Daily Task-Oriented Performance Evaluation for Commercially Available Assistive Robotic Manipulators

Cheng-Shiu Chung[1,2], Hongwu Wang[1,2], Matthew J. Hannan[1,3], Annmarie R. Kelleher[1,2] and Rory A. Cooper[1,2,*]

[1]Human Engineering Research Laboratories, Department of Veterans Affairs, Pittsburgh PA

[2]Department of Rehabilitation Science and Technology, University of Pittsburgh, PA

[3]College of General Studies, University of Pittsburgh, Pittsburgh, PA

Abstract: This preliminary study investigates the performance and cognitive loading of the two commercial wheelchair-mounted assistive robotic manipulators (ARMs) with their original user interfaces (UIs). This study of 20 able-bodied individuals evaluated the performance of two user interfaces, keypad and joystick, using six tasks on an activities of daily living (ADL) task board with environment-independent measures, self-reported cognitive loading and questionnaires. Participants performed tasks with two commercial arms with their original UIs in a randomized order of arm and the six tasks on the adl task board. Performance was evaluated using completion time, throughput, and trajectory parameters. Self-reported measures of workload and questionnaires were also administered. Statistical performance differences were found in the translational tasks ($p<0.05$) in task completion time, throughput, and difficulty. The keypad showed faster performance on the knob turning task. Similar responses were reported in the perceived workload with both UI. Participants rated the UI's low on frustration and physical workload, but higher on mental effort. The findings of this study provide a preliminary comparison between two commercial ARMs with their original UIs. Barriers and recommendations for training and evaluation for first time users were discovered. The results provide information to help develop ARM UI and recommendations for clinicians and health service providers to develop better training and evaluation for arm users.

Keywords: Robotic manipulators, wheelchairs, performance evaluation, assistive technology.

1. INTRODUCTION

In the United States, about 19.9 million people age of 15 years and older have difficulties with physical tasks relating to upper extremity functioning, including lifting, grasping, pushing/pulling, reaching, dressing and eating [1]. In addition, it is estimated that older adults with moderate to severe disabilities will increase from 10 million in 2000 to 24.6 million people in 2040 [2]. However, the increasing need for ADL assistance may not be entirely supported by the current care giving system [3]. While a caregiver is not on site, assistive robotic manipulators (ARMs) may enhance assistance and increase independence of people with impairments in completing activities of daily living (ADL) [4-6]. It is estimated that about 150,000 people can benefit from using ARMs in the United States [7].

Two studies were conducted to evaluate the improvement of the kinematic capability in performing ADL with developing and commercial ARMs [8, 9] in order to match user's needs [10]. The cost-effectiveness and long-term usage of the ARMs was also evaluated [11-13]. A review article [5] evaluated ARM studies from 1970 to 2012 with their UIs and measurement tools for demonstrating clinical significance. In addition, this review article includes International Classification of Functioning (ICF) as a criterion for functional assessment between different study results. As indicated in this article, although there were significant improvements in task completion time and success rate using a single user interface, however, due to a lack of standard evaluation tools and measurements, there was no way to systematically integrate or compare the results from different UIs and manipulators. In addition, although task completion time is a widely utilized outcome measure for ARM performance, the distance, speed, initial position, or target size and shape may vary under different environmental settings or tasks among studies. Moreover, the pick-and-place task results may be biased by the shape and size of the target object. Consequently, the environmental and grasping variance would change task completion time. Thus, in order to standardize ARM performance evaluation and compare ARM user interfaces with minimal variation due to task objects and setup, Chung *et al*. [14] developed an ADL task board and tested the reliability of a task-invariant comparable indicator, throughput (TP) from the international standard testing of physical input deveices (ISO 9241-9) [15] This ADL task board has specific start and finish locations and restricted target orientations so that the variation in task completion can be minimized. In addition, no study has been published that assesses user performance with TP in conjunction with perceived workload. User

*Address correspondence to this author at the 6425 Penn Avenue, Suite 400, Pittsburgh, PA 15206; Tel: 412-822-3700; Fax: 412-822-3699;
E-mail: rcooper@pitt.edu

E-ISSN: 2409-9694/16

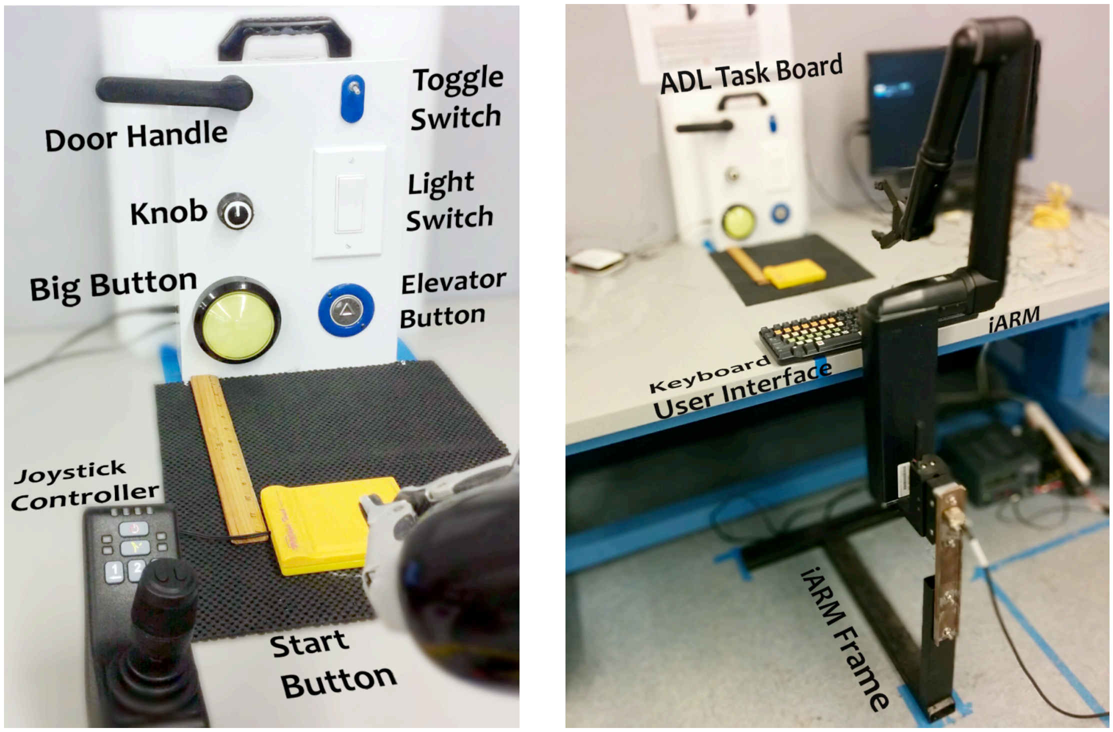

Figure 1: ADL task board with the JACO arm (left) and the mounting location of the iARM (right).

opinion and workload between user interfaces have not been considered in previous studies.

Therefore, we conducted a study using a standardized ADL task board [14] to evaluate two commercially available ARMs and their original UIs. The purpose of this study was to undertake an initial investigation of ARM performance with their original UIs. In this paper, we are trying to answer the following research questions:

- Are there differences in ARM performance, throughput of the user interfaces, and trajectory parameters between two user interfaces in completing the same ADL task?
- Is there any difference between two user interfaces in the self-reported workload and questionnaire?

In this paper, we will introduce the study design and environment-independent outcome measurements that can describe the differences between user interfaces and quantify performance. We investigated differences in trajectory characteristics between UIs. The results provide researchers and clinicians fundamental data on the ARM performance with their original UI. The results may help clinicians and service providers to develop appropriate training for ARM users.

2. METHOD

2.1. Testing Assistive Robotic Manipulators

This study focused on two commercially available ARMs that can be mounted on a wheelchair: iARM by Exact Dynamics (Netherland) and the JACO robotic manipulator by Kinova (Canada). The iARM is a six DOF robotic arm with a two-fingered hand. Its original user interface (UI) is a 4x4 keypad to control the iARM moving direction and speed incrementally [16]. The JACO manipulator is composed of six interlinked segments with a three-fingered hand. The hand can grasp objects using either two or three fingers. Its original UI is a 3-axis joystick which controls translational, rotational, and finger movement with two mode-switching buttons on top of the joystick knob [17, 18]. Other than these two ARMs, a review article [19] compared 19 commercial and developing ARM with five criteria: interaction safety, shock robustness, adaptability, energy, and position control. These robotic manipulators were compared by their functionalities and specifications. However, their costs, UIs, and clinical evidences to benefit people with disabilities were not taken into consideration.

2.2. ADL Task Board Performance Evaluation Tool

The ADL task board system used for performance evaluation consists of six daily electronic components.

These six components were selected from commonly performed ADL tasks, including one large size circular button similar to a door opener (Big Button), one small size circular elevator button (Elev. Button), one rectangular shape rocker light switch (Light Switch), one toggle switch (Toggle Switch), one door handle (Door Handle), and one knob (Knob). These components simulate common home and community activities such as turning door handles, turning knobs on the oven, using the elevator or door opener, and turning on or off light switches. The door handle and knob are attached to potentiometers to measure their angular changes. Aside from these six components on the task board, a square sized button is located one foot (twelve inches) in front of the task board as the test starting location where both ARMs can reach to [14]. The task starts from lifting the robotic manipulator from the start button and ends by activating the target button or rotating to the target angle. Figure **1** left shows a photograph of the ADL task board system. These ADL tasks can be divided by the motion change between approach and activation. Activating direction on the tasks Big Button, Elev. Button, and Light Switch are similar to the direction of the approaching motion. The other three tasks, Toggle Switch, Door Handle, and Knob, require a secondary movement to either move or twist to test the performance.

2.3. User Interfaces and Recording System

Two ARMs with their original user interfaces were utilized in this study because these are the most common user interfaces in previous studies [17, 18, 20, 21]. The JACO utilizes a 3-axis joystick interface (Figure **2** right) with a different variety of mode selection buttons, while the iARM uses a 4x4 keypad to control the device. Both ARMs are located in front of the ADL task board center within a pre-defined location to ensure the consistency between participants. The iARM was mounted on an L-shaped frame in a fixed location on the ground (Figure **1** rigtht). The JACO is mounted on a standard office table (30-inch-wide) edge fixed with two C-clamps (Figure **1** left). The ADL task board was located at the opposite edge. Both ARMs are pre-set as right-handed. The participants sit on the left side of ARMs with a comfortable distance. The UIs were placed either on the table or lap based on each participant's preference. The ARMs were connected to a data acquisition computer to record the joint and hand positions. In order to control and record the iARM under transparent mode through the CAN-Bus connection, which enables the communication from the iARM to the data acquisition program, we replaced the original 4x4 keypad with a similar color labeled regular sized keyboard (Figure **2** left). The original keypad was designed to control the iARM with two control modes: joint and Cartesian. We mapped the function keys of these two control modes to different key locations on a standard keyboard allowing users to manipulate the iARM with their preferred control mode. Although each ARM can be replaced with other types of UIs and fine-tuned control parameters for each user, we tested the original UI and factory default settings in order to maximize the consistency between participants and minimize modifications to the system. The in-house developed robot data logger was used to continuously (sampling rate: 20Hz) record the internal data from the ARMs, including joint angles, gripper Cartesian position and rotation, finger position, torques, and forces when provided.

2.4. Participants and Study Protocol

The participants were recruited from the University of Pittsburgh. In order to minimize differences due to degrees of impairment and experience with ARMs, we used only able-bodied participants in this study as our goal was to evaluate the tool. Able-bodied individuals were evaluated to provide a homogeneous sample without complex impairments. This study needs to be

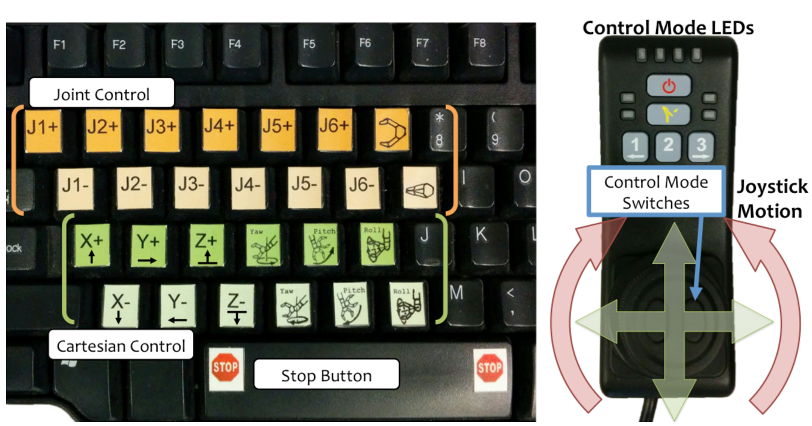

Figure 2: Keypad (left) and 3-axis joystick (right) user interface with the training instructions.

expanded for open in the future with a larger sample of intended users. Once the testing system is validated as safe, reliable, and effective, powered wheelchair users will be recruited for a future study. All the able-bodied participants were 18 years and older with no hand or wrist pain. The study was conducted in a controlled environment at the Human Engineering Research Laboratories (HERL). The study protocol was approved by the Institutional Review Board (IRB) of the University of Pittsburgh. After informed consent was obtained, it was verified that each participant met the inclusion and exclusion criteria. General demographic (i.e. age, gender, ethnicity) and educational background were then recorded. After having finished the demographic information, participants were randomly assigned and introduced to the first UI. This introduction lasted approximately 30 minutes and included a demonstration and hands on practice period with each ARM. A questionnaire regarding perspective of the ARM UI was collected before hands on practice and testing. Each paritcipant was asked to complete up to 6 tasks on the ADL task board 3 times each with both UI for a maximum total of 36 trials. The order of the UI used and tasks completed were randomized. Each task typically took less than 2 minutes. However, if the time exceeded 5 minutes or the participant expressed frustration, the task was terminated. It was considered a protocol deviation if a participant was unable to complete all tasks within the given time frame. Time to complete each task and the trajectory of the robot were recorded during testing. Participants were asked to complete the NASA TLX [22] and a questionnaire regarding the UI after the completion of each series of tasks using one UI. Following the completion of all 6 tasks with both UIs, a brief questionnaire and open interview related to use of the interfaces were then conducted.

2.5. Outcome Measurements

Throughput (TP) is defined by the international standard requirements and evaluation of physical input devices (ISO 9241-9) [23] as a parameter for user interface evaluation. It is based on the Fitts' law, the correlation between task difficulty and task completion time. Fitts' law is widely used to model the performance of human-computer interaction in rapid movement as a function of distance to the target and the size of the target. This function generalizes the relationship between the task difficulty and target distance and size. The index of difficulty is independent of the environmental setting and is easily comparable between UIs. The TP for the user interface in bits per second (bits/s) is defined by;

$$Throughput = ID/MT$$

MT is the movement time or the task completion time in seconds and *ID* is the index of difficulty in bits. For all trials within the same condition, and, different from the *ID* in Fitts' Law, the *ID* in ISO 9241-9 is defined as [15];

$$ID = \log_2\left(\frac{D}{W_e} + 1\right) \quad where\ W_e = 4.133 \times SD ,$$

We computed the distance to the target, *D*, and target effective width, W_e, using the Cartesian points from the recorded trajectory of the ARMs. W_e is defined as the effective width computed from the standard deviation *SD* of task endpoints. *SD* is the standard deviation of the target endpoints, which is modified for 3D tasks as [24];

$$SD = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(x_i - \bar{x}\right)^2 + \left(y_i - \bar{y}\right)^2 + \left(z_i - \bar{z}\right)^2\right]}{n-1}}$$

For tasks such as a toggle switch, door handle, and turning knob that require a secondary movement to complete the task, the throughput could not be calculated because Fitts' law cannot be expanded to movements with multiple steps.

Self-reported measures included the NASA-TLX and a questionnaire about user impressions and opinions of the ARM UIs. The NASA-TLX is a subjective multidimensional assessment tool that evaluates the perceived workload on six sub-scales: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Weighted workload was calculated through a pairwise comparison process that allowed each user to refine their previous answers. The questionnaire has 14 Linkert items from 0 to 10 levels (Table **3** and **4**). A higher score means stronger agreement. All subjects completed the first nine items immediately after being introduced to a UI. These nine items assess the ease, attractiveness, and capabilities of the UI. After hands on practice and completion of the six tasks using the UI, participants then completed the first nine items again and an additional five items. The five additional items assessed perspective on using the ARMs.

2.6. Trajectory Analysis

In order to describe the characteristics of the ARM movement and user's familiarity with the UI, several trajectory parameters were computed: pause percentage, number of pauses, average pause time, roughness, and average speed. The recorded trajectories were analyzed using Matlab R2014a (Mathworks).

The parameters related to pauses help to discover the continuity and potential difficulties that the user encountered such as re-planning for error correction, fine movement, verification of end-effector position, searching another key on the keypad UI, and switching modes on the joystick UI. An increasing number of pauses may be caused by more errors or more keys or modes switching in the trial. A pause was defined as the time without movement in the recorded trajectory. The number of pauses was the number of stops that occurred within one trial. Pause percentage indicates

Table 1: ADL Task Board Testing Results of Two User Interfaces (Mean ± Standard Deviation)

Parameter	Task	Keypad	Joystick	P-value
Average Completion Time (Second)	Big button	24.1±13.9	5.8±2.5	<0.001*
	Elevator button	22.1±8.3	15.5±8.3	0.021
	Light switch	29.9±16.6	12.9±7.1	<0.001*
	Toggle switch	47.5±20.3	15.4±7.4	<0.001*
	Door handle	47.9±23.0	29.9±31.6	0.008*
	Turning knob	48.6±22.5	52.8±35.0	0.640
The Fastest Trial Completion Time (Second)	Big button	14.4±6.5	4.1±1.7	<0.001*
	Elevator button	16.0±5.0	10.1±6.7	0.004*
	Light switch	20.6±10.9	8.5±5.4	<0.001*
	Toggle switch	32.8±11.3	9.1±4.0	<0.001*
	Door handle	31.6±17.0	20.7±25.7	0.017
	Turning knob	37.7±20.0	32.0±18.5	0.253
Throughput (bit/Second)	Big button	0.091±0.046	0.362±0.153	<0.001*
	Elevator button	0.081±0.257	0.132±0.089	0.021
	Light switch	0.097±0.050	0.173±0.098	0.002*
Roughness (mm)	Big button	49.2±11.8	30.4±12.9	<0.001*
	Elevator button	35.8±9.5	42.9±16.0	0.010*
	Light switch	58.7±21.6	50.6±17.9	0.061
Average Speed (mm/Second)	Big button	63.0±23.6	62.1±20.1	0.813
	Elevator button	53.7±16.1	50.0±20.8	0.177
	Light switch	69.5±27.0	63.8±22.7	0.203
	Toggle switch	53.1±17.3	61.2±17.0	0.009*
	Door handle	69.9±29.0	61.7±30.1	0.119
	Turning knob	40.4±13.7	27.0±14.2	<0.001*
Pause Percentage (%)	Big button	16.8±8.4	12.0±10.8	0.009*
	Elevator button	15.0±7.0	20.6±12.7	0.009*
	Light switch	15.8±8.4	16.1±14.3	0.901
	Toggle switch	25.6±8.9	18.8±11.5	<0.001*
	Door handle	21.6±9.7	25.1±15.9	0.155
	Turning knob	22.9±8.5	42.3±13.4	<0.001*
Number of Pauses	Big button	5.5±4.7	2.7±2.3	0.003*
	Elevator button	5.0±2.5	9.8±9.0	<0.001*
	Light switch	5.9±4.2	6.0±5.3	0.867
	Toggle switch	11.8±7.3	8.4±6.9	0.015*
	Door handle	9.2±5.6	12.0±11.8	0.105
	Turning knob	10.2±4.4	23.4±13.5	<0.001*
Average Pause Time (Second)	Big button	0.333±0.214	0.297±0.246	0.418
	Elevator button	0.286±0.135	0.435±0.338	0.006*
	Light switch	0.322±0.187	0.388±0.408	0.246
	Toggle switch	0.473±0.180	0.409±0.312	0.146
	Door handle	0.483±0.236	0.657±0.434	0.011*
	Turning knob	0.504±0.296	0.933±0.446	<0.001*

*Significant after Bonferonni adjustment.

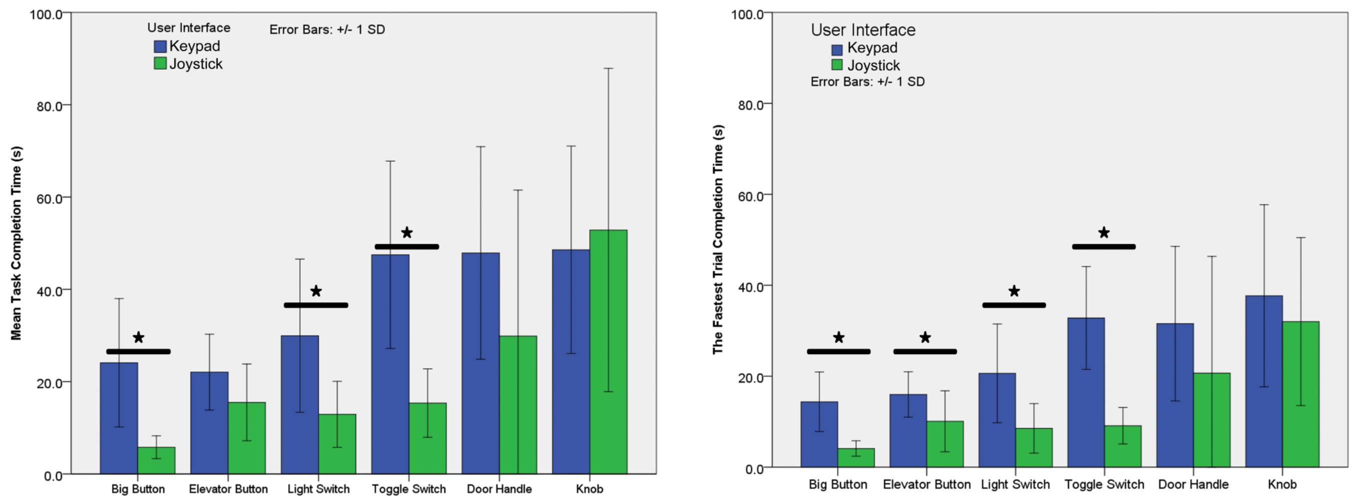

Figure 3: Left: the mean task completion time of the six tasks on the ADL task board; Right: the minimum task completion time of each participant on the ADL task board.

the time when the ARM is not moving in a trial, which is computed as the total pause time over an entire task completion time. The average pause time is computed as:

Average Pause Time = Total pause time / number of pauses

Pauses may reflect the user's decision processing time while completing a task.

The parameter, roughness, is to quantify the maneuvering while in approaching target. It is described as the trajectory's variance compared to a straight line from start to end of a task, which is computed as;

$$Roughness = \frac{\sum_{i=1}^{n} dist(p_i, l)}{n}$$

The p_i is the position of the i point on the trajectory. The n is the total number of trajectory points. The dist(p_i, l) is the distance from the point p_i to a straight line from the start to the end point. Roughness is used to analyze tasks without secondary movement, such as Big Button, Elev. Button, and Light Switch.

2.7. Statistic Methods

A Bonferroni adjusted paired-sample t-test was conducted to compare the difference between the UIs with respect to the task completion time, number of errors, ISO 9241-9 throughput, trajectory parameters, NASA-TLX, and questionnaire items. All data were examined for normality by Shapiro-Wilk test and checked for outliers using Q-Q plots. If the continuous data were not normally distributed, then non-parametric Wilcoxon Signed-rank Test was applied. Bonferroni adjustments were applied to the alpha levels to control for inflation of error associated with multiple comparison. All statistical analyses were performed in SPSS (SPSS Inc.8, Chicago, IL).

3. RESULTS

Twenty able-bodied individuals (mean age: 26.7 years old, range: 18-35 years old, 14 males) were enrolled in the study and were tested with two ARM UIs. Two participants had missing trajectory data because of technical difficulties. There were no safety issues exposed with the trajectory recording software.

3.1. ADL Task Board Results

The results of the task board measurements are listed in Table **1**. The joystick UI was statistically faster than the keypad UI in four tasks except for elevator button and knob turning. The joystick is slower than the keypad in the average completion time in the knob turning task (Figure **3**). By only comparing the fastest trial (Figure **3**) of each participant in each task, the joystick UI was statistically faster than keypad UI for the same five tasks. Conversely, the average of the fastest trial task completion time with the joystick in Knob was slightly faster than the keypad. On average, about 30-40% of completion time was reduced in the fastest trial of each task. Even though we provided time to practice with the ADL task board, we still found a learning effect with improved performance or small variations among the three trials on the each task (Figure **4**). The average speed of both UIs was around 50-70mm/Sec. The keypad was statistically faster with the Knob ($p<0.001$) but statistically slower with the Toggle Switch ($p=0.009$).

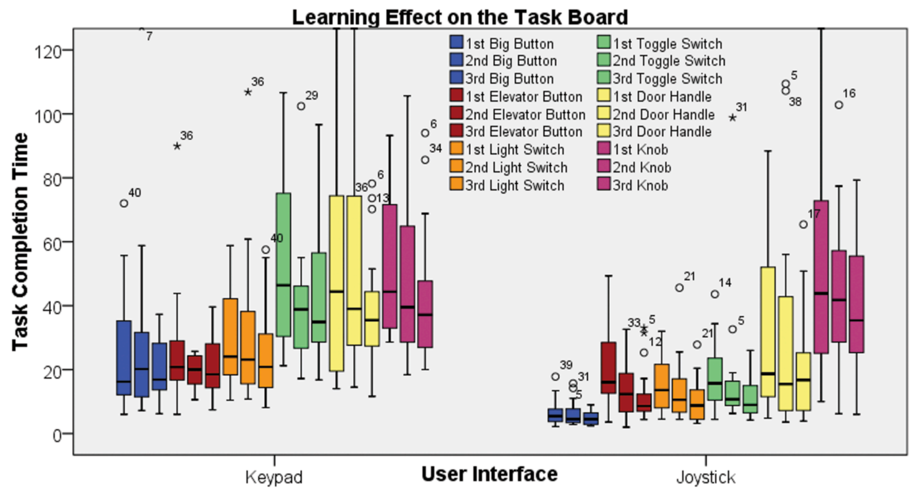

Figure 4: Learning effect of the three trials on the ADL task board.

For the single motion tasks such as Big Button, Elev. Button, and Light Switch, we compared these tasks using two characteristics: TP, and roughness. In these tasks, the joystick UI showed statistically faster motion. Similarly, its TP is statistically higher than the keypad UI (Table **1**). However, the TP of the JACO manipulator is lower than previously reported results [14] with one experienced ARM user.

3.2. Trajectory Characteristics

The joystick UI showed statistically smaller roughness in the Big Button and Elev. Button tasks (Table **1**). The keypad UI has statistically lower roughness in the elevator button task. The roughness of joystick is statistically lower in the Big Button, but it is statistically larger in the Elev. Button.

The pause percentage shows that overall the users stopped from 12-42% of the time during task performance (Table **1**). The pause percentage and number of pauses using the keypad was statistically lower in the Elev. Button and Knob and statistically larger in the Big Button and Toggle Switch. In the comparison of the average pause time, the keypad was under 0.5 seconds except the turning knob (0.504 sec) and the joystick is 0.3-0.9 second in all tasks. The keypad was statistically lower in the Elev. Button, Door Handle, and Knob. There were no statistical differences found in the other three tasks. In the comparison to the number of pauses, there were statistically more frequent pauses found in the Elev. Button, and Knob using the joystick UI, but statistically less frequent pauses in the Big Button and Toggle Switch.

In order to further examine the motion differences between the two UIs, we extracted the fastest 10 trails in the Big Button and plotted them in Figure **5**. This is the easiest task on the ADL task board which helps to discover how users plan to hit a target. In the trajectory figure, the dark blue indicates the starting location and the red is the end point. The trajectories show that the

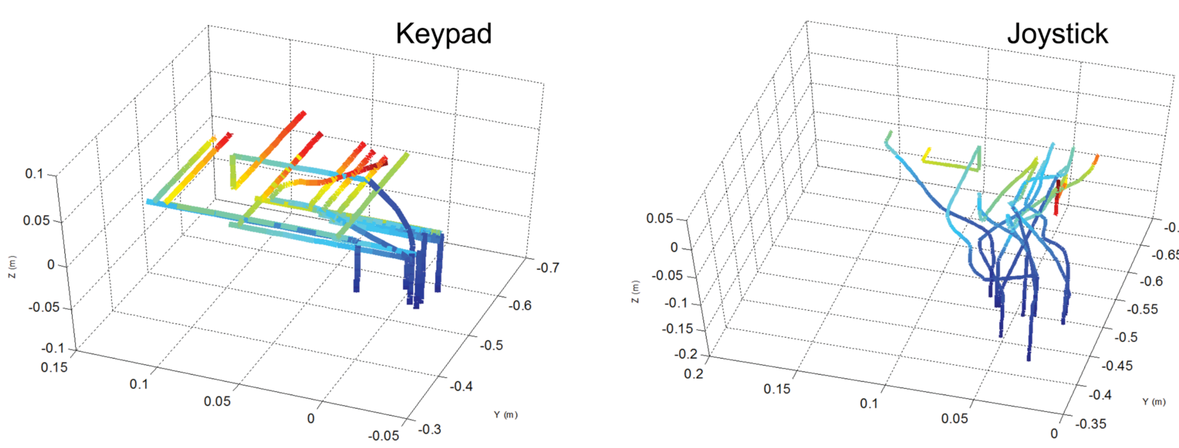

Figure 5: The ARM trajectory of the fastest 10 trials in the big button task using both UIs (left: keypad, right: joystick). Trajectories are colored from the dark blue (start position) to the red (finish position).

Table 2: NASA-TLX and Weighted Workload Index of Two User Interfaces (Mean ± Standard Deviation)

Perceived Loading		Keypad	Joystick	P-value
NASA-TLX		42.2±20.2	38.3±15.8	0.532
Weighted Workload	Mental	12.9±8.6	13.9±8.5	0.690
	Physical	1.5±2.2	2.1±2.6	0.460
	Temporal	9.0±7.8	6.7±6.4	0.325
	Perform	4.3±3.0	5.1±3.9	0.474
	Effort	9.3±6.9	9.3±7.1	0.981
	Frustration	5.2±10.4	1.3±2.3	0.123

ARM moves along one axis at a time using a keypad UI, but moves diagonally toward the target using a joystick UI.

3.3. Workload and Interview Questionnaires

Although there were statistical differences in the task completion time and throughput between keypad and joystick, we did not find statistical differences in the NASA-TLX and weighted workload, Table **2**. However, Figure **6** indicates that both UIs show similar patterns in the workload subscales. Almost all participants rated lowest on the frustration and physical and highest on the mental and effort workload. There were four users who rated high frustration in keypad UI and two on the joystick UI.

Table **3** shows the response of the UIs comparison before and after using the ARMs. We found statistically higher in the ease at learning, usage, and lower confusion in usage with the keypad UI (item 1, 2, and 4). However, participants felt statistically less embarssing (item 9) after using the joystick UI. Both UIs have similar ratings among these nine pre-post items. Participants reported being less anxious and embarrassed about using the UIs and ARMs.

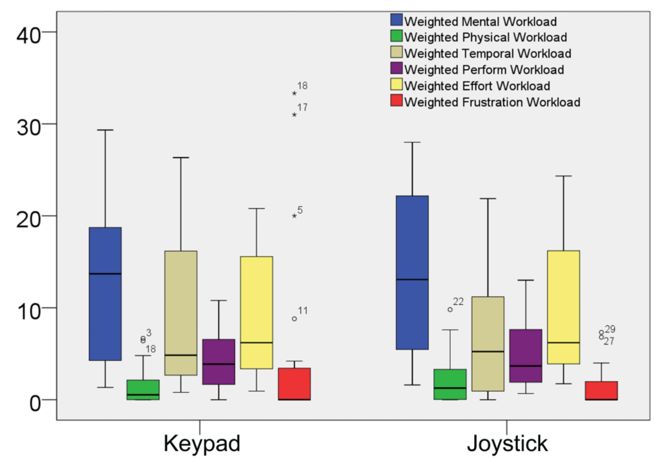

Figure 6: Weighted workload index of two user interfaces.

Table 3: Questionnaire Items Interviewed before and after the Practice and Testing of each User Interface (Mean ± Standard Deviation)

Interview Question	Keypad			Joystick		
	Pre	Post	P-value	Pre	Post	P-value
1. Learning to use ARM will be/was easy for me	4.85±2.72	7.00±2.47	0.003*	7.55±1.91	6.37±3.17	0.065
2. It will be/was easy to get ARM to do what I want it to do	5.55±2.74	6.95±2.44	0.072*	7.50±1.24	6.47±2.44	0.064
3. I am anxious about using ARM	5.05±2.84	4.40±3.00	0.382	4.75±2.79	4.60±3.03	0.836
4. It will be/is confusing for me to use ARM correctly	5.00±2.29	3.65±2.83	0.041*	4.75±2.55	5.05±3.46	0.744
5. Using an ARM will /would make my life easier	3.55±2.06	3.70±2.68	0.772	5.15±2.76	5.45±3.03	0.732
6. Using an ARM will/would help me to achieve important goals	3.8±2.38	3.75±2.65	0.924	4.55±2.69	4.85±3.35	0.690
7. It would be easier to just get another person to help rather than using ARM	5.55±2.86	4.90±3.14	0.222	5.25±2.86	4.32±3.11	0.245
8. ARM is attractive from a physical standpoint	5.30±2.01	4.80±1.99	0.248	6.35±2.96	5.90±3.02	0.529
9. It will/would be embarrassing to be seen using ARM	2.90±1.65	3.00±2.08	0.781	4.15±2.58	2.35±1.84	0.007*

Table 4: Questionnaire Items Interviewed after Completion of ADL Task Board Testing with each User Interface (Mean ± Standard Deviation)

Interview Question	Keypad	Joystick	P-value
10. The benefits ARM will provide are worth the cost of the device	5.45±2.28	4.80±2.17	0.213
11. ARM sometimes doesn't work properly	4.40±3.00	2.80±1.91	0.049*
12. ARM seems too flimsy, like it might break	3.60±2.33	3.90±2.55	0.700
13. If ARM needs repairs, I could probably fix it myself	2.80±2.38	3.20±2.95	0.631
14. ARM is just as good as newer things on the market	5.74±2.51	6.40±2.39	0.366

Table **4** shows the difference in interview items between the two UIs. There was no statistical difference found except three items (item 11). Statistally higher rating was found in the working inproperlly item (item 11).

4. DISCUSSION

This is the first study comparing commercially available ARMs and their original UIs with standardized ADL tasks. The study also investigated the performance evaluation and self-reported workload and impressions. One of the goals of this study was to evaluate the ARM performance on standardized ADL tasks with two UIs. The second goal of this study was to evaluate the user perceived workload and perspectives toward robotic manipulation assistance.

4.1. Difference between ARM User Interfaces

One of the differences between the two UIs is the incremental speed control in the keypad UI and the proportional control in the joystick UI. While using the keypad UI, some participants tried to keep pressing down the key to speed up the ARM motion, but sometimes overshot the target. Consequently, more time was spent in the re-planning from an overshoot error. In the joystick UI testing, participants accelerated when the ARM was moving in fixed space and smoothly slowed down when approaching the target to prevent overshoot. However, instead of using a proportional speed control on the joystick, it was noted that some participants used the joystick as bang-bang controller for small movements by quickly pushing the joystick knob to its extreme bound and then releasing it immediately. This observation from testing suggests that training on the familiarity with speed control on both UIs may improve the performance by reducing overshooting and producing accurate fine movements.

Switching between translation and rotation modes is another one of the differences between the two UIs. The keypad UI has all the translational, rotational, and grasping function keys within one keypad. The original 4x4 keypad provides joint and Cartesian control mode. Participants preferred to use Cartesian control mode most. The joystick UI, default setting is Cartesian motion, has to switch modes between the translation, rotation, and gripper modes. Participants spent more time in translation mode. It is worth noting that there were two different techniques used with the Door Handle and Knob. Most participants used wrist rotation mode in the Door Handle and Knob for both UIs, but we observed some participants completed these tasks with translation mode only. Door Handle can be completed either using only translation mode or with a combination of translation and rotation. During the testing, some participants switched between translation and rotation modes in the first or second trial and only used translational mode in the third trial to complete the task faster. Knob requires more accurate alignment of the knob rotation axis for better grasping and wrist spinning motion. The majority of participants first moved to the proximity of the knob, then rotate the wrist to align with the turning axis, moved forward and grasped the knob, and finally spun to the target angle. This complicated series of motion results in statistically higher pauses frequencies and times for switching modes and re-planning in alignment. However, we found a few participants successfully completed Knob by using only translational mode. The technique utilized sliding the ARM fingers on the edge of the knob a few times. Therefore, although the Door Handle and Knob are rotational motion tasks, performance may be improved by incorporating different techniques by reducing mode switching.

Affordability in managing multiple axis movement is the other difference. As shown in the trajectories of the best ten trials of the light switch task (Figure **5**), we can clearly see two types of maneuvering techniques: multiple axis and single axis. The multiple axis was in the joystick UI controls the ARM move directly toward the target. Conversely, the keypad UI moves the ARM one axis at a time. Although the keypad UI has the capability to maneuver the ARM with more than one axis simultaneously by pressing more than one key. We observed that some participants tried to apply this technique in completing an ADL task. However, it was difficult to accurately guide the ARM toward the target with incremental speed control; and consequently, the user had to stop and go back to single axis control. These findings suggest that better performance can be improved by shortening the movement trajectory. Conversely, the findings also suggest that the training in the keypad UI can start with single axis control.

4.2. Cognitive Loading and Users' Feedback

Even though performance was different between both UIs, there was no statistical difference found in the cognitive loading and user's impression. Although joystick UI shows lower average NASA-TLX score, the variation between participants did not yield a statistical difference. It is worth noting that the weighted workload shows that participants felt low frustration and physical loading in both UIs, but struggled most with mental effort workloads. These results reveal that while maneuvering the ARM, it requires a significant amount of loading in calculating, planning, looking, and searching. However, low frustration and physical loading indicate that it was not physically difficult to maneuver the ARM to the target. These results suggest that the difficulty in the ARM UI is to make a feasible trajectory plan and translate the plan into key strokes or joystick movements.

The questionnaire reveals that participants viewed these two UIs as easy to learn and use. Participants rated the keypad UI as easier and less confusing after the training and testing. This may suggest that the keypad UI may look complicated from the first impression. However, after practicing and testing with the task board, the participants perceived it as easier and less confusing. These results suggest that the amount of time in training and practicing is essential for the keypad UI. With enough time in training, the users' perception for ease was increased. In the overall opinion questions, most participants had positive responses to the ARMs and UIs. The participants' preference was for more on the independent control with the keypad UI rather than controlled by a remote caregiver. In the open interview questions, participants reported that they liked most was "efficiency, smoothness, easiness to use, independence, reach, attractiveness." What they like least was "fragility, mobility, cost, and fear in hurting objects or myself." What the participants would like to change were "force sensing fingers, stop operation on an impact, reducing flipping modes, doing heavy duty tasks (>10kg), more safety protection, spinning joystick to open/close fingers, more joints, and more accurate control." During the testing, three participants expressed that their experiences in video games for years helped to learn the UIs. This suggests that gaming experience may be a factor in learning UIs which could be considered in the future studies.

This study has limitations. First, the participants are all first time users with limited experience in the ARM control. Therefore the performance may vary among trials. However, this helps to discover the major bariors to first time users. Second, the tasks in this study did not include complex tasks or small objects. Manipulation in complex tasks is difficult to test because strategies may be different between users. Small objects are another difficult task for manipulation. Third, the priori sample size calculation was not conducted with TP and roughness. It was not possible to conduct an accurate a priori sample size estimation because this was the first study comparing two ARMs with their original UIs using TP and roughness. Therefore, there are risks of type II errors. Nevertheless, the results of this study enable a priori sample size calculation for future studies.

CONCLUSION AND FUTURE WORK

As the needs for ADL assistance are increasing among people with upper extremity impairment and older adults, assistive robotic manipulators (ARMs) have shown enhanced assistance and increased independence in completing ADL tasks. This study introduces environment-independent performance evaluation outcome measurements: TP and roughness. Two commercial ARMs were evaluated with their UIs.

There were statistical performance differences between the two user interfaces found among the simple translational tasks for completion time, TP, and trajectory parameters. Similar user responses were reported in the perceived workload with both UI's. Participants rated the UI's low on frustration and physical workload, but higher on mental effort. The results provide preliminary evidence of ARM performance. These results may help clinicians to develop appropriate training and guide researchers to develop ARM UIs to better fit users' needs.

Future work will include the testing with power wheelchair users to further identify normative performance and barriers between user interfaces. The comparison with clinical accepted performance based functional assessment tools, such as Jebsen-Tayler Hand Function Test or Wolf Motor Function Test, will help to identify the concurrent validity of the ADL task board.

ACKNOWLEDGEMENTS

This material is based upon work supported by Quality of Life Technology Engineering Research 369 Center, National Science Foundation (Grant #0540865). The contents of this paper do not represent the 370 views of the Department of Veterans Affairs or the United States Government.

REFERENCES

[1] Brault MW. Americans with disabilities: 2010 household economic studies [Internet]. US Census Bureau. 2012. Available from: http://www.census.gov/newsroom/cspan/disability/20120726_cspan_disability_slides.pdf.

[2] Johnson R. Meeting the long-term care needs of the baby boomers: How changing families will affect paid helpers and institutions [Internet]. Washington, DC 2007. Available from: http://www.urban.org/publications/url.cfm?renderforprint=1&ID=311451&buildstatic=1.

[3] Stone R. Who will care for us? Addressing the long-term care workforce crisis. Rep US Dep Heal Hum Serv [Internet]. Washington DC 2001 [cited 2013 Dec 15]; Available from: http://research.urban.org/UploadedPDF/Who_will_Care_for_Us.pdf.

[4] Wang H, Grindle GG, Candiotti J, Chung CS, Shino M, Houston E, *et al*. The Personal Mobility and Manipulation Appliance (PerMMA): A robotic wheelchair with advanced mobility and manipulation. 34th Annual International Conference of the IEEE EMBS [Internet]. San Diego California USA; 2012. p. 3324-7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23366637.

[5] Chung CS, Wang H and Cooper RA. Functional assessment and performance evaluation for assistive robotic manipulators: Literature review. J Spinal Cord Med [Internet]. 2013 Jul [cited 2014 Feb 12]; 36(4): 273-89. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23820143.

[6] Prior SD and Warner PR. Wheelchair-mounted robots for the home environment. Proceedings of 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '93) [Internet]. IEEE; 1993 [cited 2013 Nov 14]. 1194-200. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=583355.

[7] Laffont I, Biard N, Chalubert G, Delahoche L, Marhic B, Boyer FC, *et al*. Evaluation of a Graphic Interface to Control a Robotic Grasping Arm: a Multicenter Study. Arch Phys Med Rehabil [Internet]. Elsevier Inc.; 2009 Oct [cited 2012 Oct 16]; 90(10): 1740-8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19801065.

[8] Alqasemi R and Dubey R. Kinematics, Control and Redundancy Resolution of a 9-DoF Wheelchair-Mounted Robotic Arm System for ADL tasks. 2009 6th International Symposium on Mechatronics and its Applications [Internet]. Ieee; 2009. 1-7. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5164834.

[9] Capille JW, Carey S, Alqasemi RM and Dubey R. Kinematic Evaluation of Commercial Wheelchair-Mounted Robotic Arms. 2011 IEEE International Conference on Systems Man and Cybernetics. IEEE 2011; 711-6. http://dx.doi.org/10.1109/ICSMC.2011.6083725

[10] Harwin WS, Rahman T and Foulds RA. A Review of Design Issues in Rehabilitation Robotics with Reference to North American Research. Rehabil Eng IEEE Trans [Internet]. 1995 [cited 2012 Nov 6]; 3(1): 3-13. Available from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=372887.

[11] Oess NP, Wanek J and Curt A. Design and evaluation of a low-cost instrumented glove for hand function assessment. J Neuroeng Rehabil [Internet]. BioMed Central Ltd; 2012 Jan [cited 2012 Oct 12]; 9(1): 2. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3305482&tool=pmcentrez&rendertype=abstract.

[12] Romer GRBE, Stuyt HJ and Peters A. Cost-Savings and Economic Benefits Due to the Assistive Robotic Manipulator (ARM). 9th International Conference on Rehabilitation Robotics 2005, ICORR 2005 [Internet]. Ieee 2005; 201-4. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1501084.

[13] Bach JR, Zeelenberg AP and Winter C. Wheelchair-mounted robot manipulators. Long term use by patients with Duchenne muscular dystrophy. Am J Phys Med Rehabil Assoc Acad Physiatr [Internet] 1990; 69(2): 55-9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12212433.

[14] Chung CS, Wang H, Kelleher A and Cooper RA. Development of a Standardized Performance Evaluation ADL Task Board for Assistive Robotic Manipulators. Proceedings of the Rehabilitation Engineering and Assistive Technology Society of North America Conference. Seattle, WA 2013.

[15] Douglas S, Kirkpatrick A and MacKenzie I. Testing Pointing Device Performance and User Assessment with the IS0 9241, Part 9 Standard. Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit [Internet]. 1999 [cited 2012 Dec 5] 215-22. Available from: http://dl.acm.org/citation.cfm?id=303042.

[16] Chung CS, Hannan MJ, Wang H, Kelleher AR and Cooper RA. Adapted Wolf Motor Function Test for Assistive Robotic Manipulators User Interfaces : A Pilot Study. RESNA Annual Conference 2014.

[17] Maheu V, Archambault PS, Frappier J and Routhier F. Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities. 2011 IEEE International Conference on Rehabilitation Robotics [Internet]. IEEE 2011; 1-5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22275600.

[18] Routhier F and Archambault PS. Usability of a joystick-controlled six degree-of-freedom robotic manipulator. RESNA Annual Conference. Las Vegas, Nevada 2010; 1-7.

[19] Groothuis S, Stramigioli S and Carloni R. Lending a Helping Hand: Toward Novel Assistive Robotic Arms. IEEE Robot Autom Mag [Internet]. 2013 Mar [cited 2013 Mar 27]; 20(1): 20-9. Available from: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6476697.

[20] Driessen BJF, Evers HG and Van Woerden JA. MANUS--a wheelchair-mounted rehabilitation robot. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of

engineering in medicine [Internet] 2001; 285-90. http://dx.doi.org/10.1243/0954411011535876

[21] Tijsma HA, Liefhebber F and Herder JL. Evaluation of New User Interface Features for the MANUS Robot Arm. 9th International Conference on Rehabilitation Robotics, 2005 ICORR 2005 [Internet]. Ieee 2005; 258-63. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1501097.

[22] Hart S and Staveland L. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Hum Ment Workload [Internet]. 1988 [cited 2013 Dec 6]; 1(3): 139-83. Available from: http://humanfactors.arc.nasa.gov/groups/TLX/downloads/NASA-TLXChapter.pdf.

[23] Teather RJ, Natapov D and Jenkin M. Evaluating haptic feedback in virtual environments using ISO 9241-9. 2010 IEEE Virtual Reality Conference (VR) [Internet]. IEEE; 2010 [cited 2014 Apr 14] 307-8. Available from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5444753.

[24] Ha T and Woo W. An empirical evaluation of virtual hand techniques for 3D object manipulation in a tangible augmented reality environment. 2010 IEEE Symposium on 3D User Interfaces (3DUI) [Internet]. IEEE; 2010 [cited 2014 Apr 14] 91-8. Available from: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5444713.

Received on 09-03-2016 Accepted on 02-07-2016 Published on 31-07-2016

DOI: http://dx.doi.org/10.15377/2409-9694.2016.03.01.3