Usability and acceptability of the TransKinect application for the assessment of independent wheelchair transfer technique with novice therapists

Ahlad Neti B.E.[1,2], Celia Davis[1,2], Cheng-Shiu Chung PhD.[2], Nithin Ayiluri B.S.[2],
Brooke Slavens PhD.[3], Alicia M. Koontz PhD.[1,2]

[1]University of Pittsburgh, Swanson School of Engineering, Department of Bioengineering, [2]Human Engineering Research Laboratories, VA Pittsburgh HealthCare System, [3]University of Wisconsin Milwaukee, College of Health Sciences, Department of Rehabilitation Sciences and Technology

INTRODUCTION

Throughout a typical day, a wheelchair user may perform between 15 to 40 sitting-pivot-transfers (SPTs), depending on his or her lifestyle and activities of daily living [1]. Unfortunately, it is common among wheelchair users to develop chronic upper body pain or injuries stemming from poor transfer technique or inadequate training. In one study, it was shown that of the 800 spinal cord injury (SCI) patients that participated, 72.7% reported some degree of chronic wrist and/or shoulder pain that was significantly linked to SPTs [2]. Because wheelchair users rely primarily on their upper body for mobility, any upper body pain or injury poses a serious risk of reducing autonomy and quality of life [3]. Therefore, it is imperative that good technique is used during SPTs to avoid predisposition to upper extremity pain or discomfort. One method to analyze and evaluate SPT technique is the Transfer Assessment Instrument (TAI), a reliable and validated scale that measures the ergonomic and biomechanical quality of SPTs [4]. The TAI was developed to facilitate proper transfer technique that reduces the risk of upper extremity pain; however, because the body of knowledge on proper transfer technique has evolved so quickly, many practicing therapists have not adopted using the TAI for SPT assessment. While the TAI is a useful tool to identify transfer deficits, it is not currently optimized for use in the clinic and requires familiarity with TAI principles and a subjective analysis of biomechanics. To address these shortcomings, we developed the TransKinect, a software application utilizing a Microsoft Kinect sensor and machine learning algorithms to track the kinematics of a wheelchair transfer, identify deficiencies in the transfer, and provide relevant, clinical feedback to the clinician and the wheelchair user. The goal of this study was to test the usability and satisfaction of the TransKinect application with therapists who were not involved in the development process. We hypothesized that the results from the System Usability Scale [5] (SUS) and the Questionnaire for User Interface Satisfaction (QUIS) [6] would show high usability (>68%) and high satisfaction (>70%), respectively. The results of this study will be used to further iterate on the development of the application and improve its use in a clinical setting.

METHODS

This project was approved by the Department of Veterans Affairs Pittsburgh HealthCare System (VAPHS) IRB and R&D committee and each participant was consented prior to completing any part of the study. Subjects were included in the study if they were a currently a licensed physical or occupational therapist, currently employed in direct patient care, and have had a minimum of one-year direct patient care experience. Recruitment was accomplished through the VAPHS and University of Pittsburgh Medical Center. The therapists were asked to perform four different modules related to the usage of the TransKinect application: (1) Setup of Equipment and Evaluation area, (2) Patient Intake and Preparation, (3) Transfer Assessment and Evaluation, and (4) Patient Reevaluation and Report Comparison. Screenshots of the application are shown in Figure 1. Preceding each module, the participant was shown an instructional video guide that described the steps required for that specific

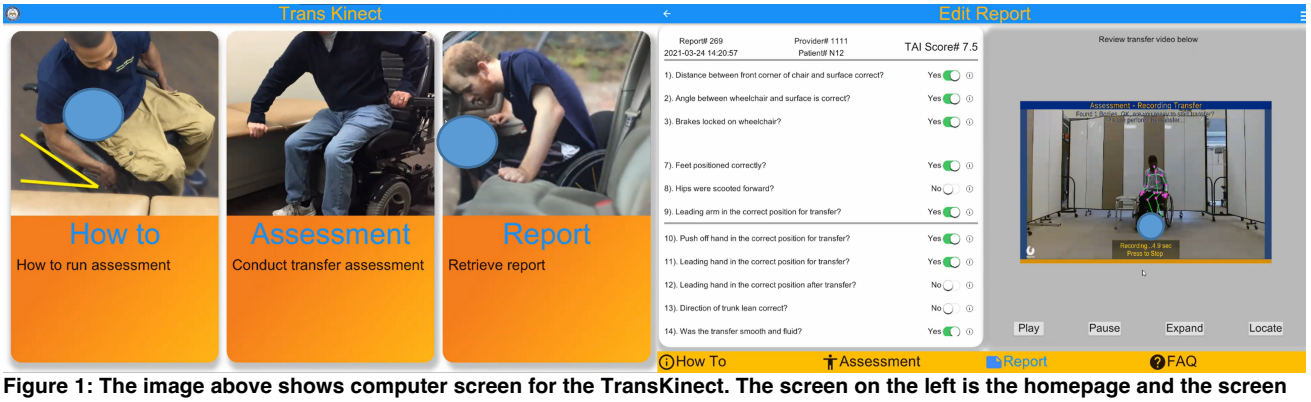

Figure 1: The image above shows computer screen for the TransKinect. The screen on the left is the homepage and the screen on the right shows the Edit Report page.

module. Furthermore, a one-page reference sheet was provided that listed the module tasks and briefly how to perform each one. Module 1 consisted of four tasks that 1) asked the participant to set up the Kinect sensor at the proper height and 2) distance from the transfer, 3) connect the Kinect sensor cables to the laptop and power supply, and 4) ensure that the evaluation area was clear of any obstacles. Module 2 involved 5 tasks: 1) logging in with provided user ID and password, 2) greeting a pseudo-patient (e.g. study team member who simulated a wheelchair user with transfer deficits), 3) explaining TransKinect to them, 4) collecting and entering patient demographic information (age, race, gender, type of disability, type of footrests/armrests used (removable or fixed) and transfer board usage), and 5) instructing the pseudo-patient to prepare for a transfer. Module 3 involved the actual transfer assessment and had a total of twelve tasks: 1) ensuring body tracking is in progress and clicking the Record button, 2) asking the pseudo-patient to transfer, 3) stopping the recording when done transferring, 4) answering questions regarding the setup of the transfer, 5) verifying the decisions made by the machine learning algorithms, 6) viewing the recorded video, 7) adding comments for any items where they may have disagreed with the result and saving changes, 8) viewing the generated TAI report, 9) explaining the results to the patient, 10) providing the feedback on how to improve deficit items, 11) using recorded video to show the patient where improper technique was used, and 12) exporting the generated report to a PDF to print or email. Module 4 required the participant to repeat the Transkinect assessment with the pseudo-patient assuming that it had been two weeks since the first assessment giving the patient some time to work on fixing his/her deficits. The eight tasks for this module included: 1) starting a new assessment, 2) reloading the patient demographic information from the database, 3) conducting a new assessment, 4) evaluating the results from the machine learning, 5) viewing the new report, 6) comparing it to the first report in a side-by-side comparison, 7) determine worsened and improved factors from the first session, and 8) filtering through generated reports using specific factors (e.g. date of original assessment, patient ID, etc.). The time that each participant took to perform each module was recorded as well as the number of times they asked for or required assistance in completing a task.

For each task within each module, the participant scored a one (1) if they completed the task independently and successfully. If they completed the task but required some assistance from the investigator to do so (i.e. asking for help or being told to do a task they missed) they scored half a point (0.5). The participant scored a zero (0) if they failed to complete the task. The investigator only prompted the participant on missed tasks that were vital to the performance and progress of the TransKinect application. The score for each task was summed and then divided by the total tasks for each module and converted into a percentage. The average score over all the participants was determined for each module, and the total average score across all modules was also calculated. Additionally, the amount of time taken for each module and the number of times the participant asked for or required assistance were also averaged for the group. After completing all tasks, the System Usability Scale (SUS), Questionnaire for User Interface Satisfaction (QUIS), and a custom survey regarding the use and function of the TransKinect application (General Questionnaire) were completed. The SUS is a validated scale to assess the usability of applications, software, or other devices [5]. A slightly modified version of the SUS with eleven questions instead of ten was used in this study. The additional question asked whether the participant would prefer the TransKinect to a paper version of the TAI. The maximum score possible for the SUS is 110, indicating perfect usability; the sum of the eleven items divided by the max score gives us a measure of usability. A total score value of 68% is considered acceptable usability; higher percentages signify higher usability. Sub-scores for individual questions were also converted to percentage of usability for reporting purposes. The QUIS is another validated tool to determine user opinions and feedback on the design and implementation of graphical user interfaces [6]. The QUIS has 27 questions on a Likert scale from 0-9; the sum of all the questions is divided by the total possible score of 243 to obtain a percentage of satisfaction. A cutoff of 70% was used to determine acceptable satisfaction, with higher percentages indicating greater satisfaction. The General Questionnaire included open ended questions about their likes, dislikes, and other thoughts on using the application.

RESULTS

Fifteen occupational and physical therapists (age = 38.3 ± 11.6 years) participated in this study. More women (11) participated than men and there was an even split among occupation (7 Occupational and 8 Physical Therapists). Participants reported an average of 14 ± 11.7 years of primary patient contact work experience (range 2 – 41 years). Four therapists reported familiarity with using the paper version of the TAI. Completion of all four modules took on average 29 minutes and 22 seconds with an accuracy score of 86.7%. Subjects spent the shortest amount of time (3:47) on the Setup module and achieved the highest average accuracy (96.7%) among the four tasks. The Subject Prep module was an average of 4 minutes and 37 seconds with 86.7% average accuracy across the five tasks. The Assessment module took the longest amount of time (12:02) with 82.2% average accuracy across the twelve items. Participants also asked for help the most during this module (average of 2.2

times). The last module, Reevaluation, took 8 minutes and 56 seconds and scored the lowest accuracy (81.3%) across the eight tasks. However, participants asked for help the least during this module (average 0.9 times). Additionally, across all the modules, the participants asked for assistance 1.46 times on average.

Table 1. Summary of QUIS Results: Quantitative (Average) & Qualitative

Category	Question	Score (STD)
Overall reactions to software	*Terrible/Wonderful*	7.3 (1.3)
	Difficult/Easy	6.3 (2.2)
	Frustrating/Satisfying	5.3 (2.7)
	Inadequate/Adequate Power	6.5 (3.2)
	Dull/Stimulating	6.9 (2.3)
	Rigid/Flexible	6.1 (2.3)
Screens	Characters on computer screen are *Hard/Easy* to ready	8.4 (0.7)
	Highlighted areas on screen simplified task *Not At All/Very Much*	7.2 (1.7)
	Organization of information was *Confusing/Very Clear*	7.7 (0.8)
	Sequence of screens was *Confusing/Very Clear*	7.7 (1.0)
Terminology and System Information	Use of terms in system is *Inconsistent/Consistent*	8.5 (0.8)
	Computer terminology is *Never/Always* related to current task	7.9 (1.2)
	Position of messages on screen is *Inconsistent/Consistent*	8.3 (1.0)
	Messages on screen for user prompt are *Confusing/Clear*	7.9 (1.2)
	Computer *Never/Always* keeps you informed of what it's doing	7.7 (1.1)
	Error messages are *Unhelpful/Helpful*	7.2 (2.0)
Learning	Learning to operate the system was *Difficult/Easy*	7.1 (1.0)
	Exploring new features by trial and error is *Difficult/Easy*	7.2 (1.1)
	Remembering names and use of commands is *Difficult/Easy*	7.4 (1.0)
	Tasks can *Never/Always* be performed in a straightforward manner	7.7 (0.7)
	Help messages on screen are *Unhelpful/Helpful*	7.8 (1.5)
	Supplemental reference material is *Confusing/Clear*	7.9 (1.0)
System Capabilities	*Slow/Fast*	7.6 (1.6)
	Unreliable/Reliable	6.5 (2.0)
	Noisy/Quiet	8.4 (0.9)
	Correcting mistakes is *Difficult/Easy*	7.8 (1.20
Universal Design	All levels of user experiences are *Never/Always* taken into consideration	7.6 (1.0)
Total Score		**200**
Total Percentage		**82.3%**

The SUS results showed an overall usability of 86.2%. Highest scoring statements included: the system was easy to learn quickly (87.2%), favoring the computer system over a paper version of the TAI (87.6%) for positive questions; and finding the system not overly complex (66.6%) and not cumbersome (65.8%) for negative questions. Lowest (positive) scoring items included: feeling confident in using the system (72.8%) and would use the system frequently (77.4%). Lowest (negative) scoring items included experience required for learning curve (60.0%), and amount of guidance and support required for use (61.4%). Overall, the QUIS results were positive (Table 1); with a total average score of 200 out of 243 (82.3%) in user satisfaction. The lowest scoring QUIS items included finding the system frustrating, rigid, and unreliable. Overall positive remarks on the General Questionnaire included: high potential for use in clinical setting, ability to improve prolonged patient care over time, ability to rewatch transfer videos while scoring, clear and organized layout of GUI, detailed and accessible supplemental training material, and highly innovative technology. Limitations to the application were also noted: initial difficulty in using new technology but easy to overcome with practice, required intermediate technology familiarity, instructional videos were too fast, and some glitches and errors.

DISCUSSION

The findings from this study indicate that there is a high potential for the TransKinect application to be a useful and robust clinical tool for the evaluation of independent SPTs performed by wheelchair users. The average scores for the SUS and QUIS were above the thresholds, indicating that the TransKinect application has high usability and satisfaction. The results from this study while overall positive also provided some areas that need to be further addressed to help improve the usability and potential

application in clinical settings. The accuracy scores show high success rate of completing all targeted tasks. The lowest scoring module was Reevaluation (81.3%); however, the most commonly missed tasks in this module were loading the saved patient data from the database, comparing two reports, and using filters to find other TAI reports. It's possible that because this was the last module that needed to be completed that participants were exhibiting some mental fatigue. The Assessment module (accuracy of 82.2%) was objectively the most complex module with 12 tasks to complete. The most commonly missed tasks in this module were: not asking the participant to move around to generate a skeleton, providing comments for changed TAI items (the first time), and exporting the report to a PDF. The missed task about waiting for the skeleton to appear is vital for the performance of the system; therefore, midway into the study a feature was added to prevent the system from recording video if a body was not detected. In the Subject Prep module (accuracy of 86.7%) the most frequently missed task was greeting and explaining the procedure to the pseudo-patient. As this was a simulated clinical setting without 'real' patients it is not surprising that therapists may have missed this step assuming that the pseudo-patient knows what is going on already. The Setup module scored the highest accuracy (96.7%). The most frequently missed tasks in this module were positioning the Kinect sensor correctly according to the measurements provided. From the missed tasks from each module, we can see that the majority could be remedied by more training and experience with using the system. The SUS results showed that some participants did not feel very confident in using the system, felt it required a large learning curve and required more support and guidance. These findings could be attributed to most of the participants having no experience in using the TAI before; however, using the TransKinect will replace the need to know the paper version of the TAI. Some of the participants who scored lower accuracy scores on the evaluation and also rated SUS and QUIS low remarked that they were not very "tech savvy". This indicates that technology usage may be a barrier to the usability and acceptability of the system. We feel that with further training and experience with the system that the usability of the system could improve by improving confidence and decreasing the need of support.

There are a few limitations to the study that are important to discuss. First, an able-bodied study team member served as a pseudo-patient. Using real wheelchair users-patients who were intimately known by the therapists may have impacted how they used Transkinect. Another limitation of this study was that minor changes to the system were made during the course of this study to improve the usability as feedback was received. This included adding a pop-up to indicate that the data were being analyzed (which takes around 30 seconds to complete), adding colored buttons to indicate which items were incorrect (red) or correct (green), and adding info boxes with educational material on each TAI item. These changes appeared to enhance the usability of the system for subsequent users. There were also several errors and glitches that occurred when using the application, some of which were repeatable and could be fixed by implementing a 'patch' and some that were not. The results of these errors (which are somewhat common with any software application) included having to restart the application, back up to a previous page or step, or wait while the system tried to resolve the issue. Part of the training moving forward should include how to address these errors if/when they occur.

FUTURE DIRECTIONS/CONCLUSIONS

Overall, the TransKinect system has been well received and accepted by the end-user clinicians; the design and set up is easy to use, simple and straight forward. Because of the participants' indication for more guidance/support to use the system, we aim to add a practice module and produce revised video content in order to facilitate learning and acclimation. As the Kinect 2 sensor has been phased out, we are working to integrate newer versions of the hardware (the Kinect Azure) with the hopes of further improving accuracy. Furthermore, we plan to enhance the display settings so Transkinect can be used with different devices (e.g., laptop, desktop, tablet). Finally, we have partnered with the Milwaukee VA SCI Center to assess TransKinect's usability and utility in a 'real-world' clinical setting.

FUNDING

Funding support was provided by the VA Rehabilitation R&D Service, Project #A2794-R

REFERENCES

[1] G. Barbareschi and C. Holloway, "An investigation of factors affecting the performance of wheelchair transfers," (in eng), no. 1748-3115 (Electronic).

[2] J. V. Subbarao, J. Klopfstein, and R. Turpin, "Prevalence and Impact of Wrist and Shoulder Pain in Patients with Spinal Cord Injury," *The Journal of Spinal Cord Medicine,* vol. 18, no. 1, pp. 9-13, 1995.

[3] P. Requejo, S. Mulroy, L. L. Haubert, C. Newsam, J. Gronley, and J. Perry, "Evidence-Based Strategies to Preserve Shoulder Function in Manual Wheelchair Users with Spinal Cord Injury," *Topics in Spinal Cord Injury Rehabilitation,* vol. 13, no. 4, pp. 86-119, 2008.

[4] L. A. McClure, H. Boninger Ml Fau - Ozawa, A. Ozawa H Fau - Koontz, and A. Koontz, "Reliability and validity analysis of the transfer assessment instrument," (in eng), no. 1532-821X (Electronic).

[5] J. Brooke, *"SUS-A quick and dirty usability scale." Usability evaluation in industry*. CRC Press, 1996.

[6] N. Stanton, P. M. Salmon, and L. A. Rafferty, *Human Factors Methods: A Practical Guide for Engineering and Design*. Ashgate Publishing Limited, 2013.