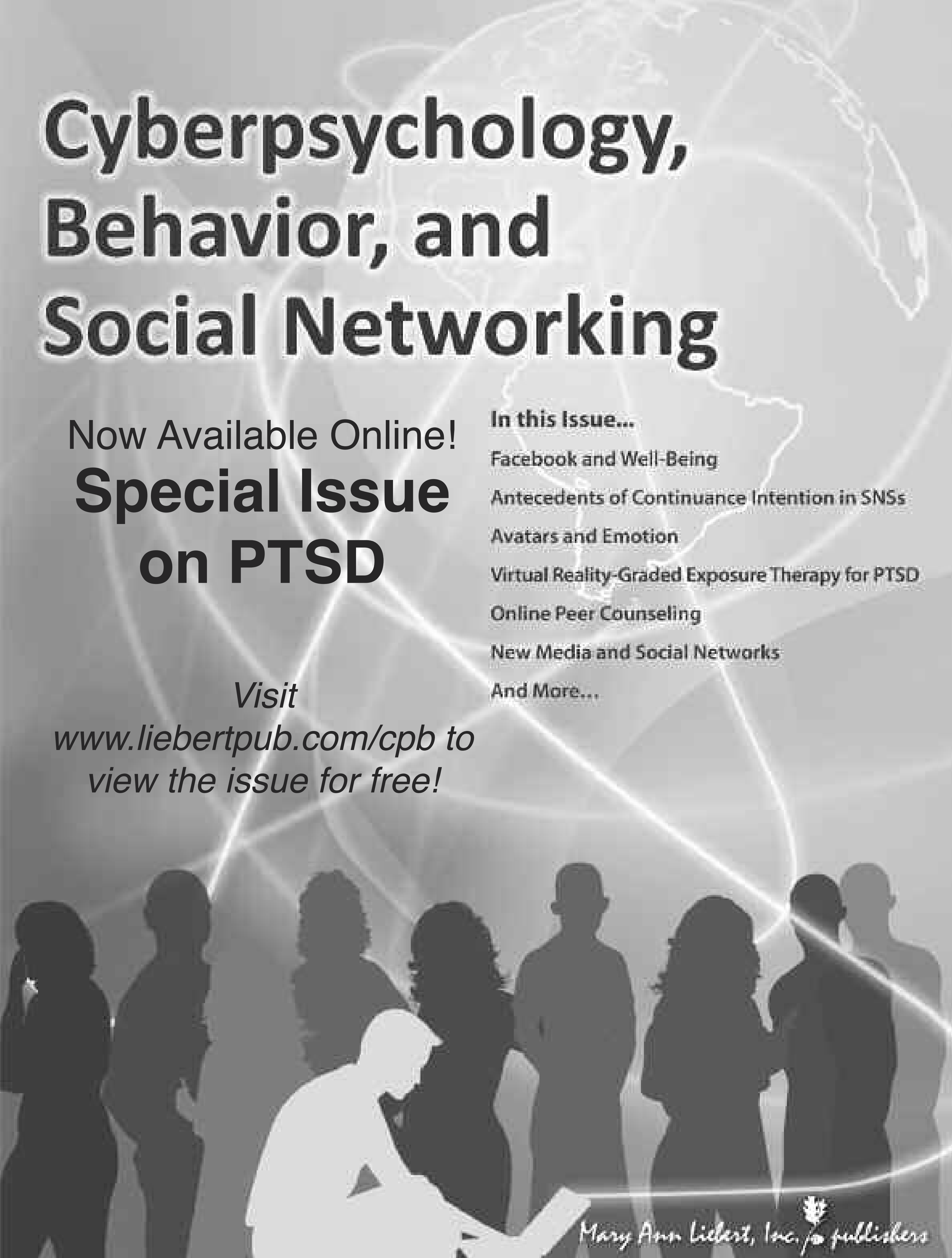
Rothbaum, B. O. (2006). Virtual Vietnam: virtual reality exposure therapy. In M. J. Roy (Ed.), *Novel approaches to the diagnosis and treatment of post-traumatic stress disorder* (pp. 205-218). Amsterdam, The Netherlands: IOS Press.

Schottenbauer, M. A., Glass, C. R., Arnkoff, D. B., Tendick, V., & Gray, S. H. (2008). Nonresponse and dropout rates in outcome studies on PTSD: Review and methodological considerations. *Psychiatry-Interpersonal and Biological Processes, 71*(2), 134-168.

Wood, D. P., Wiederhold, B. K., & Spira, J. (2010). Lessons Learned from 350 Virtual-Reality Sessions with Warriors Diagnosed with Combat-Related Posttraumatic Stress Disorder.

Cyberpsychology Behavior and Social Networking, 13(1), 3-11.

Wrzesien, M., Burkhardt, J. M., Alcañiz Raya, M., Botella, C., & Bretón López, J. M. (2010). Analysis of distributed-collaborative activity during augmented reality exposure therapy for cockroach phobia. In B. K. Wiederhold, G. Riva & S. I. Kim (Eds.), *Annual Review of Cybertherapy and Telemedicine 2010 - Advanced Technologies in Behavioral, Social and Neurosciences* (pp. 134-139). Amsterdam, The Netherlands: IOS Press.



Journal of CyberTherapy & Rehabilitation
Spring 2011, Volume 4, Issue 1

PREDEPLOYMENT STRESS INOCULATION TRAINING FOR PRIMARY PREVENTION OF COMBAT-RELATED STRESS DISORDERS

Laurel L. Hourani[1], Paul N. Kizakevich[1], Robert Hubal[1], James Spira[2], Laura B. Strange[1], David B. Holiday[1], Stephanie Bryant[1] and Amy N. McLean[1]

Predeployment stress inoculation training (PRESIT) is designed to help personnel cope with combat-related stressors and trauma exposure. PRESIT comprises education on combat and operational stress control, attentional retraining and relaxation training, and practice and assessment via a multimedia stressor environment (MSE). Heart rate variability (HRV) and a reaction time task assessed learned skills and inoculation to MSE arousal. Participants with deployment experience and who were in the experimental group demonstrated improvement, measured as greater relaxation demonstrated during the MSE of a follow-up session relative to that of a baseline session. There was also a training effect for this group, such that those participants who showed greater relaxation from a baseline HRV state during the training (i.e., on relaxation breathing and focusing) showed more improvement between sessions. In contrast, there were no significant predictive variables for the participants in training who had never deployed. Participants with more Posttraumatic Stress Disorder (PTSD) symptoms at baseline showed more capability for improvement, as was true for participants who were more anxious about their next deployment.

Keywords: Stress Relaxation, PTSD Prevention, Combat, Operational Stress

Introduction

Concern over large numbers of psychological casualties in military personnel returning from Operations Enduring Freedom and Iraqi Freedom has led to renewed impetus to identify those at risk for serious mental health problems and treat those already suffering from negative mental health consequences (e.g., Hoge et al., 2004; Office of the Surgeon Multinational Force-Iraq, 2006; APA 2007; DoD, 2007; Litz, 2007; Seal et al., 2008; Burnam et al., 2009; Marmar, 2009, Safran et al., 2009). Deployed Soldiers and Marines have an estimated prevalence rate of 16-17% for major depression, Posttraumatic Stress Disorder (PTSD), or generalized anxiety about four months after their return from Iraq (Hoge et al., 2004), and 19% one year after deployment (Hoge et al., 2006). It has been estimated that the two-year PTSD and depression-related costs for the approximately 1.6 million troops who have deployed since 2001 could range from $4.0 to 6.2 billion (Tanielian & Jaycox, 2008).

Consequently, substantial resources are being expended on the treatment of combat-related stress disorders. Although there is much ongoing research to determine the most effective *treatments* for PTSD, little is known about *preventive* efforts designed to prepare personnel to cope with potential deployment and combat-related stressors. As increasing numbers of military personnel continue to engage in combat and other stressful operational situations during deployment, attention is only beginning to shift toward predeployment efforts to reduce the potentially harmful psychological effects of traumatic exposure. Predeployment primary prevention efforts should improve combat effectiveness in the field,

Corresponding Author: Paul N. Kizakevich, RTI International, P.O. Box 12194, Research Triangle Park, NC 27709-2194, Tel: (919) 541-6639, E-mail: kiz@rti.org

[1]RTI International, Research Triangle Park, North Carolina 27709-2194 U.S.A.
[2]National Center for PTSD, Pacific Island Division, Veterans Health Administration

reduce the short-term impact of potential traumatic exposures, and ultimately reduce the risks for long-term PTSD and other stress-related injuries. Fewer symptoms and associated treatment costs of such conditions will benefit military objectives as well as improve the quality of life of the warfighters and their families.

To date, the majority of predeployment training for military personnel has focused relatively little on preparing them for the psychological impact of exposure to combat and other traumatic stressors, and is generally comprised of minimal combat stressor preparation training and the provision of educational information during boot camp (Lewis, 2006; Cohn et al., 2010). The different service branches have introduced variations of predeployment educational briefings and have begun to include some stress control techniques to reduce anxiety and increase skills to cope with stress. Examples of such training programs include the Army's Battlemind program (www.battlemind.army.mil) and the Navy's and Marines' Combat and Operational Stress Control (COSC) programs (www.manpower.usmc.mil/cosc). In contrast to prior simple predeployment briefings, the Battlemind and COSC programs have employed skill based training to prepare Soldiers and Marines for combat missions. Such training focuses on physical readiness, specific mission-relevant jobs, and teamwork. In addition, specialized forces, such as Special Operations, snipers, and West Point cadets have also trained in psychological methods intended to boost both performance and resiliency – including arousal and distraction control. However, these have not been employed in the mainstream of Army combat training, nor have they been assessed for buffering distress across the deployment cycle.

While many resources have been used to develop these programs, studies that have examined the efficacy of these predeployment programs have shown minimal post-deployment effects in the reduction of common mental health outcomes (Sharpley et al., 2008). Consequently, this pilot study was designed to improve upon current operational stress control training by applying a widely acknowledged and effective stress-reduction approach to the prevention of operational stress for military personnel in combat situations. Stress inoculation training (SIT) is an effective method of reducing arousal levels in response to powerful stressors by "inoculating" individuals to potentially traumatizing stressors (Meichenbaum, 2007). SIT derives from models that posit that stress occurs when the perceived demands of a situation exceed an individual's ability to cope (Lazarus & Folkman, 1984). SIT enables individuals to manage stressors when they occur at a low level and increases their ability to manage them under more intensive acute or chronic demands. There are three stages to SIT: *psychoeducation*, which involves learning about stress responses and the need to control them; *training*, which involves learning arousal control to mitigate negative effects of stress; and *implementation*, which involves using or practicing these skills in the stressful context.

The conceptual basis for using SIT as a preventive approach against developing stress-related symptoms after exposure to trauma is based on studies which have shown that enhanced stress resilience is associated with a protective physiologic stress response. Specifically, a reduction of anxiety within 72 hours of exposure to a traumatic event is associated with lower risk of PTSD or greater effectiveness of debriefing (Cohen et al., 2000). In addition, PTSD is often treated successfully with anxiety management techniques, including relaxation training (Schelling et al., 2001) which is a component of SIT. These studies suggest that reduction of physiological arousal during and/or shortly after trauma exposure may prevent or reduce the likelihood of development of psychological distress including PTSD symptomatology. Consequently, methods to reduce arousal levels, such as relaxation training and stress management techniques involved in SIT, play a role in reducing the risk of combat stress casualties.

A few studies have shown that SIT may help reduce the incidence of PTSD in military personnel when administered predeployment (e.g., Deahl et al., 2000), but other studies have also demonstrated useful short-term effects associated with SIT. For example, military personnel who receive SIT and are trained in a virtual stressor environment develop divided attention skills and moderate physiological responses to stress while staying focused on the task at hand (Wiederhold & Wiederhold, 2006). The ability to stay focused on the task at hand while under stress is an important short-term benefit of SIT in addition to long-term benefit of potentially reducing stress-related disorders. Thus, while providing SIT to predeploying military personnel would likely be extremely useful to prepare them to cope with stressors they will encounter during deployment, few studies have implemented and evaluated the effectiveness of predeployment SIT for preventing or mitigating stress-related negative psychological outcomes, such as PTSD.

In line with the Wiederhold & Wiederhold study and others that have focused on exposure to stressful situations (e.g., Reger at al., 2009; Spira et al., 2006), this pilot study included a virtual multimedia stressor environment (MSE) to provide practice of the SIT method, and to measure attentional skills and physiologic responses to stress (Hubal et al., 2010). The PRESIT MSE presents a scripted scenario comprised of mission objectives, anticipation of enemy engagement, ambient sounds, vigilance to in-scene cues, possible insurgents, sudden impact events (e.g., explosions, directed gunfire), depiction of casualties, and post-event chaos. In the MSE, speed and accuracy are measured against contextually relevant stimuli in a combat scenario.

The purpose of this controlled pilot study was to compare the efficacy of a newly designed SIT-based intervention called PRESIT (predeployment stress inoculation training) with COSC current best practices (CBP) by examining heart rate variability (HRV) and reaction time (RT) in response to a combat-relevant MSE. Specifically, the test was to see whether PRESIT reduces the physiological arousal associated with viewing and engaging in a MSE scenario. The hypothesis was that participants in the PRESIT group would have improved HRV, i.e., lower arousal levels and increased SDNN and faster RT scores in response to the MSE after training relative to their baseline measures. If supported, results would provide preliminary evidence suggesting that warfighters who receive PRESIT intervention learn to improve physiological control over their arousal responses in a stressful simulated combat-relevant environment. It was also predicted that participants who endorsed more symptoms of stress, PTSD, and depression on a predeployment questionnaire would show poorer, i.e., lower relative HRV at all points before and after the MSE compared to participants who reported fewer of these symptoms.

Method

Materials Development

Predeployment Baseline Survey

To establish baseline information on participants' health status, prior deployment experience, and combat experience, a 20-minute predeployment questionnaire was developed. The questionnaire emphasized brief, psychometrically sound instruments that had been normed in military populations and items that had been used successfully in previous military surveys. Key measures incorporated in the questionnaire were:

- PTSD Checklist–Civilian Version (Weathers et al., 1994). Note: The civilian version was utilized to measure potential PTSD symptoms prior to military service as well as those incurred since joining the military.
- Patient Health Questionnaire–Anxiety Subscale (Kroenke, et al., 2002)
- Center for Epidemiological Studies-Depression Scale (Tuunainen et al, 2001)
- Two-Item Conjoint Screening (Brown et al., 2001)
- Combat Exposure Scale (Keane et al., 1989)

CBP Materials

Marine Corps COSC informational trifold brochures were provided to all participants. COSC information on the Marine Corps brochures identifies potential stressors, describes signs or symptoms, recommends self-help behaviors, and provides lists of resources for seeking professional help. The skills acquisition component of PRESIT supplemented the COSC educational brochures with a presentation of two relaxation techniques taught by trained medical professionals and described in Table 1.

Multimedia Stressor Environment

The purpose of the MSE was to induce a level of psychophysiological arousal during a simulated training mission, and to assess the ability of participants to respond quickly and accurately to visual stimuli presented during the mission. The MSE consisted of a moderate-sized Iraqi village developed as a virtual scenario, including an assortment of buildings (representative houses, houses of worship), a civilian population, possible insurgents, vehicles, market stands, carts, signs, debris, and other objects (Figure 1). During the MSE, participants traveled through the village streets on a predefined path under the guidance of scripted movement. The idea was to present the notion of a somewhat active scene and to have the participants move together through the scene as if traveling in a truck or other multi-passenger military vehicle. An algorithm was written to make it appear as if the vehicle slowed down and sped up throughout the MSE, to add visualization to simulate an occasional bump or dip in the road, and to shift the view gradually from left to right and back so that the participants could see the entire rendered scene. The design of the MSE is described more fully in Hubal et al. (2010).

A large number of points along the path were established at which types of stimuli and events would occur,

Table 1
Focused Breathing Techniques

Technique 1: In-combat attentional retraining:
The following method is useful for becoming more calm and focused in the moment without excessive reactivity or arousal: • Allow your eyes to rest upon one target or area. Let that scene fill your mind completely. If other thoughts or images come into your mind, let them go and refocus on what you are seeing. • As you allow that scene to fill your mind, notice your breath flowing slowly and consistently. Allow the air to flow into your nose effortlessly. Let your breath get slower and deeper. Feel your abdomen filling and emptying. Slow your breath by letting the exhale go all the way out. • Don't close your eyes or look away. Instead, feel as if you are breathing in what you are looking at all the way into your belly. Then, let your breath extend out toward what you are looking at.
Technique 2: Post-combat recuperative training:
The following method is useful for achieving deep recuperative and restorative rest when one is in a safe environment: • Get comfortable in a chair or bed. Close your eyes, and place your hands over your lower belly. • Feel the breath effortlessly flowing into and out of your nose. Notice every quality of that breath, and how it is different flowing in and flowing out. If thoughts or feelings drift into your mind, just notice them, let them go, and refocus on the breath. (Continue this for 10 relaxed, effortless breaths). • Feel breath flowing into and out from your ribs. Allow the breath to flow in and out effortlessly. Rather, allow the breath to "breathe" you. Feel the effortless expansion and release of your body soothing and massaging you (Continue this for 10 relaxed, effortless breaths). • Feel the warmth of your hands on your belly. Allow the belly to lift and drop your hands, effortlessly. (Continue this for ten relaxed, effortless breaths.) • Feel the way your back expands against the chair or bed, and then releases. (Continue this for 10 relaxed, effortless breaths). • Anytime you notice your attention going to thoughts or feelings, just notice that distraction, let it go, and return to counting the breaths at that part of your body. • You can continue this circle of breaths as often as you like.

to which participants were instructed to be vigilant and to respond using type-specific button presses. The first type was intended to depict sudden stressful occurrences, such as explosion of an improvised explosive device (IED), gunfire, burning objects such as vehicles, and post-explosion rubble (Figure 1). The sudden stressful occurrences were meant to maintain or increase participants' arousal. The second type of event portrayed suspicious objects and events that participants were required to notice, including enemy combatants, weapons, aircraft, and suddenly moving people or vehicles. These two types of events were largely designed to achieve a strong level of psychological engagement and stress under simulated mission conditions.

The third type of event was intended to elicit reaction times from participants, using more abstract visual stimuli. These were small but visible red square targets displayed briefly around the participants' field of view (Figure 2). Equivalent targets of different colors were occasionally presented as foils. Participants' reaction time and accuracy (hits and misses) to the red targets were measured on individual workstations, as well as the number of false alarms to and correct rejections of foil-colored targets. Rumble feedback followed missed or incorrect controller responses to this type of event. We chose to use an abstract target representation to allow the same type of target to be displayed anywhere in the visual field and to minimize variation across participants in the recognition and interpretation of these measurable stimuli. We also anticipated that this abstract target would be so unlike a real target that the "combat training effect", if any, would quickly fade.

Figure 1. A post-explosion event in the MSE virtual scenario.

Figure 2. A reaction time target in the MSE virtual scenario.

The team rendered the MSE using an off-the-shelf gaming engine on a specialized simulation-based training platform (Kizakevich et al, 2007), and layered ambient noises in surround sound to accompany the simulation and enhance its realism. During pilot testing the MSE was projected onto a flat blank wall in front of groups of participants using a high-definition projector and a 5.1 channel surround sound system.

The MSE was presented in a group, theater format with up to 20 participants in a MSE viewing group (Figure 3). Although the participants were taken through the visual environment without personal navigational control, their virtual experience was only moderately immersive. Therefore, as with other theater environments, we did not expect that simulator sickness would be a concern and did not take special considerations for or measure simulator sickness effects.

Physiological Monitoring Equipment

Participants had their beat-by-beat pulse rates monitored before, during, and after the MSE exposure. An infrared LED pulse sensor (Cateye model PL-6000) was clipped to the ear lobe for sensing the peripheral pulse. The sensor was connected to miniature pulse wave monitor (Biocom Technologies model HRM-02) which was then connected to an individual workstation via a USB interface. Participants trained in stress reduction using a biofeedback training software package called Stress Sweeper Pro (Biocom Technologies). Stress Sweeper Pro provides paced breathing exercises via a graphical user interface, including display of beat-by-beat heart rate variability and a stress index for biofeedback. During the MSE, RT was measured by participants' joystick controller movements (Logitech Rumblepad 2) as they responded to stimulus events displayed in the MSE.

Figure 3. MSE session during PRESIT pilot testing.

Time Synchronization

Time synchronization is critical for the MSE and collection of physiological data. The team developed a protocol time manager program that automatically synchronized time among the MSE computer (i.e., that used to project the MSE) and all of the participant workstations via a wireless network. This application also maintained segment timing, so that segments of the protocol (e.g., baseline HRV data collection) requiring a set period of time were ensured. It also enabled synchronized logging on all workstations and aggregating measurements and other study log data to a central encrypted data store.

PRESIT Data Collection System

The integrated PRESIT system comprises an instructor's MSE computer, a wireless local area network, audiovisual presentation components, and set of trainee stations (STAs) (Figure 4). The computers are linked via the wireless network which enables time synchronization across

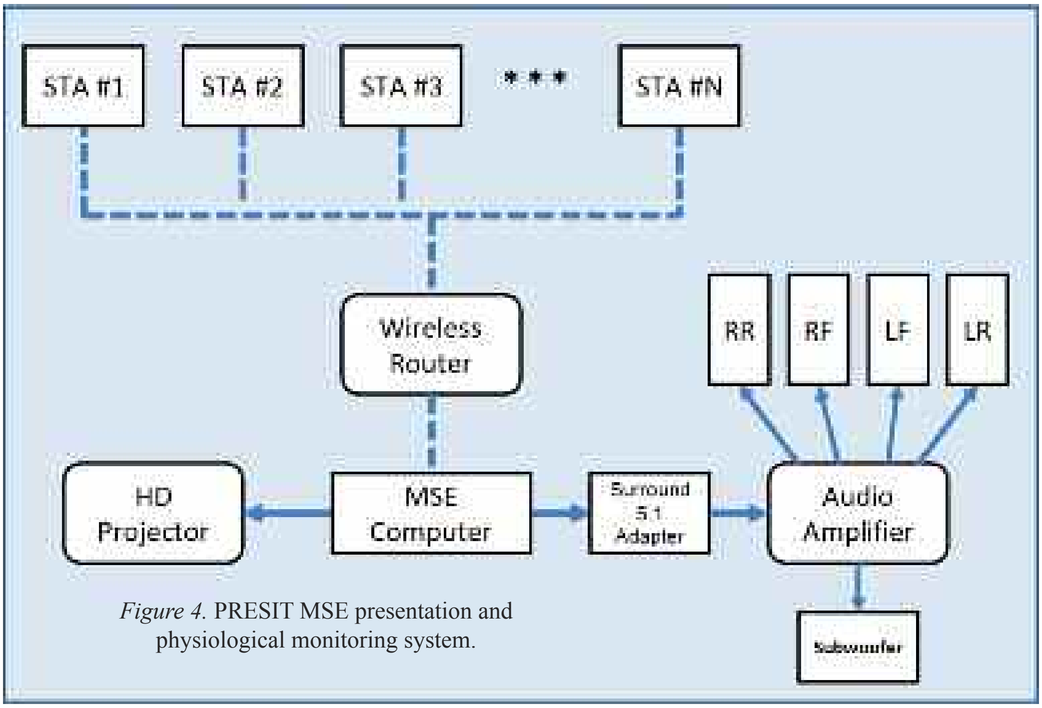


Figure 4. PRESIT MSE presentation and physiological monitoring system.

subsystems, centralized protocol sequencing, and centralized data storage. The audiovisual components comprise a SoundBlaster X-Fi Surround 5.1 USB adapter, an audio amplifier with 5.1 surround capability, four Promonitor 1000 speakers, a ProSub 800 powered subwoofer, and a Panasonic PTAE3000 high definition projector.

The MSE computer manages the sequencing of the training protocol, presents training materials via a high definition (HD) projector, presents the MSE mission scenarios via the HD projector and audio subsystem, serves as a common time base for networked computers, and provides centralized data storage for the networked computers. Software on this computer comprises a protocol management program, various training and instructional presentations, and multiple MSE scenarios. The trainee workstations (STA #1, …, STA #N) provide individualized, automated support for stress relaxation training, physiological and RT monitoring, physiological performance assessment, debriefing questionnaires, and other data collection applications. The trainee workstations are time-synchronized to the MSE computer as the onset of each PRESIT training session.

Procedures

Participants. On two occasions, pilot data were collected from Marines in infantry immersion training (IIT) at the 1 MEF Battle Simulation Center at Camp Pendleton, CA. A total of 77 Marines (43 experimental and 34 control) participated. All participants were active duty undergoing predeployment training exercises, with projected deployment to Iraq or Afghanistan; they were recruited on the basis of operational schedules. Seven squads were group randomized to receive one of two conditions, either the PRESIT protocol or CBP consisting of educational materials. Table 2 shows characteristics of the Marine participants. Despite randomization, members of the control group were more likely to have been deployed.

Groups of Marines from the same squad were gathered together in a session. A team of three to four trained individuals administered the survey, conducted PRESIT training or CBP, and ensured proper collection of physiologic measurements for each group. Data were obtained for participants in each group over two separate sessions on the same day (Figure 5). The first session (A) took place for all group members at the same time prior to specialized ground-based infantry training attended by every Marine prior to combat deployment. The second session (B) occurred for all group members at the same time after that IIT exposure and debriefing. All available Marines within the unit were eligible to participate.

Table 2
Percent Distribution of Selected Variables by Group

Variable	PRESIT	CBP	Total	*p*-value
Age (mean)	20.8	20.9	20.8	NS
Education				
HS or less	79.1	82.3	80.5	NS
Some college or trade school	20.9	17.7	19.5	NS
Paygrade				NS
E1-E3	67.4	61.8	64.9	
E4-E6	32.6	38.2	35.1	
Ever deployed	27.9	55.9	40.3	0.1
Stress in work or family ("a lot")	37.2	35.3	36.4	NS
Depression (CES-D > 16)	37.2	44.1	40.3	NS
PTSD Check-list (PCL)				
PCL>50	9.3	17.7	13.0	NS
PCL>40	34.9	35.3	35.1	NS
Received counseling	14.0	17.7	15.6	NS
Received stress management for deploy/combat	53.5	35.3	45.5	NS
Practice relaxation techniques	14.1	14.7	14.5	NS
Practice martial arts	39.5	58.8	48.0	NS
Interested in learning stress reduction	81.4	85.3	83.1	NS

The process followed during pilot testing involved several steps.

Session A

Upon entering the facility, a squad of 10 to 20 potential participants were informed about the objectives of the study, their participation, any risks and benefits, their option to consent or refuse participation, and their option to end their participation at any time during the experimental data collection. We reminded them that this study was not part of their regular combat training regimen at the IIT and that there would be no negative consequences from their command should any individual choose not to participate. However, these were warfighters trained to operate as a unit, so it was perhaps unlikely that any individual would decline consent.

After obtaining consent, the PRESIT experiment protocol was administered to the combat unit in a grouped classroom setting. To start, participants completed the ques-

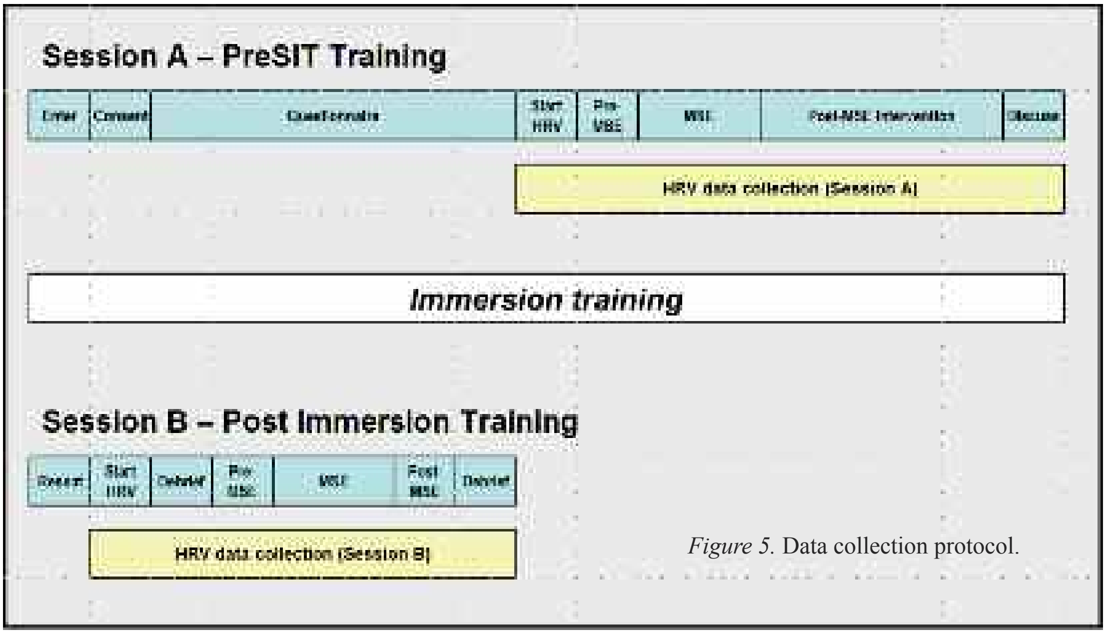


Figure 5. Data collection protocol.

tionnaire that gathered demographic, stress and coping, deployment, and other information. Following completion of this questionnaire, participants clipped on the ear pulse sensor to begin pulse recording, registering physiologic activity (i.e., pulse and HRV). Next, an orientation to the MSE was given, describing the mission within the MSE and participants' required actions. When the orientation was complete, and all participants and their controllers were synchronized, then the MSE was presented, during which team members observed the participants to ensure that they were following procedures and the quality of their physiological data. At the completion of the MSE, all participants were given a stress control intervention. The experimental group received PRESIT, which contains two phases: (1) biofeedback facilitated breathing retraining for control of autonomic reactivity, and (2) attentional control for staying fully engaged in the moment at hand, while maintaining optimal arousal levels. Participants in the control groups were treated identically in all respects except that they received CBP, which consisted of a slide presentation of COSC materials. Arousal levels were measured using HRV continuously in the Marines before, during, and after the MSE. HRV was recorded for at least three minutes prior to starting the MSE to obtain baseline data, during the MSE, during the entire intervention, and then for a five-minute segment following the intervention during the debriefing. Reaction times to the MSE tasks (i.e., responding with stimulus type-specific joystick controller movements) were recorded. After debriefing, participants were instructed to remove their ear clips. Prior to departure, they were given handouts with COSC reminder information.

Immersive Infantry Training

Following the first session, Marines entered the IIT as part of their predeployment combat training at Camp Pendleton. The IIT is live training environment designed to simulate a typical marketplace in Iraq using a - ockup of a village inside a large warehouse. Marines patrol through streets, building, and rooms while participating in infantry-training scenarios that mimic those encountered during combat deployment. IIT scenarios involve combat situations such as encountering hostile insurgents (played by live actors), ambushes, simulated IED explosions, other pyrotechnics, and other situations specifically designed to allow Marines to experience making decisions during highly stressful situations. Training objectives and scenarios are determined by their unit leaders and the IIT instructors. Marines receive debriefing by their unit leader after each scenario. The combat stress experience during these short simulated missions could be intense and vary greatly across the troop trainees. Since this training could not be coordinated with the study protocol, and was generally not observable by the study team, the potential effects of the midday IIT training on the pilot study could not be quantified. Specifically, the effects of the IIT on the second PRESIT session, particularly the physiological measurements during the second MSE presentation, were likely to be variable, both across units and individuals.

Session B

After completing the IIT scenario and debriefing, Marines returned for the second session of the pilot study. The amount of time elapsed ranged from only a few minutes after debriefing to approximately 30 minutes after Marines completed the IIT and received debriefing across groups. During the second session, Marines were provided with a refresher of either the PRESIT or CBP training (the same as earlier) and then were again exposed to the same MSE. HRV was recorded throughout the second session identical to the first session. At the end of this experience, participants were debriefed and responses to open-ended discussion of their responses to the training and MSE were elicited, their questions were answered, and their feedback was noted. After completion of the second session, Marines were provided with COSC materials before leaving. Total participation time was approximately 1.5 hours pre-IIT and 0.5 hours post-IIT.

HRV data processing. The overall data collection protocol for HRV for each experiment day is presented in Figure 5. Before, during, and after each MSE session, beat-by-beat heart rate information was acquired. After data collection, each recording session was reviewed, and a structured data file was exported comprising a list of interbeat intervals (IBI), the time and date at the end of the session, and the session duration in seconds.

Since HRV data were acquired on a separate computer for each participant, several methods were taken to synchronize timestamps between the participant's workstation and the computer that rendered the MSE. In the first pilot data collection a simple countdown procedure was used, where each participant pressed a preset key sequence on his game controller in sync with a projected presentation by the MSE computer. The time of projection was recorded on the MSE computer and the time each sequence was keyed was recorded on the participant's computer. In the second pilot data collection a wireless local area network was used, as described above, to electronically and automatically synchronize all the computers. By these means the team was able to adjust the times of recorded data, including acquired heart beats, and synchronize all data to the MSE computer and PRESIT protocol timing.

For HRV data analyses, each IBI data file was segmented and extracted to subfiles according to various analytical requirements. For example, one segmentation scheme extracted three subfiles: IBI starting 4.5 minutes prior to each MSE and ending at the start of the MSE, as a baseline measure of HRV; IBI starting one minute after the start of each MSE, and continuing for the remainder of the MSE, as a measure of HRV while engaged in the stressful environment's tasks; and IBI starting one minute after the end of the MSE and continuing for 4.5 minutes, as a measure of HRV return to baseline. Other data segmentation criteria included the use of overlapping three-minute segments starting before the MSE with a one-minute start shift; and one-minute segments occurring before and after primary MSE events (e.g., explosions). Analysis of initial pilot data were limited to time domain variables, specifically SDNN, the standard deviation of all normal IBI (i.e., no ectopic or missing beats) and pNN50, the percentage of differences between adjacent normal IBI that are greater than 50 msec over each IBI data segment.

Results

The effective¬ness of the MSE for practicing stress reduction skills depends on its ability to engage participants and elicit a physiological arousal response. To meet these objectives, the analysis (1) examined the arousal of individual experiences to the MSE, and (2) recorded the speed and accuracy of participants to the RT task within the MSE.

Descriptive analyses were conducted for summarizing the outcomes produced by the intervention across the experimental conditions and within other demographics. Summary measures were also calculated for each randomized unit (i.e., combat training squad). Simple t-tests were applied and contingency table methods to these summary measures and examined associations with SAS (SAS version 9, Cary, NC). For quantitative responses, the hierarchical modeling framework of Bryk and Raudenbush (1992) was followed. Multilevel analyses took into account the intra-class correlation of responses within the randomized units, since the participants do not behave independently. Terms considered in the model included an overall mean, effects due to services branches (combat or combat service support), the experimental conditions (PRESIT or CBP) assigned to squads, participants within squads, and baseline and other covariates associated with squads and participants.

HRV Arousal Data

Preliminary data analyses included investigation of outliers and variable frequency distributions. Crosstabs and correlations were conducted on all study variables. Initial t-tests between experimental and control group differences on HRV variables were run separately for the SDNN and pNN50 variables and again for previously deployed and non-deployed groups. Results from the HRV analyses are presented in Figures 6 and 7, in which higher values along the y axis denote greater relaxation. In Figure 6, The difference between the before and during MSE values in the pre-IIT session is <.05 for both PRESIT and control groups. The difference between the during MSE values between pre-IIT and post-IIT for the PRESIT group is also <.05. These data suggest the MSE effectively increased arousal, as measured by SDNN, during both the baseline and post-IIT training sessions. Furthermore, the p-value between experimental and control groups is nearly significant at <.08, with the control group surprisingly showing greater overall relaxation after training in Session B.

However, for pNN50 (Figure 7), the MSE arousal was only demonstrated during the post-IIT sessions. The results showed an interaction between experimental group and HRV score changes from the first session during the MSE to the second session during the MSE. Partly this is because the experimental group showed significantly more arousal than the control group during the post-IIT session as measured by pNN50. Partly, too, the data showed unexpected and/or uncontrolled variability between experimental and control groups. For example, 13% of participants screened positive for PTSD, but the control group far exceeded the PRESIT group (18% control, 9% PRESIT). An additional 34% scored as having possible PTSD (PCL>40) despite upcoming deployment. During Session A, two participants fell asleep and at least one had drunk a large caffeinated beverage immediately prior to SIT. In contrast to our expected results, the PRESIT group showed more stress after the IIT than the control group, possibly as a result of different IIT experiences.

To further examine potential covariates and confounding variables, additional multivariate analyses were conducted. For these, logistic models were created on HRV outcome variables (reduced arousal or increased HRV from session A MSE to session B MSE) on the total sample and stratified by having ever been deployed. The team began by focusing on improvement, measured as greater

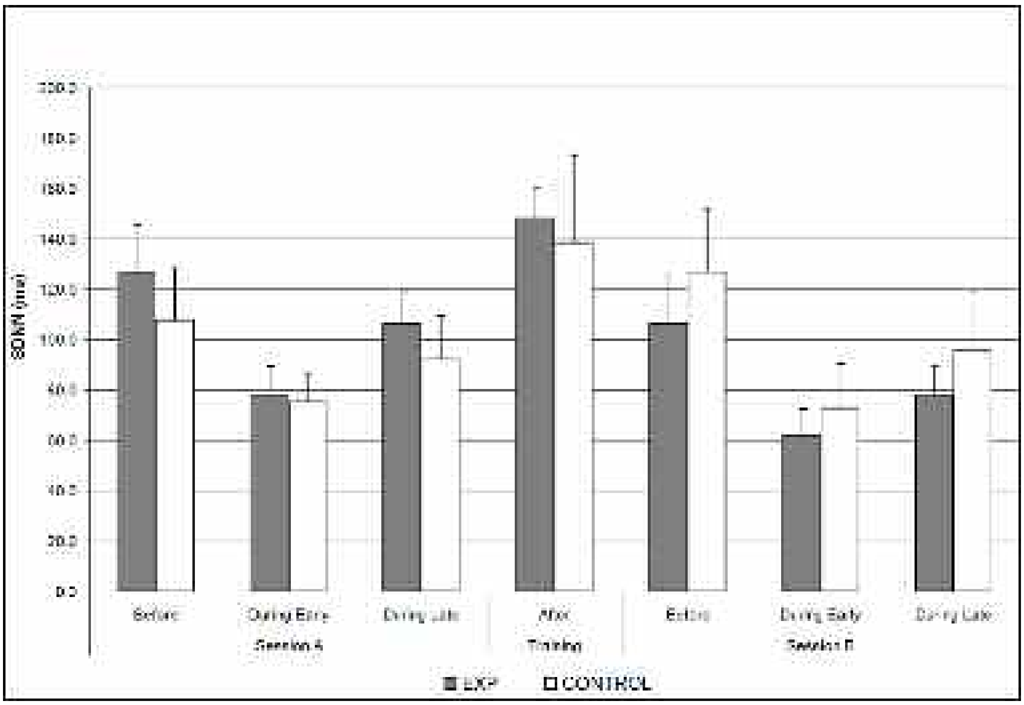


Figure 6. Heart rate variability (SDNN) during PRESIT study conditions.

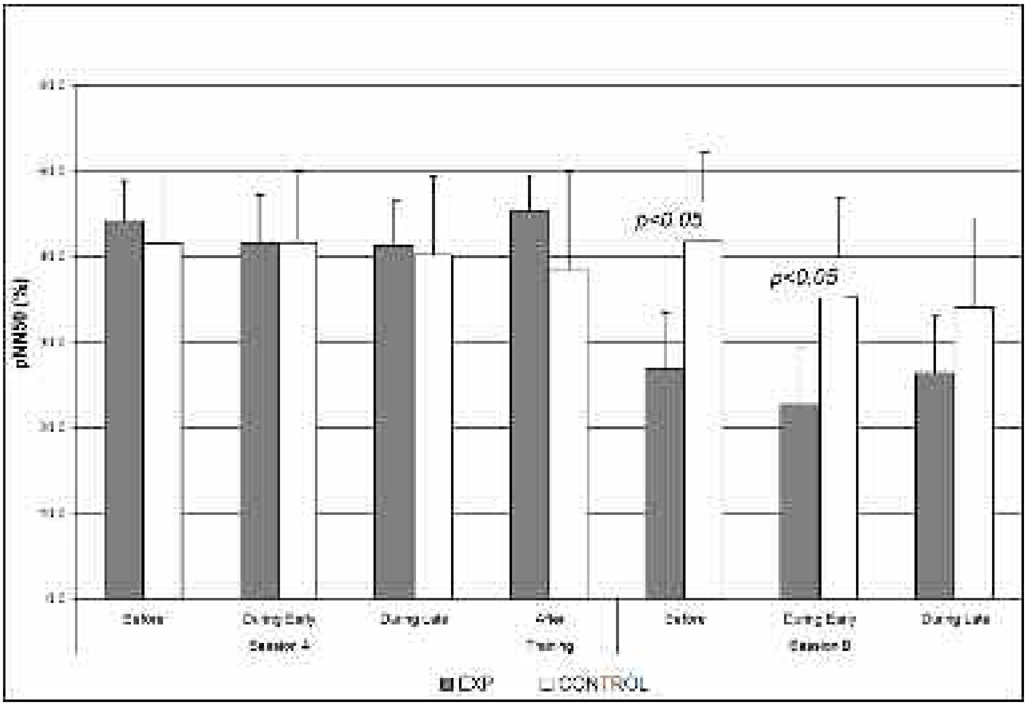


Figure 7. Heart rate variability (pNN50) during PRESIT study conditions.

relaxation demonstrated during the MSE of the second session than the MSE of the first session, considering the participants' experimental group (PRESIT or CBP) as well as characteristics such as prior deployment and PTSD checklist scale score that could act as covariates or confounds. Although logistic models were not productive, multiple regressions using degree of HRV improvement as a continuous variable did yield findings of interest (see Tables 3 & 4). For instance, those participants who had had prior deployment experience and who were in the experimental group demonstrated significant improvement from session A to B, despite a small sample size. There was also a training effect for this group, such that those participants who showed more improvement (greater relaxation) from a baseline HRV state during the training (i.e., on relaxation breathing and focusing) showed more improvement between sessions. In contrast, there were no significant predictive variables for the participants in training who had never deployed. The analyses also found that those previously deployed experimental participants who had high baseline arousal showed no improvement between sessions, whereas those who had low baseline arousal showed improvement but less so than the participants who were relaxed at baseline. Finally, a number of additional covariates were predictive of improvement though the small sample size showed the model to be saturated. For example, participants with more PTSD symptoms showed significantly more capability for improvement, as was true for participants who were more anxious about their next deployment. Furthermore, participants with martial arts training and prior stress relaxation training showed improvement, perhaps related to their prior experience in relaxation and self-awareness training.

Reaction Time Data

The MSE involved not only responding to suspicious and dangerous objects to increase participants' arousal but also responding to reaction time (colored) targets as a measure of performance. These were important as the intent was to insure that relaxation training posed no threat to reduced performance levels and to test the hypothesis that performance would improve with training. With these targets, it was possible to objectively determine speed and accuracy of response. As a further test, a subset of foil targets, of a different color, were interspersed to gauge participants' impulsivity and discrimination. Neither the timing, size, or location of primary targets or foils were presented uniformly so that any learning effects would be minimized.

For the suspicious and dangerous stimuli, it was then possible to compare within and across participant sessions to see if there were effects of stress.

Not only was reaction time considered but also number and kind of joystick controller button presses. Hits were defined as correct button-press responses to targets, and misses as either correct responses that occurred too late (beyond some pre-established threshold) or failure to respond to targets. False alarms were button-presses to foils, and correct rejections as failure to respond to foils. During the Session A (combining across experimental groups), participants identified (hit) 88% of the targets, and missed only 12%. Only about 17% of the time did participants respond incorrectly (false alarms) to foils, correctly rejecting, on average, 83%. During Session B, experimental group participants increased their hit rate significantly to 91% ($p<.04$), though they also increased their false alarm rate (a non-significant change) to 19%, while control group participants' hit rate decreased significantly to 85% ($p<.05$) though their false alarm rate also decreased (non-significantly) to 15%. These data suggest that experimental group participants were a bit more alert in the post-IIT session, relative to pre-IIT, compared to controls.

Qualitative Data

The Marines reportedly liked the PRESIT methodology and in debriefing preferred it to the didactic educational materials. 86% of participants stated they were "interested in learning techniques to manage stress."

Table 3
Multiple Regression of Potential Predictors of HRV Effectiveness of PRESIT Training (for Experimental Group Only)

Variable	**Parameter Estimate**	**Standard Error**	**t Value**	***p*-value**
Intercept	2.42	1.76	1.38	NS
Low Baseline Arousal	-0.50	1.30	-0.38	NS
High Baseline Arousal	-5.25	1.27	-4.14	P<0.01
PTSD-CL score	0.05	0.04	1.51	NS
Anxiety for next deployment	0.49	1.16	0.42	NS
Practices martial arts	1.16	1.02	1.14	NS
Has practiced relaxation techniques	1.32	1.01	1.22	NS
Prior stress management training	-0.67	1.06	-0.63	NS

Table 4

Multiple Regression of Potential Predictors of HRV Improvement from Baseline to Follow-up MSE. (for Previously Deployed Group Only, Within the Experimental Group)

Variable	Parameter Estimate	Standard Error	t Value	*p*-value
Intercept	-20.06	1.52	-13.22	P<0.05
HRV training effect (SDNN)	1.01	0.06	16.06	P<0.05
Low Baseline Arousal	-10.26	0.92	-11.12	NS
High Baseline Arousal	0.08	0.51	0.15	NS
PTSD-CL score	0.37	0.03	14.96	P<0.05
Anxiety for next deployment	30.82	1.59	19.41	NS
Practices martial arts	-10.15	0.28	-36.56	NS
Has practiced relaxation techniques	-2.90	0.21	-14.02	P<0.05
Prior stress management training	12.92	0.77	16.98	P<0.05

Discussion

This pilot test of predeployment SIT resulted in several findings, including support for the MSE to engage and arouse warfighters, support for the physiological measuring methodology, and preliminary suggestions of benefit particularly for those already having been deployed. As shown in Figure 6, physiological arousal, inversely measured by the SDNN HRV measure, clearly increased during each MSE exposure, decreased after each MSE exposure, and decreased during stress relaxation training. Moreover, integration of PRESIT into the predeployment combat training regimen was universally supported by trainers, unit leaders, and troops.

It is unclear as to why the SDNN and pNN50 measures of HRV presented somewhat different results. As shown in Figure 6, SDNN showed a consistent pattern of increased arousal and post-MSE recovery during both MSE sessions, while pNN50 only responded during the second session (Figure 7). Furthermore, SDNN and pNN50 were correlated (r=0.81) when we constrain analysis to the post-IIT training and the MSE Session B data segments.

While SDNN and pNN50 measurements are often correlated (Massin et al, 1999), they do always respond with the same significance (Guger et al, 2004; Gerakani, et al, 2008). In a pilot study with an unselect population having varied combat experience, predeployment anxieties, heightened PTSD-CL scores, and varied IIT training scenarios, it is not surprising that their arousal response as measured by these related, yet different, methods would be somewhat inconsistent.

The results of exploratory analyses to examine potential confounders, while informative, point to a need to improve the training and stressor experiences for both those who do not respond well to training and those who have no deployment experience. They also indicate that the PRESIT regimen may need tailoring based on an individual's, or unit's, prior combat experience or deployment anxiety. In addition, these results strongly indicate the need for a larger sample size to further examine the effects of several potential covariates including PTSD symptoms, anxiety about upcoming deployment, practice of martial arts and relaxation exercises, and receipt of previous stress management training.

A larger longitudinal intervention study is planned to assess its effect on PTSD symptoms as a primary prevention intervention in deploying troops. This longitudinal study will train participants prior to their deployment to theater and evaluate their stress responses and symptom levels upon their return. A revised protocol is being developed for a more usual predeployment environment rather than the IIT, and to allow for an additional training and practice period. Following the MSE, participants in the experimental group will be encouraged to practice breathing to decrease arousal during their usual predeployment training exercises. Several days later all participants will be asked to return and run through another version of the MSE (a version of the MSE that reflects village life in Afghanistan rather than Iraqi as depicted in the pilot MSE) and then relax post-MSE, practicing those same skills. The analyses will consider the possible separate influences of relaxation breathing and focusing attention. Upon notification of return from deployment, the participants will be contacted to complete a post-deployment follow-up survey similar to the predeployment baseline survey, but including questions about compliance and summaries of trauma exposures, and the efficacy of their PRESIT training will be evaluated.

Conclusion

This pilot study has shown the potential for addressing predeployment training needs to bolster resilience in troops facing potential traumatic exposures and stressful experiences. The study used a stress inoculation approach that focuses on breathing and attentional control and includes a stressor environment in which to practice these techniques. Because the skills can be learned in a group setting in conjunction with ongoing predeployment exercises, the approach represents practical, cost-effective training that may mitigate stress-related symptoms and reduce the treatment burden of PTSD among deployed Marines. Further research in this area using a refined experimental protocol is underway.

Acknowledgements

The authors would like to thank Ellen Bishop, Carrie A. Borst, and Diglio A. Simoni of RTI International for their contribution to this work. During the conduct of this project the team interacted primarily with the 1 MEF Battle Simulation Center at Camp Pendleton. The pilot experiments were conducted at that facility in conjunction with Marines rotating through the IIT. The authors like to thank Mr. Tom Buscemi, Director of IIT for his gracious cooperation and Dr. Tom Gaskin, Head, COSC, MC Headquarters for all his help. The work was funded by the U.S. Army Medical Research and Materiel Command under award number W81XWH-08-1-0334 and by the Office of Naval Research under award number N000014-08-C-0504.

References

American Psychological Association. (2007). Presidential Task Force on Military Deployment Services for Youth, Families and Service Members. The Psychological Needs of U.S. Military Service Members and Their Families: A Preliminary Report, February 2007.

Brown, R.L., Leonard, T., Saunders, L.A., & Papasouliotis, O. (2001). A two-item conjoint screen for alcohol and other drug problems. *Journal of the American Board of Family Practice, 14*(2) 95-106.

Bryk, A.S., & Raudenbush, S.W. (1992). Hierarchical linear models: Applications and data analysis methods. Newbury Park: Sage.

Burnam, M.A., Meredith, L.S., Tanielian, T., & Jaycox, L.H. (2009). Mental health care for Iraq and Afghanistan war veterans. *Health Affairs, 28*(3), 771-782.

Cohen, H., Benjamin, J., Geva, A.B., Matar, M.A., Kaplan, Z., & Kotler, M. (2000). Autonomic dysregulation in panic disorder and in post-traumatic stress disorder: Application of power spectrum analysis of heart rate variability at rest and in response to recollection of trauma or panic attacks. *Psychiatry Research, 96*(1), 1-13.

Cohn, J., Weltman, G., Ratwani, R., Chartrand, D., & McCraty, R. (2010). Stress inoculation through cognitive and biofeedback training. In Proceedings of the Interservice/Industry Training, Simulation and Education Conference. Arlington, VA: National Defense Industrial Association.

Deahl M, et al. (2000). Preventing psychological trauma in soldiers: The role of operational stress straining and psychological debriefing. *Br J Med Psychol, 73*:77-85.

Department of Defense Task Force on Mental Health. (2007). An achievable vision: Report of the Department of Defense Task Force on Mental Health, Falls Church, VA: Defense Health Board.

Garakani, A., Martiniz, J.M., Aaronson, C.J., Voustianiouk, A., Kaufmann, H., & Gorman, J.M. (2009). Effect of Medication and Psychotherapy on Heart Rate Variability in Panic Disorder. *Depression and Anxiety* 26:251-258.

Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., Brogni, A., Friedman, D. & Slater, M. (2004). Hear-Rate Variability and Event-Related ECG in Virtual Environments. *Presence*, 240-245.

Hoge, C.W., Auchterlonie, J.L., & Milliken, C.S. (2006). Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. Journal of the American Medical Association, 295(9), 1023-1032.

Hoge, C.W., Castro, C.A., Messer, S.C., McGurk, D., Cotting, D.I., & Koffman, R.L. (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *New England Journal of Medicine, 351*(1), 13-22.

Hubal, R., Kizakevich, P., McLean, A., & Hourani, L. (2010). A multimedia environment for stressing warfighters before they deploy. In Proceedings of the Interservice/Industry Training, Simulation and Education Conference (pp. 1688-1696). Arlington, VA: National Defense Industrial Association.

Keane, T.M., Fairbank, J.A., Caddell, J.M., Zimering, R.T., & Mora, C.A. (1989). Clinical evaluation of a measure to assess combat exposure. *Psychological Assessment, 1*(1), 53-55.

Kizakevich, P. N., Culwell, A., Furberg, R., Gemeinhardt, D., Grantlin, S., Hubal, R., Stafford, A., & Dombroski, R. T. (2007). *Virtual simulation-enhanced triage training for Iraqi medical personnel.* In J. D. Westwood, R. S. Haluck, H. M. Hoffman, G. T. Mogel, R. Phillips, R. A. Robb, & K. G. Vosburgh (Eds.), In vivo, in vitro, in silico: Designing the next in medicine (pp. 223–228). Amsterdam: IOS Press.

Kroenke, K., Spitzer, R.L., & Williams, J.B.W. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine, 64*(2), 258-266.

Lazarus, R.S., & Folkman, S. (1984). Stress, Appraisal and Coping. New York: Springer.

Lewis, S.J. (2006). Combat stress control: Putting principle into practice. In A.B. Adler, C.A. Castro, & T.W. Britt (Eds.), *Military life: The psychology of serving in peace and combat* (vol. 2) (pp. 121-140). Westport, CT: Praeger Security International.

Litz, B.T. (2007). Research on the impact of military trauma: Current status and future directions. *Military Psychology, 19*(3), 217-238.

Massin, M., Derkenne, B., & von Bernuth, G. (1999). Correlations between Indices of Heart Rate Variability in Healthy Children and Children with Congenital Heart Disease. *Cardiology* 91:109-113.

Marmar, C.R. (2009). Mental health impact of Afghanistan and Iraq deployment: meeting the challenge of a new generation of veterans. *Depression and Anxiety, 26*(6), 493-497.

Meichenbaum, D. (2007). Stress inoculation training: A preventative and treatment approach. In P.M. Lehrer, R.L. Woolfolk, & W.E. Sime (Eds.), Principles and practice of stress management (pp. 497–518). New York, NY: Guilford Press.

Office of the Surgeon Multinational Force-Iraq and Office of the Surgeon General United States Army

Medical Command, Mental Health Advisory Team (MHAT) IV Operation Iraqi Freedom 05-07 Final Report, 17 November 2006.

Reger, G. M., Gahm, G. A., Rizzo, A. A., Swanson, R., & Duma, S. (2009) Soldier evaluation of the virtual reality Iraq. *Telemedicine and eHealth. 15* (1): 101-104.

Safran, M.A., Strine, T.W., Dhingra, S.S., Berry, J.T., Manderscheid, R., & Mokdad, A.H. (2009). Psychological distress and mental health treatment among persons with and without active duty military experience, Behavioral Risk Factor Surveillance System, United States, 2007. *International Journal of Public Health, 54* (Supplement 1), 61-67.

Schelling, G., Briegel, J., Roozendaal, B., Stoll, C., Rothenhäusler, H.B., & Kapfhammer, H.P. (2001). The effect of stress doses of hydrocortisone during septic shock on posttraumatic stress disorder in survivors. *Biological Psychiatry, 50*(12), 978-985.

Seal, K.H., Bertenthal, D., Maguen, S., Kristian, G., Chu, A., & Marmar, C.R. (2008). Getting beyond "Don't Ask: Don't Tell": An evaluation of US veterans administration postdeployment mental health screening of veterans returning from Iraq and Afghanistan. *American Journal of Public Health, 98*(4), 714-720.

Sharpley JG, Fear NT, Greenberg N, Jones M, Wessely S. (2008) Pre-deployment stress briefing: does it have an effect? *Occupational Medicine (Lond). Jan;58*(1):30-4.

Spira, J., Pyne, J. M., Wiederhold, B., Wiederhold, M., Graap, K. & Rizzo, A. A. (2006). Virtual reality and other experiential therapies for combat-related PTSD. *Primary Psychiatry, 13*(3), 58–64.

Tanielian, T., & Jaycox, L.H. (Eds). (2008). Invisible wounds of war: Psychological and cognitive injuries, their consequences, and services to assist recovery (RAND Center for Military Health Policy). Santa Monica, CA: RAND Corporation.

Tuunainen A, Langer RD, Klauber MR, Kripke DF. (2001) Short version of the CES-D (Burnam screen) for depression in reference to the structured psychiatric interview. *Psychiatry Research. Sep 20;103*(2-3):261-70.

Weathers, F.W., Litz, B. T.,Huska, J. A.,& Keane, T.M. (1994). PTSD Checklist. Boston: National Center for PTSD, Behavioral Science Division.

Wiederhold, B.K., & Wiederhold, M.D. (2006). *Clinical psychology and neurorehabilitation: The use of virtual reality driving simulators*. In G. Riva, M.T. Anguera, B.K. Wiederhold, & F. Mantovani (Eds.), From Communication to Presence: Cognition, Emotions and Culture towards the Ultimate Communicative Experience (pp. 262-275). Amsterdam: IOS Press.

CYBERPROJECTS

IN THIS FEATURE, we will try to describe the characteristics of current cyberpsychology and rehabilitation research. In particular, CyberProjects aims at describing the leading research groups and projects, actually running around the world, with a special focus on European research.

CORBYS : Cognitive Control Framework for Robotic Systems

Project Objectives:
CORBYS' focus is on robotic systems that have symbiotic relationships with humans. Such robotic systems have to cope with highly dynamic environments as humans are demanding, curious and often act unpredictably.

CORBYS will design and implement a cognitive robot control architecture that allows the integration of 1) high-level cognitive control modules, 2) a semantically-driven self-awareness module, and 3) a cognitive framework for anticipation of, and synergy with, human behavior based on biologically-inspired information-theoretic principles.

These modules, supported with an advanced multi-sensor system to facilitate dynamic environment perception, will endow the robotic systems with high-level cognitive capabilities such as situation-awareness, and attention control. This will enable the adaptation of robot behavior to the user's variable requirements to be directed by cognitively-adapted control parameters.

CORBYS will provide a flexible and extensible architecture to benefit a wide range of applications ranging from robotized vehicles and autonomous systems, such as robots performing object manipulation tasks in an unstructured environment, to systems where robots work in synergy with humans. The latter class of systems will be a special focus of CORBYS' innovation since important classes of critical applications exist where support for humans and robots sharing their cognitive capabilities is a particularly crucial requirement to be met.

CORBYS' control architecture will be validated within two challenging demonstrators: i) a novel mobile robot-assisted gait rehabilitation system CORBYS; ii) an existing autonomous robotic system.

The CORBYS demonstrator that will be developed during the project will be a self-aware system capable of learning and reasoning that enables it to optimally match the requirements of the user at different stages of *rehabilitation* in a wide range of gait disorders.

Contact Information
GRÄSER Axel
UNIVERSITAET BREMEN
Bibliothekstrasse 1
28359
GERMANY
Tel: +49-42121862444
Fax: +49-4212189862444
http://bitbrain.es/corbys/

Compiled by Giuseppe Riva, Ph.D.,
and Simona Raspelli, Ph.D.
Istituto Auxologico Italiano
Data provided by ICT Results
(http://cordis.europa.eu/ictresults)

CYBERFOCUS

New technologies are developing at a rapid pace. To help you stay abreast of the latest trends in advanced technologies and healthcare, this feature showcases upcoming 2011 events which will provide you with the opportunity to connect with leading experts worldwide and remain on the cutting edge of the most recent developments.

The CyberFocus column welcomes your contributions. To supply relevant information for this feature, please send an e-mail to: office@vrphobia.eu.

CyberPsychology & CyberTherapy 16
June 19-22, 2011
Gatineau, Canada
www.interactivemediainstitute.com

The Journal of CyberTherapy & Rehabilitation is the official journal of the CyberPsychology & CyberTherapy Conference (CT16). CT16 brings together researchers, clinicians, policy makers and funding agents to share and discuss advancements in the growing discipline of CyberTherapy & Rehabilitation, which includes training, education, prevention, rehabilitation, and therapy. The focus of next year's conference is two-fold–first, "Technologies as Enabling Tools" will explore the use of advanced technologies in diagnosis, assessment and prevention of mental and physical disorders. In addition, attention will be drawn to the role of interactive media in training, education, rehabilitation and therapeutic interventions. Second, CT16 will investigate the impact of new technologies such as Facebook and Twitter and how they are influencing individual behavior and societal relations. Technologies featured at the conference include VR simulations, videogames, telehealth, the Internet, robotics, brain-computer interfaces, and non-invasive physiological monitoring devices. Conference attendees have the opportunity to explore interactive exhibits at the Cyberarium as well.

2011 Conferences

IEEE-VR 2011
March 19 - 23
Singapore
http://conferences.computer.org/vr/2011/

Med-e-Tel
April 6 - 8
Luxembourg
http://www.medetel.lu/

Laval Virtual Reality Conference
April 6 - 10
Laval, France
http://www.laval-virtual.org/

World Health Care Congress
April 13 - 14
Brussels, Belgium
http://www.worldcongress.com/

American Telemedicine Association 2011
May 1 - 3
Tampa, Florida, USA
www.americantelemed.org/ATA2011/

----

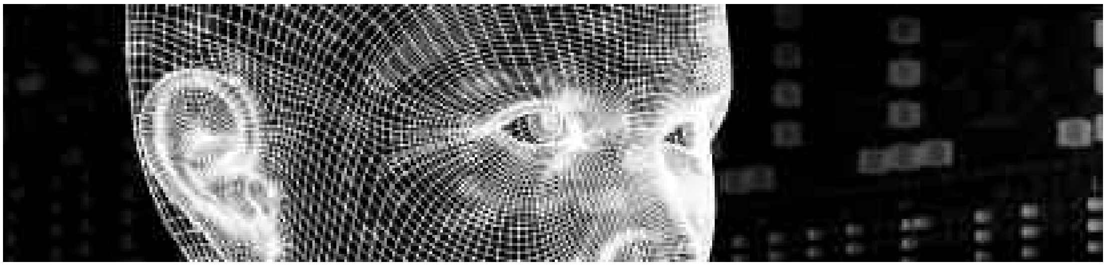
fet11 - The European Future Technologies Conference and Exhibition
May 4 - 6
Budapest, Hungary
www.fet11.eu/

Computer Human Interaction (CHI) 2011
May 7 - 12
Vancouver, British Columbia
www.chi2011.org/

World of Health IT 2011
May 10 - 13
Budapest, Hungary
http://www.worldofhealthit.org/

Pervasive 2011
June 12 - 15
San Francisco, California, USA
http://pervasiveconference.org/2011/

CyberPsychology & CyberTherapy 16
June 19-22, 2011
Gatineau, Canada
www.interactivemediainstitute.com



FOR AUTHORS

To Submit
Please submit electronic copies of your manuscript by visiting www.vrphobia.eu and clicking on the "Submit Paper" link on the right.

Manuscript Style. The language of the journal is American English. Submissions should follow American Psychological Association (APA) format. Format questions not addressed in the checklist below can be sent to the Managing Editor at office@vrphobia.eu.

Copyright. Submission of a manuscript will be held to imply that it contains original unpublished work and is not being submitted for publication elsewhere at the same time. If selected for publication, authors must sign and return a Copyright Transfer Form (available on the Web site or from the Publisher.) Submitted material will not be returned to the author, unless specifically requested.

Review. Papers will be evaluated anonymously by multiple members of the Editorial Board. They will be reviewed based on scientific merit, originality, readability, and interest to the readership.

Further Information. Typeset proofs will be sent to the author for review. This stage is only for correcting errors that may have been introduced during the production process. Prompt return of the corrected proofs, preferably within 48 hours of receipt, will minimize the risk of the paper being held over to a later issue. Two complimentary copies of the issue will be provided to the corresponding author, unless otherwise indicated. Further copies of the journal may be ordered.

MANUSCRIPT SUBMISSIONS

Format. The original manuscript should be double-spaced and formatted for A4 paper (.27in. x 11.69 in.; 21cm. x 29.7cm.) with adequate and consistent margins on all pages.

The title page, abstract, references, appendixes, author note, content footnotes, tables, figure captions, and figures must be on separate pages (with only one table or figure per page.) They should be ordered in sequence, with the text pages between the abstract and the references.

Is the author note typed on the title page, which is removed by the journal editor before review? Are all other pages free of author identification? Does each page have the paper title at the top?

All pages (except figure pages) should be numbered in sequence, starting with the title page.

Cover Letter. A cover letter must be included with the manuscript. The letter should include the author's postal address, e-mail address, telephone number, and fax number for future correspondence and a statement that the manuscript is original, not previously published, and not under concurrent consideration elsewhere.

This letter needs to inform the journal editor of the existence of any similar published manuscripts written by the author.

Title Page and Abstract. The title should be 10 to 12 words in length. The byline should reflect the institution or institutions where the work was conducted. The abstract must be between 100-150 words and up to five keywords may be included after the abstract.

Headings. The levels of headings should accurately reflect the organization of the paper, and all headings of the same level must appear in the same format.

Abbreviations. Any unnecessary abbreviations should be eliminated and any necessary ones must be explained when they first appear.

Abbreviations in tables and figures need to be explained in the table notes and figure captions or legend.

References. References must follow APA format. Please be sure references are cited both in text and in the reference list.

Text citations and reference list entries should agree both in spelling and in date, and journal titles in the reference list must be spelled out fully.References (both in the parenthetical text citations and in the reference list) are to be ordered alphabetically by the authors' surnames.

Inclusive page numbers for all articles or chapters in books must be provided in the reference list.

Notes and Footnotes. The departmental affiliation should be given for each author in the author note. The author note includes both the author's current affiliation if it is different from the byline affiliation and a current address for correspondence.

Footnotes should be avoided unless absolutely necessary. Check to make sure essential footnotes are indicated by superscript figures in the text and explanations are collected on a separate sheet at the end of the manuscript. In the text, all footnotes are to be indicated and correctly located.

Tables and Figures. Every table column must have a heading. Make sure the elements in the figures are large enough to remain legible after the figure has been reduced to no larger than 11 cm.

Lettering in a figure should not vary by more than 4 point sizes of type. Each figure must be labeled with the correct figure number, caption, and short article title.

Minimum file resolution (dots per inch) for printing:
• line art (graphs, drawings) = 1,200 dpi
• halftones (photos) = 300 dpi
• combo line/halftone = 600 dpi

Copyright and Quotations
Written permission to use previously published text, tests or portions of tests, tables, or figures must be enclosed with the manuscript.

Page or paragraph numbers need to be provided in text for all quotations.

GENERAL PRINCIPLES OF OUR JOURNAL

Ethical Standards.

It is the position of this Journal that ethical standards must be upheld and respected in the preparation and realization of research, as well as in the writing style of the papers (avoid prejudice, protect confidentiality of the patients, etc.). Please state in the cover letter that: (a) Ethical standards have been respected, and (b) participants, if any, have signed a consent form (see in the relevant ethical guidelines when written consent is not relevant and required, such as in the use of information already available in the public domain or in research that do not involve human participants). If you are not required in your country or institution to follow a specific code of ethical standards in research, please make sure that your research adheres to one of the five suggested codes: American Psychological Association, IEEE Code of Ethics, Clinical Trial Directives of the European Parliament, or Tri-Council Policy Statement of Ethical Conduct for Research Involving Humans or the Helsinki Declaration. In the paper, the authors must state that the study was conducted in accordance to ethical standards, including the use of a consent form (unless irrelevant). Ethical standards also apply to research that is not conducted with humans (e.g., animal protection protocol) and to publishing issues (e.g., plagiarism, research fraud, authorship misappropriation, duplication of publications).

Conflict of Interests.

It is the position of this Journal that potential conflicts of interest must be made available to the readers since an author's economic and commercial interests may, even involuntarily, bias objectivity. Economic and commercial interests do not necessarily constitute a conflict of interest, and conflicts of interest do not necessarily imply biased research. But the best practice is to disclose activities and relationships that, if known to others, might be viewed as a conflict of interest. Potential, or actual, conflict of interests must be disclosed by **each** author in the cover letter and in the author's note in the article.

Potential conflicts of interests include, but are not limited to: (a) funding or remuneration, including salaries and equipment, from organizations that may gain or lose financially through the publication of the paper; (b) personal financial interests, including receiving royalties or holding stocks and shares in companies that may gain or lose financially from publication; (c) holding patent and patent applications whose financial value may be affected; (d) employment by an organization that may gain or lose from publication of the paper. If you have no conflict of interest to declare, please state "The author(s) report no financial or other conflict of interest relevant to the subject of this article." All contributing authors and reviewers are encouraged to contact the Editor at any stage in the manuscript review process if they believe that a potential conflict of interest requires to be examined.

Upholding the Journal's Standards

1. In the Cover Letter.

Please, state any potential conflict of interest and confirm adherence to ethical standards. If regulatory boards exist in your country, please state that your research was conducted in accordance to which guidelines (e.g., American Psychological Association, Helsinki Declaration, IEEE Code of Ethics, Clinical Trial Directives of the European Parliament, Tri-Council Policy Statement of Ethical Conduct for Research Involving Humans).

Also confirm in the cover letter that the corresponding author assume responsibility for keeping the coauthors informed of the review process, the content of the reviews, and any revisions made. Each author on the manuscript must have a substantive intellectual contribution.

2. In the submitted manuscript and published article:

Please state in the author's note any potential conflict of interests. If you have no conflict of interest to declare, please state "The author(s) report no financial or other conflict of interest relevant to the subject of this article." Financial support for the study must always be disclosed, whether from governmental, nonprofit, or commercial sources. Also in the author's note, indicate that the study was conducted in compliance with an appropriate Ethical Code of Conduct.

The Method section must include a statement describing how informed consent was obtained from the participants (or their parents/guardians) and the project approved by a review board for ethics in research. For research with animals, please state that the study was conducted in compliance with an appropriate code of ethics for using animals in research.

Our Journal pays careful attention to the protection of a patient's anonymity in case reports and elsewhere. Identifying information such as names, initials and hospital numbers must be avoided. Also, authors should disguise identifying information when discussing patients' characteristics and personal history.

GENERAL INFORMATION

Journal of CyberTherapy & Rehabilitation
ISSN: 1784-9934
GTIN-13 (EAN): 9771784993017

Journal of CyberTherapy & Rehabilitation is published quarterly by the Virtual Reality Medical Institute, 64 Rue de l'Eglise, Bte. 3, 1150 Woluwe Saint Pierre, Brussels, Belgium and the Interactive Media Institute, 9565 Waples Street, Suite 200, San Diego, CA 92121, U.S.A.. The journal explores the uses of advanced technologies for therapy, training, education, prevention, and rehabilitation. Areas of interest include, but are not limited to, psychiatry, psychology, physical medicine and rehabilitation, neurology, occupational therapy, physical therapy, cognitive rehabilitation, neurorehabilitation, oncology, obesity, eating disorders, and autism, among many others.

Publishing House
Virtual Reality Medical Institute
64 Rue de l'Eglise, Bte. 3
1150 Woluwe-Saint-Pierre, Belgium
Telephone: +32 2 770 93 33
Fax: +32 2 762 93 33
E-mail: office@vrphobia.eu
Website: www.vrphobia.eu

Interactive Media Institute
9565 Waples Street, Suite 200
San Diego, CA 92121, U.S.A.
Telephone: +1 858 642 0267
Fax: +1 858 642 0285
E-mail: office@vrphobia.com
Website: www.interactivemediainstitute.com

Publisher
Brenda K. Wiederhold, Ph.D., MBA, BCIA

Subscription Information
Rates for subscription are for a volume of four issues.
Print
Europe: €115
International: €145

Online and Print
Europe: €135
International: €165

Online Only
All locations: €75

Subscriptions begin with the first issue of the current volume. No cancellations or refunds are available after the volume's first issue is published. There are also no refunds on single issue purchases.

Advertising
For advertising information, rates, and specifications please contact Virtual Reality Medical Institute, 64 Rue de l'Eglise, Bte. 3, 1150 Woluwe-Saint-Pierre, Belgium, Telephone: +32 2 770 93 33; Fax: +32 2 762 93 33; E-mail: office@vrphobia.eu.

Reprints
Individual article reprints are available from corresponding authors. Please contact the publisher for rates on special orders of 100 or more.

Manuscripts
Please submit electronic copies of your manuscript by visiting www.vrphobia.eu and clicking on the "Submit Paper" link on the right. Information about manuscript submission requirements and formatting can be found at the back of each issue or on the Web site.

The Journal of CyberTherapy & Rehabilitation is indexed in PsycINFO, Gale, EBSCO, Cabell's and Scopus, and is currently under review with Medline.

CONTINUING EDUCATION QUIZ

Prepared by Willem-Paul Brinkman, Ph.D.

Cognitive Engineering of a Military Multi-Modal Memory Restructuring System (Brinkman et al., 83-99)

If you answer 10 out of 12 questions correctly, you will be awarded one CE credit.

1. What type of engineering approach was followed?
a) Waterfall model
b) Rapid application development
c) Situated Cognitive Engineering
d) Participatory design

2) Which of the following functions is not a core function of the Military- 3MR system?
a) Provide remote diagnosis
b) Ensure patient trust
c) Provide a personal approach
d) Provide a structured way of storytelling

3. Which of the following statements are correct?
I) The Military – 3MR system focuses on restructuring and relearning of past events.
II) The Military – 3MR system offers exposure to generic war situations in virtual reality.
a) Both are correct
b) Statement I is correct and statement II is not correct
c) Statement I is not correct and statement II is correct
d) Both are not correct

4. Which visual element was not implemented in the evaluated Military-3MR system?
a) Photos
b) Text
c) Maps
d) Videos

5. Which of the following statements are correct?
I) The story telling experiment showed that compared to no use of the system, the use of the Military – 3MR system led to significant more time references in a story.
II) The story telling experiment showed that compared to no use of the system, the use of the Military – 3MR system led to significant more location references in a story.
a) Both are correct
b) Statement I is correct and statement II is not correct
c) Statement I is not correct and statement II is correct
d) Both are not correct

6. Which of the following statements are correct?
I) Individuals suffering from combat related PTSD always have an undistorted and complete recollection of past events.
II) The timeline (lifeline) in the Military-3MR system is a chronological representation of events in the patient's life.
a) Both are correct
b) Statement I is correct and statement II is not correct
c) Statement I is not correct and statement II is correct
d) Both are not correct

7. Which of the three main human factors were identified by the PACT analysis?
a) Trust, emotion, and cognitive task load
b) Trust, simulation sickness, and emotion
c) Simulation sickness, emotion, and cognitive task load
d) Trust, simulation sickness, and cognitive task load

8. What was not considered as a rationale for including the timeline (lifeline) in the Military-3MR system?
a) To help patients to structure their memory
b) To enhance patients' awareness of the relative small time scale of the traumatic event compared to their entire life
c) To enhance patient's awareness that the traumatic events that dominate their current life all happened in the past
d) To help patients to visualise their past treatment interventions

9. Which of the following statements are correct?
I) The focus of the evaluation of the Military-3MR system reported in the paper was to study the clinical effectiveness of the treatment.
II) The Military-3MR system can only be used in group sessions.
a) Both are correct
b) Statement I is correct and statement II is not correct
c) Statement I is not correct and statement II is correct
d) Both are not correct

10. Which of the following statements are correct?
I) The common desktop allows patients to structure their material initially in groups, and place it on specific days later on.
II) The purpose of the webcam is to record the patient during an exposure session.
a) Both are correct
b) Statement I is correct and statement II is not correct
c) Statement I is not correct and statement II is correct
d) Both are not correct

11. What was the rationale of using scenarios in the design of the Military-3MR system?
a) To record design decisions
b) To establish claims and design requirements
c) To promo the Military-3MR system among patients
d) To promo the Military-3MR system among therapists

12. How was patient trust addressed in the design of the Military-3MR system?
a) Different installation setups of the Military-3MR system
b) The use of password protection
c) Encryption of patients' material
d) Credential information of the therapist

Continuing Education Credits

To qualify for CE credits, readers will need to do the following:
• After reading an article, answer the questions above.
• Send the completed answer sheet, along with payment, to Interactive Media Institute, 9565 Waples St., Suite 200, San Diego, California 92121, U.S.A.
• If you wish to receive a receipt of payment or a copy of the correct quiz answer, please include a self-addressed, stamped envelope.
• The fee is $20 for iACToR members or subscribers and $40 for non-subscribers.
• All CE exams are open-book.
• Credits assigned are based on the length of time required to read each article.

CYBERPSYCHOLOGY & CYBERTHERAPY CONFERENCE

EVIDENCE-BASED CLINICAL APPLICATIONS OF INFORMATION TECHNOLOGY

UNIVERSITÉ DU QUÉBEC EN OUTAOUAIS (UQO), CANADA

If you, or one of your colleagues, are involved in research on the following topics, presenting and attending this international conference is worth considering:

virtual reality
mental health
anxiety
web therapy
videogames
telehealth
brain / computer interfaces
interactive media
eating disorders
human/computer interactions
stroke rehabilitation
addictions
schizophrenia
3D arts
pain and analgesia
neurocognitive assessment
MMORPG
occupational therapy
psychotherapy

ORGANIZED BY INTERACTIVE MEDIA INSTITUTE AND UQO

SUBMISSION DEADLINE:
JANUARY 15TH, 2011

http://www.interactivemediainstitute.com

EARLY REGISTRATION DEADLINE:
MAY 1ST, 2011