Archives of Physical Medicine and Rehabilitation

journal homepage: www.archives-pmr.org

Archives of Physical Medicine and Rehabilitation 2016;97:1085-92

ORIGINAL RESEARCH

Stability and Workload of the Virtual Reality–Based Simulator-2

Deepan C. Kamaraj, MD, MS,[a,b] Brad E. Dicianno, MD, MS,[a,c] Harshal P. Mahajan, PhD,[a] Alhaji M. Buhari, MA, MSIE,[a] Rory A. Cooper, PhD[a,b,c,d]

From the [a]Center of Excellence in Wheelchairs and Related Technology, Veterans Affairs Pittsburgh Healthcare System and Human Engineering Research Laboratories, Pittsburgh, PA; Departments of [b]Rehabilitation Science and Technology, and [c]Physical Medicine and Rehabilitation, University of Pittsburgh, Pittsburgh, PA; and [d]Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA.

Abstract

Objective: To assess the stability of clinicians' and users' rating of electric-powered wheelchair (EPW) driving while using 4 different human-machine interfaces (HMIs) within the Virtual Reality–based SIMulator–version 2 (VRSIM-2) and in the real world (accounting for a total of 5 unique driving conditions).

Design: Within-subjects repeated-measures design.

Setting: Simulation-based assessment in a research laboratory.

Participants: A convenience sample of EPW athletes (N=21) recruited at the 31st National Veterans Wheelchair Games.

Interventions: Not applicable.

Main Outcome Measures: Composite PMRT scores from the Power Mobility Road Test (PMRT); Raw Task Load Index; and the 6 subscale scores from the Task Load Index developed by the National Aeronautics and Space Administration (NASA-TLX).

Results: There was moderate stability (intraclass correlation coefficient between .50 and .75) in the total composite PMRT scores ($P<.001$) and the users' self-reported performance scores ($P<.001$) among the 5 driving conditions. There was a significant difference in the workload among the 5 different driving conditions as reflected by the Raw Task Load Index ($P=.009$). Subanalyses revealed this difference was due to the difference in the mental demand ($P=.007$) and frustration ($P=.007$) subscales. Post hoc analyses revealed that these differences in the NASA-TLX subscale scores were due to the differences between real-world and virtual driving scores, particularly attributable to the conditions (1 and 3) that lacked the rollers as a part of the simulation.

Conclusions: Further design improvements in the simulator to increase immersion experienced by the EPW user, along with a standardized training program for clinicians to deliver PMRT in VRSIM-2, could improve the stability between the different HMIs and real-world driving.

Archives of Physical Medicine and Rehabilitation 2016;97:1085-92

Electric-powered wheelchairs (EPWs) are vital mobility devices that provide independence and a better quality of life for people with disabilities.[1-3] However, providing training for novice EPW users or users with moderate driving skills has been a challenge for clinicians and wheelchair clinics with limited resources.[4-7] Studies have reported that nearly 40% of individuals who receive an EPW do not have the necessary skills for safe maneuverability of the EPW.[8] The lack of EPW driving training has also been stated as one of the reasons for wheelchair-related accidents and injuries within the first year of using the device.[9-11] Hence, to fill the need for an EPW driving assessment and training tool, the Virtual Reality–based SIMulator (VRSIM) was developed.[12]

The VRSIM is an EPW driving simulator with 2 displays (a set of immersive virtual reality [VR] screens or a desktop monitor [personal computer screen]) and 2 control interfaces (encoders fitted in a dual roller system or an instrumented EPW joystick that operates through custom software), providing 4 unique human-

Supported by the Department of Veteran Affairs Research and Development Merit Review Award (grant no. A6035R) with resources and facilities at the Human Engineering Research Laboratories, VA Pittsburgh HealthCare System. The contents of this publication do not represent the views of the Department of Veterans Affairs or the United States Government.

Disclosures: none.

0003-9993/16/$36 - see front matter © 2016 by the American Congress of Rehabilitation Medicine

http://dx.doi.org/10.1016/j.apmr.2016.01.032

machine interfaces (HMIs).[12] Version 1 of VRSIM was a prototype that displayed a virtual environment (VE) in which a user could navigate using a virtual EPW through different obstacles.[12] The VE of Virtual Reality−based SIMulator−version 2 (VRSIM-2) was an improved prototype of version 1, built on the Power Mobility Road Test (PMRT),[13] and simulated a real-world office space with a kitchen, lounge area, and a set of hallways lined by offices.

Prior research was conducted to determine whether clinicians with varying levels of experience would rate EPW driving performance similarly when observing individuals drive a virtual EPW in VRSIM-2. The total composite Power Mobility Road Test (TC-PMRT) scores were used as a measure of the clinicians' rating of driving performance. Initial psychometric analysis performed with 5 clinicians having varying levels of clinical experience demonstrated high interrater reliability within each of the 4 HMIs of VRSIM-2. However, stability of the TC-PMRT scores among different HMIs of VRSIM-2 (ie, the consistency of repeated measurements[14]) and the quantification of EPW user experience while using VRSIM-2 have not yet been assessed.

Higgins and Straub[14] describe stability as follows: if a sample of users (eg, EPW users with similar driving capability) was tested more than once, under similar circumstances (eg, same VE within VRSIM-2), using the same instrument (eg, PMRT), results of the tool should be similar. Stability is used with phenomena in which little or no change is expected between 2 or more trials.[14] Since VRSIM-2 used the same mathematical model of the EPW within the same VE with all its HMIs, little or no difference was expected with usage of different HMIs of VRSIM-2 within each rater group, which would then be reflected with stable TC-PMRT scores. Thus, the first aim of the present study was to assess stability of driving performance among 4 different HMIs and the real world as perceived by both raters and EPW users. To achieve this aim, 3 hypotheses were evaluated. First, the stability among the 5 driving conditions would be high (intraclass correlation coefficient [ICC] >.75), as measured by TC-PMRT scores and users' self-reported performance scores from the Task Load Index developed by the National Aeronautics and Space Administration (NASA-TLX). Second, the stability of scores among 4 HMIs would be high. Third, the stability of structured composite Power Mobility Road Test (SC-PMRT) scores and dynamic composite Power Mobility Road Test (DC-PMRT) scores would be higher among the experienced raters than the novice raters.

The second aim was to quantitatively assess user experience by measuring workload. Workload is a term that represents the cost of accomplishing a specific task's requirements for the human operator.[15] Since VRSIM-2 is intended to provide an immersive VE that is aimed to simulate a driving experience that is close to real-world EPW driving, measuring workload can provide information about whether VRSIM-2 was successful in accomplishing this goal from an EPW user's perspective. Hence, the fourth hypothesis was that workload would be higher in the 4 HMIs compared with real-world driving.

List of abbreviations:

DC-PMRT	**dynamic composite Power Mobility Road Test**
EPW	**electric-powered wheelchair**
HMI	**human-machine interface**
ICC	**intraclass correlation coefficient**
NASA-TLX	**National Aeronautics and Space Administration Task Load Index**
PMRT	**Power Mobility Road Test**
SC-PMRT	**structured composite Power Mobility Road Test**
TC-PMRT	**total composite Power Mobility Road Test**
VE	**virtual environment**
VR	**virtual reality**
VRSIM	**Virtual Reality−based SIMulator**
VRSIM-2	**Virtual Reality−based SIMulator−version 2**

Methods

The institutional review boards of the Veterans Affairs Pittsburgh Healthcare System and the University of Pittsburgh approved this research study. A within-subject repeated-measures study design was implemented, using a convenience sample of 21 EPW athletes recruited at the 31st National Veterans Wheelchair Games. Participants interacted with VRSIM-2 through 2 display options using 2 control interfaces, providing 4 possible HMIs, constituting 4 possible driving conditions (conditions 1−4: table 1; figs 1, 2). Participants performed 3 driving sessions in virtual conditions (2 sessions with the roller system and 1 session with the joystick interface) for each of the 2 display options and also performed 1 driving session in a real-world driving course (condition 5: see table 1; see figs 1, 2). Two raters were randomly selected from 2 groups of clinicians: an experienced group wherein each clinician had >5 years of performing EPW driving assessments and a novice group with <5 years of experience. The raters assessed each driving trial (see fig 2).

Outcome measures

Power Mobility Road Test

The PMRT[13] has 16 tasks with total scores ranging from 16 to 64 for 1 driving trial. The total score for each trial was calculated and expressed as a percentage of the maximum total score (64), termed the TC-PMRT score. Similarly, the scores from the structured and dynamic tasks were totaled for each trial, expressed as percentages of the maximum possible scores within each of these categories (48 and 16, respectively) and defined as SC-PMRT and DC-PMRT scores. Massengale et al[13] reported that a TC-PMRT score ≥95% would suggest that the user is a safe driver. The PMRT in VRSIM-2 was conducted in a simulated indoor office space, and an equivalent course was charted in an open space for the real-world assessment.

National Aeronautics and Space Administration Task Load Index

The NASA-TLX is a multidimensional self-report tool that estimates overall workload perceived by the user.[15,16] It has been applied extensively in studying virtual interfaces of simulators and HMIs.[17-19] In this study, the Raw Task Load Index was used as an overall measure of perceived workload,[15,16,20] calculated by averaging the scores from 6 subscales: mental demand, physical demand, temporal demand, effort, frustration, and performance.[15,16,20] Ranging between 0-100, higher Raw Task Load Index indicates higher workload.[15,16,20] Additionally, the performance subscale score from NASA-TLX was used as a user's self-reported performance measure. Participants completed the NASA-TLX after every driving condition.

A detailed description of the recruitment procedures, experimental setup, and the data collection protocol are described in an accompanying article.

Data reduction

There were 2 conditions (2 and 4) with 2 trial repetitions (see fig 2). The averages of the PMRT scores from these 2 trials were computed.