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

DEFENCE R&D DÉFENSE

Training Effectiveness of the Victoria Class Virtual Submarine:

A behavioural assessment of learning a complex task within a virtual environment

Lochlan E. Magee
Aidan A. Thompson
Brad Cain
Courtney Kersten

Defence R&D Canada
Technical Report
DRDC Toronto TM 2012-014
April 2012

Canada

Training Effectiveness of the Victoria Class Virtual Submarine:

A behavioural assessment of learning a complex task within a virtual environment

Lochlan E. Magee
Aidan A. Thompson
Brad Cain
Courtney Kersten

Defence R&D Canada – Toronto

Technical Report
DRDC Toronto TR 2012-014
April 2012

Principal Author

Original signed by Lochlan E. Magee

Lochlan E. Magee

Defence Scientist

Approved by

Original signed by Linda Bossi

Linda Bossi

Head, Human Systems Integration Section

Approved for release by

Original signed by Stergios Stergiopoulos

Stergios Stergiopoulos

Acting Chief Scientist

This work was performed to support DMTE as part of 14dn, Virtual Reality for Training

In conducting the research described in this report, the investigators adhered to the policies and procedures set out in the Tri-Council Policy Statement: Ethical conduct for research involving humans, National Council on Ethics in Human Research, Ottawa, 1998 as issued jointly by the Canadian Institutes of Health Research, the Natural Sciences and Engineering Research Council of Canada and the Social Sciences and Humanities Research Council of Canada.

Abstract

The Royal Canadian Navy eLearning Centre of Expertise (NeLCoE) in Quebec City developed the Canadian Virtual Naval Fleet (CVNF), a game-based, desk-top virtual environment (VE) to provide procedural and spatial knowledge of large vessels, which are not always available for training. This report presents the results of two experiments that assessed the training effectiveness of the Victoria Class Virtual Submarine (VCVS), one implementation of the CVNF. Each experiment assessed the ability of navy personnel to complete an emergency drill that involved isolation of a bulkhead. Initial transfers of training and improvements with practice to criterion (i.e., error-free performance) were used to compare the performances of a trained group with a novice group. However, the experiments differed in method, prior training of the groups, and the practice environment. The first experiment employed a reverse transfer of training paradigm using the VCVS as the practice environment. In this experiment, the performances of a group of ten Navy personnel qualified to perform the task aboard the submarine were compared to a group of ten novices who lacked prior submarine experience and who learned the task for the first time within the VE. The second experiment employed a more classic, forward transfer of training paradigm that used the submarine as the transfer environment. Here, the performances of the novice group that practiced the task to criterion within the VE were compared to a second group of ten novices who had the task demonstrated to them in a traditional instructional session aboard the submarine. Two days later, the members of each group individually demonstrated their ability to perform the task aboard the submarine and repeated their attempts until they could perform the task without error. Both experiments yield evidence of positive training transfer, especially for spatial knowledge.

Résumé

L’équipe du Centre d’expertise de l’apprentissage en ligne de la Marine royale canadienne à Québec a développé une flotte navale virtuelle canadienne (FNVC), c’est-à-dire un environnement virtuel pour ordinateur basé sur le jeu, et ce, afin de fournir des connaissances procédurales et spatiales sur les grands navires, qui ne sont pas toujours disponibles pour la formation. Ce rapport présente les résultats de deux expériences destinées à évaluer l’efficacité de la formation avec le sous-marin virtuel de la classe VICTORIA (SVCV), une initiative de la FNVC. Chaque expérience a évalué la capacité du personnel de la Marine à effectuer un exercice d’alerte comprenant l’isolation d’une cloison. Le transfert initial de la formation et l’amélioration des résultats grâce à la répétition (c.-à-d. jusqu’à l’obtention d’un rendement exempt d’erreur) ont été utilisés pour comparer les performances d’un groupe de participants entraînés et d’un groupe de novices. Néanmoins, les expériences différaient sur les plans de la méthode, de l’entraînement antérieur des groupes et de l’environnement d’exercice. La première expérience avait recours à un transfert inversé du paradigme de formation et utilisait le SVCV comme environnement d’exercice. Dans le cadre de cette expérience, les performances d’un groupe de dix membres de la Marine qualifiés pour réaliser la tâche à bord du sous-marin ont été comparées à celles d’un groupe de dix novices qui ne possédaient pas d’expérience antérieure à bord d’un sous-marin et qui ont appris à exécuter la tâche pour la première fois dans l’environnement virtuel. La deuxième expérience employait un paradigme de transfert de la formation plus traditionnel et utilisait le sous-marin comme environnement de transfert. Dans ce cas, les performances du groupe de novices s’étant exercés à accomplir la tâche sans commettre d’erreur dans un environnement virtuel ont été comparées avec celles d’un groupe de dix novices qui avaient assisté à une démonstration de la tâche à l’occasion d’une séance de formation traditionnelle à bord du sous-marin. Deux jours plus tard, les membres de chaque groupe ont démontré individuellement leur capacité à accomplir la tâche en question à bord du sous-marin, puis ont répété leurs tentatives jusqu’à ce qu’ils soient en mesure de la réaliser sans faire d’erreur. Les deux expériences ont prouvé l’existence de transfert de formation positif, particulièrement en ce qui a trait aux connaissances spatiales.

Executive summary

Training Effectiveness of the Victoria Class Virtual Submarine: A behavioural assessment

[Lochlan Magee; Aidan Thompson, Brad Cain; Courtney Kersten]; DRDC Toronto TR 2012-014; Defence R&D Canada – Toronto; April 2012.

Introduction: Access to naval equipment and vessels for training is diminishing in many countries. Consequently, there is international interest in the use of virtual environments for training as a supplement or replacement for traditional training methods that are reliant on the availability of operational systems. Within the Department of National Defence (DND) Canada, the Royal Canadian Navy eLearning Centre of Expertise (NeLCoE) in Quebec City developed the Canadian Virtual Naval Fleet (CVNF), a game-based, desk-top virtual environment (VE) as a solution for training navy personnel who need to gain procedural and spatial knowledge of naval vessels.

In March, 2011, the Director Maritime Training and Education (DMTE) requested an objective evaluation of the Victoria Class Virtual Submarine (VCVS), one of the first implementations of the CVNF. The main purposes of this study were to determine whether the VCVS provides effective training for a complex task or if there is a need for further development of the VE, such as the addition of a walking interface for controlling self-directed motion.

Experimental Methods: This report presents the results of two experiments that assessed the training effectiveness of the VCVS in late November and early December 2011 at Canadian Forces Base (CFB) Esquimalt. Each experiment assessed the ability of Navy personnel to complete a complex emergency procedure that was chosen by subject matter experts (SMEs) as the task for this study. Submariners need to know their way about the submarine and what to do in an emergency, such as a fire or flood, since improper procedures can be fatal. One of the many emergency procedures that must be known by all qualified submariners aboard a Victoria Class submarine is isolation of Bulkhead 35 (BH 35), which compartmentalizes the submarine. The submariners must know the locations, names, functions and operation of the valves and tools that are needed to perform this drill.

Initial transfer of training and improvement with practice to task criterion (i.e., error-free performance) were used to compare the performances of a trained group with a novice group. However, the experiments differed in method, prior training of the groups, and the practice environment. The first experiment employed a reverse transfer of training paradigm using the VE. In this experiment, the performances of ten Navy personnel qualified to perform the task aboard the submarine were compared to ten novices who lacked prior submarine experience and who learned the task for the first time within the

VE. The second experiment employed a more classic, forward transfer of training paradigm that used an actual Victoria Class submarine, Her Majesty's Canadian Ship (HMCS) Corner Brook, as the transfer environment. Here, the performances of the novice group that practiced the task to criterion within the VE were compared to a second group of ten novices who had the task demonstrated to them in a traditional instructional session aboard the submarine. Two days later, the members of each group individually demonstrated their ability to perform the task aboard the submarine and repeated their attempts until they could perform the task without error. The associations between the participants' prior experience with computer games and their abilities to perform the task within the virtual and real environments were assessed.

Results: Qualified submariners initially had difficulty completing the task to criterion within the VE due to procedural errors, but they rapidly improved with practice. The qualified submariners were able to navigate without error within the VE within only a few trials. In comparison, participants who lacked prior submarine experience required many more trials to achieve task criterion within the VE, but they too required more procedural than navigational practice to achieve error free performance. Several factors may explain the initial low, reverse transfer of training from the submarine to the VE, including subtle differences between the submarine and its computer-generated representation, the experimental rigour that we required for valve names and function, the lack of haptic cues (i.e., the "feel" of tools) within the VCVS and the decay of human memory for procedural drills. However, the difference that was found in the rate of improvement with practice of the two groups provides behavioural evidence of reverse transfer since the qualified personnel outperformed the novices within the VE.

Forward transfer of training was assessed with HMCS Corner Brook, which was docked at CFB Esquimalt. All participants were required to find, identify, and explain the function of each valve required to isolate BH 35 within the submarine. The participants who previously learned the task within the VE, and who did not receive any familiarization training aboard the submarine, completed the task aboard the submarine as well as the members of the group that received familiarization training aboard the submarine. Notably, five members (50%) of the group trained with the VCVS made no spatial error on their first attempt to perform the task aboard the submarine, whereas only one member (10%) of the group familiarized with the task aboard the submarine managed to perform as well.

No meaningful correlations between computer game experience and task performance within the VE or aboard the submarine were discovered.

Significance: Both experiments yield evidence of statistically significant positive training transfer, especially for spatial knowledge. Although some modifications of the human interface could possibly improve the transfer of procedural knowledge, there is no apparent need to consider prior gaming experience or the addition of a walking interface

to expect beneficial training from practice with the VCVS. The behavioural evidence indicates high cost effectiveness.

Sommaire

Training Effectiveness of the Victoria Class Virtual Submarine: A behavioural assessment

[Lochlan Magee; Aidan Thompson, Brad Cain; Courtney Kersten] ; DRDC Toronto TR 2012-014 ; R & D pour la défense Canada – Toronto; avril 2012.

Introduction : L'accès à l'équipement naval et aux navires pour la formation est de plus en plus limité dans de nombreux pays. Par conséquent, certains d'entre eux démontrent un intérêt à l'égard de l'utilisation d'environnements virtuels à des fins de formation, et ce, afin de compléter ou remplacer les méthodes de formation traditionnelles qui reposent sur la disponibilité des systèmes opérationnels. Au sein du ministère de la Défense nationale (MDN) du Canada, l'équipe du Centre d'expertise de l'apprentissage en ligne de la Marine royale canadienne à Québec a développé une flotte navale virtuelle canadienne (FNVC), c'est-à-dire un environnement virtuel pour ordinateur basé sur le jeu, et ce, à des fins de formation du personnel de la marine qui doit acquérir des connaissances procédurales et spatiales sur les navires.

En mars 2011, le Directeur – Instruction et éducation maritimes (DIEM) a demandé la réalisation d'une évaluation objective du sous-marin virtuel de la classe VICTORIA (SVCV), l'une des premières initiatives de la FNVC. Le principal objectif de l'étude consistait à déterminer si le SVCV offre une formation efficace en vue de réaliser une tâche complexe ou si l'environnement virtuel doit être développé davantage, par exemple grâce à l'ajout d'une interface mobile pour le contrôle des déplacements.

Méthodes expérimentales : Le présent rapport présente les résultats de deux expériences destinées à évaluer l'efficacité de la formation avec le SVCV à la fin novembre et au début décembre 2011 à la Base des Forces canadiennes (BFC) Esquimalt. Chaque expérience évaluait la capacité du personnel de la Marine à exécuter une procédure d'urgence complexe qui avait été choisie par des spécialistes en la matière en guise de tâche pour cette étude. Les sous-mariniers doivent savoir comment s'orienter dans le sous-marin et quoi faire en cas de situation d'urgence, comme un incendie ou une infiltration d'eau, car le non-respect des procédures peut être fatal. L'une des nombreuses procédures d'urgence que tous les sous-mariniers qualifiés doivent connaître à bord du SVCV est l'isolation de la cloison 35, qui compartimente le sous-marin. Les sous-mariniers doivent connaître les emplacements, les noms, les fonctions et les opérations des valves et des outils qui sont nécessaires aux fins de l'exercice.

Le transfert initial de la formation et l'amélioration des résultats grâce à la répétition (c.-à-d. jusqu'à l'obtention d'un rendement exempt d'erreur) ont été utilisés pour comparer les performances d'un groupe de participants entraînés et d'un groupe de novices. Néanmoins, les expériences différaient sur le plan de la méthode, de l'entraînement antérieur des groupes et de l'environnement d'exercice. La première expérience avait

recours à un transfert inversé du paradigme de formation et utilisait l'environnement virtuel. Dans le cadre de cette expérience, les performances d'un groupe de dix membres de la Marine qualifiés pour réaliser la tâche à bord du sous-marin ont été comparées à celles d'un groupe de dix novices qui ne possédaient pas d'expérience antérieure à bord d'un sous-marin et qui ont appris à exécuter la tâche pour la première fois dans l'environnement virtuel. La deuxième expérience employait un paradigme de transfert de la formation plus traditionnel et utilisait un véritable SVCV, soit le Navire canadien de Sa Majesté (NCSM) CORNER BROOK, comme environnement de transfert. Dans ce cas, les performances du groupe de novices s'étant exercés à accomplir la tâche sans commettre d'erreur dans un environnement virtuel ont été comparées avec celles d'un groupe de dix novices qui avaient assisté à une démonstration de la tâche à l'occasion d'une séance de formation traditionnelle à bord du sous-marin. Deux jours plus tard, les membres de chaque groupe ont démontré individuellement leur capacité à accomplir la tâche en question à bord du sous-marin, puis ont répété leurs tentatives jusqu'à ce qu'ils soient en mesure de la réaliser sans faire d'erreur. La relation entre l'expérience antérieure des participants avec les jeux électroniques et leurs capacités à exécuter la tâche dans les environnements virtuel et réel a été évaluée.

Résultats : Au début, les sous-mariniers qualifiés avaient de la difficulté à exécuter la tâche sans commettre d'erreur dans l'environnement virtuel en raison d'erreurs procédurales, mais ils se sont rapidement améliorés en s'exerçant. Les sous-mariniers qualifiés étaient en mesure de naviguer sans erreur dans l'environnement virtuel après seulement quelques essais. En comparaison, les participants qui n'avaient pas d'expérience préalable à bord d'un sous-marin ont dû effectuer beaucoup plus d'essais pour accomplir la tâche sans commettre d'erreur dans l'environnement virtuel; néanmoins, ils ont eux aussi eu besoin de s'exercer davantage sur les plans de la procédure et de la navigation pour réaliser une performance exempte d'erreur. Plusieurs facteurs peuvent expliquer les faibles résultats initiaux, le transfert inversé de formation du sous-marin vers l'environnement virtuel, y compris les différences subtiles entre le sous-marin et la représentation conçue par ordinateur, la rigueur expérimentale que nous avons exigée à l'égard du nom et de la fonction des soupapes, le manque d'indices haptiques (c.-à-d. la sensation de toucher des outils) dans le SVCV ainsi que l'altération de la mémoire humaine à l'occasion des exercices procéduraux. Néanmoins, la différence constatée chez les deux groupes en ce qui a trait au taux d'amélioration grâce à la répétition offre une preuve comportementale de transfert inversé, car le personnel qualifié a obtenu des meilleurs résultats que les novices dans l'environnement virtuel.

Le transfert de formation a été évalué à bord du NCSM CORNER BROOK, qui était ancré à la BFC Esquimalt. Tous les participants devaient trouver, identifier et expliquer la fonction de chaque soupape requise pour isoler la cloison 35 à bord du sous-marin. Les participants qui ont déjà appris à exécuter la tâche dans l'environnement virtuel et qui n'ont pas reçu de formation à bord du sous-marin ont accompli la tâche de manière aussi satisfaisante que les membres du groupe ayant suivi leur formation à bord de ce dernier. Ainsi, cinq membres (50 p. 100) du groupe formé avec le SVCV n'ont commis aucune erreur spatiale à leur première tentative d'effectuer la tâche à bord du sous-marin, alors qu'un seul membre (10 p. 100) du groupe ayant suivi une formation à bord du sous-marin a réussi à en faire autant.

Aucune corrélation significative entre l'expérience avec les jeux électroniques et l'exécution de la tâche dans l'environnement virtuel ou à bord du sous-marin n'a été constatée.

Importance : Les deux expériences ont prouvé l'existence d'un transfert de formation positif et significatif sur le plan statistique, particulièrement en ce qui a trait aux connaissances spatiales. Même si certaines modifications de l'interface humaine pourraient améliorer le transfert de connaissances procédurales, il n'apparaît pas nécessaire de tenir compte de l'expérience préalable avec les jeux électroniques ou d'ajouter une interface mobile pour profiter des avantages de la formation avec le SVCV. La preuve comportementale indique un rapport coût-efficacité élevé.

Table of contents

List of figures

Acknowledgements

The authors would like to thank the officers and crew of HMCS Corner Brook and the staff of CFB Esquimalt in British Columbia for making this project possible. In particular, we would like to thank the following: LS Nancy Daigle of MARPAC HQ for recruiting and coordinating all the personnel awaiting training (PATs) who participated in the study; HMCS Corner Brook Day Workers PO2 Will Lumsden and PO2 Jean Benoit who familiarised us with the boat and safety procedures; Executive Officer of the HMCS Corner Brook LCdr Phil Collins for his hospitality and support; the Duty Watch Supervisors PO1 Tom Levesque, PO1 Bruce Mavor, PO2 Jim Bartlett, PO2 Ron Lyle, and PO2 Rico Pitre and all the members of their duty watches for assistance in ensuring the safety of the participants and the experimenters aboard the boat and for ensuring that the conduct of the experiment progressed smoothly in harmony with the day-to-day operations of the boat. We also wish to thank the officers, crew members, and trainees of HMCS Chicoutimi who took time to provide us with additional feedback on the VCVS.

The authors wish to thank and acknowledge Major François Gilbert for providing the opportunity to assess the VCVS and for providing the administrative coordination necessary to implement the study. We thank his development team and the instructional designers of the VCVS, including Mr Guy Boulet and Mr Kamil Andrzejewski, who also provided software and technical support for the experimental implementation.

Introduction

Access to naval equipment and vessels for training is diminishing in many countries. Consequently, there is international interest in the use of virtual environments for training as a supplement or replacement for traditional training methods that are reliant upon the availability of operational systems. Within the Department of National Defence Canada (DND), the Royal Canadian Navy eLearning Centre of Expertise (NeLCoE), in Quebec City, undertook the development of the Canadian Virtual Naval Fleet (CVNF), a game-based, desk-top virtual environment (VE) as a solution for training navy personnel who need to gain procedural and spatial knowledge of large vessels.

In March, 2011, the Director Maritime Training and Education (DMTE) requested an objective evaluation of the Victoria Class Virtual Submarine (VCVS), one of the first implementations of the CVNF. The main purposes of this study were to determine whether the VCVS provides effective training or if there is a need for further development, such as the addition of a walking interface for controlling self-directed motion within the VE.

Background

Relatively inexpensive VEs can be used successfully for training (Seidel and Chatelier, 1997), although attention must be given to their human factors (Stone, 2008). Virtual environments cannot always successfully replace operational equipment or the real world for training because it is not practical or feasible to replicate all of the sensory cues that are available with the real equipment in the real environment. Effective haptic interfaces are a particularly challenging engineering hurdle (NAE, 2009) and so too are walking interfaces for exploring VEs that represent large, complex spaces (Sottilare et al., 2010).

Various devices have been invented to provide a walking interface for controlling self-directed movements within the VE as well as the immersive experience thought to be necessary for acquiring spatial knowledge of a large unfamiliar structure such as a ship, submarine or oil rig. The practical question is, “Are these devices worth the investment?” That is, “Is it necessary to invest in an expensive human interface to a VE so that trainees can learn to navigate within an unfamiliar structure?” The answer to these questions has particular relevance to the Royal Canadian Navy (RCN) because access to naval vessels for training is not always possible due to their competing operational use or maintenance.

Accidents at sea have also prompted international development and investigation of VE solutions. In 2009, Stone et al. reported plans for assessing the training effectiveness of *SubSafe*, a VE created for the United Kingdom (UK) Ministry of Defence (MOD) by the Human Factors Integration Defence Technology Centre (HFI DTC) of the University of Birmingham, from free software. *SubSafe* was designed to be a low-cost, interactive, real-time, three-dimensional model that would supplement classroom instruction by enhancing student awareness of the location of safety equipment aboard a Trafalgar Class submarine. Stone et al. (2010) later reported the abilities of three groups of trainees to locate safety equipment aboard the submarine, as judged by an instructor. The groups differed in their prior exposure to the VE. The week before testing, the members of one group received an instructor-led tour of the VE that displayed the submarine and its safety equipment on a projection screen, the members of a second group were able to interact personally with the VE through a key-board to search for the safety equipment, and the members of a third group were not provided access to the VE, although all groups had concurrent access to the submarine for familiarization training.

Stone et al. (2010) found that exposure to the VE provided a small, but statistically reliable positive benefit for performance, but no difference in performance was found between the group that explored the VE actively and the one that did not. A fourth group was later added to the study; for this group, instructor-led, VE use was integrated into early sessions of the curriculum of a basic submarine course. Additional improvement in performance aboard the submarine was found. However, the investigators were concerned about a possible misinterpretation of the findings because "… the actual increase in the number of walk-around questions answered correctly looks pitifully small!" (p. 237). The investigators speculated that concurrent access to the submarine reduced the size of the effect attributable to VE use; they expressed concern that the cost-benefit of *SubSafe* would therefore be underestimated. For unrelated reasons, the investigators were unable to carry out their plan to engage new students before they were introduced to the actual submarine (Stone, 2012). Consequently, doubt remains about the effect-size and the cost-benefit of training within a VE for the type of task and environment that *SubSafe* can simulate.

Human navigation involves route planning and the use of spatial knowledge with position and orientation information gained during travel (Loomis et al., 1993) and the integration of information from many sources (Wolbers and Hegarty, 2010). Both external and internal sources of information (cues) can help humans determine their position and orientation in the environment. External cues can be visual, auditory, tactile, or olfactory and some of these cues can be integrated over time to determine current position and orientation (spatial awareness). Internal cues associated with self-motion (over distances less than 20 m or so), which inform a process known as path integration, are also used (Wolbers and Hegarty, 2010). In addition, geometric cues associated with familiar structures (e.g., the rectangular shape of a room) can benefit active locomotion. Consequently, there is reason to question the gain in spatial knowledge that can be achieved by interacting with a VE of a large irregular structure, while seated, using a keyboard or mouse to control movement within the

VE, which is visually displayed on a small screen with no other sensory cues for learning about the spatial environment.

Grant & Magee (1998) for instance found that experience with a VE did not shorten later paths to disparate objects within a large, irregular space if a joystick was used to explore the VE. In comparison, they found that exploration of the VE controlled by movements of the feet, which mimicked walking, led to reliably shorter paths. Importantly, this finding was found for a navigational task that required the participants to use the spatial knowledge that they gained from the VE to find short cuts in the real one. Grant and Magee varied the start location and order in which the objects needed to be found in the real world; hence, the participants could not rely on a memorized sequence of point to point routes and instead needed to possess a survey representation (Loomis et al., 1993; Werner et al., 1997) of the spatial layout of the environment to minimize path length. The ability to use spatial knowledge this way is a necessary objective for training emergency drills within a VE because operational personnel will not always be at an ideal location in the real world when an alarm sounds.

Many important military tasks in need of training require not only the application of spatial knowledge, but procedural knowledge as well. Knowing the whereabouts and best route to a location (spatial knowledge) and knowing what to do and how to do things when you get there (procedural knowledge) are often necessary for successful task completion. An every day example is knowledge of the location of the nearest fire extinguisher in an office building and what to do with the extinguisher when a fire occurs. A more complex example, both spatially and procedurally, is the response to a fire aboard a submarine or ship.

Although the ability to perform complex spatial and procedural tasks is often acquired through practice with feedback in the real world, there are often practical constraints on the use of operational equipment that limit its utility for training, even if the operational platforms are available. Concurrent maintenance work or potential safety hazards associated with repairs can interfere with access to onboard systems or their use. The advancement of the technologies that enable VEs further encourage their use as alternative training environments. Some recent examples include the use of virtual reality for training surgical procedures (Johnson et al., 2011), maintenance tasks (Gavish et al., 2011) and welding (Stone et al., 2011).

While the training effectiveness of high fidelity flight simulators is generally accepted, relatively little is known about the effectiveness of VEs for training complex tasks (Stone et al., 2011). The lack of behavioural evidence about the training effectiveness of instructional systems that depart physically from the real world is especially problematic since early psychological theory attributed to Thorndike (1906), naïve views (e.g., Adams, 1972; Smallman and St. John, 2005) and glossy sales brochures have mistakenly entrenched the notion that high physical fidelity is needed for positive training transfer.

Training Transfer

Training transfer is a process that is revealed when knowledge, skills and abilities acquired in a training environment subsequently affect performance of a task performed in another environment (Ellis, 1965; Roscoe and Williges, 1980). The objective for training is usually a positive benefit on subsequent performance of an operational task. Although objective evaluations of the benefits of a games-based training environment, or its specific features, can help inform a procurement decision, few studies have provided behavioural information obtained with forward transfer of training (FTOT) to an operational environment (Alexander et al., 2005), even though it is the preferred method for generating the needed information (Caro, 1977).

One obvious barrier to an assessment of FTOT are restrictions on the use of operational equipment; the very same reason why a simulator or VE might be needed in the first place. Other reasons why objective evaluations of a training environment are often not conducted for military tasks are because they are often costly, time-consuming, difficult to implement, or hazardous (Taylor et al., 2001, McCauley, 2006). Consequently, alternative, more feasible methods have often been used. The Advisory Group for Aerospace Research & Development (1980) examined many methods for determining whether or not a training simulator has an effect on training performance. This group of experts rejected measures of user opinion, physical or dynamic fidelity, and how much a simulator is used as reliable indices and summarized the pros and cons of several alternative methods for assessing training transfer. One appealing method is known as reverse (or backward) transfer of training (RTOT). This method relies on the use of the training environment, with no need to access the operational equipment. The method is relatively inexpensive, easy to implement and safe. Added benefits include control of the experimental conditions and the avoidance of the unwanted influences of environmental or operational factors associated with testing in the field.

The RTOT paradigm involves the comparison of the performances of at least two groups of participants, one already qualified to perform the task in the real world and another that is inexperienced with the task. This method tests the validity of the VE by predicting that the qualified group will initially be able to demonstrate a high level of performance on the task within the VE. When combined with a skill acquisition phase, it also predicts that the qualified participants should show little or rapid improvement with practice (since the members of this group already know how to perform the task) and that the novice group will perform more poorly at the start and improve more slowly with practice (since the members of this group are learning the task for the first time). This experimental method aims to reveal learning by the novice group. Consequently, an RTOT experiment can also help determine whether or not an FTOT experiment should be assessed with the operational equipment or platform since learning is a prerequisite for training transfer. If nothing is learned, there is no training to transfer.

In this study, training was assessed by both methods since it seemed useful to gather the behavioural evidence about the effectiveness of the VCVS that an RTOT analysis could afford and thereby reduce our reliance on access to the submarine and the negative impact that an FTOT analysis could have on operations. This approach also seemed prudent for the following reasons: (1) limited access to the submarine prevented the conduct of a pilot study for FTOT, (2) a limited number of novice participants was available to participate in an FTOT experiment, and (3) lessons learned about the relationship between the findings of RTOT and FTOT could benefit future investigations of the VCVS or other training investigations constrained by the availability of operational equipment or participants.

Experiment 1: Reverse Transfer of Training (RTOT)

Method

Task

A complex emergency procedure was chosen by subject matter experts (SMEs) as the task for this study. Submariners need to know their way about the submarine and what to do in an emergency, such as a fire or flood, since failure to perform the procedures properly could be fatal. One of the many emergency procedures that must be known by all qualified submariners aboard a Victoria Class submarine is isolation of Bulkhead 35 (BH 35), which is the main barrier between the forward and aft compartments of the submarine. Submariners must know the locations, names, functions and operation of the valves and tools that are needed to isolate this bulkhead.

BH 35 is isolated in different ways for different situations; it requires the shutting or checking of up to fifteen valves (e.g., for attack, counter attack, preparation for collision, response to flood or fire). For this study, the six valve shutdown used in the event of a fire was selected. As the name implies, the task involves six valves, which are located on two decks (two valves on 1 Deck and four valves on 2 Deck). The valves can be checked or operated from either side of the bulkhead and can be checked or operated in any order for successful completion of the task. However, the task on the aft side of the bulkhead is not a simple mirror image of the task on the forward side of the bulkhead. The surroundings are different, and the location of the valve controls, their direction of operation, their appearance, and the location and type of tools that might be needed to check or shut the valves differ unsystematically. For instance, to shut Ventilation Valve 803 (VV803) it must be turned clockwise from the forward side of BH 35 and counter clockwise from the aft side, or to shut Ventilation Valve 801 (VV801), a ratchet must be used on the forward side, but not on the aft side, of the bulkhead.

Trainees are normally taught the task aboard the submarine (as a small group or individually) and are individually assessed two days later. To demonstrate proficiency, the trainee must be able to perform the task without error.

Participants

Twenty healthy male volunteers medically fit for duty and free of the signs and symptoms of acute illness volunteered as participants. The sample populations included ten Ordinary Seamen (OS) of the RCN awaiting training who were unfamiliar with the layout of a submarine and who were recruited by the Personnel Coordination Centre (PCC) Maritime Forces Pacific (MARPAC), and ten task-qualified submariners serving aboard HMCS Corner Brook who were recruited on site by the experimenters. The participants in the experimental group ranged in age between 18 and 35 years (mean: 23.4

± 6.5 years). The participants in the qualified group ranged in age between 26 and 50 years (mean: 39.6 ± 8.5 years) and possessed an average of 13.2 years of experience within the Navy (with a mean of 9.8 years of experience in submarine service).

All participants were informed fully of the details, discomforts, risks and potential benefits associated with the experimental protocol. They provided informed consent, and all were compensated for their participation, in accord with the study protocol (Magee et al., 2011) approved by the Human Research Ethics Committee (HREC) of Defence Research and Development Canada (DRDC).

Apparatus

Virtual Environment (VE) – The VCVS was used as the VE. The developers used the Autodesk Media and Entertainment (Montreal, Quebec) 3D Studio Max + Unity game engine (version 3.4.2) and $C^{\#}$ for modeling and programming. The VE was hosted by a laptop computer running Windows XP. The graphics were generated by a Direct 9.0 compatible video card that provided realistic images of the inside and outside of HMCS Corner Brook. The images were displayed on a 15-inch liquid crystal display (LCD). The computer-generated models of the inside and outside of the submarine, including the surrounding environment and dock, consisted of more than 3000 objects, with more than 500 textures, tuned for real-time performance (i.e., scene updates at 60 Hz). Sample imagery of the VE is shown in *Figure 1*.

The instructional designers provided options for navigating within the VE, consistent with industry standard controls for first person shooter (FPS) videogames. One option used the letters w, s, a and d of the laptop's QWERTY keyboard to move forward, back, left or right, and the other made use of the directional arrow keys on the keyboard for these movements. Either option could be used. The shift key, the control (Ctrl) key, and the space bar were used to run, crouch or jump, while actions to cross a bulkhead, climb a ladder, or use a flashlight were controlled with keys e, r, and f, respectively. The user's point of view (POV) within the VE was controlled with a computer mouse. Forward movement of the mouse tilted the view downward, backward movements tilted the view upward, and movements to the right or left moved the view to the right or left. The gains on these controls were adjusted for easy use.

The instructional designers also provided a plan, or map, view (bird's eye) of the submarine that could be toggled on or off by pressing the m key. The map appeared in place of the observer's immersive view and showed the current location and orientation of the observer within the VE. In the map mode, the submarine could be enlarged or reduced in size with the mouse, by the click-wheel, or by the + and – keys on the keyboard; it could also be rotated about a vertical axis that was centered on the eye point of the observer within the VE.

Figure 1: A view of BH 35 from the aft passageway of 2 Deck of HMCS Corner Brook as displayed on the LCD of the laptop computer.

In addition, the visual angle of the immersive view and its orientation within the VE was represented by a triangular icon. Use of the movement keys in this mode permitted the user to reorient within the VE. These visual features were provided as aids for navigation and the acquisition of spatial knowledge; they were added to redress difficulties experienced by SMEs in early implementations of the VE and they were implemented to match the human-computer interactions employed by keyboard-based computer games.

The VCVS software provided pop-up instructions and labels. The upper left corner of the screen indicated where the player was located on the submarine at all times (e.g., FWD 35 > 2 Deck > Heads; or AFT 56 > 1 Deck > Upper Motor Room) and this information was refreshed with every location change. Task-specific, instructional flags would appear in the lower left of the screen as needed. For example, when approaching a ladder the instruction "Use ladder 'R'" would appear, or when approaching a bulkhead door "Go through Bulkhead 'E'" would appear. When the learning event (LE) Isolate BH 35 (FWD of AFT) was being run a list of the valves that need to be shut would appear on the right side of the screen. These valves would then turn green when the POV (indicated by crosshairs at the centre of the screen) was directed at an appropriate valve. Once the valve was clicked (i.e., to shut the valve) then it would appear red, its name would be removed from the list, and a message indicating the successful shutting of that valve would appear beneath the list (e.g., "VV608 FWD Battery Ventilation Isolating Valve SHUT"). When ratchets are required to operate the valve, as needed for VV608 forward and aft, VV801 forward, and LPB803 aft, the message "Locate Ratchet" would appear at the right side of the screen. If the valve was selected before the ratchet was selected, the message “Wrong Action” would appear on the right side of the screen. Similar messages (followed by “Wrong Action” messages when necessary) were displayed for opening and shutting cabinet doors and for replacing the ratchets after shutting the valves.

Design and Procedure

A between-groups experimental design was employed. The members of the experimental group received familiarization training with the interface to the VE and the task, and then practiced the task to criterion. The members of the qualified group received familiarization training with the interface only and then performed the task to criterion within the VE.

The members of the experimental group used the VE within an empty classroom. They were informed that their task was to learn the locations, names, functions and operation of the valves needed to isolate BH 35 from both forward and aft approaches. While comfortably seated in front of a laptop, the participants were guided individually through the VE. An experimenter provided verbal guidance to each participant to help him find their way within the VE to each valve of the submarine that needed to be operated in order to perform the task. During this familiarization session, the participants also learned how to control their movements with the mouse and keyboard and were told the name, method of operation, and function of each valve that needed to be controlled. After familiarization, the members of the experimental group were given a fifteen minute rest. They were told that they would subsequently need to perform the task perfectly to reach criterion, but that they could ask for help at any time and that they would be provided corrective feedback if they made a mistake. They were told that they would continue to practice the task until they could do it from beginning to end without error or need for help. The practice tasks were delivered in blocks of six trials with a 15-minute rest period in between. A maximum of three blocks (i.e., 18 trials) was allowed.

The members of the qualified group were individually familiarized with the interface and tested alone aboard the submarine. They were seated in the submarine’s wardroom, or one of the mess rooms, with the laptop and mouse in front of them on a table. Each submariner in the qualified group was given the printed menu of controls, familiarised with the VE, and asked to perform the task to criterion within the VE. The delivery of the trials was the same as for the experimental group.

All participants were told that it was important to take the shortest route between valves so that the total path length would be a minimal distance. This instruction was provided to train wayfinding by encouraging the development of survey knowledge, rather than landmark or route knowledge of the spatial environment (see Lapeyre et al., 2011, for definitions of these three types of spatial knowledge and their effects on wayfinding performance). The participants were also told that the total amount of time needed to complete the task was not important; this instruction was provided to avoid subsequent dangers associated with haste aboard the submarine, such as falls while climbing a ladder.

Although this instruction was provided primarily for the safety of the members of the experimental group who would later be asked to perform the task aboard the submarine in Experiment 2, it may seem to be inappropriate for training a response to a real emergency. In fact, it is not normal for a serving seaman to complete the task alone aboard the submarine. Although all qualified submariners aboard a Victoria Class submarine must know how to complete the task on their own, standard operating procedures dedicate a watch keeper to manage a small team of submariners who are individually responsible for one or more valves near their normal work place.

All participants began each trial at one of the six starting positions chosen randomly. The locations were typical work or rest locations not near a valve. The set of start locations (three FWD and three AFT of BH 35) included the following: control room (AFT, 1 Deck), passageway (AFT, 2 Deck), junior rates mess (AFT, 2 Deck), heads (i.e., 1 and 2 trap; FWD, 2 Deck), senior accommodations/bunks (FWD, 2 Deck), and the weapons stowage compartment (WSC; FWD, 1 Deck).

Two experimenters independently recorded the errors made by a participant on each trial. The errors were classified as either a procedural or a spatial mistake. Procedural errors included failure to properly identify or explain the operation or function of a particular valve. Spatial errors included wrong turns and failure to take the shortest route to the next valve. To achieve the performance criterion the participants needed to complete the task perfectly once, without either a procedural or spatial error. Both experimenters needed to concur that the criterion had been achieved. Afterward, they resolved any discrepancy between their recordings for previous trials. Very few discrepancies needed to be resolved with the recorded errors having very high (and statistically significant) correlations between the two experimenters for both the experimental and qualified groups ($r = 0.99$ and $r = 0.96$, respectively). One of the objectives for having the two researchers monitor the performance of the participants at the same time was to achieve marking consistency since often only one experimenter was able to monitor a participant in Experiment 2, where the same task performance criterion was applied.

All participants were asked about their experience with computer games; they were asked how many years of gaming experience they had, how many hours per week they currently spend gaming, the types of games they typically play, and what platform they game on (i.e., personal computer (PC) versus console).

Results

Trials to Criterion

Figure 2 provides a bar graph that cumulates the number of participants in each group that achieved task criterion with practice. The plot for the qualified group shows that 50% of its members achieved criterion by the second attempt and that all ten members of this group achieved criterion by the sixth attempt. The median number needed to reach criterion for the qualified group was 2.5 trials. In comparison, only 50% of the participants in the experimental group achieved criterion on the sixth trial and only nine of the ten participants in the experimental group achieved criterion by Trial 18. The median number needed by the experimental group to reach criterion was 8.0 trials. A Mann-Whitney *U* test ($U = 90.0$, $p < .001$) indicates that the difference in the median number of trials required to reach criterion by the two groups is statistically reliable. This outcome indicates that the qualified submariners were able to demonstrate their knowledge of the task within the VE.

In the planning of this experiment, without benefit of a prior pilot study, we estimated that 18 trials would be sufficient for all members of the experimental group to achieve criterion. Since one member of the experimental group was unable to achieve criterion within 18 trials (due to procedural errors), we estimated the number of trials that should be required to achieve this goal by linear regression ($y = 0.5588x + 0.2353$) and found that all ten members of a randomly selected group of ten PATs should be able to reach criterion by Trial 18. This estimate is more conservative than one obtained by exponential regression ($y = 1.2622e^{0.133x}$), which estimates that a random group of ten PATs should complete the task to criterion in fewer than 16 trials. Hence, 18 trials should be a sufficient amount of practice for PATs to learn the six-valve shut down of BH 35.[1] Linear and exponential regressions performed on the data provided by the qualified group both suggest that four trials should be sufficient for qualified submariners to achieve criterion with the VCVS. This result suggests that a relatively small amount of practice within the VE would be sufficient to refresh the training of qualified submariners.

[1] It is tempting to reason that individuals unable to learn the task within 18 trials could be unsuitable for submarine service. However, the participant who was unable to complete the task to criterion within the VE was later able to perform the task successfully aboard the submarine within four attempts, comparing favourably with others within his group, as described in Experiment 2.

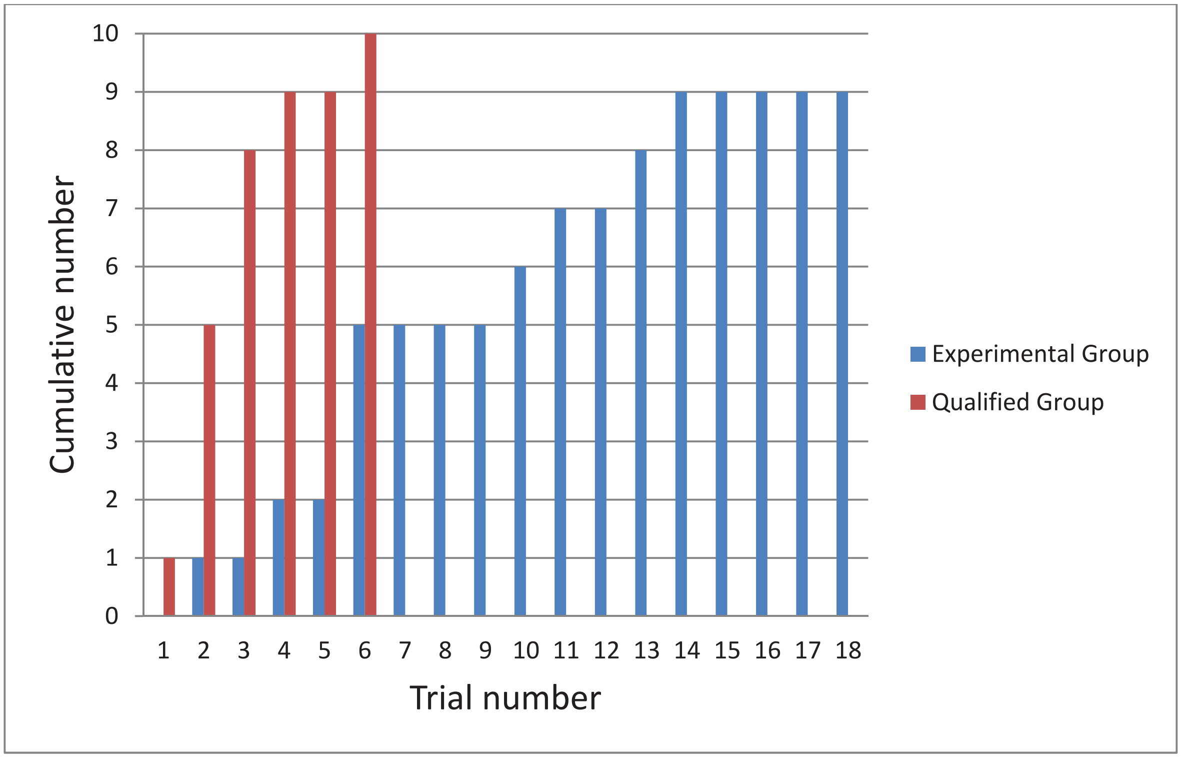

Figure 2: Trials to criterion within the VE

Procedural errors

The median number of procedural errors made by the experimental group on their first attempt at the task was 5.0, whereas the median number for the qualified group was 2.0 errors. Significantly more procedural errors were made by the experimental group on their first attempt at the task within the VE (Mann-Whitney $U = 81.5$, $p < .01$, one-tailed). Similarly, the median number of procedural errors made by the experimental group prior to achieving task criterion was 16.0 which is much greater than the median 3.0 for the qualified group; this difference in performance is highly significant (Mann-Whitney $U = 92.0$, $p < .0005$, one-tailed).

It is instructive to know that the pattern of results shown in *Figure 2* is determined almost exclusively by procedural errors; there were very few instances in which one or more spatial errors alone caused a failure to achieve criterion on a particular trial. It is also instructive to know that the qualified group did not demonstrate a high degree of proficiency on the first attempt at the task. Only one qualified submariner performed the task correctly the first time within the VE. On its own, this result would indicate a very low amount of reverse transfer of training and it would imply problems with the simulation. However, it is clear from the practice results that reverse transfer of training is occurring since there is reliable evidence that prior experience with the submarine allowed the qualified group to achieve criterion with many fewer attempts than the experimental group.

Possible explanations for this apparent contradiction of outcomes (i.e., low initial performance but high rate of adaptation by the qualified group) include the following:

1. The experimenters were very particular about the specific words and explanations that they sought as evidence for successful performance. Greater latitude is acceptable aboard the submarine. Cain et al. (2011) have noted similar practices for maritime helicopter deck landing and observed similar results when experienced Landing Safety Officers are tested within a VE.
2. There are small differences between the VE and the real submarine that can be distracting and that would be noticed only by the experienced submariners. For instance, the location of a tool needed to operate a valve is different in the VE than aboard the boat.
3. The computer interface to the simulation and the methods of interacting with the simulation interfere with performance. This explanation is substantiated by the observation that some procedural errors, such as forgetting to return a tool to its proper holder, contributed to failures. This kind of mistake is easy to commit in the simulated environment because the haptic cues (e.g., the physical cues associated with the grasp of a wrench) that might guide behaviour in the real world are missing in the simulated one.
4. Memory of the drill has decayed. This explanation is substantiated by voluntary admissions made by 6 (60%) of the qualified submariners; they indicated that competing duties aboard the submarine did not always permit them to maintain the level of task proficiency that duty requires.

Spatial Errors

As shown in *Figure 3*, 8 submariners in the qualified group (i.e., 80%) made no spatial error on their first attempt at isolating the bulkhead within the VE. In comparison, no member of the experimental group was able to perform the task without spatial error on his first attempt. The median number of spatial errors for the experimental group on their first attempt at the task was 1.5 errors. This difference in performance between the groups is highly significant (Mann-Whitney $U = 95.0$, $p < .001$, one-tailed). The comparison indicates positive reverse transfer of spatial knowledge, from the submarine to the simulation and application of the knowledge for spatial awareness within the VE. It also suggests that the components of the computer interface for controlling movement within the VE (e.g., keyboard controls) do not contribute to the difficultly that the experienced submariners initially have with the VCVS.

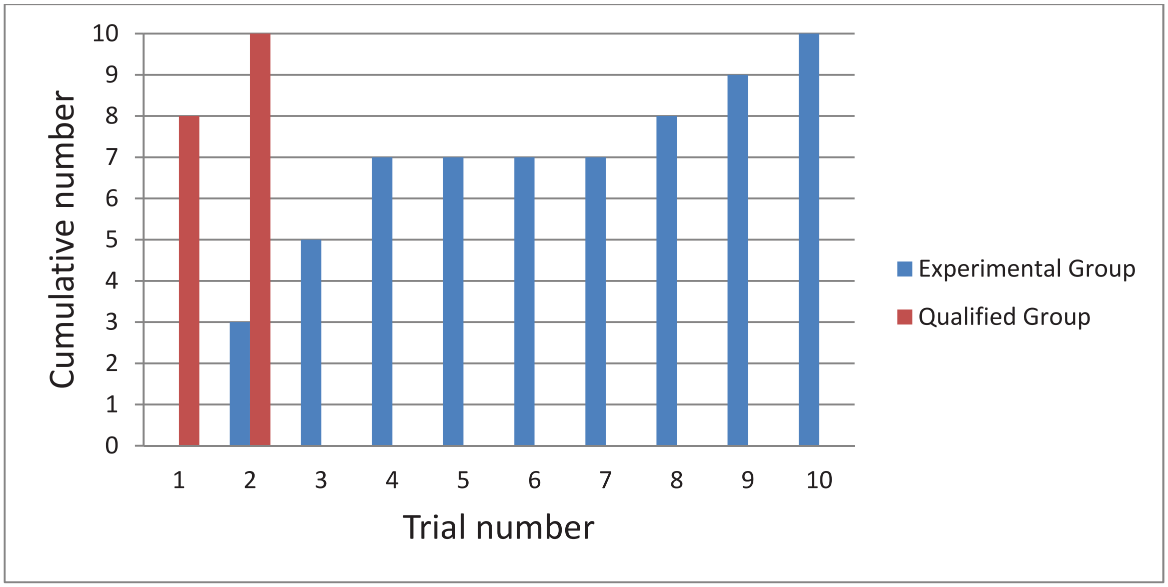

Figure 3: Trials to error-free navigation within the VE

Figure 3 provides a bar graph that cumulates the number of participants in each group that was able to locate valves without error as trials were provided. The bar graph shows that all (100%) members of the qualified group were able to find their way to the valves without error on their second attempt. In contrast, only three (30%) members of the experimental group could find their way without error on their second attempt and up to ten trials were needed by other members of this group to achieve criterion. The median number of spatial errors committed by the experimental group prior to achieving task criterion was 5.5 errors. A Mann-Whitney U test contrasting the overall spatial performance of the experimental group with the qualified group yields $U = 97.0$, $p < .0001$. These outcomes indicate convincingly that the qualified submariners were able to apply their spatial knowledge of the submarine within the VE and that they had very little difficulty establishing spatial awareness within the VE.

Gaming experience

Members of the experimental group reported an average of 8.9 years of computer game experience and an average of 5.6 hours of game play each week. Members of the qualified group did not differ significantly from the experimental group in years of experience, $t(18) = -0.95$, $p > .3$, or hours per week gaming, $t(18) = 0.68$, $p > .5$, with an average of 12.9 years experience and an average of 3.5 hours each week.

These measures of gaming experience were correlated with the task performance measures. For each group, years of experience and hours per week were correlated with number of trials to criterion, total errors committed, total procedural errors committed, and total spatial errors committed. While some correlations were statistically significant, none were meaningful or insightful. It seems that the level of performance in learning the isolation of BH 35 in the VE is not associated with gaming experience – that is, having more gaming experience does not equate to better performance in the VCVS. Practically speaking, this means that an individual with no gaming experience could equally benefit from training with the VE.

However, while some studies classify an expert gamer as an individual with four hours per week dedicated to gaming (e.g., Green and Bavelier, 2003; Granek et al., 2010), others classify expert gamers as those with seven or more hours per week dedicated to gaming (e.g., Boot et al., 2008). Consequently, the participants in our study might not have sufficient gaming experience to affect performance in the VE; that is, individuals with seven hours or more of gaming time every week may perform better, but our small sample was inadequate to find a reliable correlation. The few participants in this study with seven or more hours gaming time per week do not perform statistically better than the other participants – indeed, the participant with the greatest number of gaming hours per week reported (i.e., 14 – 21 hours) was the only one unable to achieve criterion within 18 trials. Thus, we conclude that gaming experience is not needed and does not benefit the use of the VCVS as a training medium.

Discussion

The empirical results associated with initial reverse transfer and learning with practice provide convincing behavioural evidence of the positive training benefits of the VE. The superior performance of the qualified group in achieving criterion within the VE indicates that experienced submariners are able to make use of the sensory cues provided by the simulation and that they can adapt readily to its human-computer interface. On this basis, there seems to be no apparent need to improve the fidelity of the visual cues (e.g., larger field-of-view) or need to add other sensory cues to the VE (e.g., sounds). There is also no apparent need to consider a more immersive interface, such as a walking device, to control navigation within the VE. However, there does seem to be a need to improve the simulation of interactions with tools since the participants sometimes failed to return a tool to its holder; this failure would likely not occur often in the real world where the tool would provide both visual and haptic cues as reminders.

As shown in *Figure 2*, individuals differ widely in their ability to acquire spatial knowledge from the VE; this outcome is found for virtual and real environments (Wolbers and Hegarty, 2010). The results indicate that up to ten practice trials with the VE could be needed to learn the locations of the valves and to gain the spatial knowledge needed to perform the task alone aboard the submarine. We estimate that six to eight additional practice trials with the VE could be needed to learn the procedural aspects of the task as well. In sum, we estimate that about 90 minutes of familiarization training and practice will be sufficient to train novices to perform this complex task within the VE.

On the basis of the clear differences in the performance of the qualified and novice participants, we predict that the knowledge gained by the novices will transfer positively to the submarine. In the next experiment, we test this prediction, knowing that experienced operators may be able to perform better than less experienced operators in a deprived environment that does not provide all of the cues that a novice may seek while learning.

Experiment 2: Forward Transfer of Training (FTOT)

Method

Task

This experiment employed the same task as Experiment 1, but it was performed aboard the submarine. The participants were required to find, identify, and explain the function of each valve required to isolate BH 35. They were not required to physically operate the valves or use any tools within the submarine. Physical operation of the valves was avoided so that they remained as set by the watch keepers on duty aboard the boat. This practice was followed as a safety precaution. However, the participants were required to indicate verbally the operation that they would perform on each valve as they encountered it.

Participants

All ten members of the experimental group of Experiment 1 participated in this experiment and ten, new recruits medically fit for duty and free of the signs and symptoms of acute illness formed a control group. The control group consisted of eight OS and two LS of the RCN awaiting training who were unfamiliar with the layout of a submarine. The members of the control group were recruited by the PCC of MARPAC. The participants in the control group (all male) ranged in age between 20 and 49 years (mean: 28.5 ± 9.6 years). All participants were informed fully of the details, discomforts, risks and potential benefits associated with the experimental protocol. They provided informed consent, and all were compensated for their participation according to rate schedules of DRDC Toronto that were approved by the HREC.

Apparatus

HMCS Corner Brook was used as the test environment for assessing training transfer. The submarine was alongside C Jetty at CFB Esquimalt, British Columbia. The boat is about 70 metres long. It has two principal decks, with a warren of narrow passageways that connect the compartments on each deck. Three ladders connect the decks (i.e., one forward and one aft of BH 35 and one aft of BH 56). The passageways and the ladders were unobstructed and lit normally (i.e., dimly). The tally plates for the valves were unaltered for the experiment (i.e., they were treated "as is").

Design and Procedure

A between-groups experimental design was employed. The groups differed in the type of initial training that they received prior to performance of the task aboard the submarine. As described in Experiment 1, the members of the experimental group received familiarization training and practice within the VE until they could perform the task to criterion. The members of the control group received familiarization training aboard the submarine. The members of both groups then performed the task to criterion aboard the submarine two days after initial training.

All participants were met by an experimenter on the jetty and escorted to the submarine where they received a safety briefing by a member of the duty watch. The safety briefing included a demonstration of the use of the boat's Emergency Breathing System (EBS), an EBS mask, an Emergency Escape Breathing Device (EEBD) and instructions for evacuating the submarine in the case of a real emergency.

The experimental trials began at one of six randomly chosen starting positions as described in Experiment 1. At the start of each trial, the participant was guided to the start location and was consequently afforded an opportunity for incidental learning that was not provided with the VE (there were also opportunities for incidental learning while signing in with the duty watch and during the safety briefing). As in Experiment 1, all participants were told that the amount of time needed to complete the task was not important; this instruction was provided for safety purposes.

During the conduct of the experiment, a researcher followed each participant as he completed the task. The researcher provided help when requested, provided corrective feedback when necessary, and recorded the procedural and path errors made by the participants. As in Experiment 1, the trials were conducted in blocks of six with a 15 minute rest period in between. A maximum of three blocks (i.e., 18 trials) was allowed.

Results

Trials to Criterion

Figure 4 provides a bar graph that cumulates the number of participants in each group that achieved criterion with practice. The plot for the experimental group shows that two of its members (20%) achieved criterion on their first attempt. In comparison, none of the participants in the control group achieved criterion on their first attempt. However, this difference in counts is not statistically significant (χ^2 (1) = 2.2, $p > .05$).

The plot further indicates the apparent advantage that the experimental group might have for the first three trials is lost with exposure to the submarine. The median number needed for all members of the experimental group to reach criterion is 3.0 trials. In comparison, the median number needed by the control group to reach criterion is 4.0 trials; only one more than in the experimental group. A Mann-Whitney *U* test performed on these data (U = 65.5.0, $p < .125$, one-tailed) indicates that this difference in performance between the two groups, although consistent with prediction, is not statistically reliable if the normal .05 level of probability is used as the criterion for rejecting the null hypothesis that the experimental group is not better prepared than the control group.

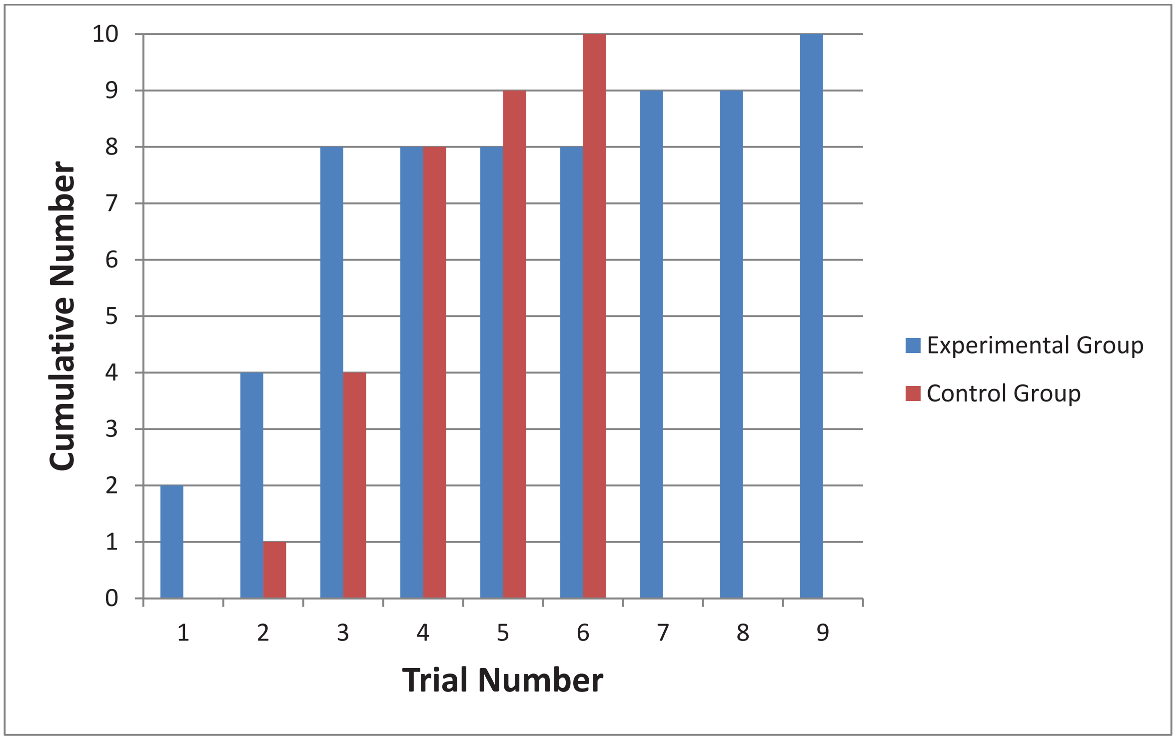

Figure 4: Trials to criterion aboard the submarine

Transfer Effectiveness Ratio (TER)

There are many ways to measure training transfer (Blaiwes et al., 1973; Gagne et al., 1948; Murdock, 1957). A useful measure of training transfer from a simulator to an operational setting is the TER, which expresses the operational savings as a proportion of simulator exposure (Povenmire and Roscoe, 1971). The operational savings can be measured by time, trials, or errors to a criterion (Blaiwes et al., 1973).

The TER based on the trials to criterion in this experiment and the previous one can be computed as shown below.

$$TER = \frac{Tc - Te}{Ts},$$

where

Tc is the median number of trials taken to reach criterion on the submarine by the Control Group
Te is the median number of trials taken to reach criterion on the submarine by the Experimental Group
Ts is the median number of trials taken to reach criterion on the VCVS by the Experimental Group

Therefore,

$$TER = \frac{(4 - 3)}{8} = 0.125.$$

This value is relatively low in comparison to TERs that have been computed for early flight simulators (Orlansky and String, 1980) and military training devices (Fletcher & Orlansky, 1989). Nevertheless, this TER could represent a significant financial savings if the costs of training with the VCVS and a Victoria Class Long Range Patrol Submarine (SSK) are compared. Note too, that this calculation of transfer effectiveness does not include task familiarization time as part of the calculation. The experimental group received no familiarization time aboard the submarine. Thus, several hours of submarine time (for a group) could be saved beyond the amounts suggested here. Furthermore, each trial takes about six or seven minutes to complete aboard the submarine and less than half that time with the simulator (i.e., two to three minutes). Thus, a TER calculated in terms of time savings would be larger. Note that we did not use time, which is a more traditional unit of measurement for calculating the TER, because we wanted to avoid haste aboard the submarine (for safety) as well as the complications that are associated with time as a measure, such as the speed (and hence time) of movement within the VE, which is determined jointly by the participant's control inputs and the software that links control inputs to the rate of travel within the VE (this rate was a nominal amount).

Procedural errors

The median number of procedural errors made by the experimental group on the first attempt at task completion was 3.0. In comparison, the median number of procedural errors made by the control group on their first attempt was 4.0 errors. This difference is not statistically reliable ($U = 59.0$, $p > .25$, one-tailed). The median number of procedural errors committed by the experimental group prior to achieving task criterion is 6.0 errors, whereas the median number of procedural errors committed by the control group prior to achieving criterion is 7.0 errors. This observed difference is also not reliable ($U = 58.5$, $p > .25$. one-tailed).

As in Experiment 1, the pattern of procedural errors largely determined the overall pattern of results shown in *Figure 4*, meaning that there were again very few instances where spatial errors alone caused a failure to complete the task to criterion. In fact, no participants in the experimental group committed a spatial error alone, and only two participants from the control group failed to achieve criterion due to spatial errors alone. Although these statistical analyses fail to reveal a reliable benefit for training procedures with the VE, it is useful to note that the participants in the experimental group did not perform any worse than the control group. In other words, the prior training that the experimental group received with the VCVS seems to be as good as the familiarization training that the control group received aboard the submarine.

Spatial errors

Figure 5 provides a bar graph showing the cumulative number of participants in each group that succeeded in finding their way to all valves without spatial error.

Five members (50%) of the experimental group made no spatial error on their first attempt to perform the task aboard the submarine, whereas only one member of the control group (10%) managed to perform as well. The median number of spatial errors made by the experimental group on their first attempt at the task is 0.5 errors, whereas the corresponding median of the control group is 2.0 errors. This difference between the performances of the two groups is statistically significant ($U = 77.5$, $p < .02$, one-tailed). The TER based on these values, and the spatial data obtained in Experiment 1, is 0.27, which indicates that about four trials with the VCVS save one trial with the submarine.

The median number of spatial errors committed by the experimental group prior to achieving task criterion is also 0.5 errors, whereas the median number of spatial errors committed by the control group prior to criterion is 3.0 errors. This difference in performance is marginally reliable ($U = 69.5$, $p = .07$, one-tailed) and suggests that the advantage of prior training with the VE, as demonstrated by the comparison of performances by the two groups on their first attempt at the task is soon lost with additional exposure to the submarine. Nevertheless, the difference has practical importance since the total number of trials needed by the experimental group (20 trials) was less than the total for the control group (26 trials). This represents a 23% reduction in the number of trials needed to train spatial awareness aboard the submarine, which has limited access for training and is very costly as a training environment.

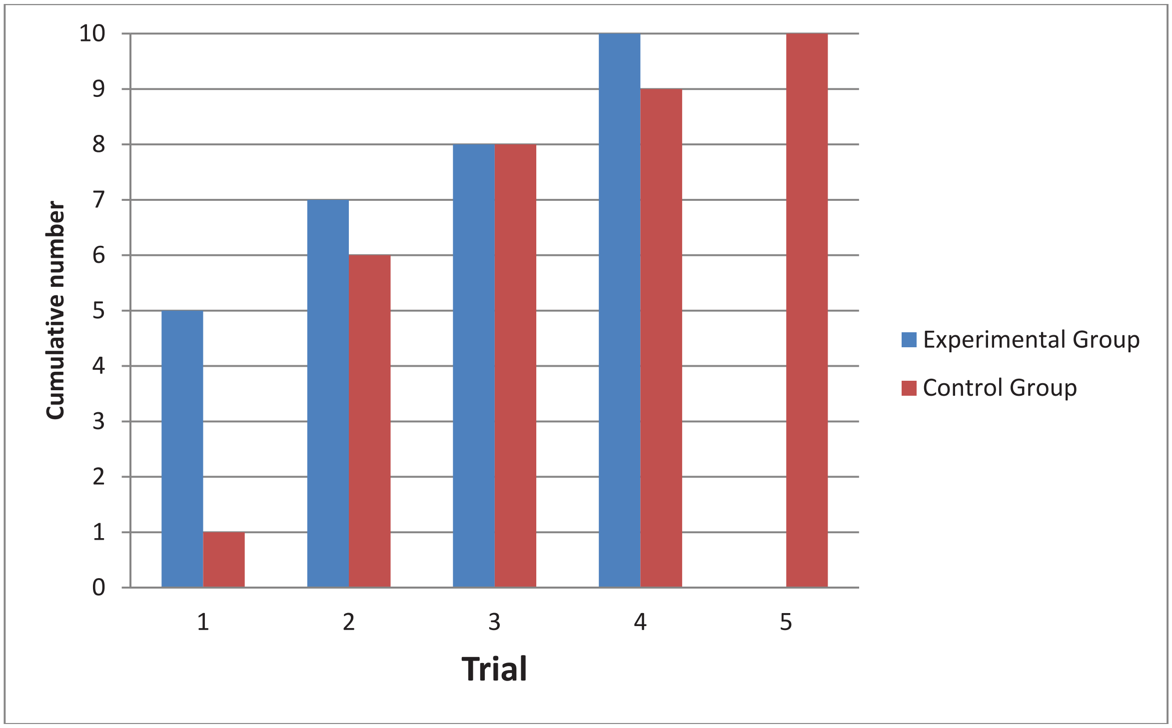

Figure 5: Trials to error free navigation within the submarine

Gaming Experience

While this experiment did not make use of the VCVS for training the control group, it has been demonstrated in functional magnetic resonance imaging (fMRI) studies that experienced gamers have additional pre-frontal cortex activation when planning complex eye-hand coordination tasks that are distinct from actual game play (Granek et al., 2010). These findings suggest that a more extensive (and perhaps more efficient) visual-motor control network exists in gamers that is commonly believed to be related to increased online control and spatial attention. Extensive videogame play may indeed alter the basic cortical network involved in visually-guided (and perhaps memory-guided) action. Therefore, gaming experience is also considered in this experiment as it might relate to learning or performance of the task on the boat.

Participants in the control group had an average of 10.1 years of gaming experience and spent an average of 3.2 hours per week gaming, which does not differ significantly from the gaming experience of the experimental group (i.e., 8.9 years, and 5.6 hours/week; $t(18) = -0.35$, $p > .7$ and $t(18) = 1.0$, $p > .3$, respectively). The same comparisons were made here as in Experiment 1, and again we found no meaningful relationship between gaming experience and task performance.

Observations made by the experimenters

The experimenters noticed a number of differences between the experimental and control groups during the conduct of the experiment, in both the initial training and testing sessions. After reading the instructions to participants on the jetty, the researchers sensed that the participants in the control group seemed generally overwhelmed by the task. After familiarization training many of the participants in the control group expressed concern about their ability to remember the procedure and asked if a refresher course or opportunity to practice before the testing session would be provided. Most of the participants in the control group were still uneasy about their ability to recall the layout of the boat and the names and functions of the valves when they returned for testing two days later.

The spatial awareness of the control group appeared weaker than that of the experimental group. The visible and often verbalised lack of confidence expressed by the participants in the control group contrasted with the members of the experimental group who showed more confidence when they encountered the submarine for the first time. In further contrast, the members of the experimental group neither exhibited nor stated any apprehension or uncertainty after reading the instructions in the classroom. The implication of these observations is that the VE seems to be a less intimidating learning environment than the submarine.

Furthermore, after training on the VCVS, most members of the experimental group stated that they were confident in their ability to isolate BH 35 aboard the submarine. When the experimental participants reported to the jetty they were excited to get aboard the boat to see how it compared to the VE. The experimental participants also seemed to be much more comfortable aboard the boat than the participants in the control group; they immediately seemed to have a greater understanding of where they were on the boat and its layout.

A number of participants from the experimental group commented that the VE had prepared them well for the task, and they indicated that the VCVS gave them a good idea of where the valves were and what they looked liked. Although the VE appears to be much cleaner and less cluttered than the actual submarine, a difference was found to be distracting; within the VE the ratchet for Low Pressure Blower 803 (LPB 803) is located inside a cabinet, but aboard the submarine, it is stowed on the cabinet door (as shown in *Figure 6*). One member of the experimental group struggled to locate the ratchet as a result. Another notable difference between the two environments is that the door of the sonar cabinet space is always open in the VE, but shut on the boat. Consequently, the area surrounding the forward side of the BH 35 door appears quite different on the boat than in the VE, and caused a few participants in the experimental group difficulty in locating VV803. At least one participant committed a spatial error by walking forward down the passageway past this door; he might have completed the task to criterion in fewer attempts if this difference did not exist. These observations reinforce the conclusion that the experimental participants acquired a good deal of spatial knowledge from the VE, but they can also help explain why the experimental group did not outperform the control group on the task.

Figure 6: Subtle differences between the VE and submarine, such as the location of the ratchet for LPB 803, sometimes confuse participants

There were also a few noteworthy difficulties experienced in the VE. One is that the user is unable to navigate between the table and the wall in the Senior Rates/Chief's and Petty Officer's (PO's) Mess, making it impossible to read the tally plate for VV803 from the wall. In comparison, the tally plate is much more accessible on the boat (*Figure 7*). Furthermore, this valve (VV803) must be turned in opposite directions when being operated from the forward and aft sides of the bulkhead. Given the importance of turning this valve in the correct direction, it would be useful to add an instructional feature to the VE that would require trainees to indicate the direction that the valve must be turned.

Figure 7: VV803 is accessible from the bench of the Senior Rates Mess in the submarine. In the VE it is only possible to get as close as shown on the right. The valve can still be shut by clicking the mouse at this distance, but the associated tally plate is not readable

Most of the qualified participants and a number of submariners who were not participants in the experiment were curious about the VE. They said that they were impressed by the quality of the visual representation of the boat. Many of these individuals admitted that they were sceptical about the VE as a training tool either because they had seen an earlier version that they were not happy with or because they did not like simulation training in general, based on past experiences. The opinion that the VCVS could be a useful tool with capability for training a large number of tasks was unanimous, especially considering the lack of time to train on the boat itself. A number of qualified submariners and trainees also commented that the VCVS would be useful beyond basic submariner qualification (BSQ) training.[2]

While all signage aboard the HMCS Corner Brook is in English only, and while all operations aboard the submarine are conducted in English, it might be useful to add bilingual functionality to the VCVS. We make this recommendation because francophone participants, who lacked facility with English, seemed to struggle more when learning the task within the VE, although this did not seem to affect training transfer to the submarine. One francophone participant required many attempts to learn the task in the VE, but successfully isolated BH 35 on the boat in a single attempt. He explained that he had started to learn English when he enlisted, just months prior to his participation in this experiment, and that he had not yet received any formal English language training. He had never before seen some of the words that labelled the valves (e.g., ventilation) and, consequently, he had to memorise the labels without understanding their meaning. Since the labels describe valve function, understanding their meaning should make them easier

[2] Personnel who made suggestions were asked to provide them to NeLCoE through their chain of command.

to recall and explain. This implies that a French translation would likely help francophone learn with the VCVS, even though they need to perform their submarine duties later in English.

Discussion and Conclusion

The behavioural results obtained in the two experiments that form this study provide converging evidence that the VCVS affords positive training benefits for isolating BH 35, a task that was chosen by SMEs to represent the spatial and procedural drills that qualified submariners need to know in order to respond to an emergency aboard a Victoria Class submarine. The principal benefits of the VCVS appear to be its ability to convey spatial information. Notably, five members (50%) of the group trained with the VCVS made no spatial errors on their first attempt to perform the task aboard the submarine, whereas only one member (10%) of the group familiarized with the task aboard the submarine managed to perform as well. In both experiments, spatial performances differed reliably between groups consistent with positive training transfer. The transfer of procedural knowledge was not as evident.

The anecdotal observations indicate that trainees who are provided prior training with the VE will be more confident and less stressed when they board the submarine for the first time than those who are not given VE training. These observations are consistent with the observations of Tate et al. (1997) who found that Navy firefighters expressed increased confidence in their spatial knowledge of the decommissioned United States Ship (USS) Shadwell after they received familiarization training within a VE. They are also consistent with the observations of Boulet (in preparation), who asked naval reservists to rate the potential usefulness of a three-dimensional model of a Kingston Class maritime coastal defence vessel (MCDV), Boulet and Gilbert (2010), who asked staff and students of a BSQ course to rate the alpha version of the VCVS, and Garrett et al. (2008), who report that trainees exposed to a VE representing the main generator room of a Royal Australian Navy (RAN) Collins Class submarine show greater confidence when performing a simplified point safety round aboard the submarine than trainees given paper-based familiarization training.

It seems reasonable to conclude that the VCVS could be used to train other drills that involve procedural and spatial knowledge of the submarine and that practice with this task within the VE will reduce the amount of time needed to learn similar tasks within the VE or aboard the submarine if the tasks share knowledge components. Spatial knowledge about specific areas of the submarine, including the location of particular valves and procedural knowledge about valve function and operation are likely to transfer positively from one task to another and from one training environment to another.

An important feature of the CVNF developed by NeLCoE is that a large number of tasks and several vessels are simulated. Only one task and only parts of the submarine were used to assess training effectiveness in this study. While we think that isolation of BH 35 was a challenging task, including both spatial and procedural components common to other naval tasks, and while we think that HMCS Corner Brook provided a challenging environment for behavioural testing of the efficacy of the VCVS, the results of this study cannot be extended blindly to other tasks or other spaces. We expect that the amount of training transfer is likely to differ for each type of task, since the training effectiveness of a flight simulator has been found to vary for different types of flight manoeuvres (Cross, 1992), and we expect that the exploration of other spaces will alter the relationship between spatial learning in the VE and wayfinding in the real environment that we found in this study. We also expect that other spaces, especially larger ones, will affect the use of instruction aids for spatial awareness, such as the plan (map) view that was included in the VCVS.

We are also uncertain about the extent to which training in the VE could transfer to the stressful circumstances of a real emergency if training is conducted as described in this study. We advise caution since we have concerns about two factors that could influence training transfer to a real emergency. One concern is the impact of stress on learning and task performance; the other is the effect of the learning criterion upon retention, particularly retention under stress.

It is important to note that while we evaluated transfer of training for an emergency procedure, we conducted learning and transfer assessments without a real or purposely-induced stress.[3] The need to respond to unexpected events is a defining feature of high-reliability occupations.[4] Submarine duty can be considered a high-reliability occupation, since tasks are performed in a stressful situation, including a complex environment, uncertainty, time pressure and severe consequences for mistakes. It is well known that stress hinders human performance (see McCleron et al., 2011 for examples). Furthermore, training conducted without stress does not often transfer well to stressful conditions (Driskell and Johnston, 1998). Hence, we cannot predict the benefits of the VCVS when the task is performed under the stress of a real emergency and we suggest that stress exposure training should be considered along with the use of the VCVS.

The benefits of VE training for later performance under stress could possibly be enhanced by carefully adding stress during learning. McClernon et al. (2011) recently investigated the effect of stress during training on subsequent flight performance by exposing novices to a stressor while they learned with a desktop simulator that represented the primary flight display for instrumented control of an aircraft. An experimental group was exposed to cold as a stressor; the members of this group were provided instructions on how to mitigate the effects of stress on their performance while they mentally rehearsed the task. Afterward, they practiced the task with the simulator while exposed to the cold. The transfer task involved performance of the

[3] This is also true for traditional training aboard the submarine, however, the inherent level of stress may be greater for training aboard the submarine when conducted by Navy instructors.
[4] See Baumann et al. (2011) for definition, examples and citations for high-reliability occupations.

task in an actual aircraft, which provided the stress associated with flying an aircraft for the first time. Participants in the experimental group flew better than participants in a control group who did not receive stress training. The importance of this finding is that stress during training benefits later performance under stress, and that a different type of stressor (e.g., one suited to use with training in a VE) can be used as a substitute for the stresses of the operational situation in some situations.

Another important consideration in an analysis of the cost-benefits of the VCVS is the criterion for learning that was used in Experiment 1 and its consequences for retention, including retention under stress. The learning criterion was one errorless trial, which is the same level of training used to qualify personnel for submarine duty. Participants in the experimental group practiced the task within the VE until they could perform it on their own without making a mistake or request for help. Overlearning, that is, deliberate practice beyond this criterion could possibly help prevent memory decay and performance loss under stress.

Driskell et al. (1992) conducted a meta-analysis to determine the amount of overlearning necessary to produce a significant benefit for retention and the duration of the benefit of over learning. They found that adding 50% more trials, beyond the number of trials needed to reach criterion, produces a reliable, positive benefit for subsequent performance and that further increases in overlearning produce larger benefits. These outcomes were found for both physical and cognitive tasks, although the benefits of overlearning were found to be stronger for cognitive tasks. Thus, retention is generally much better when practice continues after the task is learned. Driskell et al. found that the benefit was generally halved after 19 days and was gone by 38 days. Since the VCVS affords the opportunity to practice drills repeatedly, novices could overlearn the task initially and then refresh their training with the VCVS three or four weeks later and possibly thereafter, at minimal cost and without reliance on access to the submarine. In addition, the VCVS could be used to help qualified submariners maintain their knowledge of vital drills, as many qualified submariners who took part in this study suggested voluntarily.

The benefits of overlearning and refresher training with the VCVS have special importance for emergency procedures that are not practiced frequently aboard the submarine because the first attempt to perform the task during an emergency could be the critical one. Overlearning is often thought to be a means to counteract the negative effect of stress on performance in the real world, but can also lead to behavioural inflexibility and negative transfer if the overlearned task does not closely match its real world counterpart (Driskell et al., 1992). Hence, the VCVS could potentially provide an inexpensive resource to build resistance to memory decay and stress, but further behavioural research is needed to learn if these positive benefits can be accrued without penalty.

The present study employed two experimental approaches, each including immediate transfer and practice phases, to acquire as much behavioural evidence about the validity of the VCVS as possible. It afforded an opportunity to compare the outcomes of each experimental approach and their predictions for training transfer. The reason for doing this was pragmatic, from a methodological perspective, since the classic transfer of training method, as employed in Experiment 2 for the first trial, is often very expensive, difficult to implement and sometimes dangerous to conduct (see McCauley, 2006). For these reasons, it is useful to consider alternative experimental methods that can reduce uncertainty about the effectiveness of the simulation as much as possible before performing a classic transfer of training experiment (Cross, 1992). The classic method typically compares the performances of two or more groups on the operational task within an operational setting, with each participant training to proficiency, where participants in the experimental group(s) are previously trained with a simulator and those within the control group are not. The classic method is a forward transfer of training paradigm and requires use of the operational resources. It permits an assessment of training effectiveness often based on the reduction of time required to achieve task proficiency with the operational equipment. Cross (1992) argues that one shortcoming of the classic method for assessing flight simulators is that it yields a composite measure of training transfer, and that the effectiveness of a flight simulator will likely be different for different manoeuvres. He reasons that a composite measure of training effectiveness will tend to underestimate a simulator's training effectiveness. Similarly, we suspected that an overall measure of the training effectiveness of the VCVS for isolating BH 35 may provide a misleading indication of the training effectiveness of its principal components, in this case spatial and procedural knowledge. For this reason, we looked at the composite measure of performance and its spatial and procedural components. The TER for the spatial component was almost twice as high as the composite measure.

Backward transfer was chosen as an alternative method for Experiment 1 because it does not require operational equipment. Backward transfer measures the extent to which knowledge and skills training with the operational equipment benefits performance in the simulation. Backward or reverse transfer methods have been used to assess flight training simulators (Stewart, 1994), and desktop training devices (Goettl and Shute, 1996). The conduct of a backward transfer study is recommended prior to the conduct of a forward transfer study since the results of a backward transfer study can help researchers predict the failure of a forward transfer study if the results of the backward study are poor and this method can possibly predict success if the results are good (Cross, 1992). The methods for backward transfer that we describe in Experiment 1 did not evaluate the proficiency of the qualified personnel aboard the submarine prior to their practice with the VCVS. Cross (1992) has noted that this is a useful, but not essential, step and we did not take it for two reasons: (1) it is the responsibility of active duty personnel aboard the submarine to maintain proficiency on the task, and (2) it would require access to the submarine, which we wanted to avoid because it would impinge on operations.

In both experiments, we combined a transfer paradigm with a skill acquisition paradigm. There were several reasons for this. One reason is that training transfer is a phenomenon that can influence the acquisition, performance or relearning of a behaviour or skill (Keirl and Hall, 1993). It is not uncommon for experienced personnel to require a few trials in the simulated environment to adapt to its peculiarities and it is clear from the performance data provided by the qualified group in Experiment 1 that the prior experience of this group with the submarine is evident shortly after initial exposure to the VE. The very clear findings that the experienced submariners in the qualified group required many fewer trials in the VE, and made many fewer errors on their way to criterion, than the novices in the experimental group, led us to expect a high degree of positive training transfer to the submarine. In Experiment 2, the results for positive forward training transfer were evident, but less pronounced than anticipated by the reverse training transfer statistics of Experiment 1. This comparison suggests that the positive outcome of a reverse transfer of training effectiveness paradigm may over predict forward transfer, and it suggests that the amount of experience may be an important factor. The qualified submariners had many years of experience to practice the task whereas the experimental participants learned it only once to criterion. Consequently, additional training within the VE would likely increase its positive benefits.

A second reason for combining transfer paradigms in this study was to obtain an estimate of the amounts of training time that could be needed to train novices and how much training time aboard the submarine could be saved. The results indicate that about 90 minutes would be required to train the task within the VE. We think that less time would be required for additional tasks, if they involve overlapping knowledge. The results also indicate that the VCVS can save time training on the submarine. Although the TER is relatively low in comparison to flight simulators that aim to achieve high fidelity, the cost effectiveness ratio is high because the cost of running the VCVS is extremely low in comparison to the costs of a docked submarine. The DND Cost Factors Manual (2011-2012) indicates that the cost of a SSK approaches one million dollars a day while docked; this is many orders of magnitude greater than the costs of the VCVS.

This page intentionally left blank.

References

Adams, J. A. (1972), Research and the future of engineering psychology, *American Psychologist*, 27 (7), 615-622.

Advisory Group for Aerospace Research & Development (1980), *Fidelity of simulation for pilot training*. AGARD-AR-159.

Alexander, A., Brunyé, T., Sidman, J., and Weil, S.A. (2005), From Gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in PC-based simulation and games, DARWARS Technical Report. http://www.aptima.com/publications/2005_Alexander_Brunye_Sidman_Weil.pdf.

Baumann, M. R., Gohm, C. L., and Bonner, B. L. (2011), Phased training for high-reliability occupations: Live-fire exercises for civilian firefighters, *Human Factors*, 53 (5), 548-557.

Blaiews, A. S., Puig, J. A., and Regan, J. J. (1973), Transfer of training and the measurement of training effectiveness, *Human Factors*, 15, 523-533.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., and Gratton, G. (2008), The effects of video game playing on attention, memory, and executive control, *Acta Psychologica*, 129 (3), 387-398.

Boulet, G. (in preparation), Serious games: How are they perceived by the end-user?

Boulet, G., and Gilbert, F. (2010), Alpha version trial report Victoria Class Virtual Submarine (VCVS) (RDIMS # 204282), 4115-1 (NeLID), 19 Oct 2010.

Cain, B., Magee, L.E., and Kersten, C. (2011), Validation of a virtual environment incorporating virtual operators for procedural learning, (DRDC TR 2011-132) Defence R&D Canada – Toronto.

Caro, P. W. (1977), Some factors influencing simulator training effectiveness in the U.S Air Force, HumRRO FR-ED-77-18, Alexandria, VA: Human Resources Research Organizaton.

Cross, K.D. (1992), Training effectiveness assessment: Methodological problems and issues, In *Proceedings of NASA/FAA Helicopter Simulator Workshop*, 77-90, NASA Ames Research Center.

Department of National Defence, Cost Factors Manual 2011-2012. http://admfincs.mil.ca/subjects/fin_docs/cfm-mcs_11/CFM-MCS_2011-12_e.

Driskell, J. E., and Johnston, J. H. (1998), Stress exposure training. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training,*pp.191-217, Washington, DC: American Psychological Association.

Driskell, J. E., Willis, R. P., and Copper, C. (1992), Effect of overlearning on retention. *Journal of Applied Psychology*, 77 (5), 615-622.

Ellis, H.C. (1965), *The transfer of learning*, New York: NY: Macmillan.

Fletcher, D. J., and Orlansky, J. (1989), Recent studies on the cost-effectiveness of military training in TTCP countries, (IDA Paper 613). Institute for Defense Analyses - Alexandria, Virginia.

Gagne, R. M., Foster, H., and Crowley, M. E. (1948), The measurement of transfer of training. *Psychological Bulletin*, 45, 97-130.

Garrett, M., McMahon, M., and Widdis, A. (2008), Serious games for defence training: Using the Location And Scenario Training System to enhance the spatial awareness of trainee submariners. In proceedings of SimTecT 2008, Melbourne, Australia.

Goettl, B. P., and Shute, V. J. (1996), Analysis of part-task training using the backward-transfer technique, *Journal of Experimental Psychology: Applied*, 2 (3), 227-249.

Granek, J. A., Gorbet, D. J., and Sergio, L. E. (2010), Extensive video-game experience alters cortical networks for complex visuomotor transformations, *Cortex,* 46 (9), 1165-1177.

Grant, S.C., and Magee, L.E. (1998), Contributions of proprioception to navigation in virtual environments, *Human Factors*, 40 (3), 489-497.

Gavish, N., Gutierrez, T., Webel, S., Rodriguez, J., and Tecchia, F. (2011), Design guidelines for the development of virtual reality and augmented reality training systems. Paper presented at the *The International Conference SKILLS 2011*, Montpellier.

Green, C. S., and Bavelier, D. (2003), Action video game modifies visual selective attention. *Nature*. 423 (6939), 534-537.

Johnson, S. J., Guediri, S. M., Kilkenny, C., and Clough, P. J. (2011), Development and validation of a virtual reality simulator: Human factors input to interventional radiology training, *Human Factors*, 53 (6), 612-625.

Keirl, J. M., and Hall, J. R. (1993), Measures of training effectiveness leading to the cost benefit of simulation based, aircrew training (DRA/AS/FS/CR93048/1.0), Defence Research Agency - Farnborough.

Lapeyre, B., Hourlier, S., Servantie, X., N'Kaoua, B., and Sauzéon, H. (2011), Using the landmark-route-survey framework to evaluate spatial knowledge obtained from synthetic vision systems, *Human Factors,* 53 (6), 647-661.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., and Phyllis, A. F. (1993), Nonvisual navigation by the blind and sighted: Assessment of path integration ability, *Journal of Experimental Psychology: General*, 122, 73-91.

McClernon, C. K., McCauley, M. E., O'Connor, P. E., and Warm, J. S. (2011), Stress training improves performance during a stressful flight, *Human Factors*, 53 (3), 207-218.

Magee, L.E., Cain, J.B, and Thompson, A.A. (2011), Training effectiveness of the Victoria Class Virtual Submarine, Human Research Ethics Committee (HREC) Protocol L-909, Defence R&D Canada - Toronto.

McCauley, M. E. (2006), Do army helicopter training simulators need motion bases? (Technical Report 1176), United States Army Research Institute for the Behavioral and Social Sciences.

Murdock, B. B., Jr. (1957), Transfer designs and formulas, *Psychological Bulletin*, 54, 313-326.

National Academy of Engineering (2009), *Enhance virtual reality,* www.engineeringchallenges.org/cms/8996/9140.aspx.

Orlansky, J., and String, J. (1980), Reaping the benefits of flight simulation. *Defence Management Journal*, Fourth Quarter, 6-13.

Povenmire, H. K., and Roscoe, S. N. (1971), An evaluation of ground-based flight trainers in routine primary flight training. *Human Factors*, 13, 109-116.

Roscoe, S. N., and Williges, B. H. (1980). Measurement of transfer of training. In S. N. Roscoe (Ed.), *Aviation Psychology*. Ames: Iowa State University Press.

Smallman, H. S., and St. John, M. (2005), Naive realism: Misplaced faith in realistic displays, *Ergonomics in design*, 15 (3), 6-13.

Sottilare, R., Alexander, T., Andrews, D., Goldberg, S., Hourlier, S., Koerhuis, C.L., Magee, L., and Roessingh, J.J. (Eds) (2009). Human dimensions in embedded virtual simulation (AC/323(HFM-169)/TP/364), RTO Meeting Proceedings MP-HFM-169.

Seidel, R. J., and Chatelier, P. R. (Eds.) (1997), *Virtual reality, training's future?: Perspectives on virtual reality and related emerging technologies*. New York, Plenum Press.

Stewart II, J. E. (1994). Using the backward transfer paradigm to validate the AH-64 simulator training advanced testbed for aviation. US Army Research Institute Report 1666.

Stone, R. J. (2008), Human factors guidelines for interactive 3D and games-based training systems design. Birmingham, UK, Human Factors Integration Defence Technology Centre Document**:** 86.

Stone, R.J. (2012), Personal communication with Magee,L.E., 8 February.

Stone, R.J, Caird-Daley, A., and Bessell, K. (2009), *SubSafe*: a games-based training system for submarine safety and spatial awareness (Part 1), *Virtual Reality*, 13, 3-12.

Stone, R.J., Caird-Daley, A., and Bessell, K. (2010), Human factors evaluation of a submarine spatial awareness training tool. In *Proceedings of the Human Performance at Sea (HPAS) 2010 Conference*, pp. 231-241, Glasgow.

Stone, R.T., Watts, K., Zhong, P., and Wei, C. (2011), Physical and cognitive effects of virtual reality integrated training, *Human Factors*, 52,(5), 558-572,

Tate, D. L., Sibert, L., and King, T. (1997), Using virtual environments to train firefighters, *Computer Graphics and Applications, IEEE*, 17(6), 23-29.

Taylor, H. L., Lintern, G., and Koonce, J.M. (2001), Quasi-transfer as a predictor of transfer from simulator to airplane, *The Journal of General Psychology,* 120 (3), 257-276.

Thorndike, E. L. (1906), *Principles of teaching*, New York, NY: Seiler.

Werner, S., Kreig-Brückner, B., Mallot, H. A., Schweizer, K., and Freska, C. (1997), Spatial cognition: The role of landmark, route, and survey knowledge in human and robot navigation. In M. Jarke, K. Pasedach & K. Pohl (Eds.), *Informatik '97*, pp. 41-50, Berlin: Springer-Verlag.

Wolbers, T., and Hegarty, M. (2010), What determines our navigational abilities? *Trends in Cognitive Science*, 14, 138-146.

List of symbols/abbreviations/acronyms/initialisms

AFT	Aft
BH 35	Bulkhead 35
BSQ	Basic Submariner Qualification
CFB	Canadian Forces Base
Ctrl	Control
CVNF	Canadian Virtual Naval Fleet
DMTE	Director Maritime Training and Education
DND	Department of National Defence
DRDC	Defence Research and Development Canada
EBS	Emergency breathing system
EEBD	Emergency escape breathing device
fMRI	Functional magnetic resonance imaging
FPS	First person shooter
FTOT	Forward transfer of training
FWD	Forward
HFI DTC	Human Factors Integration Defence Technology Centre
HMCS	Her Majesty's Canadian Ship
HREC	Human Research Ethics Committee
LCD	Liquid Crystal Display
LCdr	Lieutenant-Commander
LE	Learning event
LPB	Low pressure blower
LS	Leading Seaman
MARPAC	Maritime Forces Pacific
MCDV	Maritime Coastal Defence Vessel
MOD	Ministry of Defence
NAE	National Academy of Engineering
NeLCoE	Navy eLearning Centre of Expertise
OS	Ordinary Seaman

PAT	Personnel Awaiting Training
PC	Personal computer
PCC	Personnel Coordination Centre
PO1	Petty Officer 1st Class
PO2	Petty Officer 2nd Class
POV	Point of view
RAN	Royal Australian Navy
RCN	Royal Canadian Navy
RTOT	Reverse transfer of training
SME	Subject Matter Expert
SSK	Long range patrol submarine
TER	Transfer effectiveness ratio
UK	United Kingdom
USS	United States Ship
VCVS	Victoria Class Virtual Submarine
VE	Virtual Environment
VV	Ventilation valve
WSC	Weapons stowage compartment

Distribution list

Document No.: DRDC Toronto TR 2012-014

LIST PART 1: Internal Distribution by Centre

DMTE
DMRS
DMSS
Canadian Submarine Force (CANSUBFOR)
CFNOS / CO
HMCS Corner Brook / XO
CFFSQ / CO
NOTC / CO
CFFSE / CO
CFNES / CO
MARPAC HQ / PCC
CDA Kingston
DST (P)

13 TOTAL LIST PART 1

LIST PART 2: External Distribution by DRDKIM

1 Library and Archives Canada

TOTAL LIST PART 2

TOTAL COPIES REQUIRED

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

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 R&D Canada – Toronto 1133 Sheppard Avenue West P.O. Box 2000 Toronto, Ontario M3K 2C9	2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED (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.)

Training Effectiveness of the Victoria Class Virtual Submarine: A behavioural assessment

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

Magee, L.E., Thompson, A.A; Cain, B.; Kersten, C.

5. DATE OF PUBLICATION (Month and year of publication of document.) March 2012	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 56	6b. NO. OF REFS (Total cited in document.) 49

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 R&D Canada – Toronto
1133 Sheppard Avenue West
P.O. Box 2000
Toronto, Ontario M3M 3B9

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.) 14dn	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
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.) DRDC Toronto TR 2012-014	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

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.))

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.)

The Royal Canadian Navy eLearning Centre of Expertise (NeLCoE) in Quebec City developed the Canadian Virtual Naval Fleet (CVNF), a game-based, desk-top virtual environment (VE) to provide procedural and spatial knowledge of large vessels, which are not always available for training. This report presents the results of two experiments that assessed the training effectiveness of the Victoria Class Virtual Submarine (VCVS), one implementation of the CVNF. Each experiment assessed the ability of Navy personnel to complete an emergency drill that involved isolation of a bulkhead. Initial transfers of training and improvements with practice to criterion (i.e., error free performance) were used to compare the performances of a trained group with a novice group. However, the experiments differed in method, prior training of the groups, and the practice environment. The first experiment employed a reverse transfer of training paradigm using the VCVS as the practice environment. In this experiment the performances of a group of ten navy personnel qualified to perform the task aboard the submarine were compared to a group of ten novices who lacked prior submarine experience and who learned the task for the first time within the VE. The second experiment employed a more classic, forward transfer of training paradigm that used the submarine as the transfer environment. Here, the performances of the novice group that practiced the task to criterion within the VE were compared to a second group of ten novices who had the task demonstrated to them in a traditional instructional session aboard the submarine. Two days later, the members of each group individually demonstrated their ability to perform the task aboard the submarine and repeated their attempts until they could perform the task without error. Both experiments yield evidence of positive training transfer, especially for spatial knowledge.

L'équipe du Centre d'expertise de l'apprentissage en ligne de la Marine royale canadienne à Québec a développé une flotte navale virtuelle canadienne (FNVC), c'est-à-dire un environnement virtuel pour ordinateur basé sur le jeu, et ce, afin de fournir des connaissances procédurales et spatiales sur les grands navires, qui ne sont pas toujours disponibles pour la formation. Ce rapport présente les résultats de deux expériences destinées à évaluer l'efficacité de la formation avec le sous-marin virtuel de la classe VICTORIA (SVCV), une initiative de la FNVC. Chaque expérience a évalué la capacité du personnel de la Marine à effectuer un exercice d'alerte comprenant l'isolation d'une cloison. Le transfert initial de la formation et l'amélioration des résultats grâce à la répétition (c.-à-d. jusqu'à l'obtention d'un rendement exempt d'erreur) ont été utilisés pour comparer les performances d'un groupe de participants entraînés et d'un groupe de novices. Néanmoins, les expériences différaient sur les plans de la méthode, de l'entraînement antérieur des groupes et de l'environnement d'exercice. La première expérience avait recours à un transfert inversé du paradigme de formation et utilisait le SVCV comme environnement d'exercice. Dans le cadre de cette expérience, les performances d'un groupe de dix membres de la Marine qualifiés pour réaliser la tâche à bord du sous-marin ont été comparées à celles d'un groupe de dix novices qui ne possédaient pas d'expérience antérieure à bord d'un sous-marin et qui ont appris à exécuter la tâche pour la première fois dans l'environnement virtuel. La deuxième expérience employait un paradigme de transfert de la formation plus traditionnel et utilisait le sous-marin comme environnement de transfert. Dans ce cas, les performances du groupe de novices s'étant exercés à accomplir la tâche sans commettre d'erreur dans un environnement virtuel ont été comparées avec celles d'un groupe de dix novices qui avaient assisté à une démonstration de la tâche à l'occasion d'une séance de formation traditionnelle à bord du sous-marin. Deux jours plus tard, les membres de chaque groupe ont démontré individuellement leur capacité à accomplir la tâche en question à bord du sous-marin, puis ont répété leurs tentatives jusqu'à ce qu'ils soient en mesure de la réaliser sans faire d'erreur. Les deux expériences ont prouvé l'existence de transfert de formation positif, particulièrement en ce qui a trait aux connaissances spatiales.

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.)

Virtual environments; simulators; training effectiveness; submarine