Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE DÉFENSE

Initial steps in inducing stress to teach stress management skills to soldiers

Testing stressors, immersive technologies, and avatars

Stéphane Bouchard
(Université du Québec en Outaouais)

François Bernier
Éric Boivin
(DRDC Valcartier)

Defence Research and Development Canada – Valcartier

Technical Report
DRDC Valcartier TR 2011-195
April 2012

Canada

Initial steps in inducing stress to teach stress management skills to soldiers

Testing stressors, immersive technologies, and avatars

Stéphane Bouchard
(Université du Québec en Outaouais)

François Bernier
Éric Boivin
(DRDC Valcartier)

Defence Research and Development Canada – Valcartier

Technical Report
DRDC Valcartier TR 2011-195
April 2012

IMPORTANT INFORMATIVE STATEMENTS

Abstract

Two studies were conducted to test whether: (a) it is possible to stress soldiers using virtual reality, and (b) it is worth developing virtual characters resembling real people experiencing strong pain in order to elicit stress. The first study was conducted with 47 soldiers who played two 3D horror games using three different types of immersive technologies. The second study was conducted with 42 civilians that were familiar with one of two virtual characters experiencing pain. Both studies were effective in inducing stress among participants. Results were supporting of our work devising an effective low-cost and high buy-in approach to assist in teaching and practicing stress management skills.

Résumé

Deux études ont été menées afin de tester si : (a) il s'avère possible de stresser des soldats à l'aide de la réalité virtuelle, et (b) il s'avère pertinent comme stresseur de développer des personnages virtuels ressemblant à des personnes connues par les soldats et ressentant une forte douleur. La première étude repose sur un échantillon de 47 soldats qui ont joué à des jeux d'horreur 3D sur trois différentes plateformes immersives. La seconde étude a été menée avec 42 civils qui étaient familiers ou non avec un personnage virtuel ressentant de la douleur. Les résultats valident les démarches pour développer une solution efficace, peu coûteuse et attrayante pour supporter l'entraînement aux stratégies de gestion du stress.

This page intentionally left blank.

Executive summary

Initial steps in inducing stress to teach stress management skills to soldiers: Testing stressors, immersive technologies and avatars

Stéphane Bouchard; François Bernier; Éric Boivin; DRDC Valcartier TR 2011-195; Defence R&D Canada – Valcartier; April 2012.

Introduction and background. This report describes two studies performed as part of a general research agenda to better train soldiers in applying stress management skills. The long-term goal is to devise a solution where military personnel could be immersed in virtual reality and subjected to moderate levels of stress so they can practice basic coping skills under supervision. Two broad research questions were asked:

(a) Is it possible to stress soldiers using off-the-shelf 3D games and immersive technologies?

(b) When seeing a virtual character expressing pain, are people more stressed when it is the representation of a person they are familiar with than when it is not?

Developing a stress management training (SMT) system and protocol for soldiers would help them cope better with stress experienced in theatres of operation. Using 3D horror games in virtual reality represents an attractive method for soldiers. Study 1 tested the efficacy of two commercial 3D video games (*Killing Floor* and *Left 4 Dead*) to elicit a significant level of stress. It also compared three immersive stereoscopic technologies: a 22-inch stereoscopic monitor, a 73-inch stereoscopic TV, and a CAVE. A sample of 47 soldiers returning from Afghanistan were recruited and randomly assigned to one of five conditions in which they played either the 3D games *Killing Floor* or *Left 4 Dead* while immersed using the different technologies. As a control and reference comparison of induced stress, participants were exposed to a standardised stressful procedure, the Trier Social Stress Test (TSST; Kirschbaum, Pirke & Hellhammer, 1993).

A scenario where the trainee would be helping a severely injured unit member may be an attractive strategy to teach and practice SMT to pre-deployed soldiers. It may allow stressing them to a more optimal level and yet fits well in a culture where virility and “Army Strong” mentality prevail. Study 2 tested whether acute pain experienced by a highly realistic avatar (virtual person) of a person familiar to the participant elicits stronger reactions than an unknown avatar. The 42 participants were immersed in virtual reality using a CAVE-Like system. The first immersion (baseline/control) was with a virtual animal, followed (after random assignment) by immersions involving discussing with a known avatar (i.e., a person they were familiar with) and with an unknown avatar. During the verbal exchanges, the human avatars experienced acute and very strong pain.

Results of Study 1. Repeated measures of analyses of variance (ANOVAs) revealed statistically significant increases in heart rate and respiration rate while playing the 3D games and during the TSST. No significant interactions were found. Increases in physiological arousal were significant when comparing the baseline to the immersion and to the TSST, but not when comparing both stressors. Immersion in 3D games is proposed as a practical and cost-effective option to practice SMT.

Results of Study 2. Repeated measures analyses of variance revealed significant reactions to the stressor of observing an avatar expressing pain, no matter if the avatar represents someone known or unknown to the participant. The impact of the stressor had no lasting effect on the immersions in Virtual Reality (VR). Participants nevertheless found the known avatar more likable and were more empathic toward his pain. The results do not support the added value of adapting avatars to someone familiar to the military personnel being trained in SMT.

Significance and future plans. Results of research conducted under the present contract lead to a more refined and grounded proposition to develop a tool allowing practicing SMT more efficiently. Off-the-shelf 3D horror games can be used with affordable immersive technologies to elicit stress. Yet, new stressors should be included in the stressful scenario and take into account the need to bring trainees outside of the comfort zones they have been extensively trained into. Having to care for someone (a civilian or a wounded soldier) may increase the level of stress, even if the avatar depicts a person the trainee is not familiar with. Further studies are needed to: (a) develop a biofeedback system allowing trainers and trainees to monitor signs of stress while immersed in virtual reality, (b) adapt some SMT currently taught to military personnel to a training protocol in virtual reality, and (c) test preliminarily the efficacy of a program based on an improved virtual reality horror game to both stress military personnel and help them control their stress.

Sommaire

Initial steps in inducing stress to teach stress management skills to soldiers : Testing stressors, immersive technologies and avatars

Stéphane Bouchard; François Bernier ; Éric Boivin ; DRDC Valcartier
TR 2011-195 ; R & D pour la défense Canada – Valcartier ; avril 2012.

Introduction et contexte: Le présent rapport décrit deux études effectuées dans le cadre d'une démarche de recherche plus large visant à entraîner les militaires à utiliser des stratégies de gestion du stress de façon efficace. L'objectif à long terme est le développement d'une méthode utilisant la réalité virtuelle pour induire un stress modéré et permettre aux militaires de pratiquer, sous supervision, leurs habiletés de gestion du stress (*Stress Management Training*, SMT en anglais). Deux questions de recherche globales ont été posées, à savoir :

(a) Est-il possible de stresser des soldats à l'aide de jeux 3D disponibles sur le marché et de technologies immersives ?

(b) Lorsqu'ils voient un personnage virtuel exprimant une forte douleur, les sujets sont-ils plus stressés si le personnage ressemble à une personne qu'ils connaissent que dans le cas contraire?

Enseigner des techniques de gestion du stress aux militaires pourrait les aider à mieux s'adapter au stress vécu en théâtre opérationnel ? L'utilisation des jeux vidéo d'horreur 3D en réalité virtuelle représente une méthode attrayante pour les soldats. L'Étude 1 visait à tester l'efficacité de deux jeux vidéo 3D (*Killing Floor* et *Left 4 Dead*) à induire un niveau modéré de stress lorsqu'utilisés sur trois différentes plateformes immersives stéréoscopiques : un moniteur stéréoscopique de 22 pouces, un téléviseur stéréoscopique de 73 pouces, et un *cave automatic virtual environment* (CAVE). Un échantillon de 47 soldats récemment de retour d'Afghanistan ont été recrutés et assignés aléatoirement à l'une des cinq conditions où ils jouaient à l'un ou l'autre des deux jeux vidéo d'horreur 3D sur l'une des trois plateformes immersives. Afin de contrôler les différences individuelles liées au stress et d'obtenir un point de repère pour comparer le stress ressenti par les participants, ces derniers ont été exposés à une procédure standardisée d'induction de stress (*Trier Social Stress Test*, TSST ; Kirschbaum, Pirke & Hellhammer, 1993).

Créer un scénario où l'on aide une personne grièvement blessée pourrait être une stratégie attrayante pour apprendre et pratiquer en pré-déploiement le SMT. Cette méthode permettrait possiblement d'induire un seuil de stress plus optimal et s'adapte bien à la culture militaire où un esprit de virilité demeure prédominant. L'Étude 2 visait à évaluer si la douleur ressentie par un personnage virtuel très réaliste représentant une personne connue induirait des réactions émotionnelles plus intenses qu'un personnage virtuel inconnu. Au total, 42 participants civils ont été immergés en réalité virtuelle en utilisant un

système immersif ressemblant à un CAVE. La première immersion (niveau de base/contrôle) impliquait un animal virtuel, et était suivie (après assignation aléatoire) d'immersions impliquant une conversation avec un personnage virtuel connu (c.-à-d. une personne qui est familière au participant) ou inconnue. Toutefois, durant les échanges verbaux, l'avatar exprime soudainement une douleur aiguë et intense.

Résultats de l'Étude 1. Les Analyses de variance (ANOVAs) à mesures répétées révèlent une augmentation significative des rythmes cardiaque et respiratoire lorsque les participants jouent au jeu d'horreur 3D et lors de l'entrevue TSST. Toutefois, aucune interaction entre les conditions n'est significative. Les analyses de contrastes confirment que l'augmentation des rythmes cardiaque et respiratoire est significative lorsque l'on compare le niveau de base aux jeux 3D et au TSST mais pas pour la comparaison entre le TSST et les jeux 3D. Ces résultats valident le potentiel des immersions stéréoscopiques dans des jeux d'horreur 3D comme moyens pratiques, efficaces et peu coûteux pour supporter l'entraînement aux stratégies de gestion du stress.

Résultats de l'Étude 2. Les ANOVAs à mesures répétées révèlent une réaction significative au stresseur de voir un avatar exprimer de la douleur, et ce peu importe si l'avatar représente une personne connue ou non. L'impact de ce stresseur ne perdure pas durant l'immersion. Les participants considèrent néanmoins l'avatar représentant une personne connue comme plus sympathique et sont plus empathiques à sa douleur. Ces résultats mettent en doute la pertinence de l'adaptation d'un personnage virtuel à quelqu'un de familier aux militaires en cours d'entraînement à l'utilisation des habiletés de gestion du stress.

Importance et perspectives. Les résultats de ces deux études conduisent à une proposition raffinée et reposant sur des informations empiriques dans le but de développer un outil de formation permettant de pratiquer de manière plus efficace le SMT. Les jeux commerciaux d'horreur 3D peuvent être utilisés avec des technologies immersives abordables afin d'induire du stress. Néanmoins, de nouveaux stresseurs devraient être ajoutés au scénario stressant afin d'amener les soldats à pratiquer le SMT en dehors de la zone de confort dans laquelle ils ont été bien formés. Avoir à prendre soin d'une personne (un civil ou un militaire blessé) pourrait induire un stress additionnel même si l'avatar représente une personne inconnue. D'autres études sont nécessaires afin de : (a) développer un système de biorétroaction permettant aux militaires et aux formateurs d'identifier les signes de stress lorsqu'immergés dans un environnement virtuel, (b) adapter certaines techniques présentement enseignées au personnel militaire dans un protocole de formation pratique en réalité virtuelle, et (c) tester de façon préliminaire l'efficacité d'un programme basé sur un jeu d'horreur amélioré et joué en réalité virtuelle pour à la fois stresser les militaires et les aider à contrôler leur stress.

Table of contents

List of figures

List of tables

Acknowledgements

The authors would like to thank the following people who contributed significantly to the project with their time, their intellectual input, or their devotion (listed in alphabetical order): Guy Barabé, Guillaume Bertinet, Dominic Boulanger, Frédéric Drolet, Stéphanie Dumoulin, Tanya Guitard, Mylène Laforest, Brigitte Morin, Geneviève Robillard, Craig Underhill, Christian Villemaire and Stanley Wany. We also wish to extend our gratitude to the people, soldiers, and civilians, who gave of their time to participate in the studies described in this report.

1 Introduction

Canadian Forces (CF) members, while being deployed in theatres of operation, have been subjected to the physiological and psychological consequences of chronic stressors and, for some of them, to potentially traumatic events. This stress could impede their operational effectiveness and potentially lead to operational stress injuries such as post-traumatic stress disorder (PTSD). A report of the Standing Committee on National Defence (Bernier, 2009) revealed that 13% of the 8222 CF members returning from Afghanistan responded to the questionnaires in a manner consistent with mental health diagnostic and 4% responded in a manner consistent with PTSD diagnostic. Military personnel could be better prepared to cope with combat stress. In a similar manner, a CF report (Zamorski, 2010) on suicide prevention points out the potential of stress inoculation therapy (training) to prevent the effect of work stress, especially prior to an upcoming anxiety-provoking event.

Training in the use of coping skills is important, and not only because it might reduce the risk of developing PTSD and other psychological injuries. Stress and anxiety also reduces operational effectiveness (Britt, Adler, & Castro, 2006). Anxiety, acute stress or operational stress affect information processing, including focus of attention, sensitivity to certain peripheral cues, memory recall and encoding. It also directly influences physiology (e.g., trembling) and emotion regulation, and thus has a strong impact on performance in situations requiring emotional, cognitive and behavioural control such a military operations (Thompson & McCreary, 2006).

The Stress Management Training (SMT) project being carried out by a team of scientists from Defence Research and Development Canada (DRDC) – Valcartier and the Université du Québec en Outaouais (UQO) under the work unit 14da01 is investigating the potential of virtual reality (VR) and video games in training personnel to cope with stress. More precisely, the objectives are to investigate the a) efficiency and the b) technical feasibility of VR and video games in practicing stress coping strategies. This project was initiated in 2008. This report presents the progress to date as of September 2010. First, Chapter 2 presents the SMT concept that is being investigated. Before engaging resources and efforts in a large scale applied efficacy trial where some of the SMT techniques are practiced in VR, experimental research must first answer the following questions: (a) Can military personnel be moderately stressed using off-the-shelf video games as VR scenarios so they could practice SMT? (b) Which immersive display should be used: a small stereoscopic monitor, a large stereoscopic TV or a fully immersive cave automatic virtual environment (CAVE)? (c) What kind of video games would effectively stress military personnel? (d) Assuming medical assistance to an injured person is used as a stressor, is it important for the virtually wounded to be a familiar person? Chapter 3 presents a first study, entitled *Modes of immersion and stress felt during a commercial (off-the-shelf) game (Protocol #L-717),* that addresses questions "a" to "c". Chapter 4 presents a second study, entitled *Empathetic emotion toward the virtual character (avatar) of a known or unknown person (Protocol #L-718),* that addresses question "d".

2 SMT for coping with stress

2.1 Existing programs

A few programs have been developed to improve stress management and coping strategies (also called SMT) among soldiers. With the hope to foster better mental health post-deployment, the Walter Reed Army Institute of Research in the US developed a program called *Battlemind* (Castro, Hoge, & Cox; 2006) to mentally prepare soldiers to the rigours of combat and other military deployments. This program was essentially based on providing unbiased information about what to expect in theatre of operations and maintain positive thoughts during times of adversity. Battlemind was replaced in 2010 by *Resilience Training*.

The most elaborate and theoretically sound SMT program was developed by Routhier (2007) for the SQFT (Secteur Québec de la force terrestre, CF Quebec Area) and is called *Programme d'entraînement à la résilience militaire (PERM)*. This SMT program explicitly aimed at preventing psychological injuries related to operational stress and improving resilience both in theatre of operations and in garrison. It is expected that the PERM will reduce soldiers' suffering and incidence of PTSD as well as increase their performance. It is grounded in a cognitive-behavioural / bio-psychosocial approach and also includes a spiritual dimension. The portion of the program developed for soldiers unfolds in 13 modules delivered in workshops and lectures cumulating in 13 hours of training pre-deployment. The program remains active during the mission, in the form of peer-support. Additional sessions are also delivered post-deployment. The program is structured in such a way that some coping strategies can be used *in situ,* during a life-threatening situation, and others can be used on more diverse occasions. Some structured practice is planned in three modules delivered during an operational exercise. The PERM has been recently replaced/nationalised by the *Road to Mental Readiness* (Chief Military Personnel, 2010) program. It aims at improving the short-term performance during deployment and the long-term mental health outcomes. This program retains many elements of the PERM, including some coping strategies referred to as the "big four". In line with US Navy SEAL programs, the big four comprises goal seeking, mental rehearsal, self-talk, and arousal reduction. One technique is taught to control arousal: diaphragmatic breathing (also referred to as tactical breathing).

Preliminary data on the Battlemind program and the PERM suggest they may have been effective (MHAT-V, 2008; Routhier, 2009). Although data were not collected in controlled trials, they provide empirical grounds to support what theoretical papers and controlled trials have shown in other contexts (Bouchard, 2009). Unfortunately, there are no empirical data on how socially acceptable it may be for military personnel to actually practice the types of stress management strategies taught in SMT programs. Because of virility or potential "Army Strong" mentality, some military personnel may be reluctant to practice emotion regulation techniques. Routhier's (2009) data suggest that too few military personnel used the PERM strategies while they were deployed.

Basic principles of behaviour change (O'Donohue & Fisher, 2009; Goldfried & Davison, 1994; Wolpe, 1990) have extensively shown that simply teaching psychological coping skills during lectures, even in a 13-hour workshop, is not sufficient to lead to effective use of the techniques, especially in stressful situations. Like any behavioural skill, learning in theory how to use a skill transfers very poorly to actual behaviour change in stressful operational situations. Practice under stressful situations is essential.

2.2 Practicing SMT in VR

The current project is based on an extensive literature review (Bouchard, 2009) and explores solutions that fit within the general context of enhancing mental readiness (Armfield, 1994; Thompson & McCreary, 2006), and this should, in its finality, allow training military personnel to be more efficient in psychological coping skills to control their emotions, cognitions and behaviour.

In a previous report (Bouchard, 2009), our team suggested a hybrid strategy that might be cost-effective to teach SMT. The general idea is that basic training assisted with biofeedback and supervision could be conducted in VR using a specially tailored environment, while frequent practice in more challenging and stressful situations could use non-immersive computer games. While providing biofeedback in the head-mounted display or in the peripheral field of view (FOV) in the case of CAVE systems, military personnel can be coached and practice some PERM or SMT strategies in a virtual environment until they are mastered and their perceived self-efficacy is high. The advantages of using VR environments include: increasing motivation, allowing coaching, using manageable stress levels and maximising the generalisability of the training by using stressors that share some similarities with what is experienced in theatre of operations. The combined use of SMT and biofeedback during VR immersion is illustrated in Figure 1.

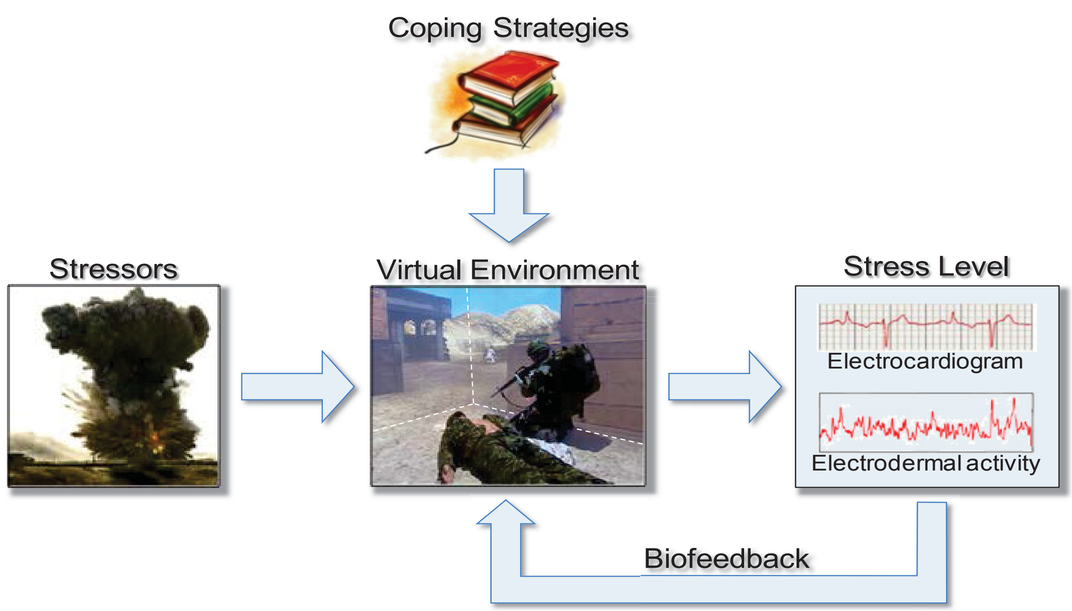


Figure 1: Using biofeedback to practice SMT while immersed in a stressful virtual environment.

The second step takes advantage of already stressful and engaging off-the-shelf 3D games (a buy-in factor for soldiers) to continue applying PERM strategies, sustained by biofeedback information provided in the 3D game console. The military personnel could then bring these “coping-skills enhanced 3D games” with them on missions. On the base, after having performed their duties (during times of rest and recreation), they could play 3D games they already like to play, but with the additions of a biofeedback device that modulates the game based on their stress level. Controlling their stress level while enjoying playing 3D games would allow regular practice of PERM coping skills while in theatre of operations, without having to overtly practice skills that can be perceived as less virile or associated with being weak. Such a strategy may also allow avoiding the problem of habituation to a specific stressful SMT training scenario.

2.3 Biofeedback in video games

For this project, biofeedback consists in inferring the player's emotion from physiological measures such as electrocardiogram (ECG), galvanic skin response (GSR), or respiration changes, and modifying the game accordingly. The objective is threefold. First, a video game could increase the stress of the player by exploiting his or her emotional state. For instance, it has been used to increase the horror affordance in horror video games (Dekker & Champion, 2007). Secondly, the game could be designed in such way that the player would have to manage his or her stress efficiently in order to win, similar to the "relax to win" game created by Bersak et al. (2001). Finally, biofeedback could inform the trainees in real-time about the efficiency of their coping strategies.

The biofeedback could be designed to be implicit or explicit (Fairclough, 2009; Kuikkaniemi et al., 2010) to the player. In implicit feedback, the user may not be aware of the effect of such feedback or cannot identify clearly how it impacts the game. For instance, the trainee's proficiency and ability to aim in first-person shooter (FPS) games or to drive in racing car games can be slightly reduced as his or her stress level increases above a certain limit. By contrast, explicit feedback implements obvious and perceivable cues to the trainee so he or she can learn to influence the game. An instance of explicit feedback could be the player seeing an increase in strength as the sound of his or her heart rate lowers.

2.4 PTSD and SMT

In order to receive a diagnosis of PTSD, an individual must, among other criteria, have been exposed to a traumatic event. According to the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV, American Psychological Association, 2000), an event is considered potentially traumatic when two criteria are met. First, the individual must have "experienced, witnessed, or been confronted with an event that involves actual or threatened death or injury, or a threat to the physical integrity of self or others […]",which is representative of a portion of the military population. Second, the individual must react with a subjective emotional response which "involved intense fear, helplessness, or horror". For military personnel, the trauma can be related either to combat or peacekeeping missions in very difficult and stressful conditions.

Additional organisational and individual factors that have an impact on the mental health of deployed soldiers (peacekeepers and combat veterans) include the pace of military operations (Castro & Adler, 2005), deployment length (e.g., Adler & Huffman, 2005), first-deployment (e.g., Adler & Huffman, 2005), past potentially traumatic deployments (e.g., MHAT-V, 2008; Stetz, Long, Schober, Cardillo, & Wildzunas, 2007), exposure to combat (MHAT-V, 2008), powerlessness (Bartone & Adler, 1994), adverse childhood experiences (e.g., Cabrera, Hoge, Bliese, Castro, & Messer, 2007; Belik, Stein, Asmundson, & Sareen, 2009), coping styles (e.g., Bliese & Castro, 2003; Gil & Caspi, 2006), and gender (e.g., Adler & Huffman, 2005). However, it has been extensively demonstrated that it is not stressors *per se,* but how people react to them (Barlow; 2002; Thompson, Pastò, & McCreary, 2002) that leads to significant psychological injuries, such as PTSD. In addition, soldiers that are seeking psychological help are *perceived* as "weak" (Hoge et al., 2004), so only few soldiers will seek help.

Two potential approaches for reducing the incidence of PTSD are treatment and prevention. The current project focuses on the second option, preventing psychological injuries by increasing resilience. Our approach is to find ways to effectively train military personnel to use psychological coping tools, as opposed to other programs that focus on reducing stress by dealing with logistic issues, mission rehearsal, enforcement of sleep discipline, building confidence that supply and equipment are dependable, ensure that family automobiles are in good repair or resolve major legal issues before leaving for deployment, clearly defining lines of communications and maintaining a good communication flow, etc.

2.5 VR

By definition, VR is "an application that lets users navigate and interact with a three-dimensional computer-generated (and computer-maintained) environment in real time" (Pratt, Zyda, & Kelleher, 1995, p. 17). A distinctive characteristic of VR systems is their ability to immerse the user in an artificial reality and create the sense of presence (feeling of being "there" in the virtual environment; Lombard & Ditton, 1997). VR offers a solution to create standardised situations to practice SMT. Military, medical and entertainment industries were instrumental in the early development of VR applications. The technology gained popularity in clinical psychology in the early 1990s (North, North & Coble, 1996; Stanney, 2002) and psychologists are now using VR for more complex disorders such as PTSD, either as a treatment tool (Wiederhold & Wiederhold, 2008) or to venture into teaching SMT (Stetz, Long, Wiederhold, & Turner, 2008).

A detailed analysis of the published research reports on the use of VR for SMT (Bouchard, 2009) shows that a couple of researchers are beginning to test its usefulness, that our research efforts are timely, and that there is a need for additional experimental research before engaging in applied efficacy trials. The innovative studies by Wiederhold (2005) suggest that VR can be used to train military personnel in performing their duties (e.g., house clearing) in stressful situations. Unfortunately, it is difficult to reach any conclusion from their results given serious methodological and analytical weaknesses. The work of Stetz et al. (2008) aimed at stressing medical personnel rather than at teaching them to cope with stress. Results suggest that VR may induce hostility and anxiety, and that prior deployment may be a factor increasing the impact of the virtual environment. Again, methodological considerations preclude any firm conclusions.

Robillard, Bouchard, Fournier, & Renaud (2003) found a strong relationship between anxiety and presence resulting from the use of VR. This sense of presence could be exploited to induce the stress required for practicing coping strategies. An extensive review of 391 papers of factors influencing presence (Youngblut, 2007) has identified the main factors contributing to the sense of presence: more immersive technologies, higher visual realism, higher resolution, more interactivity, more efficient navigation, and higher quality and spatialised audio. Usually, these characteristics are available in more expensive systems as illustrated in Figure 2. However, this relationship is not linear. For example, a strong feeling of presence might be obtained even when users are immersed in low-quality virtual systems. There may also be plateau and ceiling effects where, before and beyond certain points, the relationship between realism and presence may be weaker.

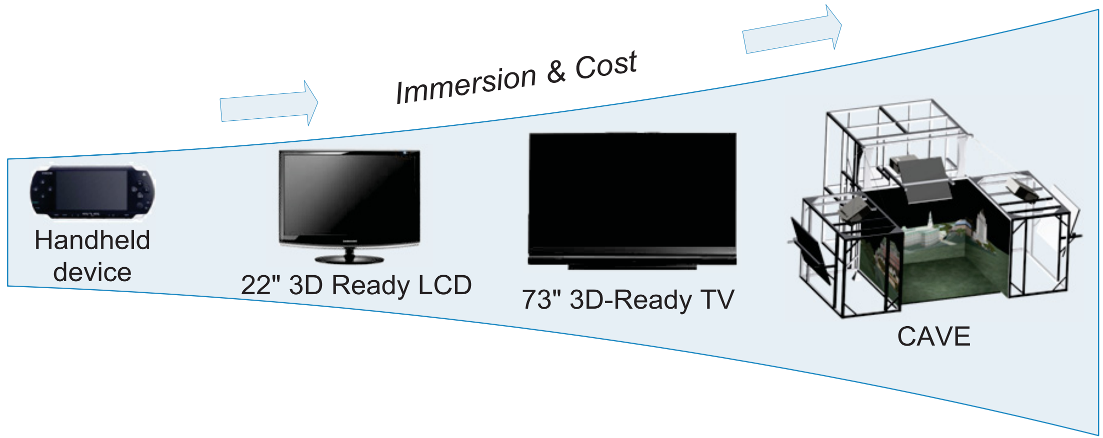


Figure 2: Increased sense of presence with an increased level of immersion and cost.

2.6 Video games

Using VR for SMT implies putting users in a 3D environment so they can feel an experience that is highly engaging yet tailored to a specific goal, such as inducing anxiety (Robillard et al., 2003), therapeutic exposure to anxiety-provoking stimuli (Bouchard, St-Jacques, Robillard, Côté, & Renaud, 2003) and many more. The creation of this environment can be costly in time and money. Creating an environment stressful enough to properly practice SMT strategies is quite a challenge. First, it has to induce a sufficient level of stress. Secondly, the stress must not be so high as to induce undesirable side effects. One must also measure the relative scope of the immersion (that is, the type of technology required) based on the graphic quality of the environment and its capacity to induce stress levels that can be used constructively. It is particularly interesting to see that the 3D game industry is investing enormous sums to reach the objective of stressing the users while walking the thin line between high and reasonable stress.

Games have traditionally been used to facilitate learning (Abt, 1970). More recently, video games have been exploited for education and training in many application domains including health care, emergency services, education, police, military, etc. The military community has been particularly prolific in term of serious gaming with dozens of games like *America's Army* (Zyda, 2005) for recruitment, *Ambush!* (Morrissette & Spaulding, 2004) for tactics, techniques, and procedure learning, the Tactical Language and Culture Training System (Johnson, Marsella, & Vilhjalmsson, 2004) for acquiring communication skills, *Virtual Iraq* (Pair et al., 2006) for treating patients suffering from PTSD, the Dismounted Infantry Virtual Environment (Stone, 2005) for urban combat training, and Virtual Battle Space 1 & 2 (http://www.vbs2.com) for teaching doctrine, tactics, techniques, and procedures.

Video games and more precisely the most stressful genre of video games, the horror genre, can be exploited to study how to artificially induce stress. These games aim mainly at entertaining and frightening, but incidentally they also provide a promising technology to be leveraged. The interest in these games is not in the brutal or bloody scenes they depict but their ability to subtly create suspense and elicit important emotional responses from the players. They try to implement the "art of frightening", often with suggestive staging and appropriate mood-setting. Horror movies preceded the game genre, and the literature about them is rich. However, video games differ significantly because of their interactivity and their non-linear storytelling. Their players can control, within the limits of some rules and a script, the actions of their character and the progress of the storyline(s). Also, video games implement sophisticated narrative approaches

(Adams, 2006), like branching storylines, nonlinear level design, emergent narrative (Sweetser, 2007), and sandboxing.

A survey of papers discussing the art of horror video games and how to develop fearful and stressful video games allowed the identification of some of the most stressful approaches. These have been categorised in three groups as listed in Table 1: 1) the elements populating the video games and the environment, 2) the staging of these elements in a scenario, and 3) specific player-game interaction.

Table 1: Categories of stressors found in video games.

Elements objectively in the game	Static characters and objects: human bodies and faces, monsters, corpses, injuries, human remains, blood, etc.
	Animations: pain, fear, screaming, flight, bleeding, vomiting, mutilating, fighting, shooting, injuries, etc.
	Sounds: cries, moans, environmental sounds, bombs, music
	Scene: dark, unfriendly, creepy, etc.
Staging	Forewarning, surprise, sudden burst, horrific premise, unexpected event, helplessness, revolting moment, altered reality, weird things
Game-player interactions	Lack of control of the character Real world effects Limited fighting abilities Limited FOV

Firstly, video games characterised as more stressful (e.g. horror) often comprise realistic and sometime horrific 3D models of corpses, expressive faces, injuries, monsters and human remains. These games also include animations of pain, fear, cries, flight, bleeding, mutilations, fights, weapon impacts and injuries. In addition, they reproduce stressful sounds like cries, moaning, bombings and music. All these elements are typically located in dark, creepy and unfriendly environments. The high level of realism of all these elements can increase the immersion in the scenario (Ivory & Kalyanaraman, 2007).

Secondly, all these elements are staged to create suspense and fear, with the goal of entertaining the player. A common method exploited in the realm of horror is the forewarning (Perron, 2005b), which consists in announcing an oncoming threat. For instance, *Left 4 Dead* (L4D) uses grunts to herald particularly threatening monsters. Experimental studies (Joanne, Kiemke, & Sparks, 1984; Hoffner & Cantor, 2006) using movies tend to support the effect of forewarning on stress, either during the anticipatory phase and/or during the threatening scene. A general horrific ambiance, often created by darkness (ex: *Vampire: The Masquerade - Bloodlines, 2004*), music and horrific premises, is also a fundamental tool of the genre. Surprise, sudden bursts or unexpected events (ex: seeing in a mirror someone going to attack the player's character from behind like in *Condemned 2*, 2008), revolting moments, weird things (e.g. an upside-down room like in *Siren*, 2004), darkness (Toet, van Welie, & Houtkamp, 2009), and altered realities are all methods exploited by horror video games and well documented in the literature (Perron, 2005a, 2005b; Taylor, 2006).

Finally, the interaction between the player and the video game itself can be exploited as a stressor. As mentioned in an editorial quoted by Perron (2004), a key element of the scare tactic is the ability of the game to "hurt" the player in the real-world by threatening the game progress or by threatening a character in whom the player has invested a lot of time. Although many methods come down to impacting the player in the real world, some are more explicit. For instance, in *Eternal Darkness* (2002), the saved games (which record and store the progress of the player) are being erased as the player's mission progresses. Also, lack of control over the character (e.g.

character freezing in front of a lethal monster like in *Clock Tower 3* (*2003*), limited FOV resulting in enemies potentially being everywhere) and limited fighting ability (e.g. lack of firepower like in *Call of Cthulhu: Dark Corners of the Earth* (2005)) are other tools exploited to create the suspense.

Before being marketed in the United States, as a guideline to prevent users from being exposed to material or stressors that would not be appropriate, games have to be approved by a body confirming they meet the principles of safety and respect for social values: the Entertainment Software Rating Board (ESRB, http://www.esrb.org). This body gives independent evaluations based on game content. Games rated Mature (more violent than those rated for teens but less violent or shocking than those rated for adults), like those belonging to the survival horror genre, are very popular in the game industry and among soldiers. The use of ESRB-approved 3D video games is thus a promising alternative to creating virtual environments because they induce stress without having recourse to stimuli that might be traumatic for certain persons (note that they are designed for a civilian population and that the ESRB is aware that teens play Mature games).

3 Study 1: Modes of immersion and stress felt during a commercial (off-the-shelf) game (Protocol #L-717)

3.1 Background

Video games are designed to elicit various emotions in their players (Freeman, 2003; Frome, 2007). The assessment of these emotions and the physiological changes induced by 3D video games has been carried on in many fields including affective gaming, psychological impacts of violent video games, psychophysiology of video games, user experiences measurements, and unconventional human-computer interfaces.

For example, a study by Sharma, Khera, Mohan, Gupta, and Ray (2006) dealt with the capacity of video games to induce anxiety in an adult population. They initially induced a standardised stressor with the Trier Social Stress Test (Kirschbaum, Pirke, & Hellhammer, 1993) in order to get a baseline level from which they could evaluate the stress-inducing power of the games tested. Since stress is known to cause physiological changes in skin conductance and heart rate, physiological measurements were taken on a sample of 43 participants while they were playing one of three games: *Tetris* (1989), *Black Hawk Striker* (2004) and *Starship Eleven* (2004). The researchers found significant increases in GSR (galvanic skin response), heart rate, and on a stress self-report questionnaire. The only measurement that did not show a physiological increase in participants when they played was blood cortisol level. The authors explained the lack of impact on the cortisol by the length of time it takes to influence cortisol levels. More recently, Nacke and Lindley (2008) measured the gameplay experience in *Half-Life 2* (2003). They reported a correlation between subjective measures and psychophysiological measures, including electroencephalography.

These two studies show that video games could induce stress and physiological arousal, which can be objectively measured physiologically. Many other similar papers documented the potential of video games to induce, and measure, stress (Arroyo-Palacios & Romano, 2010; Bersak et al., 2001; Guger et al., 2004; Herbelin, Benzaki, Riquier, Renault, & Thalmann, 2005; IJsselsteijn, de Kort, Poels, Jurgelionis, & Bellotti, 2007; Nakasone, Prendinger, & Ishizuka, 2005; Ravaja et al., 2004; Ravaja, Saari, Salminen, Laarni, & Kallinen, 2006; Ravaja & Kivikangas, 2008; Tammen & Loviscach, 2010). As noted by Weber, Behr, Tamporini, Ritterfeld, & Mathiak (2009) however, there seems to be a decrease in arousal over an extended period of play, which may explain why many experiments exposed the participants to video games during a limited period of time, usually 10 (Ivory & Kalyanaraman, 2007; Ballard & Wiest, 1996), 15 (Barlett, Harris, & Baldassaro, 2007) or 20 minutes (Anderson, 2004).

None of the previous studies help to identify specific stressors or empirically document how to design stressful video games. Three studies address these questions. First, Jeong, Biocca, & Bohil (2008) found a significant effect from the realism of violence cues, including blood colour and pain sounds, on physiological arousal. Lately, Weber et al. (2009) measured the physiological arousal of players as a function of various video games events in *Counter Strike* (2003) such as being shot, seeking enemies, etc. They found that a higher heart rate was associated with specific events, such as players running out of ammunition, while a lower heart rate was observed during other events, such as opponents appearing in the players' visual field. The increase in heart rate supports the idea that perceived threats in video games can trigger the behavioural activation system (Gray, 1975). The reduction in heart rate documents two points: (a) a stressor could cause brief reductions in heart rate (which is consistent with activation with the behavioural inhibition system, Gray, 1975); and (b) stressors such as uncertainty and lack of control of the situation may have a different impact than direct threats. The third study, from Ivory and Kalyanaraman (2007), shows higher physiological arousal (measured with skin conductance) of players in recent video games compared to those with lower visual fidelity. It is noteworthy that the aforementioned experiments did not expose participants to horror video games, nor did they include military personnel. Therefore, the impact of video game stressors remains to be tested with the horror genre and military personnel.

For the purpose of learning to apply SMT (stress management training), video games can create efficient stressors (the stressful environment/scene) and VR (virtual reality) can convey it efficiently to the player thanks to its high level of immersion. To demonstrate the assets of immersion, Rajae-Joordens (2008) compared the emotional response to video games of participants playing on a stereoscopic monitor with other participants playing on a monoscopic one. The author found that the stereoscopic monitor elicits stronger emotions and a more potent sense of presence. Similarly, Arriaga, Esteves, Carneiro and Monteiro (2008) reported a significant increase in participants' physiological arousal playing video games on a VR system versus its desktop based counterpart. Persky and Blascovich (2006) observed the same arousal in addition to an increase in aggressive feelings. However, none of these studies compared different immersive systems. Also, they did not take into account new low-cost 3D display systems developed for the home theatre, 3D gaming and handheld device markets. A first step in investigating the potential of SMT consists in assessing the effect of the level of immersion on the level of stress.

3.2 Goal of Study 1

The goal of this study is to assess the capacity of video games played with immersive technologies to induce stress in a population of military personnel. Three immersive technologies were compared: (a) a 22-inch stereoscopic monitor, (b) a 73-inch stereoscopic TV, and (c) a "4-wall" CAVE, with images projected on three adjacent walls and a floor. Two video games were also compared: *Killing Floor – Archives* (KF, http://www.killingfloorthegame.com) and *L4D (Left for Dead) – No Mercy* (http://www.valvesoftware.com/). The intrinsic design of these games plus some modifications ensured that L4D comprised more potentially stressful elements.

3.2.1 Immersive setups

Participants were exposed to three different levels of immersion, which differ essentially by their size and the FOV (field of view) of the display relative to the viewer. To a lesser extent, the brightness and the sharpness of the display, the sound system and the control devices were also different. Table 2 presents the FOV of the 22-inch stereoscopic monitor, the 73-inch stereoscopic TV, and the CAVE experimental setups. Stereoscopy was enabled for all setups.

Table 2: FOV and size of the 22-inch stereoscopic monitor, the 73-inch stereoscopic TV, and the CAVE experimental setups.

	22-inch stereoscopic monitor	73-inch stereoscopic TV	CAVE
Size	22"	73"	160" (each screen)
FOV	0.57 sr	1.21 sr	4 sr (approx.)

The following sections present the experimental setups for these three immersive systems. A more detailed description of the hardware and the configuration is given in Annex A. It should be noted that the hardware characteristics were kept as identical as possible, either by using the same model or by choosing components with the same level of performance.

1.1.1.1 22-inch stereoscopic monitor

The first (and least immersive) system was based on a Samsung model 2233RZ 22-inch stereoscopic monitor (http://www.samsung.com), an NVidia 3D Vision kit (http://www.nvidia.com) and a 5.1 Tritton AX Pro headset (http://www.trittontechnologies.com). The 3D Vision driver and the 3D video card generated frame-sequential stereo images that were displayed at 1680x1050@120 Hz by the monitor. Such a high frequency is required to prevent flickering. The participants wore liquid crystal shutter glasses included in the 3D Vision kit to filter the images for each eye. 5.1 Digital Theatre Sound (DTS) was produced by 8 speakers located in the Tritton AX Pro headset. Since the participant position was not tracked and since the 3D Vision driver does not allow modifying the default camera perspective, participants had to stay seated and centered in front of the monitor (Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992) to avoid displacement errors. Otherwise, the 3D sound effects would have been misleading and the visual stereoscopy effect could have caused eye strain, discomfort and incorrect depth cues. The video game was controlled with a standard mouse and a keyboard.

Not only was this screen the smallest among the three used for this experiment, but also the experimental setup ensured the narrowest FOV. Figure 3 shows the physical layout and a picture of the experimental setup. It should be noted that lights were turned off during all gaming sessions, and this for all experimental conditions.

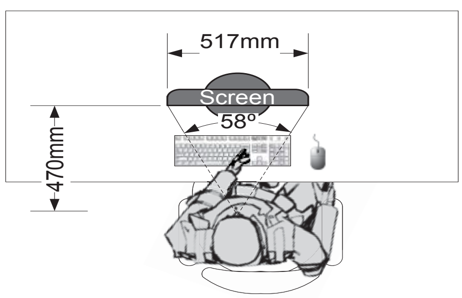


Figure 3: Physical layout (left) and picture (right) for the 22-inch stereoscopic monitor based immersive setup.

1.1.1.2 73-inch stereoscopic TV

The second system was based on a high-definition 73-inch stereoscopic Mitsubishi (model 73735) rear-projection TV and the same 3D Vision and headset models used for the 22-inch stereoscopic monitor-based system. The 73-inch stereoscopic TV requires a special Digital Light Processing (DLP)-based 3D compatible image format, i.e. 60 Hz checkerboard stereoscopic images. The 3D Vision driver implements the SmoothPicture (Hutchison, 2008) algorithm that interleaves the left and right images to produce these checkerboard images. This algorithm was designed to provide the highest image quality within the available bandwidth. However, it also creates visual artefacts as illustrated in Figure 4. Checkerboard/hatching areas are mostly visible at the edge of objects that contrast significantly with the surrounding background. The result is an obviously degraded image but is unlikely to cause any discomfort. Furthermore, this artefact becomes almost unnoticeable during fast action games. The new generation of stereoscopic TVs has solved this problem but these were not available at the time of the experiment.

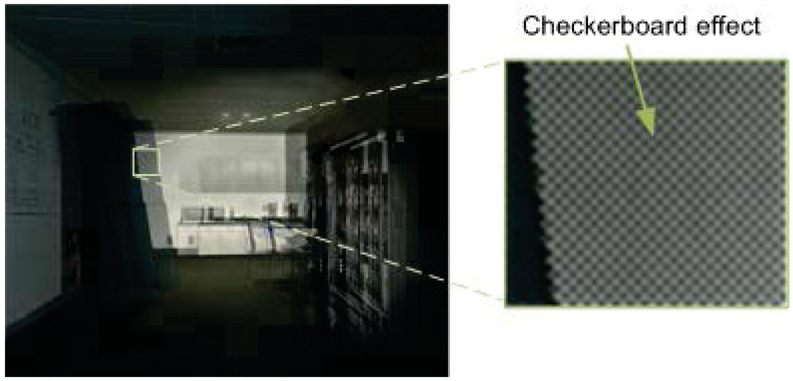


Figure 4: Checkerboard (or hatching) effect caused by the SmoothPicture algorithm.

The resulting 1920x1050 stereo images were frame-sequentially displayed by the TV at twice the frequency provided by the video card (2×60 = 120 Hz). As for the 22-inch setup, the images were filtered by the 3D Vision glasses. Like most rear-projection TVs, this stereoscopic TV suffers from a limited viewing angle. For this reason, viewers must stand at a minimal distance or further to avoid lower brightness in the screen corners. An additional constraint came from the 3D Vision driver that requires the viewer to be located at a minimal distance from the screen to avoid incorrect depth cues and visual discomfort. A third location constraint was imposed by the need to obtain a wider FOV than with the 22-inch setup. A trade-off was found by sitting the participants approximately 920 mm from the TV. Video game controls were based on a standard mouse and a keyboard. Figure 5 shows the physical layout and picture of the 73-inch stereoscopic TV experimental setup.

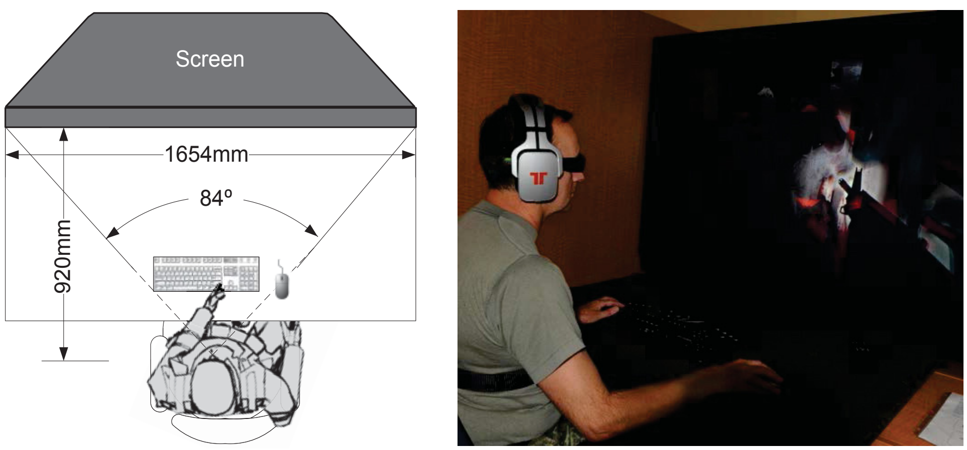


Figure 5: Physical layout (left) and picture (right) for the 73-inch stereoscopic TV based immersive setup.

1.1.1.3 CAVE

The third (and most immersive) system, the CAVE (Cruz-Neira, Sandin, & DeFanti, 1993) consists of a three-wall one-floor stereoscopic immersion environment called the Flex (http://www.mechdyne.com) and designed by Mechdyne Corporation for DRDC Valcartier. Four Digital Projection (http://www.digitalprojection.com) Highlite 8000Dsx+ DLP projectors displayed images at 8000 ANSI lumens and a native resolution of 1400x1050@96 Hz on 10’8” x 8’ screens. Figure 6 shows a back view of the CAVE.

Figure 6: Back view of the CAVE showing the structure, the projectors and the screens.

Participants wore RealD Crystal Eyes 3 (http://www.reald.com) shuttle glasses synchronised with the DLP projectors. An Intersense IS-900 VET motion tracking system (http://www.intersense.com) continuously corrected the image perspective in order to take into account the participant position in relation to the screens (including minor participant's movements).

For this experiment, the participants used an Intersense Wand to move the camera, open doors, switch weapons, jump, seize objects and shoot enemies. Figure 7 shows the physical layout and a picture of the CAVE experimental setup.

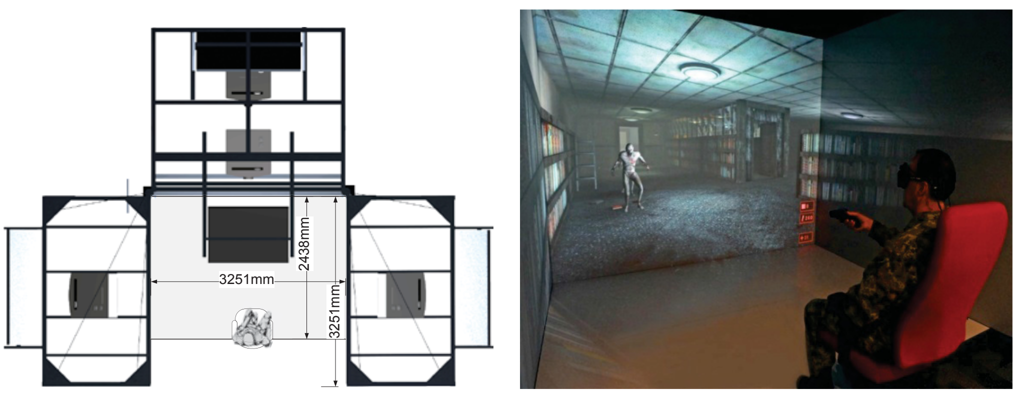


Figure 7: Physical layout (left) and picture (right) of the CAVE experimental setup.

Participants had to remain seated on a desk chair in the CAVE at a specific location but rotations were allowed. Aiming with the Wand was required in order to shoot targets. A simulated rifle, available for the experiment, would have improved the immersion. However, it was discarded because the effort needed to manipulate it would have increased the heart and respiration rates, two important measures of the experiment. Finally, sound was produced by a 700 watts 5.1 DTS system of five speakers and one sub-woofer.

3.2.2 Video game setups

Selection process

Five criteria were retained for the video game selection: (a) stressful, (b) ethically correct, (c) experiment-"friendly", (d) compatible with VR, and (e) adaptable. For the first criterion, the survivor horror genre was identified as the most stressful of all game genres. The evaluation of video games titles was based on the (subjective) ranking of the seven "top-10 scariest video games" done in the last two years according to websites and gaming magazines (About.com, Askmen.com, BrightHub.com, GossipGamers.com, JoystickDivision.com, PopCrunch.com and Telegraph.co.uk). It should be noted that the overall ranking of scariest games is probably inflated by the popularity of some titles. Nevertheless, it was considered the best approach in the absence of a more objective alternative. Second, in order to avoid ethical issues or exposing the participants to inappropriate content, games involving strong language, alcohol and drugs, very intense violence with blood and gore, or sexual themes were penalised while games involving some violence, conflict, weapons and injuries were tolerated. Experimental friendliness included the possibility to minimise the experimentation time and avoid a significant decrease of arousal over gaming time as observed in other similar experiments.

It was decided that participants would be exposed to the video games for a maximum of 20 minutes, preceded by 10 minutes of familiarisation. Consequently, the video games also had to avoid long introductory phases, which are mandatory in the adventure style. Although considered as less stressful than some horror games that do not include elements of FPS (e.g., adventure style), the combined FPS / horror style was identified as being more suitable for the experimentation.

Compatibility with the selected VR systems was an important criterion. All console video games were rejected because of their lack of stereoscopic vision support at the time of the experiment. For the PC-based video games, NVIDIA 3D Vision (http://www.nvidia.com) offers a simple solution to enable stereoscopy, without the need for special game patches. This solution is compatible with a list of certified stereoscopic displays, including the Samsung 22-inch stereoscopic monitor and the Mitsubishi 73-inch stereoscopic TV. Unfortunately, 3D vision is not compatible with the CAVE because it lacks multi-monitor support (at the time of the experiment) and off-axis rendering. As an alternative, Unreal Engine 2 (UE2, http://www.unrealtechnology.com/)-based video games can be adapted for the CAVE. Thus, at least one video game based on UE2 could be retained in order to test it in the CAVE. Finally, adaptability was essential to allow modification of the game for the experimental needs. For instance, strong language dialogs or nude characters had to be removed. Adjusting the difficulty level, the number and type of weapons, and other modifications also needed to be possible.

Reviewed titles

Table 3 shows a list of 12 titles initially identified as most promising for the experiment. Each criterion was rated on a scale ranging from 0 (least) to 5 (most). A group of three video games clearly stands out above the others: *F.E.A.R.*, *Killing Floor* (KF) and L4D. Two similar titles, KF and L4D, were selected in order to ensure uniformity in the experimentation.

Table 3: Review of a selection of video games titles.

	Criterion[A]					
Titles	Stressful	Ethical	Exp. Friendly	VR Compatible	Adaptable	**Total**
Call of Cthulhu: Dark Corners of the Earth (2005)	2	1	1	0	0	**4**
Condemned: Criminal Origins (2006)	3	2	3	0	0	**8**
Dead Space (2009)	5	3	3	2	0	**13**
F.E.A.R. (2005)	5	2	4	4	5	**20**
F.E.A.R. 2 (2009)	2	1	4	2	0	**9**
Fatal Frame II : Crimson Butterfly (2003)	4	5	1	0	0	**10**
Killing Floor (2009)	*1*	*3*	*4*	*5*	*5*	***18***
Left 4 Dead 1 (2008)	*3*	*3*	*5*	*4*	*5*	***20***
Penumbra: Black Plague (2008)	2	4	0	0	1	**7**
Resident Evil 4 (2005)	2	2	3	0	1	**8**
Silent Hill 2 (2003)	4	2	2	0	0	**8**

Selected titles

The two titles retained for the experiment, KF and L4D as illustrated in Figure 8, belong to the FPS – horror genre. The general aim and concept of the two games are rather similar. In both games, the player fights off infected humans (zombies) in various urban locales. The scenario takes place in an abandoned building for KF and a hospital skyscraper for L4D. Game engines used by KF (UE2) and L4D (Source engine, http:// http://source.valvesoftware.com/) are similar regarding graphics with may be a slight advantage for L4D. Aspects related to animation, lighting, texturing and physics are also somewhat superior in L4D. As presented in the next section, the difference lies in the "stressful factors", more present in L4D and increased by some modifications.

Figure 8: Screenshot from L4D (left) from KF (right).

KF

In KF, the player has to clean an area from waves of specimens of zombies. The zombies can bite, attack and eviscerate humans. Some even use weapons such as chainsaws, blades, machine guns, rockets, etc. The zombies wander around and attack in waves of 16 at a time. The high level of modding, available through its UE2 scripting classes, allowed removing some nude characters and strong language dialogs. Like many other UE2 based video games, KF is compatible with the

NVIDIA 3D Vision and can be qualified as excellent. Also, since KF is based on UE2, it could and was easily adapted to the CAVE setup by integrating CORE3D, a middleware developed by DRDC Valcartier. Based on a similar approach proposed by CaveUT (Jacobson & Hwang, 2002), CORE3D adds new essential functionalities for the CAVE. First, CORE3D takes into account the head-tracking location to adjust the image perspective and the Unreal player in real-time. It also synchronises this changing perspective between the four rendering computers. Finally, the module integrates various VR devices with the game.

L4D

In L4D, the player is a survivor from an apocalypse that transformed most of the population into zombies. The player has to survive in a hospital filled with zombies and "special infected". During his journey through the hospital, waves of hordes of up to 12 zombies each try to catch him while he is trying to find the exit. There are five classes of special infected: Smoker, Boomer, Hunter, Tank, and Witch, each of them with different characteristics (i.e., catches and strangles humans with its tongue, blinds opponents by covering them with vomit, climbs walls, jumps very far, incredibly strong, kills with one blow when disturbed by light, etc.). L4D is compatible with the NVIDIA 3D Vision kit and the resulting stereoscopic effect is qualified as excellent. The 3D Vision supports the 22-inch stereoscopic monitor and the 73-inch stereoscopic TV setups but not the CAVE. Consequently, one of the six conditions (L4D – CAVE) was not included in the experiment. Based on Valve's Source engine, L4D provides a good level of modding through its game console and automated script system. This allowed strong language to be removed from the original game. Adapting this engine for the CAVE setup was considered but the task turned out to be too tedious for the available timeframe.

Differences between KF and L4D

Despite the fact that KF and L4D are both *FPS Survival horror* video games, several differences exist between them. Table 4 shows a comparison of both games. It is based on categories presented in Table 1. First, these two games differ in the way their represent the zombies, those in L4D being more human-like than those in KF. Second, L4D exploits more sophisticated staging approaches (game progress, forewarning, surprise/sudden burst) to scare players.

Game configurations and mission

A common mission was defined to ensure a fair comparison between the two games. The participant must ensure that his character survives while looking for the safe room (extraction point). The player is attacked by many infected humans who can be killed with firearms (pistol or rifle). Below is the standardised description of the mission (translated from French) given to the participants at the beginning of the familiarisation phase:

> *A pandemic transformed the population into zombies. You are the only non-infected in the area. You have 20 minutes to find the extraction point located at the other end of the building. You must prevent the infected from biting or catching you. If you die, you go back at the beginning of the map and the time elapsed is lost. The zombies are aggressive: they hide anywhere and constantly try to attack. To defend, you have a submachine gun with limited munitions. To increase your chance of success, collect ammunition. As a last resort, use your pistol, which is provided with unlimited ammunition. Be careful, large hordes and particularly tough infected can attack you.*

KF settings

The game can be played in single-player mode but a modified version of the multiplayer mode was developed for the CAVE setup. Some in-game behaviours were also modified for experimental purposes. For instance, stores selling various weapons, armours and ammunition were disabled between zombie waves to ensure uniformity among participants. The delay between those waves was also reduced to maintain a permanent threat during the experimentation. Finally, the difficulty level was controlled directly in configuration files (see Annex B) by gradually increasing it between waves and keeping only two zombie species with limited abilities: the Clot and the Gorefast.

The selected map was a modified version of KF-Archives representing a two-floor building with a basement and a small suburban surrounding. All weapons and items were removed except the rifle with some ammunition and the pistol with infinite ammo. To make sure the players were always in motion in the environment (to meet enemies), we made them believe that there was a safe room to find (exit), although no such exit existed.

Table 4: Comparison of stressful factors between KF and L4D and significant modifications.

		KF	L4D
Elements objectively in the game	**Static characters & animations**	Zombies are scrawny and misshapen humans. They do not wear clothes. Their behaviour is simple: walking at a constant pace toward the player wherever his location. Zombies are of lower visual fidelity and seem less human than in L4D.	Zombies are more similar to humans and more realistic than in KF. They are based on a sophisticated artificial intelligence with a proximity awareness sense. When a zombie detects any player movement or noise, it attacks by using realistic assault techniques. Each zombie can be fast and agile. There are 5 "special infected" whose mutations grant them special powers that make them more dangerous.
	Sound	Background music was turned off in both games and language issues were resolved by disabling any survivor talk.	
	Scene	Both games/levels were selected and modified to be similar in terms of darkness and to take place in a building with many corridors and levels.	
Staging	**Game progress**	KF uses a sandbox environment where the player is free to go anywhere within the level. This should not be a contributing factor for stress.	In L4D, levels are more linear and constraining. The player must progress along a route to reach a safe room. As a result, the player is more aware of his progress and should be stressed (more at stake) when getting closer to the final objective.
	Forewarning	KF uses grunt to announce an oncoming wave but since these are spawned periodically, this does not provide much forewarning information.	Highly rhythmic music announces an oncoming wave of zombies. Also, the proximity of a special infected, more difficult to kill, is announced with a sound. Finally, the game creates mood and tension with dynamic music.
	Surprise / Sudden burst	The player must kill a fixed number of enemies in each wave. The interval between the waves is constant. Finally, the spawn location is predetermined (one of four fixed areas in the level). This approach eliminates most uncertainty and should reduce the stressfulness of the game.	An unknown number of enemies are spawned at different moments. Common enemies are placed in various locations and in quantities based on the player's current parameters (status, skill, and location). Also, attacks are triggered in order to surprise the player.
Game Player	**Limited FOV**	Sufficient light was provided in all areas so that a flashlight was not required.	The flashlight, required to navigate in dark areas, reduces the player's FOV. This contributes to the uncertainty of the location of enemies (enemies being virtually everywhere).

L4D settings

Usually played as a team of four, the game was modified by disabling the other survivor bots so that the user was alone against the zombies. To counteract this modification, the difficulty level of the game was eased by reducing the strength and the health of the special infected and by halving the size of the infected hordes spawned by the game.

The fourth episode of the No Mercy campaign (Hospital) was selected. All seizable items were removed from the original map except for the rifle firearm and the ammunition near the elevator leading to the top of the skyscraper. The extraction point remained located at the end of the level, at the top of the skyscraper. The linear design of this map combined with many game restarts following each player death ensured the gameplay uniformity for the experimentation. Audio comments and advice from the instructor were disabled. All changes were wrapped up in a script automatically loaded at each play-through (See Annex B).

3.3 Participants

The initial sample was comprised of 57 male Army soldiers from CF Base Valcartier whose unit volunteered to send participants for an experiment on virtual reality and stress. A procedure was devised to guarantee freedom to refuse to participate in the study. Once on-site for the experiment, each soldier had the right to refuse to participate. In that case, he had to stay in the research facility and read magazines. None of the volunteers referred to the study initially refused to participate, but four participants withdrew from the study during the Trier Social Stress Test (TSST). With the addition of two participants who did not show up and four who were excluded based on the selection criteria (e.g., having pre-existing migraines conditions), the final sample consisted of 47 participants immersed in virtual reality.

The mean age of the participants was 25.9 (SD = 5.2). All of them had been in a rotation in Afghanistan in the six months preceding the experiment and only one had not been exposed to combat. The ranking distribution of the participants was 44% privates, 44% corporals and 12% master corporals. A minority of the participants (18%) had not completed their secondary education degrees, 54% had a secondary degree as their highest school diploma, 26% had a college degree and 2% had a bachelor’s degree. The majority of participants were not engaged in a stable relationship with a partner (61%) and half the sample (52%) reported an annual income of less than 50 000 $ per year. Three-quarters (78%) of the sample had never played the games used in this study. The average immersive tendencies found in this sample using the Immersive Tendencies Questionnaire was 66.7 (SD = 11.43), which is representative of the general population.

The following exclusion criteria were set *a priori* for ethical reasons and, if met, led to the rejection of the participant: showing signs of suffering from schizophrenia, psychotic disorder or PTSD. This was assessed using the respective sections of the Structured Clinical Interview for DSM-IV Disorders Non-Patient Edition (SCID-NP; First, Spitzer, Gibbon, & Williams, 2000; see below for more information). These criteria were assessed by research assistants (Ph.D. candidates in psychology who had experience in administering the SCID-NP, including assessing these criteria, and who had been trained in handling clinical cases). Also, people suffering from the following physical problems were excluded since they were at risk of having significant cybersickness during the immersion (Kennedy, Lane, Berbaum, & Lilienthal, 1993): vestibular (or inner ear) problems, recurrent migraines, epilepsy, balance, cardiac or ocular problems, frequent and intense motion sickness in transports. People with no stereoscopic vision (as assessed with the Randot stereoscopy test) were excluded since people with this condition cannot perceive depth during stereoscopic immersions.

All participants were informed of the details of the study (except about the nature of the TSST, see below) and the risks associated with the experimental protocol. They were told that they were free to refuse to participate and could withdraw their consent without prejudice or hard feelings at any time, before being asked for their written informed consent. The project was approved earlier by the ethics review board of DRDC and the Université du Québec en Outaouais (UQO).

3.4 Experimental protocol

3.4.1 Baseline

The experiment was conducted within a week of intense experimentation at DRDC Valcartier and was conducted by a research assistant trained in using VR. When participants arrived, they were asked to read and sign the consent form. If they agreed to volunteer for the experiment, they filled out questionnaires providing general information (sociodemographic data, immersive tendencies, physical health, previous exposure to combat, etc.). This step took about 5 minutes:

- Sociodemographic Information (pre-test) : an information request about age, sex, socio-economical status, etc., was completed for statistical purposes.
- Immersive Tendencies Questionnaire (ITQ; Witmer & Singer, 1998): the Immersive Tendencies Questionnaire (French Canadian version validated by the Laboratoire de Cyberpsychologie of the UQO) was administered at the end of the intake interview. The ITQ has 18 items rated on a 7-point scale (1: *Never*, 7: *Often*) which provide a *Total* score and four subscale scores: *Focus* (the ability to concentrate and to ignore distractions), *Involvement* (the feeling of being caught up by stories and movies), *Emotion* (the intensity of the emotions evoked by stimuli such as movies), and *Play* (the frequency of playing video games).
- State-Trait Anxiety Inventory (STAI-Y1) - French-validated version by Gauthier & Bouchard (1993). The *State-Trait Anxiety Inventory* is divided in two sub-scales (state anxiety and trait anxiety) of 20 items each, with various questions about how the participant feels the same day or in general. Only the State Anxiety form was used in this study to measure the level of anxiety and stress at baseline, as well as before and after the TSST and the experimental immersion.
- Simulator Sickness Questionnaire (SSQ; Kennedy, Lane, Berbaum, & Lilienthal, 1993): this questionnaire assesses the level of discomfort before and after the immersions, using 16 different items related to symptoms (headache, eyestrain, nausea, etc.). It was administered with the other baseline questionnaires, after the standardised stressful procedure and post immersion. Raw total scores (i.e., not weighted) were used.

In order to test the ability to perceive binocular disparity in distance from static stimuli (stereoscopic vision for depth perception), the participant performed a stereopsis test (Randot SO-002, Stereo Optical Company Inc., Chicago, IL; www.stereooptical.com) that consisted in observing various geometric shapes (between 400 and 20 seconds of arc in apparent size) and animals (400 and 100 seconds of arc) in front of random dot backgrounds while wearing polarising glasses. Participants must recognise the stimuli they perceived as 3D (i.e., floating above the test board). This test took about 5 minutes and was scored according to the procedures described in the test manual.

Once the first set of measures was completed, a research assistant installed a physiology monitoring device on the participant's chest to take physiological measures of heart rate and respiratory rate. This step took 5 minutes, including the baseline. These measures allowed us to document objective signs of stress among the participants. Heart rate and respiration rate were set as the main dependent variables in the study since physiological data are not influenced by social desirability and the ability to recognise subjective signs of stress.

Heart rate, respiratory rate and amplitude of body movements: The most often used objective measure of stress is the heart rate, which is measured by an ECG. When an individual is confronted with a stressor, heart rate increases. Respiratory rate is also an index of stress level since respiration rate increases with stress. Both physiological parameters were collected using the Zephir Bioharness garment and device (Biopac Systems, Inc., Goleta (CA), USA, http://www.biopac.com). Data gathered with the Bioharness was recorded continuously during the entire experiment and logged in the device's internal memory (i.e., the wireless function was not used). The time in hours and seconds was manually recorded to document the beginning and end of each phase of the experiment. Data was acquired at a sampling rate of 250 MHz.

The ECG sensors in the garment detect electric pulses in the heart, which is filtered in the device with a highpass filter at 15 Hz and a low pass filter at 78 Hz. The low pass filter cut-off enabled precise heart rate measurement under vigorous activity. The Bioharness respiration sensors in the garment detect breathing rates for up to 70 breaths per minute (0.05 Hz to 1.166 Hz). Accelerometers in the device measured movement (acceleration and vector magnitude) in the X, Y and Z planes, allowing analyses of global body movements by the participants during the experiment. Peak acceleration is the maximum 3-axis acceleration magnitude achieved during one second and is measured in g (gravitational force), with values ranging between 0 and 5.7 g. Vector magnitude is the average vector magnitude achieved in the previous second and varies between 0 and 5.7 g.

Physiological data were analyzed with the AcqKnowledge Software from Biopack (version 3.9.1, Biopac Systems, Inc., Goleta (CA), USA, http://www.biopac.com). Markers were put in the series of data to define the beginning and end of each phase of the experiment, allowing the selection and analysis of specific blocks of physiological data. To provide a safe margin of error in mismatch between the Bioharness and the experimenter’s time logs, the first and last 30 seconds of each event in the experiment were excluded from the analyses. Due to technical problems during data acquisition, heart rate data provided by the software (channel 4) was recalculated for every participant using the original raw R-R intervals (interval between ventricular depolarisations). Rejection of lost data, spurious data points and artefacts for both heart rate and respiration rate were performed manually for each participant based on the available reliable data. The sample size is reported with each set of analyses to document the exact number of participants without missing data.

A baseline was recorded for both physiological parameters during three minutes before the experimentation began, while the participant was sitting on a chair with his eyes open. The physiological parameters were recorded continuously, including during the standardised stressor task and the experimental immersion.

Participant’s performance during the game was assessed by the number of zombies (i.e., any kind or enemy characters) killed during the experimental immersion and the number of times they were killed while playing. Note that when participants were killed the computers automatically launched a new game, always beginning at the same starting point. The frequency of kills and being killed were crude indices of performance since the games did not unfold in a standardised manner for each participant and due to discrepancies between games. Yet, the conditions were deemed similar enough to warrant the use of performance as an exploratory variable for post-hoc analyses.

The SCID-NP was administered to study participants to exclude participants suffering from axis-I psychiatric disorders mentioned in the exclusion criteria (i.e., schizophrenia, psychotic disorder and PTSD). This step lasted 5 minutes and was performed by a research assistant extensively trained in the use of the SCID and the management of clinical cases.

The TSST (Kirschbaum et al., 1993) was conducted after the SCID-NP. The TSST involved the following standardised steps designed to induce stress:

- the participant entered a room where three people (the interview panel) were sitting at a table, close to a small video camera (see figures 9 and 10 below) and was asked to stand in front of the panel;
- a member of the interview panel asked the participant to assume the role of a job applicant who was invited to a personal interview with the company's staff managers;
- the topic of the job interview was told to the participant (selling dermatological creams, toothbrushes, shoes, vacuum cleaners, household appliances; each of the five topics was used on a different day to avoid contamination from participants who underwent the TSST on a previous day);
- the participant was told that after a 5-minute preparation period he would have to give his speech in front of the interview panel for five minutes and convince them he was the perfect applicant for the job;
- the panel was introduced as being especially trained to monitor non-verbal behaviour and that a voice frequency analysis of non-verbal behaviours would be performed on the video recording;
- after the five minutes of preparation in another room, the participant came back standing in front of the panel for his five-minute talk;
- if the participant finished his talk before the end of his five minutes, a member of the panel said abruptly "You still have some time; please continue" or, if the participant finished a second time before his allocated time, the member of the panel asking questions said nothing for 20 seconds and then ask prepared questions;
- after the allocated five minutes, another member of the panel told the participant to count down from 2083 in increments of 13 as quickly and accurately as possible;
- on every failure the panel member said abruptly “Error! 2083!” and the participant had to start over again.

Figure 9: Trier Social Stress Test.

Figure 10: Trier Social Stress Test (interview panel).

The goal of administering the TSST was to obtain a standardised baseline and comparison point in the participant's reaction to a moderate stressor. The TSST allowed inducing stress so a repeated measures analysis could control for the impact of individual differences in each participant's reaction to stressors when comparing with the stress experimentally induced by the immersions. Since it was considered as a baseline measure of reactions under stress, it was systematically administered before the experimental immersions. This standardised procedure to measure stress is well documented (Kudielka, Hellhammer & Kirschbaum, 2007), does not induce any negative consequences because the stressor remains low, and is preferred to other methods like the cold-pressure test (i.e., immerse a hand in very cold water).

Administration of the last part of the TSST (steps h and i, where the participant must count backward) was creating significant frustration in the participants, and experimenters became suspicious that some participants may have memorised the task based on information provided by colleagues who underwent the TSST on a previous day. It was decided to modify the second part of the TSST during the course of the study. The duration was reduced from five to three minutes and the starting number was changed every day by a few units (between 2081 to 2085). Due to the lack of consistency in the second part of the TSST, the obvious frustration induced, and the very high rate of missing and spurious data, only physiological data gathered during the first portion of the TSST (the job interview in front of a panel) were analyzed.

3.4.2 Experimentation

After the TSST, participants were brought in a different room and told to which condition they had been randomly assigned. The participants were then taught how to use the controls for the game they would play. This familiarisation phase was conducted using a small 17-inch monoscopic monitor and a sheet detailing clearly the goal of the mission and how to use the peripherals (e.g., how to shoot at zombies). Participants played during 10 minutes to learn the different controls and to be introduced to the game, with the lights in the training room remaining turned on.

For the experimental immersion, the participants were brought to one of three experimental rooms and played for 20 minutes (duration based on Weber et al., 2009) at a commercial video game (L4D – No Mercy or KF – Archives) where different immersive technologies were compared with the surrounding lights in the room turned off: (a) 22-inch stereoscopic monitor (low immersion), (b) 73-inch stereoscopic TV (middle immersion), and (c) a CAVE (high immersion).

At the end of the experimental immersion, participants were asked to fill in questionnaires. This step took 10 minutes:

- State-Trait Anxiety Inventory: (see description, page 22).
- UQO Presence Questionnaire: This questionnaire has four items assessing the sense of presence after an immersion and one item assessing the level of cybersickness.
- Simulator Sickness Questionnaire (see description, page 22).
- Personal impressions from the experiment: This questionnaire was developed for the purpose of the study and asked participants to answer on a zero to ten scale the following questions: 1) "To what extent did you feel stressed during the immersion?"; 2) "Did you enjoy the immersion?"; 3) "Do you think this type of immersion could be useful to practice stress management skills?"; 4) "Which of the following items did you think was more stressful" (sounds, image quality, darkness).

3.5 Data analysis

The data were analyzed with repeated measures analyses of variance (ANOVAs) with 3 Times (Baseline, TSST/Speech, Experimental Immersion) and 5 Conditions (see Table 5 below). Participants were randomly assigned to each condition upon their arrival in the lab, prior to verification of the exclusion criteria. Mixed model repeated measures ANOVAs were in order since the proposed research design was unusual, in the sense that not all conditions were distributed among the cells. A trade-off was necessary to accommodate our need of testing the games in a very immersive environment such as a CAVE. Power analyses were conducted following Cohen (1988).

Table 5: Distribution of the 47 participants among the conditions.

Game / visual quality	***Immersiveness of the technology***		
	22-inch Monitor	***73-inch TV***	***CAVE***
L4D	$n = 8$	$n = 11$	
KF	$n = 10$	$n = 9$	
KF			$n = 9$

Three sets of repeated measures ANOVAs were conducted. First, for each set of key variables a comprehensive analysis was performed and included all five conditions. Two additional analyses were conducted to refine further the comparisons (and gain in the number degrees of freedom for the ANOVA). The second set of analyses consisted of a comparison between screen sizes (i.e., 22-inch monitor vs. 73-inch TV, with participants in the CAVE condition being excluded and game type combined within their respective monitor format condition) and the third set of analyses compared the two games (i.e., L4D vs. KF, with participants in the CAVE condition being excluded and screen format combined within their respective game condition). Several additional exploratory comparisons were performed (e.g., comparison of CAVE vs. other conditions, analysis of covariance (ANCOVAs) using vector magnitude or peak acceleration, prior experience with gaming, military rank, performance during the games): none revealed significant differences that were not detected by the main analyses. The additional exploratory comparisons are therefore not reported in this report.

3.6 Results

The main dependent variables for this study were the heart and respiration rates. Descriptive results are reported in Tables 6, 8 and 9 for all conditions, combined groups for monitor sizes, and combined groups for games played, respectively. The heart rate revealed an increase in stress levels from the baseline to the TSST and to a lesser extent from the baseline to the experimental immersion. The repeated measures ANOVA was significant (see Table 7) for the main effect of Time, and contrasts showed that the significant increase in heart rate came from differences from the baseline to the TSST and from the baseline to the experimental immersion, while playing the games was not significantly more or less stressful than the TSST [$F(1,31) = 0.04$, *ns*]. There was no significant difference between the five conditions, either on the Condition main effect or on the Interaction effect. In fact, the effect size for the interaction was small (partial eta-squared = 0.06, Cohen's f = .25) and to reach statistical significance the study would have needed 40 participants per cell. The effect size was even smaller when looking specifically at the interaction between the TSST and the experimental immersion (partial eta-squared = .03, Cohen's f = .17, 105 participants per cell needed to reach significance at the .05 level with a power of .80).

Table 6: Heart rate - means and standard deviations for all five conditions (N = 36).

	Baseline Mean (SD)	Mean (SD)	Mean (SD)
L4D – 22 inch (n=4)	85.31 (13.26)	108.96 (25.68)	98.56 (27.58)
L4D – 73 inch (n=10)	88.28 (8.41)	93.35 (13.96)	94.97 (23.10)
KF – 22 inch (n=7)	85.09 (15.55)	94.69 (22.57)	93.60 (31.58)
KF – 73 inch (n=8)	83.44 (10.17)	89.07 (14.25)	87.78 (18.73)
KF – CAVE (n=7)	80.09 (16.40)	97.06 (9.80)	103.18 (21.32)

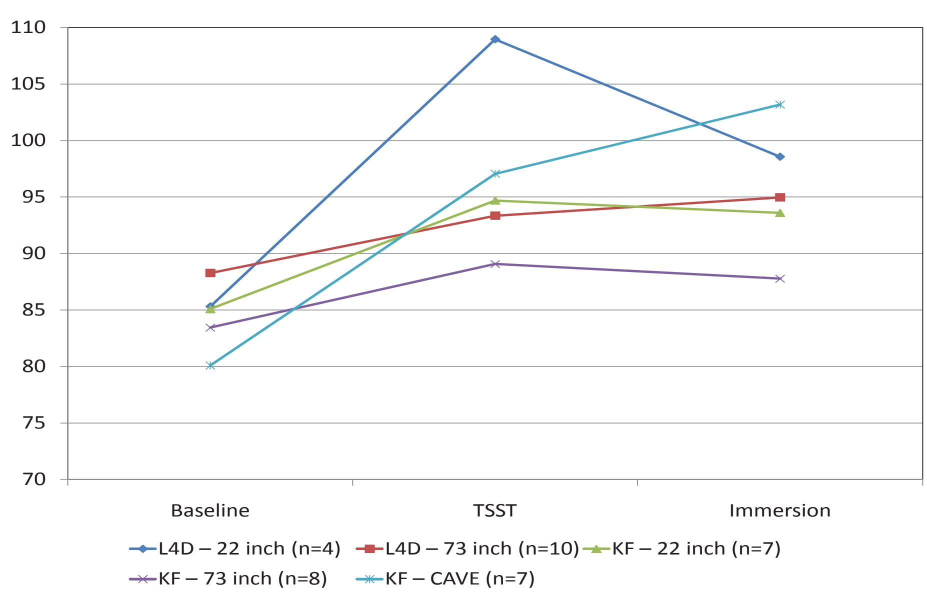


Figure 11: Heart rate for all five conditions, monitor sizes and games.

Table 7: Heart rate - ANOVAs for comparisons between all five conditions, monitor sizes and games.

	Time F	Condition F	Interaction F	Time Contrast: Baseline vs. TSST	Time Contrast: Baseline vs. Immersion
Comparison between all conditions (n=36)	4.53*	0.73	0.48	15.95***	5.21*
Comparison between monitor sizes (n=29)	2.24	0.78	0.44	7.87**	2.19
Comparison between games (CAVE excluded) (n=29)	1.92	1.52	0.05	6.27*	2.02

Note. *p < .05; **p < .01; ***p < 0.001

Additional analyses comparing monitor sizes and games (see Tables 8 and 9 for descriptive results) revealed a pattern that was consistent with the general analysis, although the sizes of the samples were too small—partly because of the large number of missing data on the heart rate measure. The contrasts confirmed the power of the TSST to generate stress, and that there was no significant effect of the monitor sizes (effect sizes: partial eta-squared = .007, Cohen's f = .07, expected n per cell > 400) or the games (effect sizes partial eta-squared = .002, Cohen's f = .05, expected n per cell > 1550) on the experimental immersion.

Table 8: Heart rate - means and standard deviations for 22-inch and 73-inch Monitor-size conditions (N=29).

	Baseline Mean (SD)	TSST Mean (SD)	Immersion Mean (SD)
22-inch (n=11)	85.17 (14.07)	99.88 (23.57)	95.41 (28.86)
73-inch (n=18)	86.13 (9.28)	91.45 (13.84)	91.78 (20.99)

Table 9: Heart rate - means and standard deviations for L4D and KF games excluding KF CAVE condition (N=29).

	Baseline Mean (SD)	TSST Mean (SD)	Immersion Mean (SD)
L4D (n=14)	87.43 (9.57)	97.81 (18.46)	95.99 (23.41)
KF (n=15)	84.21 (12.50)	91.69 (18.12)	90.50 (24.73)

The analysis of the respiration rate was more robust since it was performed on almost the entire sample. Examination of the results (see Table 10) showed an important increase in respiration rates both during the TSST and the experimental immersion. The repeated measures ANOVA confirmed a significant Time main effect for breathing rate, and contrasts revealed that the increase was related to significant differences between the baseline and the TSST as well as the baseline and the experimental immersion (see Table 11). Interestingly, the contrast between the TSST and the experimental immersion was also significant [$F_{(1,37)} = 40.8$, $p < .001$], revealing that people breathed faster while playing the games than during the TSST. The comparisons between the groups, either on the Condition main effect or on the Interaction effect, did not reveal any significant difference. The effect size of the interaction was small (partial eta-squared = .07, Cohen's f = .26; expected *n* for a significant difference at .05 with a power of .80 = 40 per cell), either when looking at the entire experiment or when contrasting the TSST with the experimental immersion.

Table 10: Respiration rate (in breaths/min) - means and standard deviations for all five conditions (N = 42).

	Baseline Mean (SD)	TSST Mean (SD)	Immersion Mean (SD)
L4D – 22 inch (n=7)	2.32 (5.92)	15.37 (3.46)	20.12 (3.62)
L4D – 73 inch (n=10)	8.45 (16.55)	12.93 (3.49)	17.29 (7.16)
KF – 22 inch (n=8)	5.58 (7.76)	15.69 (2.72)	16.96 (3.70)
KF – 73 inch (n=9)	10.43 (13.22)	14.50 (3.51)	18.67 (4.46)
KF – CAVE (n=8)	5.68 (7.20)	12.76 (4.46)	18.14 (2.86)

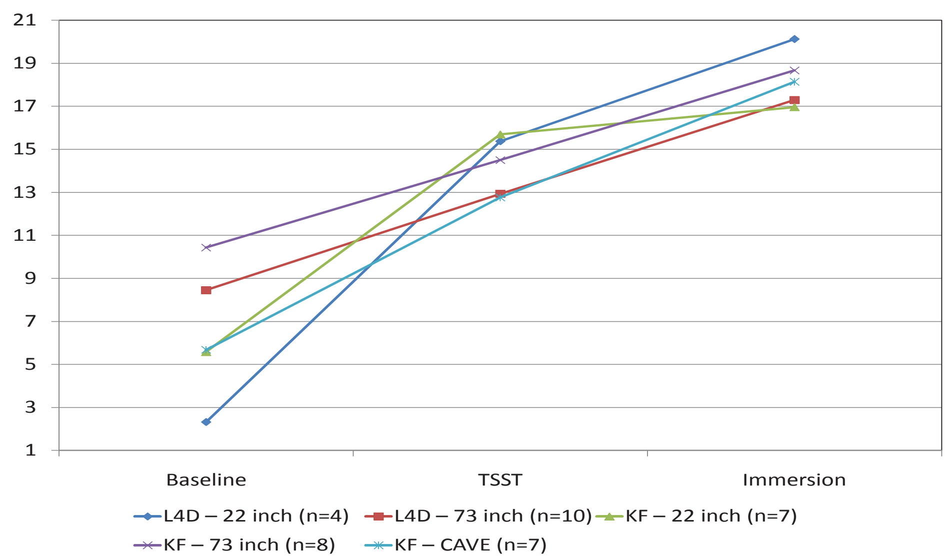


Figure 12: Respiration rate (in breaths/min) for all five conditions, monitor sizes and games.

Table 11: Respiration rate - ANOVAs for comparisons between all five conditions, monitor sizes and games.

				Time contrasts	
	Time F	**Condition F**	**Interaction F**	**Baseline vs. TSST**	**Baseline vs. Immersion**
Comparison between all conditions (n=42)	25.88***	0.40	0.73	15.7***	34.15***
Comparison between monitor sizes (n=34)	18.53***	0.24	1.56	12.55***	23.67***
Comparison between games (CAVE excluded) (n=34)	16.84***	0.38	0.26	10.46**	22.17***

Note. **p < .01; ***p < 0.001

Additional analyses comparing monitor sizes or games (see Table 11) were very consistent with the general analysis of the entire sample, supporting the contention that both the experimental immersion and the TSST induced significant stress levels (see Tables 12 and 13) but that differences in monitor sizes or games did not matter much. The contrasts for the interactions between baseline, experimental immersion, and the conditions for monitor sizes (effect sizes partial eta-squared = .03, Cohen's f = .19, expected n per cell > 90) or games (effect sizes partial eta-squared = .01, Cohen's f = .11, expected n per cell > 300) were very small.

Table 12: Respiration rate - means and standard deviations for 22-inch and 73-inch Monitor-size conditions (N=34).

	Baseline Mean (SD)	**TSST Mean (SD)**	**Immersion Mean (SD)**
22-inch (n=14)	4.35 (7.09)	15.71 (3.02)	18.33 (4.02)
73-inch (n=20)	8.92 (14.44)	13.65 (3.40)	18.04 (5.77)

Table 13: Respiration rate - means and standard deviations for L4D and KF games excluding KF CAVE condition (N=34).

	Baseline Mean (SD)	**TSST Mean (SD)**	**Immersion Mean (SD)**
L4D (n=17)	5.93 (13.30)	13.94 (3.59)	18.45 (5.98)
KF (n=17)	8.15 (10.96)	15.06 (3.12)	17.86 (4.09)

Since physical movement can affect both heart rate and breathing patterns, an examination of vector magnitude and peak acceleration in each condition (see Tables 14, 15 and 16) was in order. Interestingly, the repeated measures ANOVAs for both variables confirmed a significant decrease in motion from the TSST to the experimental conditions, with no significant differences between the conditions. Such an effect was expected since participants were standing up—and even sometimes walking—during the TSST but were seated during the experimental immersion. Yet, it was necessary to confirm that increases in heart and respiration rates were not caused by an increase in physical activity.

Table 14: Vector Magnitude (in g) - means and standard deviations for all five conditions (N=43).

	TSST Mean (SD)	**Immersion Mean (SD)**
L4D – 22 inch (n=7)	0.03 (0.02)	0.01 (0.01)
L4D – 73 inch (n=10)	0.02 (0.01)	0.01 (0.004)
KF – 22 inch (n=8)	0.02 (0.01)	0.01 (0.004)
KF – 73 inch (n=9)	0.03 (0.01)	0.01 (0.01)
KF – CAVE (n=9)	0.03 (0.01)	0.01 (0.01)

Table 15: Peak Acceleration - means and standard deviations for all five conditions (N=43).

	TSST Mean (SD)	Immersion Mean (SD)
L4D – 22 inch (n=7)	0.07 (0.03)	0.03 (0.02)
L4D – 73 inch (n=10)	0.06 (0.03)	0.03 (0.01)
KF – 22 inch (n=8)	0.06 (0.04)	0.02 (0.01)
KF – 73 inch (n=9)	0.06 (0.02)	0.03 (0.02)
KF – CAVE (n=9)	0.06 (0.02)	0.03 (0.01)

Table 16: Peak Acceleration and Vector Magnitude - ANOVAs for comparisons between games (N=43).

	Time F	Condition F	Interaction F
Vector Magnitude	55.92***	0.77	0.37
Peak Acceleration	55.32***	0.37	0.19

Note. ***p < 0.001

The analysis of subjective data collected with the State scale of the State-Trait Anxiety Inventory was interesting but must be interpreted with caution. Indeed, the average scores reported by the participants (see Table 17) were extremely low; in a range after the baseline and the immersion that is expected during relaxation exercises. The repeated measure ANOVA was generally supporting the results found with the physiological data for a significant Time main effect and no difference between the five conditions (see Table 18). The same was found when looking only at the comparison between the monitor sizes (see Table 19) or the games (see Table 20). The contrasts systematically supported the increase in self-assessed anxiety from the baseline to the TSST, but not between the baseline and the experimental immersion. In fact, the contrast between the TSST and the experimental immersion showed a significant reduction in anxiety [$F(1,41) = 47.7$, $p < .001$]. The effect size of interaction was small, with a partial eta-squared of .02 (Cohen's f = .13) and more than 200 participants needed per cell to find a significant difference with an alpha at .05 and a power of .80.

Table 17: STAI-Y1- mean and standard deviation for all five conditions (N=46).

	Baseline Mean (SD)	TSST Mean (SD)	Immersion Mean (SD)
L4D – 22 inch (n=7)	27.57 (5.38)	36.57 (7.04)	27.71 (5.02)
L4D – 73 inch (n=10)	27.00 (5.25)	36.3 (14.23)	28.00 (7.32)
KF – 22 inch (n=11)	29.82 (6.56)	37.73 (12.17)	28.91 (6.66)
KF – 73 inch (n=9)	30.67 (6.63)	37.89 (7.66)	30.67 (6.08)
KF – CAVE (n=9)	32.22 (8.97)	38.33 (9.10)	29.33 (7.28)

Table 18: STAI-Y1- ANOVAs for comparisons between all five conditions, monitor sizes and games.

	Time Contrasts

	Time F	Condition F	Interaction F	Baseline vs. TSST	Baseline vs. Immersion
Comparison between all conditions (n=46)	31.70***	0.34	0.21	36.12***	0.28
Comparison between monitor sizes (n=37)	29.28***	0.003	0.46	33.39***	0.003
Comparison between games (CAVE excluded) (n=37)	29.08***	0.89	0.21	34.22***	0.01

Note. **p≤.01; ***p≤0.001

Table 19: STAI-Y1- means and standard deviations for 22-inch and 73-inch screen-size conditions (N=37).

	Baseline Mean (SD)	TSST Mean (SD)	Immersion Mean (SD)
22-inch (n=17)	29.29 (6.06)	37.59 (10.47)	28.18 (6.01)
73-inch (n=20)	28.45 (6.05)	36.80 (11.07)	29.45 (6.59)

Table 20: State anxiety - means and standard deviations for L4D and KF games (N=37).

	Baseline Mean (SD)	TSST Mean (SD)	Immersion Mean (SD)
L4D (n=17)	27.24 (5.14)	36.41 (11.52)	27.88 (6.29)
KF –CAVE excluded (n=20)	30.20 (6.43)	37.80 (10.13)	29.70 (6.30)

When asking participants at the end of the study if they found the immersion stressful, if they appreciated it and if they considered their immersion could be useful to learn to practice stress management skills (see Table 21 for results), one-way ANOVAs found significant differences for the general appreciation of the immersion and its usefulness. Post-hoc comparisons performed with a Tukey test revealed that only two paired comparisons were significant: (a) less positive interest towards playing KF on a 22-inch monitor versus playing L4D on a 73-inch TV, and (b) considering playing KF in a CAVE as less useful to practice SMT than playing L4D on a 73-inch TV.

Table 21: Assessment of the immersion by participants – means and standard deviations for all five conditions (N=46).

Personal impressions (0-10)	**L4D-22 inch (n=7)** **Mean (SD)**	**L4D-73 inch (n=10)** **Mean (SD)**	**KF-22 inch (n=11)** **Mean (SD)**	**KF-73 inch (n=9)** **Mean (SD)**	**KF-CAVE (n=9)** **Mean (SD)**	**F**
Stress during immersion	6.00 (2.52)	4.20 (2.86)	3.55 (2.21)	3.56 (2.79)	3.44 (2.30)	1.36
Appreciation of immersion	7.67 (2.25)	9.50 (0.71)	6.27 (3.04)	7.33 (2.78)	6.78 (2.39)	2.67*
Usefulness to practice stress control	6.20 (3.56)	7.30 (2.83)	4.90 (3.64)	6.44 (2.51)	3.11 (2.76)	2.63*

Note. *$p < 0.05$.

Analyses were also performed to describe the subjective impression of being in the virtual environment (presence) and side effects induced by the immersion (Simulator Sickness Questionnaire) (see Tables 22, 23 and 24 below). Presence and simulator sickness were relatively low, although the immersion led to a significant reduction in simulator sickness.

Table 22: UQO Presence Questionnaire- means and standard deviations for all five conditions (N=46).

Presence	L4D-22 inch (n=7) Mean (SD)	L4D-73 inch (n=10) Mean (SD)	KF-22 inch (n=11) Mean (SD)	KF-73 inch (n=9) Mean (SD)	KF-CAVE (n=9) Mean (SD)	F
Total score	29.29 (18.58)	33.90 (14.47)	17.91 (21.69)	31.00 (15.92)	33.78 (20.48)	1.34
Item 1.	57.14 (33.52)	75.00 (31.06)	36.64 (31.06)	59.11 (26.94)	49.44 (29.31)	2.36

Item 1 reads " To what extend did you have the impression of being present in the virtual environment?"

Table 23: Simulator Sickness Questionnaire total score - means and standard deviations for all five conditions (N=43).

	Baseline Mean (SD)	TSST Mean (SD)	Immersion Mean (SD)
L4D – 22 inch (n=6)	3.50 (4.23)	3.00 (1.10)	1.67 (1.86)
L4D – 73 inch (n=9)	2.00 (1.12)	2.78 (1.86)	1.33 (1.50)
KF – 22 inch (n=11)	4.36 (3.17)	4.73 (3.77)	2.18 (1.99)
KF – 73 inch (n=9)	3.56 (4.56)	3.89 (2.15)	1.67 (1.32)
KF – CAVE (n=8)	3.00 (3.26)	2.88 (2.36)	2.38 (2.50)

Table 24: ANOVA for the Simulator Sickness Questionnaire for all five conditions (N=43).

				Time contrast	
	Time F	**Condition F**	**Interaction F**	**Baseline vs. TSST**	**Baseline vs. Immersion**
SSQ-Total	8.45***	1.11	.85	1.32	9.1**

Note. **p < 0.01.

Finally, participants' performances during the experimental immersion (see

Table 25) were compared with a one-way ANOVA and a Tukey post-hoc test. Although the actual number of zombies differed in each game and depending on the progression of each participant, the number of zombies killed was not statistically different between the five conditions [$F(4,46) = .86$, ns]. However, the number of times participants were killed while playing the games was significantly different [$F(4,46) = 3.16$, $p < .05$]. The Tukey revealed only one significant comparison pair: participants were killed more frequently when playing KF in the CAVE than when playing L4D with a 73-inch TV.

Table 25: Participant's performance (killing zombies or being killed) in all five conditions (N = 47).

	Number of zombies killed Mean (SD)	Number of time the participant was killed by zombies Mean (SD)
L4D – 22 inch (n=9)	175.11 (90.53)	5.56 (1.88)
L4D – 73 inch (n=10)	188.00 (111.61)	4.80 (1.32)
KF – 22 inch (n=10)	197.10 (24.35)	8.30 (4.72)
KF – 73 inch (n=9)	202.33 (32.37)	6.33 (2.65)
KF – CAVE (n=9)	147.67 (50.90)	8.67 (2.83)

3.7 Discussion for Study 1

The goal of this study was to test whether being immersed in a game could induce stress in military personnel. The study was conducted with 47 male soldiers who have recently been in Afghanistan and compared two games combining the FPS and the horror genre, two sizes of stereoscopic screens and a highly immersive CAVE. Two physiological indices of stress were set as the primary dependent variables in this study: heart rate and respiration rate. Results on the heart rate partially confirmed our hypotheses. The analysis did indeed reveal a significant increase from the baseline to the experimental immersions in the 3D games in *both* heart rate and respiration rate, therefore answering the most important research question in this study: 3D games can be used to induce stress.

More refined analyses comparing two games with different qualities of graphical realism and stressors or the three levels of immersion provided by the different technologies (i.e., 22-inch stereoscopic monitor, 73-inch stereoscopic TV and CAVE) did not reveal any significant difference between the conditions. The small number of participants, especially in some conditions, can only partially explain the lack of significant difference. Power analyses performed on the effect sizes of the interaction effects yielded expected sample sizes of 200 participants or more, suggesting that there may be differences in games and immersion format, but these differences are not very large. In sum, playing mixed FPS / horror games in stereoscopy was stressful enough to significantly increase heart and breathing rates.

A standardised induction of stress (a procedure called the TSST) was performed to account for individual differences in reactions to stress and was designed as a methodological control. Playing the 3D games was expected to be significantly more stressful than the TSST. The lack of significant difference between the TSST and the experimental immersions deserves to be discussed in more detail. Impressions felt by the experimenters during the TSST and gathered while listening to participants commenting on the study suggested that the TSST might have induced more frustration than stress. Although this is unusual since the TSST has been successfully used to specifically induce anxiety in past research (Kudielka et al., 2007), the context of the study (military personnel participating in a psychological experiment) may be different. In that case, the experimental immersions may have provided relief from frustration, hence a weaker impact of playing the horror games. However, the heart rates of the study participants during the baseline (Mean = 83.7) and the TSST (mean = 94.4) were similar to what is found in other studies conducted either with a large sample (Taylor et al., 2010), with elite sportsmen (Rimmele et al., 2007; Rimmele et al., 2008), participants measured only during the speech portion of the TSST (Fries, Hellhammer, & Hellhammer, 2006), or with young males in the same age group (Toet et al., 2009). These comparisons were reassuring as they confirmed that our sample reacted normally to the TSST stressor, yet they do not rule out the psychophysiological impact of frustration.

Toet et al. (2009) also measured the heart rate of their participants when they played a 3D FPS game on a 19-inch monoscopic monitor. The average heart rate of Toet et al. (2009) participants while playing *Counter-Strike* was 75.8 (Toet 2010). Weber et al. (2009) also obtained a heart rate of between 70 and 74 with gamers playing *Tactical Ops*. As opposed to the games used in the current study (KF and L4D), the game played in the Toet et al. (2009) and Weber et al. (2009) studies were not a mix of FPS - horror games and involved significant elements of sustained concentration (e.g., sniper, tactical analysis), which may partly explain why they were less stressful. In addition, the desktop technology used by Toet et al. (2009) was less immersive than the least immersive technology of the current study, which may contribute to the fact that their procedure (playing in a daytime or nighttime scenario) was not sufficient to induce stress. Therefore, the significant increase in heart rate found in the current study in going from the baseline to playing the games should already be considered a success. In addition, the physical demands were higher in the TSST, as participants were standing up and in some cases walking, than in the experimental immersion, where participants were seated.

Yet, in retrospective it became apparent that the task performed in the current study may be less stressful for trained soldiers who have been in combat than for the general population. The task performed in games like KF and L4D is close to the regular training of soldiers and it may be difficult to reach stress levels higher than what some of them have been through in missions. A significant difference between the games as they were played in this study and soldier's tasks such as house clearing was that the participants were facing the enemy alone and could not count on other unit members to help them. The impossibility of relying on other unit members to buffer the impact of the stressor was also felt by the experimenters during the TSST. Although this is based on subjective impressions, it may deserve more attention in further studies on the induction of stress in military personnel.

As mentioned above and in concordance with heart rate, the analyses revealed that playing the FPS - horror games while immersed in a stereoscopic environment significantly increased the participant's breathing rate. The increase in stress, as measured by breathing, was significant from the baseline to the TSST and from the baseline to the experimental immersion. Contrary to the heart rate, the analysis revealed that the further increase in respiration rate during the experimental immersions was statistically significant over and above the increase from the baseline to the TSST. Consistent with the heart rate findings, the comparisons between types of games or immersive displays was far from being significant. Comparison with other studies revealed that similar breathing rates have been observed with the TSST (Sjörs, Larsson, Dahlman, Falkmer, & Gerdle, 2009), confirming the validity of the results.

Finding a significant increase in stress from the TSST to the experimental immersion with the respiration rate but not the heart rate remained intriguing. The fact that participants were talking during the TSST may have imposed a pace on their breathing, a condition that was different during the experimental immersion. The observed difference may be due to different contingencies imposed on physiology by the TSST and the experimental contexts. Or the heart rate response of soldiers may be less sensitive to some stressors, such as close combat. Cox, Hallam, O'Connor, & Rachman (1983) found that courageous soldiers, defined in their study as bomb-disposal operators who had been decorated for gallantry, had lower cardiac rates when under stress. Although the sample in the current study was different from that of Cox et al. (1983), participants may have acquired through training the stronger control on heart rate found in fearless people when performing specific tasks. On the other hand, in another study, Parsons et al. (2009) compared the heart rates of nine civilians and six West Point cadets and did not find significant differences between the two populations in high immersion conditions like the ones used in this study.

When comparing the TSST and the game stressors, it stands out that being stressed by the TSST was not a situation participants had been trained for, as opposed to looking for and shooting at enemies. The significant reduction in physical motion (vector magnitude and peak acceleration) raised the possibility that the increase in stress between the TSST and the experimental immersion may have not been fully expressed by an increase in heart rate. The potential reduction in heart rate associated with lessened physical motion could have partly blurred the impact of the stressful experimental immersion. But despite this minor discrepancy, both physiological measures leave no doubt that playing off-the-shelf 3D horror games (with elements of FPS in the game) while immersed with stereoscopic technologies can reliably stress soldiers. The effect sizes are strong enough that the difference would remain significant despite applying a Bonferroni correction to compensate for the risk of making a Type – I error.

The self-rated assessment of stress performed with the State anxiety scale of the STAI revealed puzzling results. The average scores reported at the baseline were extremely low, in the range usually found after practicing deep-muscle relaxation (Spielberger, 1983). Such low scores raise the possibility that participants' answers on the self-report scale may not be valid. The average state anxiety scores observed in the general population and among military recruits should range between 36 and 48 (Bouchard, Gauthier, French, & Paradis, 1992). Being subjected to the TSST led to a significant increase in state anxiety compared to the baseline, but the experimental immersion did not. However, the average scores for the TSST were in the range of the normative population when not stressed. These results were inconsistent with physiological data and comments expressed by participants during and after the study. Why would state anxiety score be unreliable?

Social desirability was the first argument that came to mind. Participants may have been hesitant to express overtly their signs of stress. Social desirability may operate under explicit control from participants willing to hide information that would not be perceived as virile. It may also be more subtle, as soldiers may have learned not to be too sensitive to signs of stress. A third reason could be the influence of a CF screening process. All military personnel returning from Afghanistan must undergo the Enhanced-Post-Deployed Screening Process (EPDSP) within 180 days of their return. This screening aims at identifying mental injuries pertaining to the deployment. This could explain why some participants, back from Afghanistan five months prior to the experiment, commented on the purpose of the questionnaires in ways (i. e. "Is this related to the screening …?") that suggest they may have wanted to conceal their stress symptoms. Another likely explanation would be restricted or blunted affect (emotional numbing). Being in potentially traumatic situations might leave temporary sequels such as those found in people suffering from PTSD, including being less in tune with internal signs of stress and emotions (e.g., Asmundson et al., 2000; Erbes, Curry, & Leskela, 2009; Neugebauer, Fisher, Turner, Yamabe, Sarsfield, & Stehling-Ariza, 2009; Solomon, Mikulincer, & Arad, 1991). Participants in the current study did not report signs of PTSD, but the general levels of anhedonia were not assessed.

Other interesting data were gathered at the end of the study when participants were probed to state how much they appreciated their experience and whether they found the experience useful to practice stress management skills. In both cases playing L4D on the 73-inch TV obtained the best score. Playing KF on a 22-inch monitor was the least appreciated condition and the CAVE was considered the least useful. Data collected with people immersed in the CAVE were not supporting the use of this technology as much as expected.

Although the CAVE condition was associated with the highest heart rate and the second highest respiration rate, these differences were not statistically significant. In addition, the CAVE was the condition with the fewest number of zombies killed and where participants were significantly the most frequently killed. The performance was in accordance with the fact that manipulating the Intersense Wand in the CAVE was less intuitive for participants that using a mouse on a desktop setup. A reason could be that highly-immersive systems require more natural devices to interact efficiently with the simulated world. A simulated rifle, available for the experiment, would have probably improved the performance but such a device was discarded during the pilot study. Its weight, the same as a real M4 weapon, would have required sustained physical effort and would have increased the heart rate, interfering with the psychophysiological measures.

The results were otherwise quite consistent in not revealing any statistically different difference between the games and immersion technologies used. Indeed, the effect sizes of the difference in stress between the two games or between the immersive technologies were very small. As an example of the triviality of the differences, to find a significant difference in heart rate when comparing L4D and KF with a respectable statistical power and a significance level at .05, an experimenter would need to recruit more than 3,100 soldiers. One of the consequences of this finding is the flexibility it provides in the choice of stressors and immersive systems to be used in further studies.

Results of the Simulator Sickness Questionnaire are reassuring since they confirm the safety of the immersions. The questionnaire was designed to assess the side effects of immersion in flight simulators and is often used in research on virtual reality in general. The virtual reality induced side effects are lower than what is commonly reported in therapeutic applications to treat anxiety disorders (Bouchard, St-Jacques, Renaud, & Wiederhold, 2009) and after flight simulations (Kennedy et al., 1993). In fact, the scores were lower after the experimental immersion than when measured before the start of the study. This unexpected finding contributes to the growing worry that some SSQ symptoms may confuse signs of stress with signs of motion sickness or cybersickness (Bouchard et al., 2009). Nevertheless, the number of signs associated with nausea, disorientation and oculomotor symptoms are low enough to consider the stressful procedure safe, at least when assessed within 15 minutes post immersion.

Performance results are similar to those found in Rajea-Joordens (2008). In the experiment aiming at comparing a stereoscopic with non-stereoscopic display, participants playing *Quake III Arena* for 45 minutes performed better with the stereoscopic display but only for the first part of the experiment.

An additional thought should be given to the choice of the stressful games. The use of 3D horror games definitively represents an interesting stressor to be used for teaching and practicing SMT. Yet, a task that differs more from the regular training imposed on soldiers might be even more fruitful. One of the lessons learned during the TSST was that bringing soldiers outside their zone of comfort was very stressful. To avoid situations where soldiers' responses have been overlearned (i.e., house clearing), the stressor could be based on a scenario where the participant is not in service but gets caught in an ambush and must help someone severely wounded and in acute pain. Such a scenario could even be included within an already existing stressful horror game. For example, the game could include a team of virtual characters controlled by the software comprised of a child, a young female representative of the United Nations and a male wearing the uniform of a member of the CF. One of the three characters (randomly selected each time the game is played) could survive the ambush and become a burden for the participant due to the wounds and extreme pain reactions. If a scenario is developed along those lines, an important question would be whether being familiar with the wounded person significantly increases stress levels. Finally, according to Table 3, L4D and KF are not considered by the gaming community as being very stressful compared to some other video games. There is still room for choosing a more stressful video game but other important factors should be considered as well.

This page intentionally left blank.

4 Study 2: Empathetic emotion toward the virtual character (avatar) of a known or unknown person (Protocol #L-718)

4.1 Background

Following the results of Study 1, it appears that FPS - horror games are somewhat stressful for soldiers. However, the intensity of the stress was not as high as expected, probably because soldiers are trained to react efficiently to situations involving close combat or search and rescue missions in hostile areas. Other stressful situations reported by Bouchard, Baus, Bernier and McCreary (2010) may represent more appropriate stressors, especially if they involve being alone and disarmed. Knowing someone seriously injured, receiving artillery and small-arms fire, and being unable to help are stressors frequently encountered by Canadian soldiers. A better stressor to practice SMT while stressed might be a scenario where: (a) the user is on leave, walking unarmed with a few other people in an area surrounded with destroyed homes, (b) they are ambushed and suddenly attacked from artillery and small-arms fire, (c) most of the other people are killed, except the user and one other person who is severely wounded and in pain, (d) the user cannot return enemy fire and must provide basic medical help to the wounded. This scenario would include several stressors that are typical of what is experienced in a theatre of operations (Bouchard et al., 2010), and yet remain different from the L4D map used in Study 1 where the gameplay shares similarities with house clearing, a task for which soldiers are heavily trained. Before designing and testing such a stressor, it is important to know if being familiar with the wounded compatriot is important.

The aim of this study was to assess how users react to an avatar (graphical representation of a person) expressing pain, for possible use in a scenario close to what was described in the previous paragraph. Scientific inquiry was needed to know whether stress could be induced simply by the fact of seeing an avatar in pain or whether the avatar needs to represent a known individual.

The theoretical background of this study could be approached from a layman's point of view, which would suggest that a user may not react to pain expressed by virtual people, essentially because they are not real. However, several studies have already shown that avatars could induce emotions in users immersed in VR, even though the avatars are just pixels on a computer monitor. For instance, avatars are used to elicit emotions and assess or treat fear of social evaluations in social anxiety (Klinger et al., 2005; Pertaub, Slater, & Barker, 2002), panic attacks associated with crowded places in panic disorder with agoraphobia (Botella et al., 2007), body image issues among people suffering from eating disorders (Aimé, Cotton, & Bouchard, 2009; Gutiérrez-Maldonado, Ferrer-Garcia, Caqueo-Urizar & Letosa-Porta, 2006; Riva, Bacchetta, Baruffi, Rinaldi, & Molinari,, 1999), social pressure encountered by stutterers (Brundage, Hancock, Kiselewich, & Stallings, 2007) or persecutory delusions in schizophrenia (Fornells-Ambrojo et al., 2008; Kim et al., 2006). Other psychological applications of avatars have also been successful in areas such as forensic psychology (Renaud et al., 2007), attention deficit hyperactivity disorder (Rizzo et al., 2000) and substance abuse (Bordnick, Graap, Copp, Brooks & Ferrer, 2005).

Users immersed in VR can effectively recognise emotions expressed by avatars, including happiness, fear, anger and sadness, as demonstrated by Dyck et al. (2008). Males also modify their posture, hand movement and head movements when they interact with a female avatar in a CAVE (Pan, Gillies, & Slater, 2008). Rossen, Johnsen, Deladisma, Lind, and Lok (2008) even found that negative skin-tone biases are transferred from the physical to the virtual reality. In that study, empathy toward dark skin-tone avatars was predicted by biases towards African-

Americans. Slater et al. reproduced in 2006 the classical social psychology experiment from Stanley Milgram and illustrated quite effectively that users react with strong empathy to the pain experienced by an unknown virtual person. To explore if users can assess pain intensity in synthetic humans, Hirsh, Alduquah, Stutts, and Robinson (2009) conducted an experiment with 75 young adults who watched videos of various avatars expressing their pain through facial expressions. Participants rated the unpleasantness and intensity of avatars' pain and were able to differentiate accurately the level of pain experienced by virtual humans.

If avatars can elicit emotions and empathy in the user, would a user immersed in VR react more strongly to the avatar if the user knows the person represented by the avatar than if it is an unknown stranger? Indeed, literature on empathy (e.g., Davis, 1996; Ickes, 1997; Fitzgibbon, Giummarra, Georgiou-Karistianis, Enticott, & Bradshaw, 2010; Goubert et al., 2005) suggests that people are more likely to react to someone experiencing pain given that they can relate positively to that person (e.g., a member of the same race, faction or army). The generalisation from the physical to the virtual reality of empathic feelings for a known person deserves to be tested empirically.

The implication of knowing the person used to create the avatar is important because it will document whether or not there is a need to create highly realistic (both in terms of physical and behavioural) replicas of a potentially large number of Canadian soldiers in order to develop VR scenarios to practice SMT where an avatar is in pain.

4.2 Goal of the study

The current study investigated three research questions: (a) can pain experienced by an avatar induce reactions in a user immersed in VR? (b) Is the emotion induced by observing the avatars of a known individual stronger than that induced by observing the avatar of an unknown person? (c) Would familiarity with the individual used to create the avatar have an impact on the user's reactions?

4.3 Experimental setup

4.3.1 Hardware

All immersions were performed in the CAVE-like system of the Laboratoire de Cyberpsychologie of the Université du Québec en Outaouais (UQO; see Figure 13). This system was not designed nor purchased from Mechdyne (the current owner of the CAVE trademark; http://www.mechdyne.com), it was built instead in collaboration with VizTek (http://www.viz-tek.com), hence the expression CAVE-like. This immersive system, nicknamed Psyche because of its use in studies of psychology and mental health, consists of six projector surfaces: four sides, a ceiling and a floor. The back wall is mounted on a railing and can be closed once the participant is inside the cube. Each screen of 8.6' receives images from a VizTek 1 projector (modified from an Electrohome Marquee 8500) located 15' away (see Figure 13). Each CRT projector displays 225 ANSI lumen images at a native resolution of 1280x1024@100 Hz and delivers active stereoscopy.

Psyche is driven by a cluster of six slave computers and one master computer, all running a Virtools VPPublisher Unlimited 5.0 and built with the following specs: Intel Core 2 Quad Q6600@2.40 GHz with 4 GiB of RAM, NVidia graphics card Quadro FX 5500G with 1024 MiB of VRAM, Intel D975XBX2 motherboard, and Windows XP Pro 32-bits Service Pack 2. The rendering library uses OpenGL 2.0 Stereo. The master computer also has a Creative SoundBlaster

X-Fi sound card. The cluster of PCs includes an additional computer used to link the Intersense tracker to the virtual environment on VRPN 7.18 with a Pentium 4 3.20 GHz CPU and 512 MiB of RAM. All computers are linked in a network using a Cysco Systems Catalist 2950 100 Mbits/sec switch.

The viewer's head position is tracked by a wireless Intersense IS-900 VET system to correct for the perspective in real-time since the position of the viewer in relation to the screens changes based on the user's movement within Psyche. The user wears 3D glasses (NuVision^{TM}), wireless stereo headphones (Sennheiser RS146) and a Shure L3 wireless lapel microphone. The experimenter wears a set of Bose QuietComfort 15 noise-cancelling headphones and can talk to the participant via a microphone. Before starting the immersions, the participants were positioned in the middle of the back wall / door, two feet inside the cube (See Figure 14). They were allowed to physically move freely within the environment.

The experimenter was controlling Psyche, and the computer running the psychophysiological assessment, from outside the room hosting Psyche (see Figure 15). In this figure, the user is depicted seated during baseline physiological monitoring and the experimenter outside the room controlling the virtual environment and the physiological monitoring computer. The participant was thus left alone either sitting close to a table with ambient lights turned on but dimmed for completing questionnaires and baseline physiological monitoring, or standing up in Psyche with the ambient lights turned off.

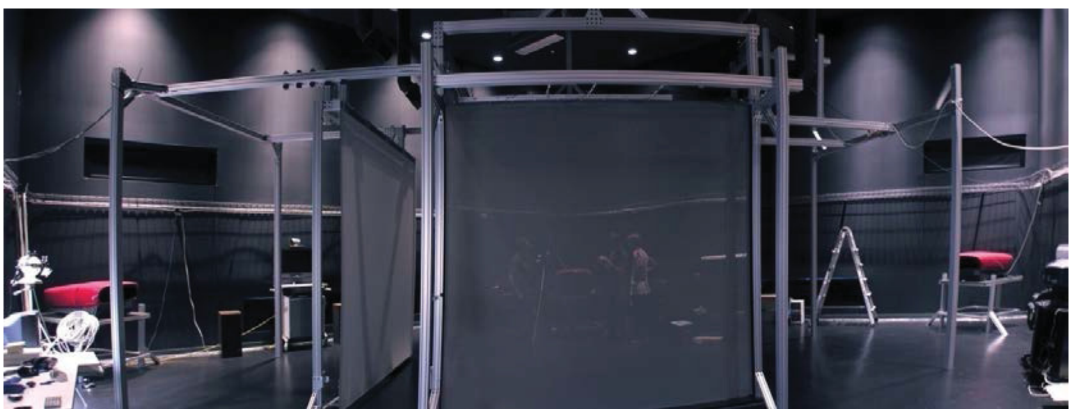


Figure 13: Side view of Psyche, with the door almost closed.

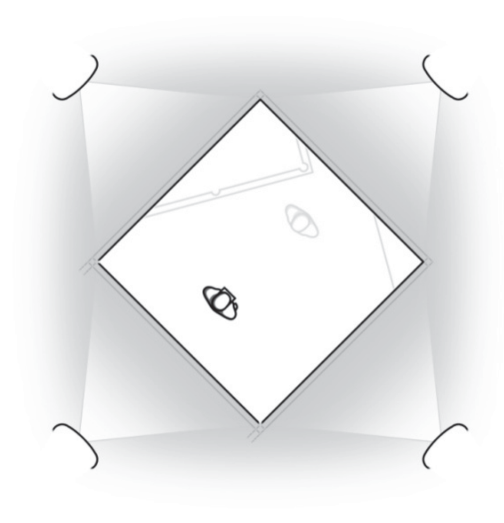


Figure 14: Physical layout of the experimental setup in Psyche.

The user is depicted in black and the avatar in grey.

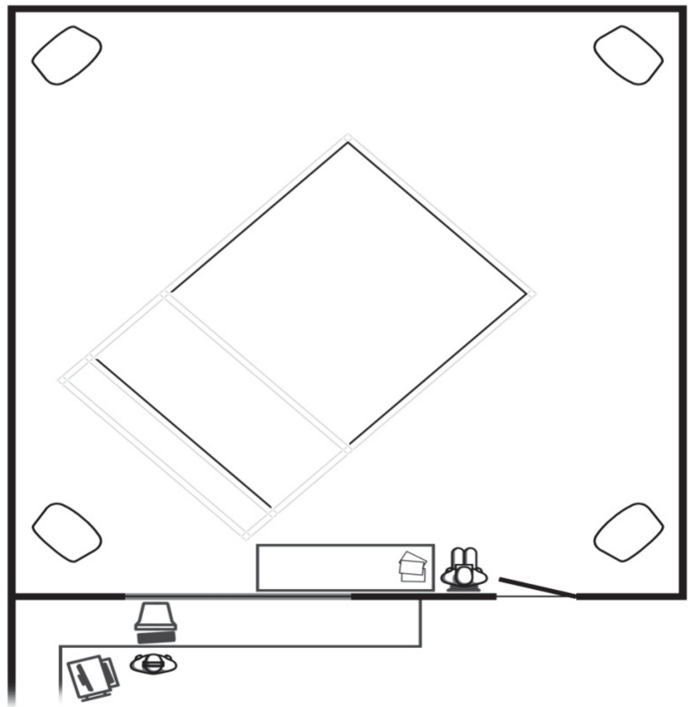


Figure 15: Physical layout of the general experimental setup.

A PC was used to monitor physiological responses continuously during the study (computer with Intel Core 2 Duo CPU @ 3 GHz, 1.96 GiB of RAM, with a NVidia GeForce 9800 GTX PCI Express with 512 MiB of RAM, running on Windows XP Service Pack 3). Heart rate and GSR were monitored using a ProComp Infinity and a wireless-enabling Tele-Infiniti Compact Flash T9600, both from Thought Technology (http://www.thoughttechnology.com). ECG electrodes were put on the participant's wrists using EKG Wrist Straps from Thought Technology and GSR electrodes were strapped on the index and ring fingers of the non-dominant hand. Physiological data were sampled at 256 Hz. While data were acquired and recorded, the research assistant put markers in the series of physiological data at the same time she triggered specific events during the experiment, defining blocks of physiological data to be analyzed later.

4.3.2 Software

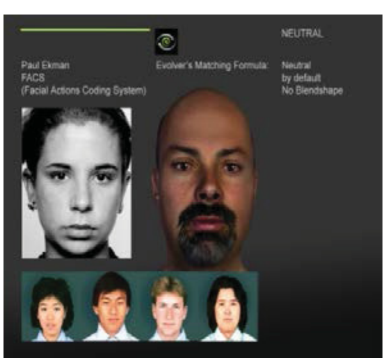
Two virtual environments were used, a training (baseline) environment and an experimental environment. The training environment consisted of an empty room with three windows, a glass door and a cat waiting on a table behind the glass door. The user can hear a relaxing breeze and birdsong coming from the windows. The purpose of this environment was twofold. First, it allowed participants to familiarise with being immersed in a CAVE-like system, including moving close to an avatar and talking to a virtual stimulus. Second, physiological data recorded during this immersion served as a baseline, controlling for the effect of being immersed in virtual reality as well as for seeing and interacting verbally with a virtual character.

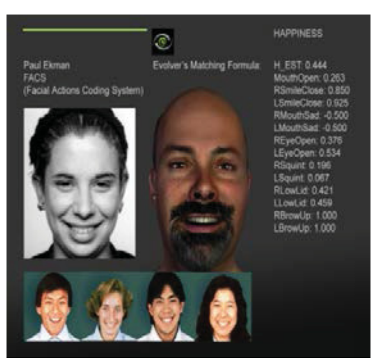
The experimental environment was a sports barroom measuring 8.5 metres in width by 13.5 metres in depth, with a pool table, several tables and chairs, posters on the walls, alcoholic beverages on the counter, sports games playing on TV monitors and rhythmic music playing in the background. The bar was empty, except for a human avatar standing near the pool table. The avatar was in a waiting position, looking vaguely at the participant. Quantitatively, both human avatars were built with 59,644 polygons and the cat avatar with 3,036 polygons. The experimental environment, including the human avatar, displayed about 259,339 polygons and the training environment was much lighter, with 53,234 polygons. The polygon counts fluctuated depending on the visible information at any given time, but the above information gives an idea of the complexity and richness of the scenes.

The three avatars used were a cat, Stéphane Bouchard (see Figure 16; the principal investigator being a university professor, it was easier to find people familiar with his physiognomy), and a virtual human created without reference to a real person. It should be noted that these images were not rendered in virtual reality but using a modelling software. Complex shader effects had been added that could not transfer to real time 3D. Stéphane Bouchard's avatar (referred in this report as the known avatar) was developed using pictures, videos and voice recordings of himself enacting the various behaviours and facial expressions needed for the study (see Figures 17 to 20). Attention was paid to animate the avatar as realistically possible, although the avatar's gaze patterns were not synchronised with the participant's head position in Psyche.

The unknown avatar was designed to have the same dimensions as the known avatar, in order to use the same animations. The audio recordings were modified (increasing the higher frequencies in the voice, raising the pitch by 8% and removing audio artefacts induced by the modifications) so the voice of the unknown avatar could carry the same emotions yet sound different from the known avatar. The two human avatars were therefore physically different but had the exact same behaviours. Slight differences still remained between the two avatars: the textures of the known avatar were more refined and nuanced and the facial animations of the known avatar were performed using blendshapes while the facial animations of the unknown avatar were created using morph targets.

Figure 16: Avatar clone of Stéphane Bouchard created by Darwin Dimension.





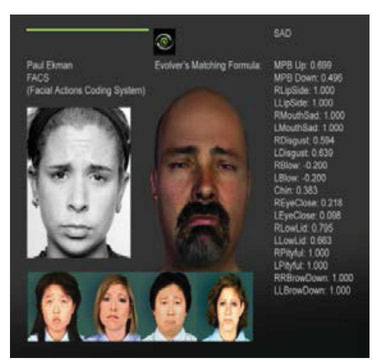


Figure 17: Comparison of facial expressions of the known avatar and the Eckman classification system

Five animations were created for the human avatars: (a) waiting in an idle position (see Figure 18, which also shows the cat in idle position), (b) welcoming the participant and inviting him or her to describe one of the best things that has happened in his/her life (see Figure 19), (c) idle / listening (d) asking for more details, and (e) pain (see Figure 20)

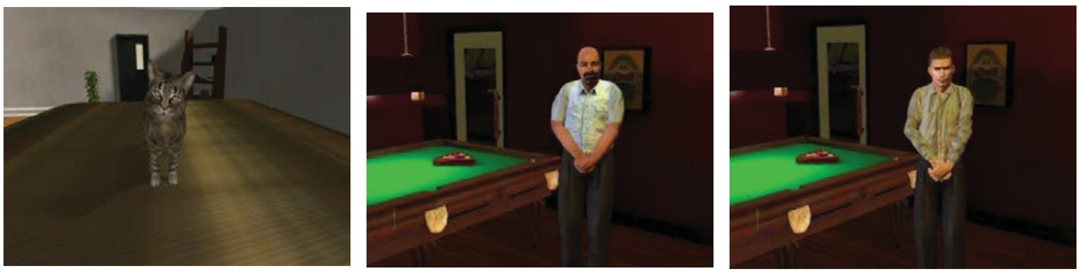


Figure 18: Screenshots of the three avatars (cat, known and unknown) in idle position waiting for the participant to interact with them.

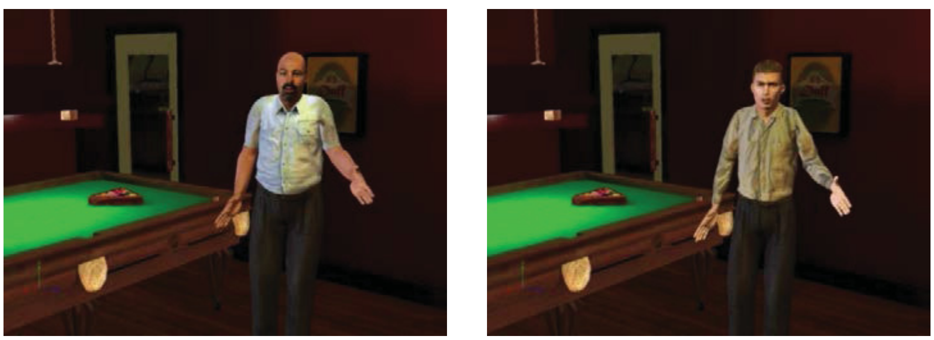


Figure 19: Screenshots of the known and unknown avatars greeting the participant.

Figure 20: Screenshots of the known and unknown avatars expressing acute and sustained pain.

The virtual environments were created using 3D StudioMax^{TM} version 2009 (Autodesk; http://en.autodesk.ca), the cat was adapted from the Vizard Complete Animals^{TM} set (WorldViz; http://www.worldviz.com) and the virtual humans were developed using the Evolver software (Darwin Dimensions Inc; http://www.evolver.com). Stephane Bouchard's clone biped was created by Darwin Dimensions (see Figure 16 rendered in Maya) and preliminary facial animations with blendshapes were tested with the Evolver software (see Figure 17 the avatar's expressions with Eckman's classification system). The unknown avatar was also developed using Evolver, but only with the tools available on the web version of the software. Import into 3D Studio Max, final facial and biped animations and synchronisation of lip movements with the audio files were performed by FacePro (FacePro LLC; http://facepro.net). Animations were transferred from the known to the unknown avatar by FacePro. Each step of the avatar creation and animation processes actively involved the development team at UQO, who also used the above software to work on all stimuli. Integration of visual stimuli, avatars, sounds and animations, as well as real-time rendering was performed using 3DVia Virtools ^{TM} 5.0 (Dassault Systèmes; http://www.3ds.com). Several pilot tests were performed in Psyche with staff and relatives to improve the realism of the avatars and animations.

4.4 Participants

The sample was composed of 42 civilian adults (18 to 60 years old, 15 males and 15 females) in good mental and physical health, as verified through telephone prescreening. Because the goal of the study was to document people's reactions to an avatar of a person they were familiar with, it was simpler to recruit participants that were familiar with Stéphane Bouchard instead of spending months developing multiple realistic avatars tailored for each participant. Recruitment was therefore conducted on the university campus among students, university staff, friends, colleagues and relatives of members of the research team. To be eligible for the study, participants had to have directly talked at least once with the person that was used as a model to create the known avatar (Stéphane Bouchard) but had to have never seen the avatars used for the project nor been aware that the avatars would express pain. Such criteria allowed recruiting participants with a wide range of familiarity with the known avatar. Participants were randomly assigned to two avatar conditions: (a) Known Avatar First (KAF; i.e., the immersion with the known avatar preceded the immersion with the unknown avatar, n = 22) or (b) Unknown Avatar First (UAF; i.e., the immersion with the unknown avatar preceded the immersion with the known avatar, n = 20).

The following exclusion criteria were set a priori for ethical reasons and, if met, led to the rejection of the participant: showing signs of suffering from: schizophrenia, psychotic disorder or PTSD. This was assessed using the respective sections of the SCID-NP. These criteria were assessed by research assistants (Ph.D. candidates in psychology who had experience in administrating the SCID-NP, including assessing these criteria, who had been trained in handling clinical cases, and who were not relatives of the participant). Also, people suffering from the following physical problems would have been excluded since they were at risk of having significant cybersickness during the immersion (Kennedy et al., 1993): vestibular (or inner ear) problems, recurrent migraines, epilepsy, balance, cardiac or ocular problems, frequent and intense motion sickness in transports. No participants needed to be excluded on the basis of these criteria. People with no stereoscopic vision (as assessed with the Randot stereoscopy test) were excluded since people with this condition cannot perceive depth during stereoscopic immersions.

All participants were informed about the details of the study (except about the details of the animation, including the expression of pain) and the risks associated with the experimental protocol. Before being asked for their written informed consent, they were told they were free to refuse to participate and could withdraw their consent without prejudice or hard feelings at any time. Experimenters insisted that the main researcher (Stéphane Bouchard) would not have access to their personal information (i.e., would only have access to anonymised information) or details about their personal experience. The project was approved by the ethics review board of DRDC and the UQO.

4.5 Experimental protocol

4.5.1 Baseline

The experimentation took place at the Laboratoire de Cyberpsychologie of the UQO. When participants arrived, they were informed about the study and asked to read and sign the consent form. After agreeing, they filled out questionnaires providing general descriptive (sociodemographic data, immersive tendencies, attachment style, physical health, etc.) and baseline information before starting the experiment. This step took about 30 minutes:

- Sociodemographic Information (pre-test): information was gathered about age, gender, socio-economical status, etc.
- Immersive Tendencies Questionnaire (ITQ; Witmer & Singer, 1998): the ITQ (French-Canadian version validated by the Laboratoire de Cyberpsychologie of the UQO) was administered at the end of the intake interview. The ITQ has 18 items rated on a 7-point scale (1: *Never*, 7: *Often*) which provide a Total score and four subscale scores: Focus (the ability to concentrate and to ignore distractions), Involvement (the feeling of being caught up by stories and movies), Emotion (the intensity of the emotions evoked by stimuli such as movies), and Play (the frequency of playing video games).
- Familiarity with the person used to create the avatar: This scale has four questions about to which extent participants were familiar with Stéphane Bouchard (the known avatar). This questionnaire was created by the Laboratoire de Cyberpsychologie of the UQO for

this study. The principal investigator also rated how much he was familiar with the participant, on a zero to 10 scale.

- Adult Attachment Questionnaire: This instrument is a French-Canadian translation of the *Adult Attachment Questionnaire* (Bouthillier, Tremblay, Hamelin, Julien, & Scherzer, 1996) aimed at measuring adult attachment style. It consists of 13 items using a scale from 1 (strongly disagree) to 7 (strongly agree). Administering this questionnaire allowed controlling for potential individual differences in attachment styles that may lead to a greater sensitivity to an avatar's pain reactions.

- Simulator Sickness Questionnaire (Kennedy et al., 1993): this questionnaire assesses the level of discomfort before and after the immersions, using 16 different items related to symptoms (headache, eyestrains, nausea, etc.). It was administered at baseline and after each immersion to document potential side effects induced by the experiment.

- State-Trait Anxiety Inventory (French-Canadian version validated by Gauthier and Bouchard, 1993). The *State-Trait Anxiety Inventory* is divided in two sub-scales (state anxiety and trait anxiety) of 20 items each, with various questions about how the participant feels that day or in general. Only the State Anxiety form was used in this study to measure the levels of anxiety and stress before and after each immersion.

- Positive And Negative Affect Schedule (PANAS): The French-Canadian version validated by Gaudreau, Sanchez & Blondin (2006) was used. It contains 20 items representing various emotions that people may feel at the present moment. Only the negative affect subscale was used in this study.

- Empathy: Empathy toward the pain of the avatar was assessed with two items embedded in the Social Presence scale. The questions, rated on a 1 to 7 scale, were "I was empathic to the pain of the virtual person" and "the pain of the virtual character was credible". These items were developed specifically for this study and were not used after the control immersion since the cat did not express any pain. There are several indices of social presence (i.e., the user's impression that there is someone else in the virtual environment) and empathy is one dimension that could shed light on the perceived realism of the virtual characters. (Bailenson, Swinth, Hoyt, Persky, Dimov & Blascovich, 2005).

- Likability: This variable was measured with three of the four items developed by Bailenson et al. (2005) to assess to what extent the user liked the avatar he met during the immersion. The end of each item was modified to adapt the question to the avatar encountered during the immersion (i.e., the cat, the known character or the unknown character). Bailenson et al.'s (2005) approach to the general concept of social presence was to differentiate how users respond to avatars from the subjective impression of being co-situated. Likability fits in the former category, as opposed to co-presence.

- Embarrassment: This variable was measured by a French-Canadian translation and adaptation of items created by Bailenson et al. (2005) to assess how people treat avatars as real people. It contains 4 items measuring the participant's willingness to perform embarrassing acts in front of the virtual character. Embarrassment addresses the participant's perception of the social influence exerted by the virtual character and is a proxy for social presence. The format was adapted and worded specifically for each immersion and avatar by referring explicitly to the cat, the avatar of the known character, or that of the unknown character. Some items were modified because the context of the current study would have made them implicitly embarrassing (e.g., willingness to change clothes in front of the known character).

- Co-presence: This instrument is a French-Canadian translation and adaptation of the perceived co-presence instrument created by Bailenson et al. (2005). It contains two items measuring the impression of being *there* in the virtual environment. The format was adapted and worded specifically for each immersion by referring explicitly in the item to the cat, the avatar of the known character, or that of the unknown character. This measure addresses the perception of being co-situated within an interpersonal environment (ibid., p. 380).
- Social Presence: This instrument is a French-Canadian translation and adaptation of the Gerhard, Moore and Hobbs (2001) questionnaire. It contains 17 questions about the feeling of presence and interactions with virtual characters rated on a 1 to 7 scale. A few items were dropped from the original scale because they were irrelevant to our virtual environment (e.g. "How aware were you of the existence of your own avatar?"; "How natural was the mechanism which controlled the actions of your avatar?"; "How responsive were the avatars of other participants to non-verbal communication that you initiated?). A global measure of the feeling of social presence was used in order to fit our results within the broad literature on social presence.

In order to test the ability to perceive binocular disparity in distance from static stimuli (stereoscopic vision for depth perception), the participant performed a stereopsis test (Randot SO-002, Stereo Optical Company Inc., Chicago, IL; http://www.stereooptical.com) that consisted in observing various geometric shapes (between 400 and 20 seconds of arc in apparent size) and animals (between 400 and 100 seconds of arc) in front of random dot backgrounds while wearing polarising glasses. Participants must recognise the stimuli they perceived as 3D (i.e., floating above the test board). This test took about 5 minutes and was scored according to the procedures described in the test manual. A PD-5 Topcon pupillometer (Topcon Medical Systems, http://www.topconmedical.com) was used to measure interpupilar distance and individually tailor the stereoscopy to each participant immersed in Psyche.

Once the first set of measures was completed, a research assistant installed the Thought Technology physiological monitoring equipment on the participant (see Figure 21 and the earlier section on hardware) to measure heart rate and GSR. Participants wore shutter glasses, head tracker, headphones, microphone and physiological monitoring (ECG and GSR).This step took 5 minutes, including baseline recording. These measures allowed us to document objective signs of stress among the participants:

- Heart rate: The most frequently used objective measure of stress is the heart rate, which is measured by an ECG (see Study 1 and the current section on hardware for details about the ECG sensors, their placement and data acquisition). Heart rate is traditionally either measured by non-invasive means when participants are not moving (e.g., while participants are seated), or by more invasive means when participants are in motion (e.g., on a treadmill for the diagnosis of heart conditions). To avoid shaving participants' chests, scratching their skin and using abrasive stickers to hold the ECG sensors in place, a non-invasive approach was taken. In order to control for body movements, it was intended to use users' motion in covariance analyses. However, the participants' body movements waere larger than expected, creating interference in the cables of the ECG sensors and invalidating all the data. Details about the ECG are described in the Methods section of this report but ECG data will not be reported.
- GSR: Electrical conductivity is commonly used to assess immediate response to stressors (Perala & Sterling, 2007). GSR measures the skin's electrical resistance; since stress changes the level of hydration by sweat glands, it allows electricity to circulate better in the human skin, thus increasing galvanic response. GSR was also measured

with the wireless ProComp Infinity (Thought Technology, see section on hardware for details about the ECG sensors, their placement and data acquisition).

- Relative position (movement) from the avatar: Data from the InterSense IS-900 VET on each participant's head position was monitored at 256 hertz during the immersions with human avatars. The distance in metres from the avatar was also recorded. This measure allowed documenting interpersonal distance and the behavioural reactions of the participants when the avatars expressed pain. To measure the impact of the immediate reaction to acute pain, distance data recorded during 12-second segments before and during the acute pain animation were also extracted. Given the range of motions within these segments of 12 seconds, data will be reported in changes in distance (in metres) from the previous time point recorded by the IS-900 VET.

A baseline measurement was taken for both physiological parameters during two minutes after having completed the questionnaires and before the first immersion. The participant was sitting on a chair with eyes open. This dataset was used to objectively quantify stress based on physical and behavioural manifestations. In this study, physiological and movement responses represent the principal dependant variables.


Figure 21: Participant wearing the equipment relevant for the study (left) and experimenter in the control room (right).

The SCID-NP was administered to the study participants to exclude those likely to suffer from axis-I psychiatric disorders mentioned in the exclusion criteria (i.e., schizophrenia, psychotic disorder and PTSD). This step lasted 5 minutes and was performed by research assistants extensively trained in the use of the SCID and the management of clinical cases.

4.5.2 Experimentation

After completing the selection process, questionnaires and consent form in the experimenter's office, participants were brought in Psyche's room. The experimenter installed the physiological monitoring and gave the first set of instructions to the participant by reciting the following text:

> *"Nous nous intéressons aux interactions sociales en réalité virtuelle. Vous allez faire trois immersions en réalité virtuelle. Une première pour vous habituer, et deux autres par la suite. Les deux immersions virtuelles que vous allez faire se passent dans un bar où se trouve un personnage virtuel connu, ou inconnu. Le personnage virtuel vous saluera et initiera la conversation en vous demandant : « Parle-moi du plus beau moment que tu as vécu dans ta vie » et vous pourrez répondre comme bon vous semble. Vous devez parler pendant un bon deux minutes alors ça vaut la peine de bien préparer ce que vous direz.*

Si vous avez plusieurs idées en tête et ne savez pas quel beau moment choisir, veuillez simplement choisir UN beau moment. C'est vous qui décidez. Je vous laisse deux minutes pour choisir le beau moment à décrire et les détails ou anecdotes que vous allez mentionner. Est-ce que c'est ok ?"

Participants were told to prepare describing "the most beautiful moment of their life", and the experimenter left the participant in the room for baseline physiological assessment during two minutes. After the baseline, participants received instructions specific to the training environment:

"Avant de débuter l'expérimentation et pour faciliter la familiarisation avec l'équipement, vous serez immergé à l'intérieur d'une pièce virtuelle où vous pourrez apprendre à vous déplacer. Veuillez simplement suivre les instructions verbales que je vais vous donner dans vos écouteurs. Vous et moi pourrons parler à l'aide de micros, bien que le but soit que vous puissiez vous concentrer sur l'expérience vécue en réalité virtuelle. Lorsque je vous demanderai d'approcher d'une porte, vous pourrez parler à un être virtuel. Vous pourrez vous en approcher. Je vous demanderai de lui dire ce qui vous passe à l'esprit, de façon naturelle. Mais gardez l'histoire du plus bel événement de votre vie pour la prochaine immersion, ce n'est pas pour ce personnage là. Lorsque l'étape de familiarisation sera terminée, les écrans de la voûte s'éteindront. À ce moment, je reviendrai ouvrir le mur arrière de la voûte et je vous demanderai d'enlever les lunettes et les écouteurs et de répondre à quelques questionnaires. Ça va ?"

The experimenter invited the participant to walk into the virtual room and say whatever was on his mind to the virtual animal they were going to see. The experimenter finished installing the VR equipment on the participant (see Figure 21 above), adjusted stereoscopy for the participant's interpupilar distance and adjusted the sound level so participants could hear well a repetitive welcoming message played at the same loudness as the voice of the virtual humans. The experimenter physically positioned the participant in the cube, closed the back door, turned off the ambient lights and launched the training environment. The experimenter encouraged the participant to explore the room for 30 seconds, then opened a door which made the cat walk toward the participant (see Figure 18). Participants were told to talk to the cat and look at him for 30 seconds. The cat was animated to look naturalistic, look around and move its tail. After this 60-second immersion, the participant was invited to walk outside the cube, remove the headphones and shutter glasses, and complete questionnaires.

The immersion in the first experimental environment began after participants were done completing the questionnaires. Based on random assignment, the experimental environment included either the known or the unknown avatar. The following instructions were recited:

"*Tel que mentionné plus tôt, les deux immersions virtuelles que vous allez faire se passent dans un bar où se trouve un personnage virtuel, connu ou inconnu. Au début, vous pourrez vous en approcher et le regarder à votre aise. Après un moment, le personnage virtuel vous saluera et initiera la conversation en vous demandant : « Parle-moi du plus beau moment que tu as vécu dans ta vie ». Vous devez lui répondre et parler pendant une bonne minute de l'événement que vous avez identifié précédemment. Le personnage pourra lui aussi vous parler un peu. L'immersion dure un peu plus de deux minutes alors quand ce sera terminé je vous parlerai dans les écouteurs. Nous prendrons alors un moment pour répondre encore à quelques questionnaires, puis nous passerons à la dernière immersion. Avez-vous des questions ?*"

After stating the instructions, the immersion was initiated according to the same sequence as for the training environment. In the virtual bar, the events occurred in the following sequence: (a) participants had up to 30 seconds to explore the surroundings and approach the avatar who was slowly looking around in the bar (see Figure 18), (b) (after the time limit has passed or the participant stopped moving or exploring) the greetings animation was launched (see Figure 19), where the avatar invited the participant to tell his or her story and remained in a listening (idle) animation while holding a positive emotional expression, (c) if the participant paused, the experimenter could launch additional animations asking for more details about the story, (d) the experimenter let the participant talk to the avatar for a total of 60 seconds and then launched the pain animation, (e) while the participant was still talking, the avatar had an acute stomach cramp (see Figure 20) during 12 seconds, excused himself and invited the participant to keep talking, and then experienced even stronger and longer pain reactions (including screams, moaning, bending and stretching, putting hands on the stomach, facial expressions, etc. – see Figure 21), (f) after the pain animation, the avatars went back to the idle position and the experimenter stopped the immersion.

Figure 22: Screenshots of facial expressions of avatars while in pain.

Following the first experimental immersion, participants followed the same procedure to complete questionnaires, received instructions again about the coming immersion (the content was identical, except for referring to meeting "the other avatar") and the second experimental immersion unfolded just like the previous one.

The following questionnaires were filled after each of the three immersions:

- Co-presence: This instrument is a French-Canadian translation and adaptation of the co-presence questionnaire created by Bailenson et al. (2005). It contains six items measuring the impression of being *there* in the virtual environment and to which extend the person treats the avatars as if they were real people. The format was adapted and worded specifically for each immersion by referring explicitly in the item to the cat, the avatar of Stéphane Bouchard, or the unknown avatar.
- Embarrassment: This instrument is a French-Canadian translation and adaptation of the co-presence questionnaire created by Bailenson et al. (2005). It contains 4 items measuring the participant's willingness to perform embarrassing acts in front of the virtual character. Embarrassment addresses the participant's perception of the social influence exerted by the virtual character and is a proxy for social presence. The format was adapted and worded specifically for each immersion and avatar by referring explicitly to the cat, the avatar of Stéphane Bouchard, or that of the unknown avatar. Some items were modified because the context of the current study would have made them implicitly embarrassing (e.g., willingness to change clothes in front of the known avatar).
- Social Presence: This instrument is a French-Canadian translation and adaptation of the Gerhard, Moore, & Hobbs (2001) questionnaire. The original version contains 17 questions about the feeling of presence and interactions with virtual characters. Nine items were dropped from the original scale because they were irrelevant to our virtual environment (e.g. "How aware were you of the existence of your own avatar?"; "How natural was the mechanism which controlled the actions of your avatar?"; "How responsive were the avatars of other participants to non-verbal communication that you initiated?).

- PANAS (see description, page 49)
- State-Trait Anxiety Inventory (see description, page 49)

At the end of the study, the experimenter asked the participants to complete additional measures:

- UQO Presence Questionnaire: (see description, page 26)
- Simulator Sickness Questionnaire:(see description, page 49)

4.6 Data Analysis

Participants were randomly assigned to KAF and UAF conditions upon their arrival in the lab, prior to verification of the exclusion criteria. Descriptive information was collected and subjected to two-way ANOVAs and chi-square tests in order to confirm that randomisation did not introduce biases in sample characteristics. The data were analyzed with repeated measures analyses of variance (ANOVAs), followed by repeated measures contrasts and power analyses (according to standard procedures described by Cohen, 1988) when necessary. Inferential analyses were first conducted to assess the immediate impact of the stressor of observing an avatar in pain, with GSR and user's distance from the avatar as dependent variables. A second set of analyses was conducted to document the impact of the stressor on the user's emotions during the immersion, with the following dependent variables: state anxiety post immersion, negative affect post immersion, GSR for the entire duration of the immersion, and user's distance from the avatar for the entire acute pain animation. Finally, the user's perceptions of the avatars were submitted to 3 X 2 repeated measures ANOVAs. Additional exploratory analyses were conducted with the data collected on each participant when immersed with the known avatar, as well as with measures of familiarity with the person used to create the avatar.

4.7 Results

Descriptive results

The sample consisted of 42 civilians, mostly Canadian citizens (86%), females (62%), who were completing (or had finished) a university degree. Only 43% of them had been immersed in VR before. All of them had met the person used to create the known avatar, 2% on a familial basis, 29% as friends, 62% on a professional basis and 7% on some other basis (e.g., met him in a lecture in class). The distribution of people who met him on a regular basis was: 26 % almost daily, 14% once a week, 17% once a month and 44% twice a year or less. The mean age of the entire sample was 37 (SD = 12.26). The number of participants who were considered by the experimenter to know him well (i.e., rated 5 and above on the familiarity scale) or not were almost equal, with 59% being considered more familiar and 41% less familiar. Age, attachment style, immersive tendencies and side effects of the immersion are reported in Table 26. Note that analyses comparing the conditions on these descriptive variables did not find any significant differences.

Table 26: Descriptive information for participants randomly assigned to the Know Avatar First (KAF) or the Unknown Avatar First (UAF) conditions (N=42).

	KAF	UAF
Age	38.05 (12.45)	36.05 (12.28)
Attachment questionnaire - Secured	21.86 (3.90)	22.35 (3.88)
Attachment questionnaire – Avoidant	12.23 (4.24)	12.10 (5.39)
Attachment questionnaire – Anxious Ambivalent	11.82 (3.50)	11.00 (2.15)
Immersive Tendencies Questionnaire	67.82 (13.99)	68.65 (12.29)
Simulator Sickness Questionnaire	3.45 (2.89)	3.50 (2.74)

Note: Mean, with standard deviations in brackets.

Physiological data reported in Table 27 show that GSR and distance from the avatars measured during the 12 seconds preceding the expression of acute pain by the avatars increased during the 12 seconds of acute pain. The stressful impact of observing pain was documented for the GSR by a significant Time effect during the first immersion [$F_{(1,40)} = 7.26$, $p < .01$] and the second immersion [$F_{(1,40)} = 5.58$, $p < .05$], as well as for distance from the avatar in the first [$F_{(1,40)} = 8.87$, $p < .01$] immersion. Change in distance from the avatars during the second immersion was much smaller and non-significant [$F_{(1,40)} = 2.25$, *ns,* partial eta-squared = .05]. These results confirmed that avatars experiencing pain can induce reactions in the observers.

Table 27: Immediate physiological and behavioural reactions to the pain stressor (N=42).

	Immersion 1		Immersion 2	
	Pre pain	During pain	Pre pain	During pain
Galvanic Skin Response				
Known Avatar First	8.10 (4.26)	8.40 (4.35)	7.74 (3.89)	7.92 (3.89)
Unknown Avatar First	8.33 (5.88)	8.65 (5.98)	8.64 (5.85)	8.77 (5.94)
Change in distance from the avatar (sampled in metres at 256 hertz)				
Known Avatar First	0.04 (0.03)	0.10 (0.10)	0.07 (0.06)	0.08 (0.09)
Unknown Avatar First	0.05 (0.06)	0.09 (0.11)	0.05 (0.03)	0.07 (0.08)

Note: Mean, with standard deviation in parentheses.

The Time×Condition interactions were not significant for each of the four 2×2 repeated measures ANOVAs: [$F_{(1,40)} = .003$, *ns*, partial eta-squared = .000] for GSR during the first immersion, [$F_{(1,40)} = 1.42$, *ns*, partial eta-squared = .004] for GSR during the second immersion, [$F_{(1,40)} = .025$, *ns*, partial eta-squared = .006] for change in distance during the first immersion, and [$F_{(1,40)} = .36$, *ns*, partial eta-squared = .009] for change in distance during the second immersion. No Conditions main effects were significant either. Results of these analyses suggests that the users' reactions were the same when observing a known or an unknown avatar. The effect-sizes (partial eta-squared) were also very small, with expected *n* for a significant difference at .05 with a power of .80 estimated as more than 395 per cell.

The GSR response during the entire study is reported in Table 28. Although GSR data were collected and analyzed on eight occasions, attention was given only to the repeated contrasts comparing before (i.e., during the discussion) and during the entire 40 seconds of pain animation for each immersion. The overall main effect of Time was significant [$F_{(7,273)} = 46.29$, $p < .001$], but not the Condition main effect [$F_{(1,39)} = .30$, *ns*, partial eta-squared = .008] nor the Time×Condition interaction [$F_{(7,273)} = .33$, *ns*, partial eta-squared = .008]. The repeated measures contrasts for the comparison before and during the entire pain animation during the first immersion were not significant for the Time main effect [$F_{(1,39)} = .39$, *ns,* partial eta-squared = .01] or the interaction [$F_{(1,39)} = .02$, *ns,* partial eta-squared = .001]. The same results were found for the second immersion for the Time effect [$F_{(1,39)} = 3.81$, *ns,* partial eta-squared = .09] and the interaction [$F_{(1,39)} = 1.56$, *ns,* partial eta-squared = .04]. The significant main effect of time was driven by increased skin conductance from the baseline to the control immersion with the cat avatar [$F_{(1,39)} = 36.37$, $p < .001$] and from the control immersion to the first experimental immersion with a human avatar [$F_{(1,39)} = 27.72$, $p < .001$]. In sum, the effect size of knowing or not the avatar was very small, with the largest partial eta-squared suggesting that at least 395 participants per condition would be required to detect a significant difference at .05 with a power of .80.

Table 28: GSR reaction lasting during the complete pain stressor (N = 42).

First Avatar	Baseline	Control (cat)	First immersion			Second immersion		
			Entrance	Discussion	Pain	Entrance	Discussion	Pain
Known Avatar First	4.44 (2.96)	6.16 (3.51)	7.74 (3.81)	7.69 (3.82)	7.73 (3.85)	7.78 (3.96)	7.84 (3.93)	7.79 (3.96)
Unknown Avatar First	5.03 (3.56)	7.36 (5.69)	8.48 (6.60)	8.38 (6.15)	8.44 (5.91)	8.85 (6.36)	8.66 (6.07)	8.45 (5.88)

Note: Mean, with standard deviation in parentheses.

Consistent with physiological reactions, behavioural changes in distance from the avatar were not supporting a lasting acute pain reaction (see Table 29 for descriptive results and Table 30 for results of the repeated measures ANOVAs and contrasts). The partial eta-squared for the key contrasts (discussion vs. pain) during the first immersion were .003 for the main Time effect and .03 for the Interaction effect, and respectively .03 and .02 for the second immersion. A minimum of 175 participants per condition would have been required to detect a significant difference with a probability of .05 and a power of .80.

Table 29: Behavioural reaction (distance in metres) lasting during the complete pain stressor (N = 42).

	First immersion			Second immersion		
	Entrance	Discussion	Pain	Entrance	Discussion	Pain
Known Avatar First	1.37 (0.30)	1.19 (0.21)	1.16 (0.23)	1.31 (0.18)	1.22 (0.19)	1.24 (0.19)
Unknown Avatar First	1.42 (0.30)	1.25 (0.33)	1.27 (0.36)	1.31 (0.20)	1.20 (0.22)	1.21 (0.25)

Note: Mean, with standard deviation in parentheses.

Table 30: F values for the ANOVAs and contrasts applied to the distance between the user and the avatars during both immersions.

	Main effects			Repeated measures contrast			
				Entrance vs. Discussion		Discussion vs. Pain	
	Time	Condition	Interact.	Time	Interact.	Time	Interact.
First immersion	21.24***	.87	.55	34.71***	.1	.13	1.22
Second immersion	17.29***	.13	.42	25.58***	.05	1.15	.78

Note: * = p < .05, ** = p < .01, *** = p < .001.

Average scores on the subjective measures of state anxiety and negative affect are reported in Table 31. The 4 Times and 2 Conditions repeated measures ANOVAs revealed a significant main effect of Time [$F_{(3,117)} = 3.51$, $p < .05$], a non-significant effect of Conditions [$F_{(1,39)} = .11$, *ns,* partial eta-squared = .003] and a non-significant interaction [$F_{(3,117)} = 2.26$, *ns,* partial eta-squared = .06]. Interestingly, the significant Time effects were driven only by a decrease in anxiety and negative affect from the baseline to the following assessment occasion. The repeated measures Time×Condition interaction contrasts between the first and second experimental immersion were non-significant for both the state anxiety measure [$F_{(1,39)} = .45$, *ns,* partial eta-squared = .01] and the negative affect measure [$F_{(1,40)} = .72$, *ns,* partial eta-squared = .02]. With 175 participants per condition, there is an 80% probability that the contrasts would have detected a significant decrease in anxiety and negative affect at a significance level of .05.

Table 31: Post-immersion impact of the pain stressor on anxiety and negative affect (N=42).

	Baseline	Control (cat)	First immersion	Second Immersion
State anxiety				
Known Avatar First	33.73	30.73	31.09	29.05
	(7.89)	(5.79)	(9.53)	(7.11)
Unknown Avatar First	32.11	31.26	32.58	31.63
	(9.51)	(6.45)	(8.17)	(7.91)
Negative affect				
Known Avatar First	13.14	11.41	11.36	10.55
	(2.95)	(1.56)	(2.08)	(1.22)
Unknown Avatar First	12.10	12.50	12.75	11.35
	(2.59)	(4.07)	(3.01)	(1.81)

Note: Mean, with standard deviation in parentheses.

The last set of analyses was performed to document the user's perceptions of the avatars (see Table 32 for average scores). Analyses were performed on the user's reactions towards the avatars being as if they were real people, followed by the subjective impression of being in the same room with them. Results on empathy and the likability of the avatars were quite interesting (see Table 33). Participants were equally empathic to the pain of the first avatar they met. However, participant's reactions were quite different in the following immersion. Those who met the known avatar first were significantly less empathic towards the pain of the unknown avatar and those who met the unknown avatar first were significantly more empathic to the pain of the known avatar in the second immersion. A similar interaction pattern was found with likability. The participants all considered the first virtual human they met as more likable than the cat avatar. However, after having encountered their first human avatar, participants' reactions went in opposite directions in the following immersion. After having been immersed with the known avatar first (KAF), participants found the unknown avatar clearly less likable, whereas those in the UAF condition clearly rated the known avatar as more likable.

Table 32: Users' perception of the avatars (N=42).

	Control (cat)	First immersion	Second immersion
Empathy			
Known Avatar First	N/A	9.09 (3.66)	7.36 (4.03)
Unknown Avatar First	N/A	9.6 (2.7)	10.7 (3.02)
Likability			
Known Avatar First	16.86 (6.46)	20.36 (4.72)	11.64 (6.84)
Unknown Avatar First	11.15 (9.43)	18.25 (6.32)	20.20 (7.50)
Embarrassment			
Known Avatar First	20.86 (11.12)	21.09 (11.03)	17.64 (10.86)
Unknown Avatar First	19.30 (11.31)	19.35 (10.78)	17.35 (12.31)
Co-presence			
Known Avatar First	6.42 (3.7)	7.86 (5.31)	7.41 (5.51)
Unknown Avatar First	7.65 (4.75)	11.1 (4.97)	13.65 (5.27)
Social presence			
Known Avatar First	35.45 (8.76)	36.18 (6.62)	33.32 (7.36)
Unknown Avatar First	38.32 (7.06)	38.0 (5.91)	42.16 (6.98)

Note: Mean, with standard deviations in parentheses.

Embarrassment significantly increased (lower scores mean being less willing to perform embarrassing actions in front of the avatars) between the first and second immersions with a virtual human, probably because participants were paying more attention to the human characteristics of the avatars after having seen them in pain in the previous immersion. No significant interaction was found on embarrassment. The partial eta-squared was .013 for the interaction, which represents a small effect size as illustrated by the large number of participants (more than 350 per condition) required to find a significant difference at .05 with a power of .80.

Results for the subjective impression of co-presence revealed a significant increase in co-presence in both conditions when moving from the virtual cat to the virtual humans, and a Condition by Time interaction when comparing the known and unknown avatar. Participants had the impression the known avatar was physically more there and aware of their existence than the unknown avatar. Social presence, measured with a more global scale, remained relatively stable from the control immersion to the first experimental immersion with a virtual human. Consistent with the co-presence findings, the interaction between the first and the second immersions revealed a clear impact of knowing or not the avatars. Participants considered the known avatar as significantly more present in the virtual environment (e.g., aware, responsive to the user, as if physically there) than an unknown avatar enacting the exact same animations.

Table 33: Summary of ANOVAs and contrasts for the users' perception of the avatars.

	Main effects			Repeated measures contrasts			
				Control vs.		Immersion 1 vs. 2	
	Time	Condition	Interact.	Time	Interact.	Time	Interact.
Empathy	.92	3.66	18.59***	N/A	N/A	N/A	N/A
Likability	7.93**	.03	15.16***	11.55**	1.33	6.53*	16.2***
Embarrassment	1.98	.16	.13	.01	.00	7.17*	.57
Co-presence	11.07***	8.16**	5.43**	8.15**	1.35	3.32	6.82**
Social presence	.29	6.17*	3.32**	.023	.15	.51	15.11***

Note: * = p < .05, ** = p < .01, *** = p < .001.

Additional analyses were conducted to further explore the impact of familiarity with the known avatar. Data were extracted to obtain the results on each instrument when users were immersed with the known avatar, no matter if it was in the first or the second experimental immersion. Scores on each variable were compared for participants who were rated by the principal investigator with a score of 5 or more on the familiarity scale, and also for participants who rarely met (i.e., less than twice a year), or did not meet, the principal investigator. None of the ANOVAs comparing participants who were considered familiar or not with the person used to create the known avatar reached significance.

4.8 Discussion for Study 2

The general purpose of the current research program was to find stressors that could be used to train military personnel to practice stress management skills. The specific aim of this study was to shed light on the usefulness of knowing or not the person represented by an avatar experiencing pain. Three results stand out from this experiment: (a) the users did react significantly to observing an avatar in pain, (b) knowing or not with the person represented by the avatar had no impact on the user's reaction to the pain stressor, and (c) participants had a more positive attitude towards the known avatar, as expressed by stronger empathy and likability.

To begin with, proxemics (the distance between people when they interact) were considered normal for social interactions. Participants were discussing with the human avatars at a distance of 1.1 to 1.2 metres. There was no difference whether they were discussing with the known or the unknown avatar. Participants related more strongly to the known avatar than the unknown one, both in terms of empathy, likability, co-presence and social presence. However, the stronger bond with the known avatar did not have any significant impact on the reactions of the users to the stress induced by seeing the avatar in pain. All participants reacted with a surge in arousal and a backward movement when the avatars first experienced acute pain. The impact of this stressor did not last and was not apparent when measured subjectively post-immersion or with physiological and behavioural indices recorded during the entire immersion.

Research on avatars and virtual humans may provide some explanation for the discrepancy between a stronger bond with the known user and the lack of difference in reactions when observing pain in the avatars. Craig, Versloot, Goubert and Vervoort (2010) argued that an observer's reaction to the pain of someone else does not only depend on perceiving pain cues. The observer's reactions are moderated by how pain is expressed by the person in pain. Automatic, involuntary and spontaneous manifestations of pain lead to different reactions in the observer than controlled expressions of pain. If the observer's reactions to painful stimuli are modulated by how they perceive the sufferer, then even a slight mismatch between what is expected from a virtual human and his or her actual behaviour and appearance may influence the observer's reactions.

There is a growing body of evidence showing that the relationship between emotional reactions and avatars' appearance and behaviour is not linear and simple. An early theory by Mori (1970) suggested that in a graph depicting how observers are increasingly empathic towards synthetic humans as the latter become more humanlike, the response of the observer will suddenly drop just before the synthetics reach a very human look, forming an "uncanny valley" caused by sensitivity and revulsion to perceived imperfections in near-humanlike forms. Empirical studies (e.g., Hanson, 2006; MacDorman, 2006) brought nuances to the theory, notably that many factors influence how observers react to the realism of a synthetic human. Experimental research on virtual humans and avatars now shows (Bailenson et al., 2005; Garau et al., 2003; Nowak & Biocca, 2003; Seyama & Nagayama, 2007) the existence of a non-linear relationship between the realism of avatars and the consequences of imperfections in—and mismatches between—their behaviour and appearance. It would appear that people can tolerate significant imperfections in behavioural realism when the avatars are not considered realistic (Nowak & Biocca, 2003). However, as avatars become more realistic, anthropomorphic and behavioural realism set up higher expectations, which in turn could lead to more disenchantment when small imperfections are perceived, either because of behavioural (Garau et al., 2003) or physical (Seyama & Nagayama, 2007) discrepancies.

Therefore, it is possible that observers familiar with the person used to create the avatar were aware of imperfections in the replication of the behaviour and appearance of the physical person. These imperfections would have diminished the impact of a potentially stronger bond between the user and the avatar. In other words, the impact of imperfections when trying to replicate a known individual in VR may be cancelling its advantages. When using a fictive unknown avatar, the user cannot be disturbed by the mismatches between the behaviour of the physical human and his or her virtual copy, since he has no knowledge of the prototype to compare the copy with.

The effect sizes for the lack of significant differences in stress between the known and the unknown avatars support the notion that if there are differences favouring the use of a copy of a known person, these differences are very small and hardly meaningful. At least with the technology currently available; the synthetic humans were of a very good visual quality, the immersions were conducted in a very powerful immersive set-up and the novelty of discussing with virtual avatars had been reduced by a control immersion. The effort required to make very realistic copies (avatars) of friends of every military personnel submitted to stress management training in VR simply cannot be justified based on our current data.

The intensity and duration of the pain expressed by the avatars deserve to be addressed briefly. It is possible that developing pain animations depicting excruciating pain for a long time, with cries, complaints and requests for help expressed directly to the user may elicit stronger and longer lasting stress reactions in the user. Aside from the potential ethical questions raised by immersing people in situations where their best friend is agonizing and they can’t help, the important question is whether it would matter to create a highly realistic copy of their best friend or if any credible avatar would have just the same effect. Given the effect sizes found in this study, and the literature on the non-linear impact of imperfections in visual and behavioural realism, it is currently doubtful that developing avatars known to the user is necessary to practice SMT. Maybe once the technology becomes more powerful and efficient, then it would be worth the effort.

5 Conclusion

The general purpose of the current research program was to find stressors that could be used to train military personnel to practice stress management skills. The information presented in this research report was gathered to answer the following questions: (a) can we moderately stress military personnel using off-the-shelf video games as VR scenarios so they could practice SMT? (b) which immersive display should be used: a small stereoscopic monitor, a large stereoscopic TV, or a fully immersive cave automatic virtual environment (CAVE)? (c) what kind of video games would effectively stress military personnel? (d) if we decide to use medical assistance to an injured person as a stressor, is it important if the virtually wounded is a familiar person?

Conducting Study 1 and Study 2 provided clear answers to all four questions: (a) yes, games combining the FPS and horror genres could stress soldiers, (b) a large stereoscopic display is sufficiently immersive, (c) the video game does not have to include all the latest technical developments and gameplay tricks to be stressful, although playing a game such as L4D may increase the buy-in factor from military personnel, and (d) to induce reactions in the user, the additional value of creating the avatar of an individual known to the user is not founded, given the level of technology currently available.

The game used in Study 1 (L4D) should probably be modified to improve its strength as a stressor. Several possibilities should be explored, from playing in multiplayer mode in order to add unpredictablility and challenge, having to wait for rescue instead of clearing an area, and being forced to attend to the needs of a civilian while in combat. A more typically military scenario could be where: (a) the user is walking with other military personnel and a civilian in an area surrounded with destroyed homes, (b) they are ambushed and under severe and repeated waves of attack, (c) the user must move to another location, wait for rescue and keep the civilian alive. Some maps of L4D could be edited and modified to provide a scenario combining the above suggestions. In any case, it does not appear necessary to worry about creating avatars representing people known to the user; having avatars wearing Canadian Forces uniforms should be sufficient.

With adequate stressors on hand, the next step in using VR in SMT is to create situations where the users could practice the techniques taught in the class or briefing room. An effective solution is to provide explicit biofeedback in the visual field of the users and coach them in practicing SMT strategies until they master them. In the coming year, our team will work on finalising the stressor, developing a biofeedback interface that can be used in the game, select a basic SMT technique (such as tactical breathing) that can be practiced while immersed, and empirically test the efficacy of the system.

This page intentionally left blank.

6 References

Abt, C. C. (1970). *Serious games,* 1st edition. New York, NY: The Viking Press.

Adams, E. (2006). *Fundamentals of game design,* 1st edition. Upper Saddle River, NJ: Prentice Hall.

Adler, A. B., & Huffman, A. H. (2005). The impact of deployment length and experience on the well-being of male and female soldiers. *Journal of Occupational Health Psychology*, *10*(2), 121–137.

Aimé, A., Cotton, K., & Bouchard, S. (2009). Reactivity to VR immersions in a subclinical sample of women concerned with their weight and shape. *Journal of Cybertherapy and Rehabilitation*.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders 4th ed.* Washington, DC: American Psychiatric Publishing, Inc.

Anderson, C. A. (2004). An update on the effects of playing violent video games. *Journal of Adolescence*, 27(1), 113–122.

Armfield. F. (1994). Preventing posttraumatic stress disorders resulting from military operations. *Military Medicine, 159*(12), 739–746.

Arriaga, P., Esteves, F., Carneiro, P., & Monteiro, M. B. (2008). Are the effects of Unreal violent video games pronounced when playing with a virtual reality system? *Aggressive Behaviour*, *34*(5), 521-538.

Arroyo-Palacios, J., & Romano, D. M. (2010). A flexible bio-affective gaming interface. *Artificial Intelligence and Simulation of Behaviour (AISB'10) AI & Game Symposium,* Leicester, United Kingdom.

Asmundson, G. J. G., Frombach, I., McQuaid, J., Pedrelli, P., Lenox, R., & Stein, M. B. (2000). Dimensionality of posttraumatic stress symptoms: A confirmatory factor analysis of DSM-IV symptom clusters and other symptom models. *Behaviour Research and Therapy, 38,* 203–214.

Bailenson, J. N., Swinth, K., Hoyt, C., Persky, S., Dimov, A., & Blascovich, J. (2005). The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence*, *14*(4), 379–393.

Ballard, M. E., & Wiest, J. R. (1996). Mortal Kombat: The effects of violent video game play on males' hostility and cardiovascular responding. *Journal of Applied Social Psychology, 26*(8), 717–730.

Barlett, C. P., Harris, R. J., & Baldassaro, R. (2007). Longer you play, the more hostile you feel: Examination of first person shooter video games and aggression during video game play. *Aggressive Behavior, 33*(6), 486–497.

Barlow, D. H. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic,* 2nd Edition. New York, NY: Guilford Press.

Bartone, P. T., & Adler, A. B. (1994). A model for soldier psychological adaptation in peacekeeping operations. *Proceedings of the 36th Annual Conference of the International Military Testing Association* (pp.33–40). Rotterdam, The Netherlands.

Belik, S.-L., Stein, M. B., Asmundson, G. J. G., & Sareen, J. (2009). Relation between traumatic events and suicide attempts in Canadian military personnel. *The Canadian Journal of Psychiatry*, *54*(2), 93–104.

Bernier, M. (Chair) (2009). *Doing well and doing better: Health services provided to Canadian Forces personnel with an emphasis on post-traumatic stress disorder*. Report of the Standing Committee on National Defence, Ottawa, ON, Canada.

Bersak, D., McDarby, G., Augenblick, N., McDarby, P., McDonnell, D., McDonald, B., & Karkun, R. (2001). Intelligent biofeedback using an immersive competitive environment. *Proceedings of Designing Ubiquitous Computing Games Workshop at UbiComp*, Atlanta, GA.

Bliese, P. D., & Castro, C. A. (2003). The soldier adaptation model (SAM): Applications to peacekeeping missions. In T. W. Britt and A. B. Adler (Eds). *The psychology of the peacekeeper: Lessons learned from the field.* (pp. 185–204). Westport, CT: Praeger.

Bordnick, P. S., Graap, K. M., Copp, H. L., Brooks, J., & Ferrer, M. (2005). VR cue reactivity assessment in cigarette smokers. *CyberPsychology & Behavior*, *(8)*5, 487–92.

Botella, C., García-Palacios, A., Villa, H., Baños, R. M., Quero, S., Alcañiz, M., & Riva, G. (2007), VR exposure in the treatment of panic disorder and agoraphobia: A controlled study. *Clinical Psychology and Psychotherapy*, 14(3), 164–175.

Bouchard, S. (2009). *Foundation for stress management training using virtual reality*. Contract Report DRDC Valcartier CR 2009-170.

Bouchard, S., Baus, O., Bernier, F., & McCreary, D. R. (2010). Selection of key stressors to develop virtual environments for practicing stress management skills with military personnel Prior to Deployment. *CyberPsychology and Behavior*, *13*(1), 83–94.

Bouchard, S., Gauthier, J., French, D. J., & Paradis, J. (1992). Normes francophones de l'Inventaire d'anxiété situationnelle et de trait d'anxiété (IASTA-Y). Congrès annuel de la Société canadienne de psychologie, Montréal, QC, Canada.

Bouchard, S., St-Jacques, J., Renaud, P., & Wiederhold, B. K. (2009). Side effects of immersions in virtual reality for anxious people. *Journal of Cybertherapy and Rehabilitation*, *2*(2), 127–137.

Bouchard, S., St-Jacques, J., Robillard, G., Côté, S., & Renaud, P. (2003). Efficacité de l'exposition en réalité virtuelle pour l'acrophobie : Une étude préliminaire. *Journal de thérapie comportementale et cognitive, 13*(3), 107–112.

Bouthillier, D., Tremblay, N., Hamelin, F., Julien, D., & Scherzer, P. (1996). Traduction et validation canadienne-française d'un questionnaire évaluant l'attachement chez l'adulte. *Revue canadienne des sciences du comportement*, *28*(1) 74–77.

Britt, T. W., Adler, A.B., & Castro, C. A. (Eds.) (2006). *Military life: The psychology of serving in peace and combat*. 4 volumes. Westport, CT: Praeger Security International.

Brundage, S. B., Hancock, A. B., Kiselewich, K., & Stallings, L. (2007). Affective, Behavioral & Cognitive Measures of Stuttering During Speeches to Virtual and Live Audiences, *12th Annual Cybertherapy Conference 2007,* Washington, DC.

Cabrera, O. A., Hoge, C. W., Bliese, P. D., Castro, C. A., & Messer, S. C. (2007). Childhood adversity and combat as predictors of depression and post-traumatic stress in deployed troops. *American Journal of Preventive Medicine*, *33*(2), 77–82.

Cantor, J. R., Ziemke, D., & Sparks, G. G. (1984). The Effect of forewarning on emotional responses to a horror film. *Journal of Broadcasting*, *28*(1), 21–31.

Castro, C. A., & Adler, A. B. (2005). Preface to the special issue. *Military Psychology*, *17*(3), 131–136.

Castro, C. A., Hoge, C. W., Cox, A. L. (2006). Battlemind Training: Building Soldier Resiliency. In Human Dimensions in Military Operations – Military Leaders' Strategies for Addressing Stress and Psychological Support. (pp. 42-1 – 42-6). *Meeting Proceedings RTO-MP-HFM-134*, Neuilly-sur-Seine, France.

Chief Military Personnel. (2010). Pre-Deployment Training – Road to Mental Readiness.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ: Lawrence Erlbaum Associates.

Cox, D. N., Hallam, R. S., O'Connor, K. P., & Rachman, S. J. (1983). An experiment analysis of fearlessness and courage. *British Journal of Psychology, 74,* 107–117.

Craig, K. D., Versloot, J., Goubert, L., & Vervoort, T. (2010). Perceiving pain in others: Automatic and controlled mechanisms. *The Journal of Pain*, *11*(2), 101–108.

Cruz-Neira, C., Sandin, D. J., & DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: the design and implementation of the CAVE. *Proceedings of the 20th Annual SIGGRAPH Conference,* pp. 135–142), Anaheim, CA.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM (Association for Computing Machinery)*, *35*(6), 64–72.

Davis, M. H. A. (1996). *Empathy: A social psychological approach*. Boulder, CO: Westview Press.

Dekker, A., & Champion, E. (2007). Please biofeed the zombies: Enhancing the gameplay and display of a horror game using biofeedback. *Proceedings of the Digital Games Research Association (DiGRA) 2007 Conference* (pp. 550–558)*,* Tokyo, Japan.

Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., Gur, R. C., & Mathiak, K. (2008). Recognition profile of emotions in natural and virtual faces. *PLoS (Public Library of Science) ONE, 3*(11), 1–8.

Erbes, C. R., Curry, K. T., & Leskela, J. (2009). Treatment presentation and adherence of Iraq/Afghanistan era veterans in outpatient care for posttraumatic stress disorder. *Psychological Services, 6*(3), 175–183.

Fairclough, S. H. (2009). BCI and physiological computing for computer games: Differences, similarities & intuitive control. *Proceedings of the CHI'08* [Computer-Human Interaction 2008] *Workshop on BCI* [Brain-Computer Interface] *and Computer Games,* Florence, Italy.

First, M. B., Spitzer, R. L, Gibbon. M., & Williams, J. B. W. (2000). *Structured clinical interview for DSM-IV axis I disorders, Research Version, Non-Patient Edition*.

Fitzgibbon, B. M., Giummarra, M. J., Georgiou-Karistianis, N., Enticott, P. G., & Bradshaw, J. L. (2010). Shared pain: From empathy to synaesthesia. *Neuroscience and Biobehavioral Reviews*, 34, 500–512.

Fornells-Ambrojo, M., Barker, C., Swapp, D., Slater, M., Antley, A., & Freeman, D. (2008). VR and persecutory delusions: Safety and feasibility. *Schizophrenia Research*, *104*, 228–236.

Freeman, D. (2003). *Creating emotion in games: The craft and art of emotioneering*, Berkeley, CA: New Riders Games.

Fries, E., Hellhammer, D. H., & Hellhammer, J. (2006). Attenuation of the hypothalamic-pituitary-adrenal axis responsivity to the Trier Social Stress Test by the benzodiazepine alprazolam. *Psychoneuroendocrinology, 31*, 1278–1288.

Frome, J. (2007). Eight ways videogames generate emotion. *Proceedings of the DiGRA 2007 International Conference.* (pp. 831–835), Tokyo, Japan.

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A. J., Sasse, M. A. (2003). The Impact of Avatar Realism and Eye Gaze Control on Perceived Quality of Communication in a Shared Immersive Virtual Environment. *Proceedings of CHI* [Computer-Human Interaction] *2000,* Ft. Lauderdale, FL.

Gaudreau, P., Sanchez, X., & Blondin, J.-P. (2006). Positive and negative affective states in a performance-related setting: Testing the factorial structure of the PANAS across two samples of French-Canadian participants. *European Journal of Psychological Assessment*, 22(4), 240–249.

Gauthier, J., & Bouchard, S. (1993). Adaptation canadienne-française de la version révisée du State-Trait Anxiety Inventory de Spielberger. *Canadian Journal of Behavioural Science, 25*(4), 559–589.

Gerhard, M., Moore, D. J., & Hobbs, D. J. (2001). Continuous presence in collaborative virtual environments: Towards the evaluation of a hybrid avatar-agent model for user representation. In A. de Antonio, R. Aylett, & D. Ballin (Eds.): *Proceedings of the 3rd International Conference on Intelligent Virtual Agents (IVA 2001),* Madrid, Spain, Lecture Notes in Computer Science (LNCS) 2190 (pp. 137–153), Berlin, Germany: Springer-Verlag.

Gil, S., & Caspi, Y. (2006). Personality traits, coping style, and perceived threat as predictors of posttraumatic stress disorder after exposure to a terrorist attack: A prospective study. *Psychosomatic Medicine*, *68*, 904–909.

Goldfried, M. R., & Davison, G. C. (1994). *Clinical behaviour therapy*. New York, NY: John Wiley & Sons, Inc.

Goubert, L., Craig, K. D., Vervoort, T., Morley, S., Sullivan, M. J. L., Williams, A. C. de C., Cano, A. & Crombez, G. (2005). Facing others in pain: The effects of empathy. *Pain, 118*, 285–288.

Gray, J. A. (1975). *Elements of a two-process theory of learning*. New York, NY: Academic Press.

Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., et al. (2004). Heart-rate variability and event-related ecg in virtual environments. *Proceedings of the 7th Annual International Workshop on Presence (PRESENCE 2004 (*pp. 240–245), Valencia, Spain.

Gutiérrez-Maldonado, J., Ferrer-García, M., Caqueo-Urízar, A., & Letosa-Porta, A. (2006). Assessment of emotional reactivity produced by exposure to virtual environments in patients with eating disorders. *CyberPsychology and Behavior*, 9(5), 507–513.

Hanson, D. (2006). Exploring the Aesthetic Range for Humanoid Robots. *Proceedings of the Conference on Cognitive Science 2006 (CogSci 06), Workshop on Android Science*, Vancouver, BC, Canada.

Herbelin, B., Benzaki, P., Riquier, F., Renault, O., & Thalmann, D. (2005). Using physiological measures for emotional assessment: A computer-aided tool for cognitive and behavioural therapy. *Proceedings of the 5th conference on Disability, Virtual Reality & Associated Technologies* (pp. 307–314), Oxford, United Kingdom.

Hirsh, A. T., Alqudah, A. F., Stutts, L. A., & Robinson, M. E. (2009). Virtual human technology: Capturing sex, race and age influences in individual pain decision policies. *Pain*, 140, 231–238.

Hoffner, C., & Cantor, J. (2006). Forewarning of a Threat and Prior Knowledge of Outcome Effects on Children's Emotional Responses to a Film Sequence. *Human Communication Research*, 16(3). 323–354.

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*, 13–22.

Hutchison, D. C. (2008). Introducing DLP 3-D TV. Texas Instruments Report.

Ickes, W. (1997). *Empathic accuracy*. New York, NY: The Guilford Press.

IJsselsteijn, W., de Kort, Y., Poels, K., Jurgelionis, A., & Bellotti, F. (2007). Characterising and measuring user experiences in digital games. *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, Hollywood, CA.

Ivory, J. D., & Kalyanaraman, S. (2007). The effects of technological advancement and violent content in video games on players' feelings of presence, involvement, physiological arousal, and aggression. *Journal of Communication, 57*(3), 532–555.

Jacobson, J. & Hwang, Z. (2002). Unreal tournament for immersive interactive theatre. *Communications of the ACM*, *45*(1), 39–42.

Jeong, E.J., Biocca, F.A., & Bohil, C.J. (2008). Effects of representational realism in 3D violent games. *Proceedings of the 2008 PRESENCE Conference* (pp. 191–200), Padova, Italy.

Joanne, C., Ziemke, D., & Sparks, G.G. (1984). The effect of forewarning on emotional responses to a horror film. *Journal of Broadcasting, 28*(1), 21-31.

Johnson, W. L., Marsella, S., & Vilhjálmsson, H. (2004). The DARWARS tactical language training system. *Proceedings of I/ITSEC 2004,* Orlando, FL.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology,3*(3), 203–220.

Kim, K., Kim, J. J., Kim, J. Ku, J., Kim, I. Y., & Kim, S. I. (2006). Investigation of VR social skill training: Social cue perception & social problem solving for schizophrenia, *11th Annual Cybertherapy 2006 Conference,* Gatineau, QC, Canada.

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*(1–2), 76–81.

Klinger, É., Bouchard, S., Légeron, P., Roy, S., Lauer, F., Chemin, I., Nugues, P. (2005). VR therapy for social phobia: A preliminary controlled study. *CyberPsychology and Behavior, 8*(1), 76–88.

Kudielka, B. M., Hellhammer, D. H., & Kirschbaum, C. (2007). Ten years of research with the Trier Social Stress Test (TSST) – Revisited. In: E. Harmon-Jones & P. Winkielman (Eds.), *Social neuroscience: Integrating biological and psychological explanations of social behavior*, chapter 54, pp. 56–83, New York, NY: The Guilford Press.

Kuikkaniemi, K., Laitinen, T., Turpeinen, M., Saari, T., Kosunen, I., & Ravaja, N. (2010). The influence of implicit and explicit biofeedback in first-person shooter games. *Proceedings of the 28th International Conference on Human Factors in Computing Systems.* (pp. 859–868), Atlanta, GA.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, *3*(2).

MacDorman, K. F. (2006). Subjective Ratings of Robot Video Clips for Human Likeness, Familarity, and Eeriness: An Exploration of the Uncanny Valley. *Proceedings of the Conference on Cognitive Science 2006 (CogSci 06), Workshop on Android Science*, Vancouver, BC, Canada.

MHAT-V (Mental Health Advisory Team V). (2008). *Operation Iraqi Freedom 06-08: Iraq /Operation Enduring Freedom 8: Afghanistan*. U.S. Department of Defense.

Mori, M. (1970). 不気味の谷現象 [Bukimi no Tani Genshō ; The Uncanny Valley]. *Energy*, *7*(4), 33–35 (Translated from the Japanese by K. F. MacDorman and T. Minato).

Morissette, G. G. & Spaulding, B. (2004). DARWARS, A Future Training Environment. *Proceedings of EURO SIW Conference*, Edinburgh, Scotland.

Nacke, L. & Lindley, C.A. (2008). Flow and Immersion in First-Person Shooters: Measuring the player's gameplay experience. *Proceedings of FuturePlay,* Toronto, ON, Canada.

Nakasone, A., Prendinger, H., & Ishizuka, M. (2005). Emotion recognition from electromyography and skin conductance. *Proceedings of the 5th International Workshop on Biosignal* Interpretation (pp. 219–222), Tokyo, Japan.

Neugebauer, R., Fisher, P. W., Turner, J. B., Yamabe, S., Sarsfield, J. A., & Stehling-Ariza, T. (2009). Post-traumatic stress reactions among Rwandan children and adolescents in the early aftermath of genocide. *International Journal of Epidemiology, 38*(4), 1033–1045.

North, M. M., North, S. M., & Coble, J. R. (1996). *Virtual reality therapy. An innovative paradigm.* Colorado Springs, CO: IPI Press.

Nowak, K. L., & Biocca, F. (2003).The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence*, *12*(5), 481–494.

O'Donohue, W. T., & Fisher, J. E. (Eds.) (2009). *General principles empirically supported techniques of cognitive behavior therapy*. Hoboken, NJ: John Wiley & Sons, Inc.

Pair, J., Allen, B., Dautricourt, M., Treskunov, A., Liewer, M., Graap, K. M., & Reger, G. (2006). A virtual reality exposure therapy application for Iraq war post. *Proceedings of the IEEE VR2006 Conference* (pp 64–71), Alexandria, VA.

Pan, X., Gillies, M., & Slater, M. (2008). Male bodily responses during an interaction with a virtual woman. In H. Prendinger, J. C. Lester, and M. Ishizuka (Eds.): *Proceedings of the 8th International Conference on Intelligent Virtual Agents (IVA 2008), Tokyo, Japan*, Lecture Notes in Computer Science (LNCS) 5208 (pp. 89–96), Berlin, Germany: Springer-Verlag.

Parsons, T. D., Courtney, C., Cosand, L., Iyer, A., Rizzo, A. A., & Oie, K. (2009). Assessment of psychophysiological differences of West Point cadets and civilian controls immersed within a virtual environment. In D. D. Schmorrow, I. V. Estabrooke, & M. Grootjen (Eds.): *Proceedings of the 5th International Conference on Foundations of Augmented Cognition (FAC 2009), Dan Diego, CA: Neuroergonomics and Operational Neuroscience*, Lecture Notes in Computer Science (LNCS) 5638 (pp. 514–523). Berlin, Germany: Springer-Verlag.

Perala, C. H., & Sterling, B. S. (2007). *Galvanic skin response as a measure of soldier stress*. Research Report ARL-TR-4114. Aberdeen Proving Ground, MD: Army Research Laboratory.

Perron, B. (2004). Sign of a Threat: The effects of warning systems in survival horror games. *Proceedings of COSIGN 2004,* Split, Croatia.

Perron, B. (2005a). A cognitive psychological approach to gameplay emotions. *Proceedings of the DiGRA 2005 International Conference,* Vancouver, BC, Canada.

Perron, B. (2005b). Coming to play at frightening yourself : Welcome to the world of horror games. *Proceedings of the Computer Game Aesthetics Conference,* Bergen, Norway.

Persky, S., & Blascovich, J. (2006). Consequences of playing violent video games in immersive virtual environments. In R. Schroeder & A. S. Axelsson (Eds.), *Avatars at Work and Play*, Computer Supported Cooperative Work, 34, ch. 8, pp. 167–186. Dordrecht, The Netherlands: Springer.

Pertaub, D.-P., Slater, M., & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments*, *11*(1), 68–78.

Pratt, D. R., Zyda, M., & Kelleher, K. (1995). VR: In the mind of the beholder. *IEEE Computer, 28*(7), 17–19.

Rajae-Joordens, R. J. E. (2008). Measuring experiences in gaming and TV applications. In J. H. D. M. Westerink, M. Ouwerkerk, T. J. M. Overbeek, W. F. Pasveer, & B. De Ruyter (Eds.), *Probing experience: From assessment of user emotions and behaviour to development of products,* 77–88. New York, NY: Springer-Verlag

Ravaja, N. & Kivikangas, M. J. (2008). Psychophysiology of digital game playing: The relationship of self-reported emotions with phasic physiological responses. *Proceedings of Measuring Behaviour* (pp. 26–29), Maastricht, The Netherlands.

Ravaja, N., Laarni, J., Saari, T., Kallinen, K., Salminen, M., Holopainen, J., Järvinen, A. (2004). Spatial presence and emotional responses to success in a video game: A psychophysiological study. *Proceedings of the PRESENCE 2004 Conference.* (pp. 112-–116), Valencia, Spain.

Ravaja, N., Saari, T., Salminen, M., Laarni, J., & Kallinen, K. (2006). Phasic emotional reactions to video game events: A psychophysiological investigation. *Media Psychology*, *8*(4), 343–367.

Renaud, P., Proulx, J., Rouleau, J. L., Bradford, J., Fedoroff, P. & Bouchard, S. (2007). Dans la perspective visuelle du déviant sexuel : l'utilisation de personnages virtuels et de technologies de suivi oculomoteur en clinique de la déviance sexuelle. *Revue québécoise de psychologie, 28*(2), 31–42.

Rimmele, U., Seiler, R., Marti, B., Wirtz, P. H., Ehlert, U., & Heinrichs, M. (2008). The level of physical activity effect adrenal and cardiovascular reactivity to psychosocial stress. *Psychoneuroendocrinology, 34*, 190–198.

Rimmele, U., Zellweger, B. C., Marti, B., Seiler, R., Mohiyeddini, C., Ehlert, U., & Heinrichs, M. (2007). Trained men show lower cortisol, heart rate and psychological responses to psychosocial stress compared with untrained men. *Pseudoneuroendocrinology, 32*(2), 627–635.

Riva, G., Bacchetta, M., Baruffi, M., Rinaldi, S., & Molinari, E. (1999). VR based experiential cognitive treatment of anorexia nervosa. *Journal of Behavior Therapy and Experimental Psychiatry, 30*(3), 221–230.

Rizzo, A. A., Buckwalter, J. G., Humphrey, L., van der Zaag, C., Bowerly, T., Chua, C., et al. (2000). The virtual classroom: A VR environment for assessment and rehabilitation of attention deficits, *CyberPsychology and Behavior*, *3*(3), 483–499.

Robillard, G., Bouchard, S., Fournier, T., & Renaud, P. (2003). Anxiety and presence during VR immersion: A comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games. *CyberPsychology and Behavior, 6*(5), 467–476.

Rossen, B., Johnsen, K. J., Deladisma, A., Lind, S., & Lok, B. (2008). Virtual humans elicit skin-tone bias consistent with real-world skin-tone biases. In H. Prendinger, J. C. Lester, and M. Ishizuka (Eds.): *Proceedings of the 8th International Conference on Intelligent Virtual Agents (IVA 2008), Tokyo, Japan*, Lecture Notes in Computer Science (LNCS) 5208 (pp. 237–244), Berlin, Germany: Springer-Verlag.

Routhier, C. (2007). *Programme d'entraînement à la résilience militaire (PERM)*. Unpublished document: SQFT (Secteur Québec de la force terrestre).

Routhier, C. (2009). Military resilience training program: Executive report. Presented at the *Mental Health* [MH] *& Operational Stress Injury* [OSI] *Joint Speakers Bureau* [JSB] *Harmonization of MH Education Initiative*.

Seyama, J., & Nagayama. R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence, 16*(4), 337–351.

Sharma, R., Khera, S., Mohan, A., Gupta, N., & Ray, R. B. (2006). Assessment of computer games as a psychological stressor. *Indian Journal of Physiology and Pharmacology, 50*(4), 367–374.

Sjörs, A., Larsson, B., Dahlman, J., Falkmer, T., & Gerdle, B. (2009). Physiological responses to low-force work and psychosocial stress in women with chronic trapezius myalgia. *BMC* [BioMed Central] *Musculoskeletal Disorders, 10*(63).

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., et al. (2006). A virtual reprise of the Stanley Milgram obedience experiments. *PLoS One – Public Library of Science*, *1*(1), e39.

Solomon, Z., Mikulincer, M., & Arad, R. (1991). Monitoring and blunting: Implications for combat-related post-traumatic stress disorder. *Journal of Traumatic Stress*, *4*(2), 209–221.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory, (Form Y) (Self-Evaluation Questionnaire)*. Palo Alto, CA: Consulting Psychologist Press.

Stanney, K. M. (Ed.) (2002). *Handbook of virtual environments. design, implementation, and applications.* Mahwah, NJ: Lawrence Erlbaum Associates.

Stetz, M. C., Long, C. P., Schober, W. V., Cardillo, C. G., & Wildzunas, R. M. (2007). Stress assessment and management while medics take care of the VR wounded. *Annual Review of CyberTherapy and Telemedicine*, *5*, 165–171.

Stetz, M. C., Long, C. P., Wiederhold, B. K., & Turner, D. D. (2008). Combat scenarios and relaxations training to harden medics against stress. *Journal of CyberTherapy and Rehabilitation*, *1*(3), 239–246.

Stone, B. (2005). Serious gaming. *Defence Management Journal*, *31*.

Sweetser, P. (2007). *Emergence in games*. Hingham, MA: Charles River Media.

Tammen, H., & Loviscach, J. (2010). Emotion in video games: Quantitative studies? In C. Peter, E. Crane, M. Fabri, H. Agius, & L. Axelrod (Eds.) *Proceedings of the 4th International Workshop on Emotions in Human Computer Interaction: Designing for People, Liverpool, United Kingdom*, 25–29. Stuttgart, Germany: Fraunhofer-Verlag.

Taylor, L. N. (2006). *Not of woman born: Monstrous interfaces and monstrosity in video games*. Ph.D. Thesis dissertation. University of Florida.

Taylor, S. E., Seeman, T. E., Eisenberger, N. I., Kozanian, T. A., Moore, A. N., & Moons, W. G. (2010). Effects of a supportive or an unsupportive audience on biological and psychological responses to stress. *Journal of Personality and Social Psychology, 98*(1), 47–56.

Thompson, M. M., & McCreary, D. R. (2006). Enhancing Mental readiness in military personnel. In. T. W. Britt et al. (Eds.) *Military Life: The psychology of serving in peace and combat, 2*, 54–79.

Thompson, M. M., Pastò, L., & McCreary, D. R. (2002). Deployment stressors. coping. and psychological well-being among peacekeepers. Presented at the *7th International Command and Control Research and Technology Symposium,* Quebec City, QC, Canada.

Toet, A. (2010). Personal communication, March 2010.

Toet, A., van Welie, M., & Houtkamp, J. (2009). Is a dark virtual environment scary? *CyberPsychology & Behavior, 12*(4), 363–371.

Weber, R., Behr, K.-M., Tamborini, R., Ritterfeld, U., & Mathiak, K. (2009). What do we really know about first-person-shooter games? An event-related, high-resolution content analysis. *Journal of Computer-Mediated Communication*, 14(4), 1016–1037.

Wiederhold, B. K., & Wiederhold, M. D. (2008). VR for posttraumatic stress disorder and stress inoculation training. *Journal of CyberTherapy and Rehabilitation, 1*(1), 23-35.

Wiederhold, M. D. (2005). *Physiological monitoring during simulation training and testing (Final report)*. U.S. Army Aviation and Missile Command. San Diego, CA: Virtual Reality Medical Center.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225–240.

Wolpe, J. (1990). *The practice of behavior therapy (4th Ed.)*. New York, NY: Pergamon Press.

Youngblut, C. (2007). *What a* decade of experiments reveals about factors that influence the sense of presence: latest findings. Alexandria, VA: Institute for Defense Analyses.

Zamorski, M. A. (2010). *Report of the Canadian Forces expert panel on suicide prevention*. CS09-0407. Ottawa, ON, Canada: Canadian Forces Health Services Group Headquarters.

Zyda, M. (2005). From visual simulation to virtual reality to games. *Computer*, *38*(9), 25–32.

Bibliography

Black Hawk Striker (2004)

Call of Cthulhu: Dark Corners of the Earth (2005)

Clock Tower 3 (2003)

Condemned: Criminal Origins (2006)

Condemned 2 (2008)

Counter-Strike (2003)

Dead Space (2009)

Eternal Darkness (2002)

Fatal Frame II: Crimson Butterfly (2003)

F.E.A.R. (2005)

F.E.A.R. 2 (2009)

Half-Life 2 (2004)

Killing Floor (2009)

Left 4 Dead (2008)

Penumbra: Black Plague (2008)

Resident Evil 4 (2005)

Silent Hill 2 (2003)

Siren (2004)

Starship Eleven (2004)

Tetris (1989)

Vampire: The Masquerade – Bloodlines (2004)

This page intentionally left blank.

Annex A Experimental setup

		Setup 1 (L4D 22)	Setup 2 (KF 22)	Setup 3 (L4D 73)	Setup 4 (KF 73)	Setup 5 (KF CAVE)
Hardware	Display	Samsung 22" 120 Hz Model 2233RZ	Samsung 22" 120 Hz Model 2233RZ	HDTV Mitsubishi 73" (Rear projection DLP) Model 73735	HDTV Mitsubishi 73" (Rear projection DLP) Model 73735	4-side Flex from Fakespace Projectors: DP Highlight 8000Dsx (8000 lumens)
	Resolution	1680x1050@120Hz	1680x1050@120Hz	1920x1080@120Hz	1920x1080@120Hz	1400x1050@96Hz
	Video card	GeForce GTX 285 1024MB Benchmark: 115 fps	GeForce GTX 285 1024MB Benchmark: 115 fps	GeForce GTX 260 1792MB Benchmark: 106.1 fps	GeForce GTX 260 1792MB Benchmark: 106.1 fps	2 Quadro FX 5600 1500 MB
	Video quality cong.	4AA, 8AF, High quality	4AA, 8AF, High quality	4AA, 8AF, High quality	4AA, 8AF, High quality	4AA, 8AF, High quality
	Stereo glasses	3D Vision	3D Vision	3D Vision	3D Vision	Crystal Eyes 3
	Head Tracking	N/A	N/A	N/A	N/A	Intersence IS-900 VET
	Audio	Headset Tritton AX Pro	Headset Tritton AX Pro	Headset Tritton AX Pro	Headset Tritton AX Pro	700 Watts 5.1 Speakers System
	Sound card	Integration 7.1 Audio	Integration 7.1 Audio	Sound Blaster X-Fi Titanium	Sound Blaster X-Fi Titanium	Sound Blaster X-Fi Platinum
	Controller Interface	Keyboard + Mouse	Keyboard + Mouse	Keyboard + Mouse	Keyboard + Mouse	Intersense Wand
	Computer	Dell XPS 630	Dell XPS 630	Dell Studio XPS 9000	Dell Studio XPS 9000	BOXX 7500
	CPU	Core2 Q9550 2.83 Ghz	Core2 Q9550 2.83 Ghz	Core i7-920 2.66 Ghz	Core i7-920 2.66 Ghz	Dual Opt. 2222 (3.0 Ghz)
	Memory	8 GB	8 GB	12 GB	12 GB	4 GB
Software	Game	L4D (04-12-2009)	KF 3339 (128.29)	L4D (04-12-2009)	KF 3339 (128.29)	KF 3339 (128.29)
	Map/Level(s)	NoMercy - Hospital 04	Archives (modified)	NoMercy - Hospital 04	Archives (modified)	Archives (modified)
	Graphics drivers	Forceware 190.62	Forceware 190.62	Forceware 190.62	Forceware 190.62	Forceware 162.65
	Stereo drivers	3D Vision CD v1.11	3D Vision CD v1.11	3D Vision CD v1.11	3D Vision CD v1.11	VirtualUT
	Stereo setting	25%	25%	25%	25%	Eye separation: 4 cm
	Rendering library	DirectX 9	OpenGL 2.1	DirectX 9	OpenGL 2.1	OpenGL 2.1
	Operating system	Windows Vista 64	Windows Vista 64	Windows Vista 64	Windows Vista 64	Windows XP 32

Figure 23: Hardware and software configurations for all five experimental conditions

This page intentionally left blank.

Annex B KF configuration files

Many characteristics of the original KF game had to be modified for the experimentation. Some of these modifications were directly applied to the original game scripts but most of them were defined in a new game mod: `VirtualKF`. For instance, the entire game cycle (store mechanism between zombie waves, zombies' behaviour, automatic game restart after death, etc.) was handled in a new game type, `SMTGameType`, inherited from the original game type, `KFGameType`. With a specific configuration, it also allows the game to be rendered in a CAVE environment. This mod also modifies weapon handling and some visual effects such as motion blur and overlays. However, some game parameters such as game difficulty, game length and wave definitions (zombie count, species and squads) can be directly configured through configuration files.

Available zombie species were defined in the `[KFmod.KFGameType]` section of the `KillingFloor.ini` file. The `Mid` value gives the alias letter used to define the squads later in the configuration:

```
...
[KFmod.KFGameType]
...
MonsterClasses=(MClassName="KFChar.ZombieClot",Mid="A")
MonsterClasses=(MClassName="KFChar.ZombieCrawler",Mid="B")
MonsterClasses=(MClassName="KFChar.ZombieGoreFast",Mid="C")
MonsterClasses=(MClassName="KFChar.ZombieStalker",Mid="D")
MonsterClasses=(MClassName="KFChar.ZombieScrake",Mid="E")
MonsterClasses=(MClassName="KFChar.ZombieFleshpound",Mid="F")
MonsterClasses=(MClassName="KFChar.ZombieBloat",Mid="G")
...
```

In the same section, next to the species definition, zombie squads could be formed using a combination of digits and the corresponding alias letter:

```
[KFmod.KFGameType]
...
MonsterClasses=...
...
MonsterSquad=1A  //define squad 0 with 1 Clot
MonsterSquad=2A  //define squad 1 with 2 Clots
MonsterSquad=1A1C      //define squad 2 with 1 Clot AND 1 Gorefast
...
```

Squads can be formed of many species using multiple combinations in a single string (e.g. `1A1C`). The digit cannot exceed 9, but a combination like `5A5A` (for a total of 10 zombies) is allowed. Since only the Clot and Gorefast species were used, only the alias letters A and C were retained in the squad definitions. There can be as many squads as needed for species diversity and game difficulty. Each call to `MonsterSquad=#@` will create a new squad with a specific index. The order is important here as the squad index will be used in the wave definition.

Following the squad definition in the `[KFmod.KFGameType]` section, zombie waves can be defined using four parameters: the maximum zombie count, the duration in seconds, the difficulty level and the wave mask. The first two parameters are self explanatory. The difficulty level is a floating-point value (by default 1.0) scaling zombie capabilities such as health, speed and damage. The wave mask is a bitwise combination of squad indices allowing specific zombie squads during a wave. To include specific squads in a wave, one must specify the mask as the following sum:

$$Wavemask = \sum_i 2^i \; Wavemask = \sum_i \quad 2i \quad \text{, where } i \in \{\text{ allowed squad indices set }\}$$

During familiarisation, only the first 2 squads are used ($2^0 + 2^1 = 3$) with a constant difficulty level:

```
Waves[0]=    (WaveMask=3,WaveMaxMonsters=20,WaveDuration=255,
             WaveDifficulty=1.000000)
...
Waves[15]=   (WaveMask=3,WaveMaxMonsters=20,WaveDuration=255,
             WaveDifficulty=1.000000)
```

In the experimentation phase, the player can meet up to 24 different squads in 15 waves gradually increasing in difficulty:

```
Waves[0]=    (WaveMask=1,WaveMaxMonsters=10,WaveDuration=45,
             WaveDifficulty=1.000000)
Waves[1]=    (WaveMask=255,WaveMaxMonsters=15,WaveDuration=45,
             WaveDifficulty=1.000000)
Waves[2]=    (WaveMask=511,WaveMaxMonsters=20,WaveDuration=90,
             WaveDifficulty=1.000000)
Waves[3]=    (WaveMask=383,WaveMaxMonsters=20,WaveDuration=90,
             WaveDifficulty=1.000000)
Waves[4]=    (WaveMask=1023,WaveMaxMonsters=22,WaveDuration=90,
             WaveDifficulty=1.050000)
Waves[5]=    (WaveMask=1023,WaveMaxMonsters=24,WaveDuration=90,
             WaveDifficulty=1.100000)
Waves[6]=    (WaveMask=1023,WaveMaxMonsters=26,WaveDuration=90,
             WaveDifficulty=1.150000)
Waves[7]=    (WaveMask=16383,WaveMaxMonsters=28,WaveDuration=90,
             WaveDifficulty=1.200000)
Waves[8]=    (WaveMask=16383,WaveMaxMonsters=30,WaveDuration=90,
             WaveDifficulty=1.250000)
Waves[9]=    (WaveMask=32767,WaveMaxMonsters=30,WaveDuration=90,
             WaveDifficulty=1.30000)
Waves[10]=   (WaveMask=32767,WaveMaxMonsters=30,WaveDuration=90,
             WaveDifficulty=1.000000)
Waves[11]=   (WaveMask=32767,WaveMaxMonsters=30,WaveDuration=90,
             WaveDifficulty=1.100000)
Waves[12]=   (WaveMask=8388607,WaveMaxMonsters=30,WaveDuration=90,
             WaveDifficulty=1.200000)
Waves[13]=   (WaveMask=8388607,WaveMaxMonsters=30,WaveDuration=90,
             WaveDifficulty=1.300000)
Waves[14]=   (WaveMask=16777215,WaveMaxMonsters=30,WaveDuration=90,
             WaveDifficulty=1.400000)
Waves[15]=   (WaveMask=16777215,WaveMaxMonsters=30,WaveDuration=90,
             WaveDifficulty=1.500000)
```

Finally, the new game type `SMTGameType` was configured through `KillingFloor.ini` to modify initial health, speed, damage and range of attack for Clot and Gorefast zombies and make them more or less resistant to headshots.

An easy mode was defined for the familiarisation phase:

```
[VirtualKF.SMTGameType]
InitialCountDown=60
InitialClotSpeed=50
InitialClotHealth=10
```

```
InitialClotDamage=0
InitialClotRange=100
ClotHeadshotRobust=false
InitialGorefastSpeed=50
InitialGorefastHealth=10
InitialGorefastDamage=0
InitialGorefastRange=100
GorefastHeadshotRobust=false
```

During the experimentation, all enemies were more robust in every way:

```
[VirtualKF.SMTGameType]
InitialCountDown=15
InitialClotSpeed=120
InitialClotHealth=200
InitialClotDamage=8
InitialClotRange=80
ClotHeadshotRobust=true
InitialGorefastSpeed=150
InitialGorefastHealth=250
InitialGorefastDamage=8
InitialGorefastRange=100
GorefastHeadshotRobust=true
```

This page intentionally left blank.

Annex C L4D configuration scripts

Common script (Lines added to `autoexec.cfg`)

Description	*Required code lines*
// Enable cheat mode	`sv_cheats 1`
// Remove survivor bots	`director_no_survivor_bots 1`
// Disable enemies attack features	`survivor_no_pounce_or_hang 1`
// Crosshair enhanced	`cl_crosshair_alpha 255`
// Remove voice	`voice_enable 0` `gameinstructor_enable 0`

// Log	`version`
	`log on`
	`player_debug_print_damage 1`

// Execute right scenario	`exec autoexec_exp`

Add-on script for familiarisation (`autoexec_fam.cfg`)

Description	*Required code lines*
// No Fantastic infected	`director_no_specials 1`
// Blind infected	`nb_blind 1`
// Zombie hordes will not spawn	`director_no_mobs 1`
// Default difficulty level	`z_difficulty EASY`

Add-on script for experimentation (`autoexec_exp.cfg`)

Description	*Required code lines*
// No Fantastic infected	`director_no_specials 0`
// No Blind infected	`nb_blind 0`
// Zombie hordes will be spawned	`director_no_mobs 0`
// Reduce mob size	`z_mob_spawn_min_size 5` // from 10 `z_mob_spawn_max_size 15` // form 30
// Default difficulty level	`z_difficulty NORMAL`
// Reduce hit dmg of special zombies	`hunter_pz_claw_dmg 5` // from 10 `z_witch_damage 10` // from 30
// Reduce health of special zombies	`z_witch_health 250` // from 1000 `z_tank_health 500` // from 4000 `z_tank_incapacitated_health 700` // from 5000

This page intentionally left blank.

Annex D Game selections: detailed criteria rating

Table 34: Video games titles reviewed: stressful criterion

Titles	About[A] (2009)	Ask Men[B] (2010)	Bright Hub[C] (2010)	Gossip Gamers[D] (2009)	Joystick Division[E] (2010)	Pop Crunch[F] (2010)	Telegraph[G] (2009)	Total[H]
	Websites							
Call of Cthulhu Dark Corners of the Earth (2005)	1	-	1	-	-	-	-	2
Condemned: Criminal Origins (2006)	-	1	-	1	1	-	-	3
Dead Space (2009)	1	1	1	-	1	1	1	(6) 5
Fatal Frame II: Crimson Butterfly (2003)	-	-	-	1	1	1	1	4
F.E.A.R. (2005)	1	-	1	1	1	1	-	5
F.E.A.R. 2 (2009)	1	1	-	-	-	-	-	2
Killing Floor (2009)	1	-	-	-	-	-	-	1
Left 4 Dead (2008)	1	1	1	-	-	-	-	3
Penumbra: Black Plague (2008)	1	-	1	-	-	-	-	2
Resident Evil 4 (2005)	1	1	-	-	-	-	-	2
Silent Hill 2 (2003)	-	-	-	1	1	1	1	4

Notes
A – http://compactiongames.about.com/od/topgames/tp/scarygames.htm
B – http://www.askmen.com/top_10/entertainment/top-10-scariest-video-games.html
C – http://www.brighthub.com/video-games/pc/articles/46528.aspx
D – http://www.gossipgamers.com/top-10-scariest-games-of-all-time/
E – http://www.joystickdivision.com/2010/02/top_ten_scariest_games_of_all.php
F – http://www.popcrunch.com/the-13-scariest-video-games-of-all-time/
G – http://www.telegraph.co.uk/technology/video-games/6455534/Top-10-scariest-video-games.html
H – Maximum rating = 5

Table 35: Video games titles reviewed: ethical criterion

	ESRB Initial Rating[A]	ESRB Detailed Rating

Titles	Mature (+5)	Language (0) or Strong Language (-1)	Use of Drugs and Alcohol (-1)	Violence (0) or Intense Violence (-1)	Blood (0) or Blood and Gore (-1)	Partial Nudity (0) or Sexual Theme (-1)	Total
Call of Cthulhu: Dark Corners of the Earth (2005)	5	-1	-1	-1	-1	-	**1**
Condemned: Criminal Origins (2006)	5	-1	-	-1	-1	-	**2**
Dead Space (2009)	5	-1	-	-1	-1	-	**3**
Fatal Frame II: Crimson Butterfly (2003)	5	-	-	0	0	-	**5**
F.E.A.R. (2005)	5	-1	-	-1	-1	-	**2**
F.E.A.R. 2 (2009)	5	-1	-	-1	-1	-1	**1**
Killing Floor (2009)	5	-1[B]	-	-1	-1	0[B]	**3**
Left 4 Dead (2008)	5	0[B]	-	-1	-1	-	**3**
Penumbra: Black Plague (2008)	5	0	-	0	-1	-	**4**
Resident Evil 4 (2005)	5	-1	-	-1	-1	-	**2**
Silent Hill 2 (2003)	5	0	-	-1	-1	-1	**2**

Notes:

A – Entertainment Software Rating Board – http://www.esrb.org/

B – Because of game modding capabilities, these elements could be removed

Table 36: Video games titles reviewed: experiment friendly criterion

Titles	Game genre(s)[A]				Features		Total
	First Person Shooter (+3)	Third Person Shooter (+2)	Action (+1)	Adventure and/or RPG (0)	Short introduction (+1)	Repeatability: linear (+1) vs. sandbox	
Call of Cthulhu: Dark Corners of the Earth (2005)	-	-	1	0[B]	-	-	**1**
Condemned: Criminal Origins (2006)	-	-	1	-	1	1	**3**
Dead Space (2009)	-	2	-	-	-	1	**3**
Fatal Frame II: Crimson Butterfly (2003)	-	-	1	0[B]	-	-	**1**
F.E.A.R. (2005)	3	-	-	-	-	1	**4**
F.E.A.R. 2 (2009)	3	-	-	-	-	1	**4**
Killing Floor (2009)	3	-	-	-	1	-	**4**
Left 4 Dead (2008)	3	-	-	-	1	1	**5**
Penumbra: Black Plague (2008)	-	-	-	0[C]	-	-	**0**
Resident Evil 4 (2005)	-	2	-	-	-	1	**3**
Silent Hill 2 (2003)	-	-	1	0[B]	1	-	**2**

Notes:

A – Game genres extracted from http://en.wikipedia.org

B – Action-Adventure genre

C – Graphical Adventure genre

Table 37: Video games titles reviewed: VR compatible criterion

Titles	Experimental Setups		Total
	22-inch monitor & 73-inch TV with NVIDIA 3D Vision[A]	CAVE	
Call of Cthulhu: Dark Corners of the Earth (2005)	0[B]	0	**0**
Condemned: Criminal Origins (2006)	0[B]	0	**0**
Dead Space (2009)	2[A]	0	**2**
Fatal Frame II: Crimson Butterfly (2003)	0[A]	0	**0**
F.E.A.R. (2005)	4[A]	0	**4**
F.E.A.R. 2 (2009)	2[A]	0	**2**
Killing Floor (2009)	4[A]	1	**5**
Left 4 Dead (2008-2010)	4[A]	0[C]	**4**
Penumbra: Black Plague (2008)	0[B]	0 [C]	**0**
Resident Evil 4 (2005)	0[B]	0	**0**
Silent Hill 2 (2003)	0[B]	0	**0**

Notes:

A – NVIDIA 3D vision rating (Fully supported or Excellent = 4, Good = 3, Fair = 2, Poor = 1 & Unsupported = 0) http://www.nvidia.com/object/3d-vision-3d-games.html

B – Unrated

C – Not implemented, but Scripts, SDK and open source code suggest it could be adapted to a CAVE system

Table 38: Video games titles reviewed: adaptable criterion

Titles	Modding capabilities[A]						Total
	File replacement (+1)	Config and/or (+1)	Script files (+1)	C++ Interface files (+1)	In-game console (+1)	Level editor (+1)	
Call of Cthulhu: Dark Corners of the Earth (2005)	-	-	-	-	-	-	**0**
Condemned: Criminal Origins (2006)	-	-	-	-	-	-	**0**
Dead Space (2009)	-	-	-	-	-	-	**0**
Fatal Frame II: Crimson Butterfly (2003)	-	-	-	-	-	-	**0**
F.E.A.R. (2005)	1	1	1	1	-	1	**5**
F.E.A.R. 2 (2009)	-	-	-	-	-	-	**0**
Killing Floor (2009)	1	1	1	-	1	1	**5**
Left 4 Dead (2008)	1	1	1	1	1	1	**6(5)**
Penumbra: Black Plague (2008)[B]	1	-	-	-	-	-	**1**
Resident Evil 4 (2005)	1	-	-	-	-	-	**1**
Silent Hill 2 (2003)	-	-	-	-	-	-	**0**

Notes:

A – No selected title is available in "open source"

B – The first of the series, *Penumbra: Overture* (2007), is now available in open source

This page intentionally left blank.

List of symbols/abbreviations/acronyms/initialisms

3D	Three-Dimensional
AI	Artificial Intelligence
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ANSI	American National Standards Institute
CAVE	Cave Automatic Virtual Environment
CF	Canadian Forces
CPU	Central Processor Unit
DLP	Digital Light Processing
DRDC	Defence Research & Development Canada
DRDKIM	Director Research and Development Knowledge and Information Management
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders Fourth Edition
DTS	Digital Theatre Sound
ECG	Electrocardiogram
EKG	Electrocardiogram
EPDSP	Enhanced Post Deployed Screening Process
ESRB	Entertainment Software Rating Board
FAC	Foundations of Augmented Cognition
FOV	Field of View
FPS	First-Person Shooter
GSR	Galvanic Skin Response
I/ITSEC	Interservice/Industry Training, Simulation and Education Conference
IKM	Ignorance and Knowledge Management
KAF	Known Avatar First
KF	Killing Floor
ITQ	Immersive Tendencies Questionnaire
L4D	Left 4 Dead
LNCS	Lecture Notes in Computer Science
MHAT-V	Mental Health Advisory Team V
ns	Not Significant
PANAS	Positive And Negative Affect Schedule
PERM	Programme d’entraînement à la résilience militaire
PTSD	Post-traumatic Stress Disorder
RPG	Role Playing Game

R&D	Research & Development
R2MR	Road to Mental Readiness
S&T	Science & Technology
SCID-NP	Structures Clinical Interview for DSM Disorders – Non Patient Edition
SEAL	SEa-Air-Land
SMT	Stress Management Training
SoS	System of Systems
SQFT	Secteur Québec de la force terrestre
SSQ	Simulator Sickness Questionnaire
STAI-Y1	State-Trait Anxiety Inventory Y1
TSST	Trier Social Stress Test
TV	Television
UE2	Unreal Engine 2
UAF	Unknown Avatar First
UQO	Université du Québec en Outaouais
US	United States
VET	Virtual Environment Tracker
VR	Virtual Reality

DOCUMENT CONTROL DATA

(Security markings for the title, abstract and indexing annotation must be entered when the document is Classified or Designated)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence Research and Development Canada – Valcartier 2459 Pie-XI Blvd North Quebec (Quebec) G3J 1X5 Canada	2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) UNCLASSIFIED 2b. CONTROLLED GOODS (NON-CONTROLLED GOODS) DMC A REVIEW: GCEC JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Initial steps in inducing stress to teach stress management skills to soldiers : Testing stressors, immersive technologies and avatars

4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used)

Bouchard, S., Bernier, F., Boivin, É.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
April 2012	110	160

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence Research and Development Canada – Valcartier
2459 Pie-XI Blvd North
Quebec (Quebec)
G3J 1X5 Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
14da	

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
DRDC Valcartier TR 2011-195	

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

Unlimited

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

Two studies were conducted to test whether: (a) it is possible to stress soldiers using virtual reality, and (b) it is worth developing virtual characters resembling real people experiencing strong pain in order to elicit stress. The first study was conducted with 47 soldiers who played two 3D horror games using three different types of immersive technologies. The second study was conducted with 42 civilians that were familiar with one of two virtual characters experiencing pain. Both studies were effective in inducing stress among participants. Results were supporting of our work devising an effective low-cost and high buy-in approach to assist in teaching and practicing stress management skills.

Deux études ont été menées afin de tester si : (a) il s’avère possible de stresser des soldats à l’aide de la réalité virtuelle, et (b) il s’avère pertinent comme stresseur de développer des personnages virtuels ressemblant à des personnes connues par les soldats et ressentant une forte douleur. La première étude repose sur un échantillon de 47 soldats qui ont joué à des jeux d’horreur 3D sur trois différentes plateformes immersives. La seconde étude a été menée avec 42 civils qui étaient familiers ou non avec un personnage virtuel ressentant de la douleur. Les résultats valident les démarches pour développer une solution efficace, peu coûteuse et attrayante pour supporter l’entraînement aux stratégies de gestion du stress.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

stress; video games; virtual reality; immersion; avatars; empathy

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE R&D DÉFENSE

www.drdc-rddc.gc.ca