Proceedings of the IISE Annual Conference & Expo 2024
A. Brown Greer, C. Contardo, J.-M. Frayret, eds.

Predicting Usability of Prosthetic Devices in a Virtual Reality Setting

Junho Park[a] and Maryam Zahabi[b]

[a]Department of General Engineering, Santa Clara University, CA, 95053
[a]Wm Michael Barnes '64 Department of Industrial & Systems Engineering, Texas A&M University, TX, 77843

Virtual reality (VR) simulations can be used for prosthetic device training and patient rehabilitation. Prior studies found that using VR simulations can reduce mental workload and increase perceived usability of prosthetic devices. Although previous studies have evaluated the usability of prostheses in VR settings, they mainly relied on user-testing which require a functional device or prototype. However, assessing the usability of prosthetic devices in early stages of the design process can be cost effective and help developers design a more usable prosthetic device. Therefore, the objective of this study was to present an approach to predict usability of prostheses in VR training settings. A human-subject study with 20 participants was conducted. Two prosthetic device configurations (i.e., direct control and pattern recognition) and a daily living task (i.e., clothespin relocation test) were simulated in a VR setting. The results suggested that the model outcomes were similar to the results of the human-subject experiment. The tool was able to predict the usability dimensions based on a few input parameters (e.g., device calibration quality, first impression of the device) and using a graphical user interface. The findings provided a quick and practical tool for a prototype-level usability analysis of prostheses. Furthermore, the tool could eventually help clinicians find, test, and recommend prosthetic devices that better fit the needs of amputees.

Keywords
Usability, virtual reality, prosthesis, prediction, human performance model

1. Introduction
Virtual reality (VR) offers immersive training for complex situations [1] and allows for customizing prosthetic devices to individual needs, enabling personalized training [2, 3]. Furthermore, training in VR gives immediate feedback on user movements, aiding in movement adjustment and prosthesis control [3]. It also enables safe and repetitive practice, critical for developing muscle memory and improving prosthesis handling [4]. Lack of usability in prostheses can lead to device underuse or rejection [5]. Existing questionnaires like System Usability Scale (SUS) [6] and Usefulness, Satisfaction, and Ease of Use (USE) [7] assess the usability of prosthetic devices but are typically used in later design stages, requiring functional devices and costly, time-consuming human subject studies [8]. These methods also have self-report biases. Therefore, this study aims to present an approach to predict usability of upper-limb prostheses during VR-based training (i.e., **H**uman **P**erformance **M**odel for **Up**per Limbs; HPM-UP).

2. Human Performance Model for Upper Limb (HPM-UP)
HPM-UP predicts six dimensions of usability including learnability, error rate, efficiency, memorability, satisfaction, and cognitive workload. The model is developed based on human performance models and a machine learning approach. To calculate learnability, (1) the adaptive learning curve formulation [9] was used; (2) based on the learnability formula, error rate was calculated using a natural exponential function, (3) memorability was calculated based on the ACT-R declarative module [10], (4) efficiency was formulated based on Cognitive-Perceptual-Motor Goals, Operators, Methods, and Selection rules (CPM-GOMS) [11] and motion-time measurement (MTM) [12], (5) satisfaction was formulated based on the expectation confirmation theory [13], and (6) cognitive workload classification model was developed using the Naïve Bayes algorithm [14, 15]. CPM-GOMS extends the original GOMS framework by incorporating a detailed model of human cognitive, perceptual, and motor processes, as described in the Cognitive, Perceptual, and Motor (CPM) theory [16]. The model is released to Github including instructions and other researchers can modify or update it. The computational details of the model can be found in Park [17]. Figures 1 and 2 illustrate the input parameters and output of HPM-UP respectively. To calculate the usability dimensions, the model requires some inputs including: the control mode (physical or virtual prosthetic device), task (e.g., clothespin relocation tasks; CRT), control scheme (e.g., pattern recognition; PR), minimum and maximum training time duration (based on analysts' previous knowledge or pilot test results), device calibration quality (0-1), first impression (a number between -1 and 1), and effort (0-1) based on the end users' interaction with the device. To classify cognitive workload, input measures including percentage change in pupil size (PCPS) and blink rate (collected using eye-tracking device embedded in the VR goggles), and task performance (collected from training trials) need to be added. Once all the input parameters are added, analysts can see the outcomes in terms of the six usability dimensions (Figure 2). More information regarding the interpretation of outputs is provided in [17].

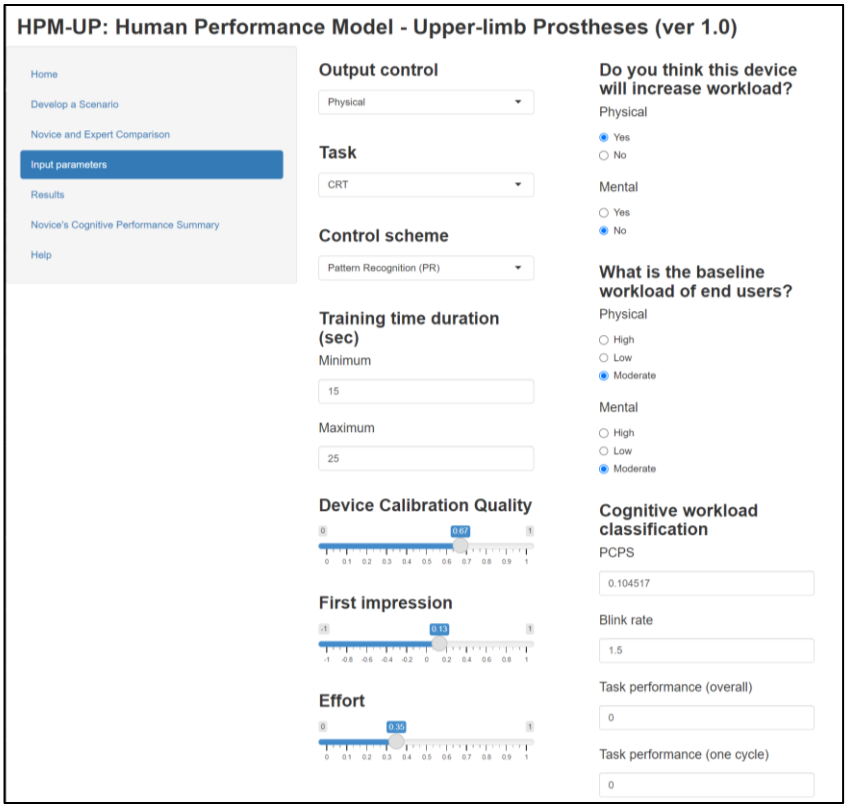

Figure 1: Input parameters for HPM-UP

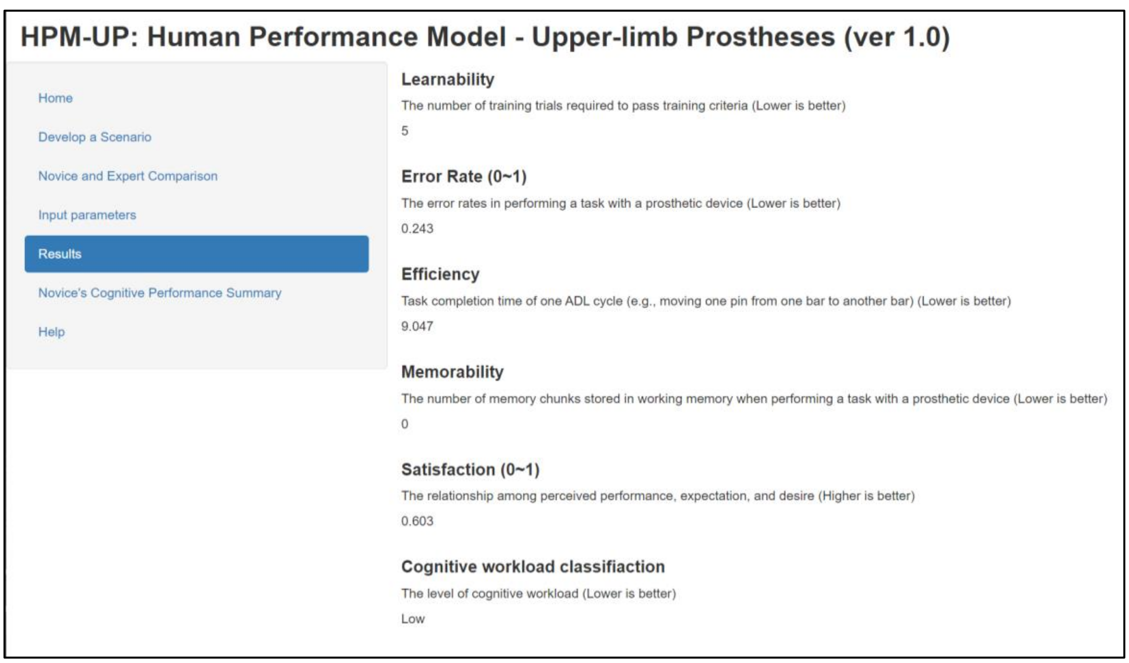

Figure 2: Predicted usability dimensions from HPM-UP

3. Model Validation

This experiment aimed to collect data from human subjects using VR for activities for daily living (ADL), validating HPM-UP results. It involved 20 participants (13 males, 7 females, average age 26.9 years) at Texas A&M University, all with normal or corrected vision and no experience with prosthetic or myoelectric upper limb devices. The study, approved by Texas A&M IRB (IRB2021-0990D), used direct control (DC) and patter recognition (PR) configurations.

The experiment setup consisted of three modules: (1) EMG/kinematic data collection and processing, (2) a server module, and (3) a VR module, with detailed input/output data formats for compatibility. The VR simulation used the HTC VIVE Pro Eye head-mounted interface which included a built-in eye tracking, and was developed on the Unity Game Engine v2019.4.28f [18]. The virtual prosthesis was modeled after the Fillauer Motion Control Electric Terminal Device 2. EMG signals were collected using a Delsys Trigno Wireless Biofeedback system and four Trigno Avanti Sensors, set to a sampling rate of 1,111 Hz. For the DC configuration, which requires EMG from an agonist-antagonist muscle pair [19], sensors were placed on the flexor carpi radialis and extensor carpi radialis longus. The open/close and pronation/supination motions could be controlled via wrist flexion and extension in the DC mode. To switch between the modes (e.g., open/close to pronation/supination), the user must co-contract their muscles, i.e., make a fist. In the PR mode, the open/close and pronation/supination motions are controlled via their natural hand motions (e.g., prosthesis pronation is achieved by pronation of the intact hand for able-bodied subjects or imagined pronation of the missing hand for amputees).

The VR application simulated the Clothespin Relocation Test (CRT) (Figure 3). CRT is a commonly applied ADL for assessing upper limb prostheses [20, 21]. It requires participants to move as many pins as possible from one bar to another within 2 minutes. The experiment included three trials. Between each trial, there was a 2-minute rest. The virtual prosthesis must be in the open position and close enough to a clothespin to see the highlighted outline cue to pick up a clothespin in the VR environment. This yellow outline is a visual indicator that the virtual prosthesis is close enough to grip a clothespin. Visual cues for interaction are necessary features as there is no tactile feedback afforded by the VR environment. The participant must then generate the command to close the hand to grip the clothespin. Clothespins in hand can be released by generating another open command. If a clothespin is released in a position in which it clamps onto any one of the bars of the base station, it will lock to that position until it is gripped again. If a clothespin is released anywhere other than onto one of the four bars, it will automatically respawn in the last valid position in which it was placed. If a clothespin is dropped immediately after removing it from the start position, it will return to the starting position.

The usability measures were collected after the experiment to be used as a ground truth and to compare with the HPM-UP outcomes. For example, the USE questionnaire was used at the end of experiment and the number of training trials to achieve mastery was counted for each participant. A benchmark model was developed using the CPM-GOMS method with Cogulator software [22] to be compared with HPM-UP and human-subject experiment outcomes. We chose Cogulator as it is open-sourced, has been continuously updated, and has CPM-GOMS/ ACT-R logics [23].

Figure 3: Clothespin Relocation Test in Virtual Reality

Due to the limited number of data points for each device configuration, nonparametric analysis was conducted to assess the differences in usability dimensions among the human subject data, HPM-UP, and benchmark model. For the comparison between two sets of data, Wilcoxon rank sum test was conducted [24]. The Wilcoxon test statistic "W" was used to determine the significance of the difference. All the statistical analysis was conducted using R 4.0.5. Effect size for Wilcoxon signed-rank test was calculated with $r = \frac{Z}{\sqrt{n}}$, where Z-score is a test statistic and standardized score of U-value calculated from Mann-Whitney U-test [25] and n is the total number of observations. The effect size of Kruskal-Wallis test was calculated using Eta-squared [26].

3.2. Results

Table 1 summarizes the descriptive statistics. All research hypotheses were supported by the data (Table 2). The benchmark model does not provide learnability, error rate, satisfaction, and cognitive workload estimates, and therefore, these cells are marked with "N/A" in Table 1. However, HPM-UP was able to generate all six dimensions.

Table 1. Descriptive statistics from the experiment (mean (sd))

Usability Dimensions	Human subject experiment		HPM-UP		Benchmark model	
	DC	PR	DC	PR	DC	PR
Learnability	5.5 (2.33)	3.6 (0.49)	3.4 (0.92)	4.3 (1.10)	N/A	N/A
Error rate	N/A		0.25 (0.06)	0.30 (0.07)	N/A	
Memorability	N/A		3.48 (0.13)	0	3.60 (0.00)	0
Efficiency	26.27 (10.19)	9.05 (3.06)	17.83 (1.67)	8.42 (0.94)	13.3 (0.00)	6.9 (0.00)
Satisfaction	0.68 (0.16)	0.74 (0.14)	0.74 (0.12)	0.75 (0.11)	N/A	N/A
Accuracy of Cognitive Workload Classification (%)	N/A		65.00	80.00	N/A	

Table 2. Summary hypothesis test results

Hypothesis ID	Hypothesis	Test Result	Test statistics, *p*-value, effect size
H1 (Learnability)	The results of HPM-UP learnability dimension would be similar to the human-subject data.	Cannot reject	$W = 436.5$, $p = .85$, $r = .03$
H2 (Memorability)	The results of HPM-UP memorability dimension would be similar to the benchmark model.	Cannot reject	$W = 65$, $p = .23$, $r = .28$
H3 (Efficiency)	The results of HPM-UP efficiency dimension would be similar to the human subject data	Cannot reject	$Z = 0.26$, $p = .79$, $r = .03$
H4 (Satisfaction)	The results of HPM-UP satisfaction dimension would be similar to the human-subject data	Cannot reject	$W = 413$, $p = .59$, $r = .07$

4. Discussion

All research hypotheses were supported by the data. The model's formulations were based on well-established human performance models and theories, such as CPM-GOMS and MTM. Furthermore, unlike previous prosthetic device studies that used 2D-displays [27, 28], this study used an immersive VR (i.e., VR headset) for training. This could provide potential benefits to researchers who are planning to conduct VR-based trainings for amputees.

The findings suggested that the model can estimate usability of prosthetic devices by simulating an ADL in VR settings. Once the analyst knows the tasks, device control modes, and users' characteristics as input measures, they can get the usability estimate for the prosthesis and compare that with the existing thresholds for acceptable usability [17]. Using these estimates, the analyst can provide personalized training programs in VR with diverse prostheses, adjusting control schemes and settings to suit individual users' needs, thereby facilitating user's learning [3, 29].

Using these dimensions, the analysts can conduct a holistic assessment of prosthetic devices. For example, the learnability dimension can help improve the quality of the training. By adjusting the device calibration, the analyst or VR developers can work together to reduce the number of training trials required for prostheses' users to get familiarized with the device. The learnability dimension has impact on other usability dimensions. This was to emphasize the prostheses' foremost functionality as an assistive device, which means if the device is hard to be used or learned for users, it does not provide good efficiency or satisfaction [17]. Therefore, the analyst can test their prostheses with emphasis on having reasonable learnability. The designers can improve the intuitiveness of the control scheme by minimizing the memory chunks required to use the prosthetic devices, which can impact users' memory load. sing this model, the VR developers can compare several prostheses and their control modes in VR. This could lead to a more efficient design and development cycle, as compared to conducting user testing studies with each prosthetic device. The model not only helps in customizing the prosthesis to the user's specific needs [30], but also can reduce the error rate, memory load, and thereby increasing efficiency and satisfaction for the user. Furthermore, the model can make the VR training environment more engaging for users, which is one of the important factors for training in VR [29].

This study also contributes to the field of human performance modeling (HPM) by highlighting the importance of timing in predicting the usability of the final product. Traditionally, HPMs have been used during the early phases of product or service development [16, 31]. However, with the advent of VR development and testing, it is possible to create HPMs using VR prototypes before initiating the prosthetic device development process.

There are several aspects of this study that may limit the generalizability of findings. First, HPM-UP has some free parameters, especially in learnability and satisfaction dimensions. Second, decision to work with an able-bodied population was made due to the limited number of trans-radial amputees in the surrounding area. Third, with the GUI, analysts can develop scenarios only with mouse clicks however, they need to have basic knowledge of human performance modeling. Lastly, although HPM-UP provides estimates of device usability, it cannot guarantee the fitness or feeling of embodiment of a prosthesis to amputees.

5. Conclusion

This study proposed an approach to predict usability of upper-limb prostheses in VR settings, which is expected to save time and effort for VR designers, experimenters, or prosthetic device developers. The model can be modified and used to estimate usability for future prostheses with new control schemes.

Acknowledgement

This project was funded by the National Science Foundation (No. IIS- 1856676). The opinions expressed in this report are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Zahabi, M. and A.M. Abdul Razak, *Adaptive virtual reality-based training: a systematic literature review and framework.* Virtual Reality, 2020. **24**(4): p. 725-752.
2. Biddiss, E.A. and T.T. Chau, *Upper limb prosthesis use and abandonment: a survey of the last 25 years.* Prosthetics and orthotics international, 2007. **31**(3): p. 236-257.
3. Park, J., A. Music, D. Delgado, J. Berman, A. Dodson, Y. Liu, J. Ruiz, H. Huang, D. Kaber, and M. Zahabi. *Cognitive Workload and Usability of Virtual Reality Simulation for Prosthesis Training*. in *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2023. IEEE.
4. Broeren, J., L. Claesson, D. Goude, M. Rydmark, and K.S. Sunnerhagen, *Virtual rehabilitation in an activity centre for community-dwelling persons with stroke.* Cerebrovascular Diseases, 2008. **26**(3): p. 289-296.
5. Park, J., J. Berman, A. Dodson, Y. Liu, A. Matthew, H. Huang, D. Kaber, J. Ruiz, and M. Zahabi. *Cognitive Workload Classification of Upper-limb Prosthetic Devices*. in *2022 IEEE 3rd International Conference on Human-Machine Systems (ICHMS)*. 2022. IEEE.
6. Brooke, J., *Sus: a "quick and dirty'usability.* Usability evaluation in industry, 1996. **189**.
7. Lund, A.M., *Measuring usability with the use questionnaire12.* Usability interface, 2001. **8**(2): p. 3-6.
8. Park, J. and M. Zahabi. *Comparison of cognitive workload assessment techniques in emg-based prosthetic device studies*. in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2020. IEEE.
9. Zhang, W., W. Ma, M. Brandao, D.B. Kaber, P. Bloomfield, and M. Swangnetr, *Biometric validation of a virtual reality-based psychomotor test for motor skill training.* Assistive Technology, 2016. **28**(4): p. 233-241.
10. Bothell, D., *ACT-R 7.21+ reference manual*. 2020, Technical Report. Technical Report, Carnegie Mellon University.
11. John, B.E. *Extensions of GOMS Analyses to Expert Performance Requiring Perception of Dynamic Visual and Auditory*. in *Empowering People: CHI'90 Conference Proceedings [on] Human Factors in Computing Systems: Seatle, Washington, April 1-5, 1990*. 1990. Addison-Wesley.
12. Maynard, H.B., G.J. Stegemerten, and J.L. Schwab, *Methods-time measurement.* 1948.
13. Oliver, R.L., *A cognitive model of the antecedents and consequences of satisfaction decisions.* Journal of marketing research, 1980. **17**(4): p. 460-469.

14. Majka, M., *naivebayes: High performance implementation of the Naive Bayes algorithm. R package version 0.9. 2*. 2018.
15. Park, J., J. Berman, A. Dodson, Y. Liu, M. Armstrong, H. Huang, D. Kaber, J. Ruiz, and M. Zahabi, *Assessing workload in using electromyography (EMG)-based prostheses.* Ergonomics, 2023(just-accepted): p. 1-24.
16. Park, J. and M. Zahabi, *A review of human performance models for prediction of driver behavior and interactions with in-vehicle technology.* Human factors, 2022: p. 00187208221132740.
17. Park, J., *Human Performance Modeling of Upper Limb Prosthetic Devices*. 2023.
18. Music, C.A., *ELECTROMYOGRAPHY-BASED ASSISTIVE VIRTUAL REALITY HUMAN-MACHINE INTERFACE.* UNIVERSITY OF FLORIDA, 2022.
19. Resnik, L., H.H. Huang, A. Winslow, D.L. Crouch, F. Zhang, and N. Wolk, *Evaluation of EMG pattern recognition for upper limb prosthesis control: a case study in comparison with direct myoelectric control.* Journal of neuroengineering and rehabilitation, 2018. **15**(1): p. 23.
20. Zahabi, M., M.M. White, W. Zhang, A.T. Winslow, F. Zhang, H. Huang, and D.B. Kaber, *Application of Cognitive Task Performance Modeling for Assessing Usability of Transradial Prostheses.* IEEE Transactions on Human-Machine Systems, 2019. **49**(4): p. 381-387.
21. Stubblefield, K., R. Lipschutz, M. Phillips, C. Heckathorne, and T. Kuiken. *Occupational therapy outcomes with targeted hyper-reinnervation nerve transfer surgery: Two case studies*. in *Proceedings of the MyoElectric Controls/Powered Prosthetics Symposium*. 2005.
22. Estes, S., *Cogulator.* The MITRE Corporation, 2017.
23. Estes, S., *Cogulator - A Cognitive Modeling Calculator.* 2021.
24. Wilcoxon, F., *Individual comparisons by ranking methods*, in *Breakthroughs in statistics*. 1992, Springer. p. 196-202.
25. Tomczak, M. and E. Tomczak, *The need to report effect size estimates revisited. An overview of some recommended measures of effect size.* Trends in sport sciences, 2014. **21**(1).
26. Rosenthal, R., *Meta-Analytic Procedures for Social Science Research Sage Publications: Beverly Hills, 1984, 148 pp.* Educational Researcher, 1986. **15**(8): p. 18-20.
27. Rezazadeh, I.M., S. Firoozabadi, S.H. Golpayegani, and H. Hu. *Controlling a virtual forehand prosthesis using an adaptive and affective Human-Machine Interface*. in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2011. IEEE.
28. Deeny, S., C. Chicoine, L. Hargrove, T. Parrish, and A. Jayaraman, *A Simple ERP Method for Quantitative Analysis of Cognitive Workload in Myoelectric Prosthesis Control and Human-Machine Interaction.* Plos One, 2014. **9**(11).
29. Phelan, I., M. Arden, M. Matsangidou, A. Carrion-Plaza, and S. Lindley. *Designing a virtual reality myoelectric prosthesis training system for amputees*. in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021.
30. Kluger, D.T., J.S. Joyner, S.M. Wendelken, T.S. Davis, J.A. George, D.M. Page, D.T. Hutchinson, H.L. Benz, and G.A. Clark, *Virtual reality provides an effective platform for functional evaluations of closed-loop neuromyoelectric control.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2019. **27**(5): p. 876-886.
31. Park, J. and M. Zahabi. *A novel approach for usability evaluation of mobile applications*. in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2021. SAGE Publications Sage CA: Los Angeles, CA.