Feasibility of Immersive Virtual Reality and Customized Robotics with Wearable Sensors for Upper Extremity Training

Behdokht Kiafar KIAFAR@UDEL.EDU
Pinar Kullu PKULLU@UDEL.EDU
Rakshith Lokesh LOKERAK@UDEL.EDU
Amit Chaudhari AMITC@UDEL.EDU
Qile Wang KYLEWANG@UDEL.EDU
Shayla Sharmin SHAYLA@UDEL.EDU
Sagar M. Doshi SMDOSHI@UDEL.EDU
Elham Bakhshipour ELHAMBAK@UDEL.EDU
Erik Thostenson THOSTEN@UDEL.EDU
Joshua Cashaback JOSHCASH@UDEL.EDU
Roghayeh Leila Barmaki RLB@UDEL.EDU
University of Delaware, Newark, DE, USA

Abstract

Upper limb impairment significantly impacts daily activities and quality of life. Traditional robotic systems have been widely used in neurological rehabilitation applications. However, its adoption has been limited to laboratory and clinical settings due to cost constraints. Our study aimed to assess the feasibility and usability of a cost-effective virtual reality (VR) system for home-based upper limb training. We used a customized wearable sleeve sensor to assess the hand and elbow joint movements objectively. A pilot user study ($n = 16$) with healthy participants involved evaluating system usability, task load, and presence within two conditions of *VR alone* and *VR combined with a customized inverse kinematics robot arm (KinArm)*. Results of statistical analysis using a two-way repeated measure (ANOVA) revealed no significant difference between conditions in task completion time. However, significant differences were observed in the normalized number of mistakes and recorded elbow joint angles between tasks. Our findings highlight the potential advantages of an immersive and multi-sensory approach towards performance assessment. This study explores avenues for the development of potentially cost-effective, tailored, and engaging environments for home-based therapy applications.

Data and Code Availability The data used in this study consists of objective performance metrics, subjective questionnaire responses, and resistance measurements obtained from a wearable sleeve sensor, detailed in our experimental procedures. This data is publicly accessible via the following GitHub repository link.

Institutional Review Board (IRB) The Institutional Review Board (IRB) of the *University of Delaware* (Protocol #:1982585-1) has approved the research, confirming that it meets all ethical standards for studies involving human participants. This includes the protection of participant confidentiality and reducing any possible harm.

1. Introduction

The upper extremity impairment induces disabilities for everyday tasks and activities, e.g., reaching, holding, and picking objects (Hatem et al., 2016). Moreover, previous research has shown that regaining lost functions in the upper extremities is often more complicated than those in the lower extremities (Iruthayarajah et al., 2016; Welmer et al., 2008). Physical therapy and therapeutic exercises are highly recommended for rehabilitation and improving quality of life (Ingram et al., 2021). Physical therapy is usually performed under the supervision of a trained therapist and in a clinical setting. The patients are also expected to perform regular exercises at home. How-

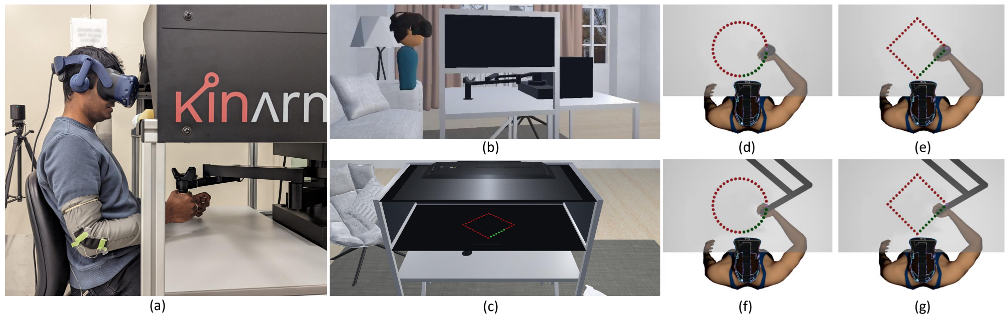

Figure 1: Overview of the therapeutic system for upper extremity rehabilitation using immersive VR and end-point robotics (KinArm): (a) technical setup for study conditions, (b) the immersive VR environment with virtual avatar and virtual robotics with inverse kinematics, (c) the first-person view in VR, and conceptual views of the conditions and tasks: VR condition with the (d) circle and (e) diamond tasks, and VR KinArm robotic condition with the (f) circle and (g) diamond tasks.

ever, studies have shown that the patients' adherence levels to the prescribed regimens are often low because of the inability to track progress, lack of motivation, slow recovery progress, and limited understanding of the benefits of treatment (Emmelkamp and Meyerbröker, 2021; Wiederhold and Riva, 2019).

Advances in technology and computing have increased the use of virtual reality (VR) in physical therapy and have shown potential to improve patient adherence. Patients with neuro-musculoskeletal injuries, including stroke, need to perform reaching and stretching movements to recover mobility in their upper limbs. The use of VR in upper limb therapy offers several benefits, such as increased patient engagement through game-like characteristics, portable setup for home therapy, reduced outside pressures and distractions, and the ease of measuring and evaluating performance (Juan et al., 2022).

Compared to traditional physical therapy, robotic rehabilitation also promotes remarkable improvement in motor and physiological skills (Wu et al., 2021; Frisoli et al., 2007). However, However, the integration of immersive VR with end-point robots for upper extremity rehabilitation remains underexplored (Tarnita et al., 2022; Mubin et al., 2019). Additionally, the current upper extremity VR rehabilitation tools mainly focus on the analysis of hand movements (Alexandre et al., 2019). Thus, further assessment is needed on how VR-based therapy impacts upper limb exercise performance and user engagement.

In this work, we propose a framework that leverages immersive VR to enhance the effectiveness of upper extremity rehabilitation, focusing on the feasibility of VR for home-based therapy. We compared two conditions: *VR alone* and *VR KinArm*, the latter incorporating both VR and customized end-point robotics (*KinArm*, BKIN Technologies). This comparison aims to assess the interchangeability of these setups and their potential benefits.

We developed a VR environment with two reaching tasks, replicating the end-point robot in the virtual setting. A knit fabric-based nanocomposite wearable sensor was used to assess elbow joint movement in addition to the hand movement data from VR and KinArm for a holistic assessment of the upper extremities. This sensor was sewn into a one-size-fits-all sleeve for the user to wear on their dominant arm.

We conducted a pilot user study (n = 16) to assess the effectiveness of our framework from the perspective of user performance, usability, task load, and sense of presence. Since it is an early prototype, the aim was to evaluate the feasibility, usability, and qualitative feedback with non-clinical, healthy subjects. The results provide insights regarding the potential benefits of our *VR* system for upper extremity rehabilitation. They also show future research directions for virtual therapy and multi-sensor movement sensing for more precise and data-driven therapeutic assessment. Our contributions are the following:

- Development of a therapeutic VR framework offering a cost-effective, tailored, and personalized solution for upper limb therapy applications.

- Evaluative comparison of the VR with a setup combining VR with KinArm robotics, using an in-house wearable sensor to capture comprehensive limb and hand movement during task performance, aiming to provide insights into the interchangeability and feasibility of our VR framework as a potential scalable upper limb rehabilitation solution.

- Presenting the results of a pilot user study ($n = 16$) to assess system usability, task load, presence, and user performance as a preliminary system assessment before extensive evaluation of clinical studies.

- Exploratory analysis of participants' qualitative feedback to determine advantages, limitations, and potential improvements and research directions.

2. Related Work

In this section, we describe related work regarding the use of VR, wearable sensors, and end-point robotics for upper extremity rehabilitation.

2.1. Virtual Therapy and Wearable Sensors

Research shows that virtual systems can improve healthcare training (Kiafar et al., 2025), increase patient adherence and support their mental health (Høeg et al., 2023). A recent review reported the feasibility of using VR in treating obsessive-compulsive disorder, psychosis, autism spectrum disorder, attention deficit hyperactivity disorder, post-traumatic stress disorder, and eating disorders (Emmelkamp and Meyerbröker, 2021).

VR-based physical therapy gives patients support and flexibility to exercise in a home environment with fewer visits to clinics (Høeg et al., 2020; Baron et al., 2021). Cui et al. (2019) proposed a self-guided healthcare system, with a wireless inertial sensor combined with VR interactive technology. This system provides a frozen shoulder joint mobility self-measurement system that can measure shoulder joint movements at-home without the intervention of a specialist. A clinical Rehab-Immersive (RI) framework has been developed to support the rehabilitation of patients with spinal cord injuries addressing upper limb motor impairments (Herrera et al., 2023). They found that RI has the potential to enhance the rehabilitation experience for both therapists and patients. Similarly, (Aderinto et al., 2023) demonstrated the effectiveness of VR for stroke recovery by improving patient engagement and functional independence.

VR has also shown potential as a promising tool for upper limb rehabilitation after stroke (Tokgöz et al., 2023; Ross et al., 2023). Subramanian et al. (2019) characterized the outcomes of 100 existing studies utilized in assessing VR interventions for post-stroke upper limb motor improvement. They highlighted the importance of including both kinematic and motor performance measures to capture all dimensions of improvement in stroke rehabilitation. Dewil et al. (2023) investigated how VR can improve upper-extremity motor rehabilitation after neurotrauma, emphasizing the integration of sensory feedback and psychophysiological elements.

Wearable sensors are essential in VR-based physical therapy systems for monitoring physical activity and providing remote feedback on exercises performed at home (Rawashdeh et al., 2022). Jurioli et al. (2020) introduced a low-cost wearable device integrated with VR to enhance therapeutic sessions, making the rehabilitation process for patients with motor disabilities more engaging and effective.

Integration of wearable sensors with a virtual physical therapy system gives an edge over traditional rehabilitation methods by offering real-time feedback to the patient, making it more engaging and enabling remote rehabilitation in a home setting (Baron et al., 2023). Alexandre et al. (2019) used IoT wearable devices developed by embedding smart sensors in headbands and two gloves and proposed a VR-based solution for the physical rehabilitation of upper limbs. Additionally, recent work has demonstrated the potential of combining VR environments, wearable sensing technologies, and machine learning models to predict upper limb motion intentions with high accuracy, offering more personalized and objective rehabilitation assessment (Ravva et al., 2024).

2.2. End-Point Robotics for Upper Extremity Rehabilitation

Robot-mediated neurorehabilitation has attracted significant attention in the last decade due to advancements in the field of robotics and neuroscience. Walker et al. (2022) provided a review of systems using robots with VR and augmented reality (AR) to enhance efficiency and safety in healthcare and also being utilized in therapeutic and assistive robots. Computational motor learning principles and advanced training algorithms have also been integrated to provide customized support to each patient (Marchal-Crespo and Reinkensmeyer, 2009).

In comparison to conventional physical therapy, rehabilitation robots can be used to administer consistent and high-intensity training for a longer duration of time (Huang and Krakauer, 2009; Lo and Xie, 2012). Several robotic rehabilitation studies have

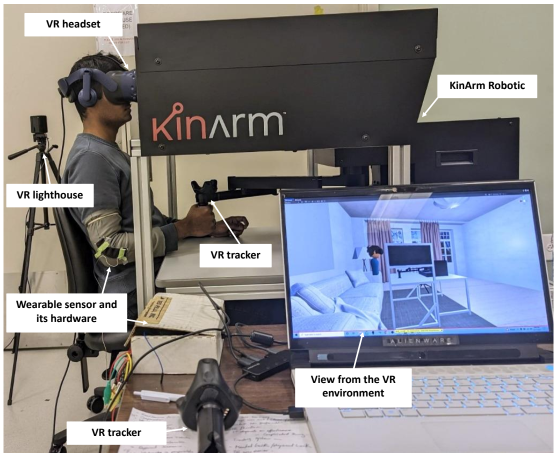

Figure 2: Overview of the technical setup of the proposed framework. We customized the KinArm robotics with a VR setup and wearable sleeve sensors to monitor elbow movements. VR trackers were used to provide initial position and calibration of the robot handle in VR. A third-person viewpoint is displayed on the computer screen.

reported significant improvements in motor control and physiological measures in individuals affected by stroke (Kwakkel et al., 2008; Wu et al., 2021). Hussain et al. (2019) developed a rehabilitation system using adaptive robotics with semi-immersive VR simulations, specifically designed to assist children with upper extremity hemiplegia.

Wonsick and Padir (2020) presented a systematic review indicating the rapid growth and synergy between VR interfaces and the field of robotics. Particularly, individuals suffering from motor impairments caused by stroke or Parkinson's disease can greatly benefit from such technology. For patients with Parkinson's disease, a meta-analysis found that VR-based therapy can enhance balance and gait (Wang et al., 2019). Mohammadi et al. (2018) presented a framework for using VR and robots to enable juggling games for patients in motor rehabilitation to release the burden on therapists and improve the effectiveness of physiotherapy. Their system combines real-time motion generation based on targeting human motion, real-time robot control, and VR for therapeutic juggling. The study highlights that there are still crucial questions to be answered, and future studies are needed to advance our knowledge about human-robot interaction.

Table 1: Participant background and characteristics ($n = 16$).

Characteristics	Value	Mean
Age	[22-37]	27.38 ± 4.13
Gender		
Male	8	(50.00%)
Female	8	(50.00%)
Video game experience		
None	7	(43.75%)
Several times a year	6	(37.50%)
Daily	3	(18.75%)
VR experience		
None	9	(56.25%)
Several times a year	7	(43.75%)

3. Materials and Methods

We describe the participants, apparatus, study procedure, and study design in the following sections.

Our key research questions are the following:

RQ1 How do *VR alone* setup and *VR KinArm*, which is a setup with robotics conditions influence the user performance for common therapeutic tasks in upper extremity rehabilitation training?

RQ2 How are subjective measures in terms of usability, task load, and presence in the virtual environment associated with both *VR alone* and *VR KinArm* conditions?

RQ3 What are the applicability, perceived advantages, and limitations of each condition according to qualitative feedback?

3.1. Participants

We conducted *a priori* power analysis to determine the sample size for interaction effects with ANOVA analysis (Tabachnick and Fidell, 2007) (repeated measures, within factors). A statistical G*Power analysis program was used with an effect size $\eta_p^2 : 0.14$ for two groups, resulting in a total sample size of 16 (Faul et al., 2007). Therefore, 16 participants were recruited for the study. Participants were gender matched (8 males and 8 females) and all were right-handed. Table 1 summarizes the background and characteristics of the participants.

3.2. Apparatus

A virtual environment for upper extremity rehabilitation was developed using the *Unity* game engine (version 2021.3.10f1). We used *VRTK* toolkit for basic interactions in VR, such as teleportation. The built-in *OpenXR* package in Unity ensured compatibility with several VR headsets. In this study, we used *HTC Vive Pro Eye* headset, along with its tracker, controllers, and other components. In the virtual environment, a living room and additional 3D models from *Sketchfab* were customized and integrated. The user is represented by an avatar with virtual hands in the VR environment, designed to provide a comfortable and motivating experience.

For evaluative comparison, we designed two conditions: *VR setup alone without robotics* and *VR setup with KinArm robotics.* For the VR KinArm setup, we used a KinArm end-point robot (*BKIN Technologies Ltd.*, Canada) integrated with our customized VR environment (see Fig. 2). The end-point robot is a planar manipulandum that consists of a cylindrical handle. Participants could grasp the handle and move it freely in the horizontal plane (i.e. 2 dimensions, lateral and forward). A virtual model for the KinArm robot was designed to mimic the physical one in the virtual environment (see also Fig. 1). VR allows for the creation of a controlled environment that can be customized for the user. This control is harder to achieve in real-world settings, where environmental variables can fluctuate and affect the outcomes. The reason of replicating the KinArm in the virtual environment is to provide the user with a sense of robotics in front of them while performing in VR and to avoid safety concerns. This replica facilitates the accurate tracking and reliable measurements of different metrics in a controlled environment. An HTC Vive tracker was attached to the KinArm handle to calibrate and capture its movement positions as well as simulate the inverse kinematics for the VR KinArm setup. Thus, when the user moves the KinArm robotic handle, the movement of the virtual KinArm is also simulated.

To evaluate the joint movement at their elbow, a wearable fabric-based carbon nanotube sensor (Doshi et al., 2022) was used. The sensor was made of knit fabric, and it was integrated into a fabric sleeve. This sensor is piezo-resistive and resistance across the electrode changes with the stretch/strain in the sensor for elbow flexion/extension. We used this wearable sensor on the participant's elbow for all the conditions (see also Fig. 2). The resistance was recorded by an Arduino-based voltage divider circuit and data was streamed via a *Microsoft Excel* streamer. The sensor provided changes in electrical resistance corresponding to elbow flexion/extension during the performance.

3.3. Study Procedure

An overview of the study procedure is shown in Fig. 3. The participants were welcomed with a brief overview of the project and study. They were asked to complete a demographic questionnaire. Then, they were informed about the detailed study procedure and were instructed in the use of the KinArm robot, VR headset, and sleeve sensor. The entire session for each participant was planned for one hour.

The procedure started with a training in a different virtual environment, and they were asked to perform a simple priming task. After proper wearing and positioning of the sleeve sensor on the participant's right elbow, the participants were asked to perform a full flexion/extension to collect their baseline resistance change data. A calibration curve was generated for the participants wearing the sleeve. They were also asked to explore all the features and interaction possibilities. The tasks and the measurements were explained. If the participants stated that they understood the procedure, they were connected to the actual VR environment with the first condition, and their performance data was recorded. The order of the conditions was counterbalanced.

After completing the first condition, the participants were asked to take off the VR headset and answer mid-questionnaires, that included usability, presence, and task load index questionnaires. They were asked to rest for a few minutes before starting the second session. After that, the procedure for the second condition was initiated, and the same training and experimental protocols were followed. Lastly, after the participants completed all the conditions, we conducted a semi-structured interview to collect feedback regarding their experience with the system.

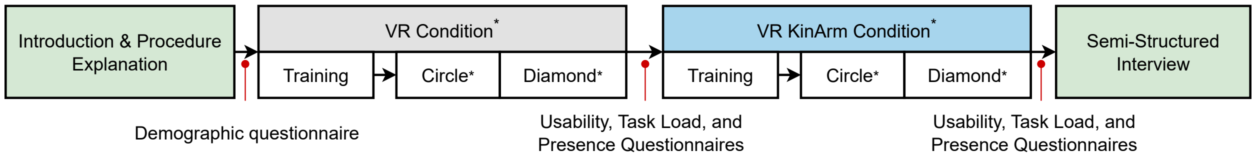

Figure 3: Overview of the study procedure. The order of the conditions (marked with *) was counterbalanced.

3.4. Study Design

This study is a 2 × 2 with-in subjects design. Here we describe the design of our user study, including independent variables, dependent variables, subjective measures from questionnaires, and semi-structured interviews.

Independent Variables The user study was designed as a within-subject experiment with a two-factor test. These two factors were defined as independent variables: *Condition* and *Task* (see Fig. 1). The *Condition* refers to the therapeutic setup, and there were two conditions: *VR alone* and *VR KinArm* (see also Fig. 1).

- VR alone: Participantsperformed the tasks by holding and moving an HTC Vive tracker on a table. Two lighthouses were used to track the position of the VR tracker and headset, and to calibrate the participants in the virtual environment.

- VR KinArm: This condition integrates the VR setup with the KinArm robot. The participants were asked to use the KinArm robot's handle to perform the tasks. A VR tracker was placed on the KinArm handle to calibrate and track its position and to simulate the virtual KinArm in the virtual environment.

In this study, we created a controlled environment that can be standardized for all participants, ensuring consistency in the experimental conditions. This level of control is challenging to achieve in real-world settings, where environmental variables can fluctuate unpredictably and significantly affect outcomes.

Each condition included two *Tasks*, *Circle* and *Diamond*, with each task consisting of 40 drawing dots. These tasks are common in therapeutic training and enable us to accurately assess fundamental motor control and precision. Their simplicity ensures a clear and focused evaluation of the impact of the system on basic yet essential motor skills, offering a foundational understanding of its potential benefits.

The participants followed a pink, blinking dot (see Fig. 1). Upon selection, it turns green while the remaining dots stay red. To avoid bias due to the learning effect, we followed a balanced Latin square (Keedwell and Dénes, 2015). Therefore, the order of the tasks and the direction of drawing (clockwise and counterclockwise) were counterbalanced.

An abstract outline aims to provide a simple task but a better assessment of the wearable sleeve sensor for resistance changes from elbow stretches. The diameter of each dot was 5 cm, and it was identical for both tasks for consistency. Both tasks encourage broad arm movements to mimic upper extremity exercises; the circle task encourages more continuous arm movement, while the diamond encourages more rigid movements given the right angles in the shape.

Dependent Variables Three measurements to objectively evaluate the drawing performance are task completion time, number of mistakes, and resistance change from the wearable sleeve sensor.

- Task Completion Time: the completion time of the drawing task was calculated based on the starting time of the task until the participant hit all dots.

- Normalized Number of Mistakes: the number of mistakes relative to the task completion time (i.e., number of mistakes per second). For example, if successive dots were not hit, the number of missed dots was divided by the task completion time.

- Resistance Change: the change in electrical resistance of the sleeve sensor was recorded throughout the task as the participant flexed and extended their elbow joint.

Questionnaires In addition to the performance data, we also collected the questionnaire data as subjective measures. All questionnaires were designed using the Qualtrics survey platform.

- Usability: the system usability scale (SUS) questionnaire was used to assess usability (Brooke, 1995). This questionnaire includes ten questions with a 5-point Likert-scale ranging from *strongly disagree* to *strongly agree*. The SUS score was then converted to a range between 0–100% (0–50%: not acceptable, 51–67%: poor, 68%: okay, 69–80%: good, 81–100%: excellent) (Bangor et al., 2009).

- Task Load: *NASA TLX* questionnaire was used as an indicator to evaluate the subjective task load of the system interactions and conditions (Hart, 2006). The questionnaire includes mental demand, physical demand, temporal demand, performance, effort, and frustration.

- Presence: to evaluate the sense of presence in the immersive environment, *igroup presence questionnaire* (IPQ) was used (Schubert et al., 2001; Schwind et al., 2019). The questionnaire contains 14 questions with 7-point Likert-scale ranging from *strongly disagree* to *strongly agree*. It is divided into four categories: general presence, spatial presence, involvement, and experienced realism.

Data Analysis *RStudio* with R was used for data analysis. A two-way analysis of variance (ANOVA) was conducted for the dependent variables: *task completion time*, *normalized number of mistakes*, and *resistance change*. We performed a Shapiro-Wilk test to determine the normal distribution of the data. The significane level was set at 0.05, and pairwise comparisons were assessed with Bonferroni correction. It is noteworthy that the resistance change data from the wearable sleeve sensor for one participant was excluded from the analysis due to very high values of resistance, possibly due to error in connection or measurement. The performance and questionnaire data were analyzed with pairwise t-tests and the Bonferroni correction method to determine differences between the conditions.

3.5. Hypotheses

We formulated the following hypotheses to evaluate the effectiveness of the framework:

H1 The *VR KinArm* setup will perform better in therapeutic tasks with the support of the robotics compared to the *VR alone* setup condition.

H2 The *VR alone* condition will be rated higher in terms of subjective measures, including usability and task load, due to its flexibility and the absence of the physical presence of the robotics.

H3 Participants will exhibit high engagement in the VR environment for rehabilitation, potentially increasing motivation through the immersive experiences and their potential for cost-effective and personalized experience.

4. Results

In the following sections, we describe the results of statistical analysis for user performance, questionnaire results, and qualitative participant feedback.

4.1. User Performance Results (RQ1)

The summary of descriptive results for objective measures of user performance is shown in Table 2 and Fig. 4. The statistical results are listed in Table 3.

Task Completion Time (TCT) We did not find any significant differences in the task completion time on their main effects for both factors: *Condition (VR and VR KinArm)* and *Task (Circle and Diamond)*. There was also no significant effect for the interaction effect. It could indicate that both conditions are comparable and interchangeable regarding the task completion time. For descriptive results, the *VR KinArm* condition ($M = 24.10$, $SD = 8.59$) was on average faster than the *VR alone* condition ($M = 26.60$, $SD = 14.10$). However, the results show a small difference between these conditions.

Normalized Number of Mistakes We found no significant difference in the *Condition* factor ($F(1, 15) = 4.290$, $p = 0.055$). However, there was a significant difference of the *Task* factor between *Circle* and *Diamond* ($F(1, 15) = 4.983$, $\boldsymbol{p = 0.041}$, $\eta_p^2 = 0.017$; small effect). We further analyzed the data with pairwise t-test, and found the significant effect (t = -1.92, $df = 15$, $\boldsymbol{p = 0.037}$) between the *Circle* ($M = 0.088$, $SD = 0.04$) and *Diamond* ($M = 0.180$, $SD = 0.05$) in the *VR KinArm* condition – the raw number of mistakes before normalization: *Circle* ($M = 0.812$, $SD = 1.51$) and *Diamond* ($M = 1.688$, $SD = 1.49$). This could indicate that the participants made more mistakes with the *Diamond Task* in the *VR KinArm* condition. It is also noteworthy that the task completion time for this task was faster than others. Thus, it could show that the participants tried to perform this task quickly but made more mistakes. There was no significant effect on the interaction effect of both *Condition* and *Task* factors.

Resistance Change There was no significant difference on the *Condition* factor. In the *Task* factor, however, we found a significant effect ($F(1, 14) = 5.319$, $\boldsymbol{p = 0.036}$, $\eta_p^2 = 0.016$; small effect). The results of pairwise t-test show a significant difference (t = -2.35, $df = 14$, $\boldsymbol{p = 0.034}$) between the *Circle* (M

Table 2: Summary of descriptive results of user performance.

Variable	Task Completion Time (s)	Normalized Number of Mistakes	Resistance Change (%)
VR	26.60 (14.10) [2.50]	0.088 (0.16) [0.03]	605.11 (546.43) [99.76]
Circle	26.29 (15.85) [3.96]	0.088 (0.10) [0.02]	486.99 (354.49) [91.53]
Diamond	26.99 (12.66) [3.16]	0.089 (0.21) [0.05]	723.23 (680.40) [175.68]
VR KinArm	24.10 (8.59) [1.52]	0.134 (0.19) [0.03]	579.69 (525.03) [95.85]
Circle	24.67 (8.84) [2.21]	0.088 (0.17) [0.04]	562.62 (529.04) [136.60]
Diamond	23.45 (8.56) [2.14]	0.180 (0.20) [0.05]	596.76 (538.97) [139.16]

All entities are in the format: mean value (standard deviation) [standard error].

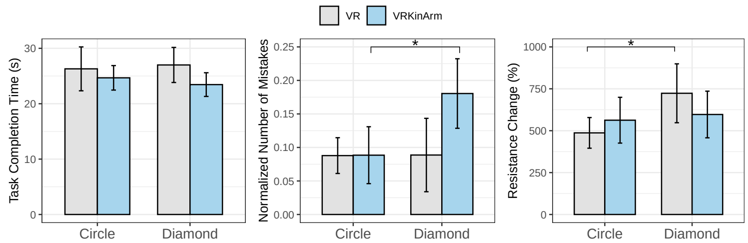

Figure 4: Results of objective measures resulted from the user performance (n=16) on task completion time (left), normalized number of mistakes (middle), and resistance change (right). * denotes significance.

Table 3: Summary of statistical results ($p < .05$).

Variable	DFn	DFd	F	p	Sig.	η_p^2
Task Completion Time						
Condition	1	15	0.921	0.352		0.012
Task	1	15	0.039	0.845		0.0001
Condition * Task	1	15	0.484	0.496		0.001
Norm. Number of Mistakes						
Condition	1	15	4.290	0.055		0.017
Task	1	15	4.983	0.041	*	0.017
Condition * Task	1	15	1.098	0.311		0.016
Resistance Change						
Condition	1	14	0.037	0.849		0.0005
Task	1	14	5.319	0.036	*	0.016
Condition * Task	1	14	4.287	0.057		0.009

= 486.99, SD = 354.49) and *Diamond* (M = 723.23, SD = 680.40) *Tasks* in the *VR alone* condition.

This indicates that during the *Diamond* task, the participants bent their elbow more, leading to a higher resistance change, when compared to the *Circle* task in *VR alone* condition. The changes in resistance values are dependent on the type of movement each task involves. Specifically, the *Diamond* task involves sharper direction changes, which require more frequent elbow bending and straightening, resulting in greater resistance fluctuations. Conversely, the *Circle* task with smoother movements shows lower resistance changes. However, in the *VR KinArm* condition, however, there was a small difference between the *Circle* (M = 562.62, SD = 529.04) and *Diamond* (M = 596.76, SD = 538.97). This might be due to the mechanical support provided by the KinArm, which stabilizes movements and reduces variation in resistance. We did not observe a significant effect on the interaction effect between the factors.

4.2. Questionnaire Results (RQ2)

We describe the subjective results from the questionnaires in the following sections.

Usability The results of system usability using SUS questionnaire from all participants show an average score of (M = 76.88, SD = 10.10) for *VR alone* and (M = 74.84, SD = 10.18) for *VR KinArm* Condition (see Fig. 5). Descriptive results show that there was a small difference between both *Conditions*. However, we did not find any significant differences on the main effect and their interaction effect. The SUS scores for both *Conditions* higher than 68 are above average, which could indicate their potential benefits in terms of usability (Bangor et al., 2009).

Task Load The subjective task load among *Conditions* was assessed using unweighted (raw) NASA-TLX questionnaire. Descriptive results for overall scores show an average of (M = 17.60, SD = 17.68) for *VR alone* and (M = 16.61, SD = 18.44) for *VR KinArm* condition. Moreover, descriptive results show that the average scores of frustration, mental, physical, and temporal demands in the *VR alone* are on average higher, while performance and effort are

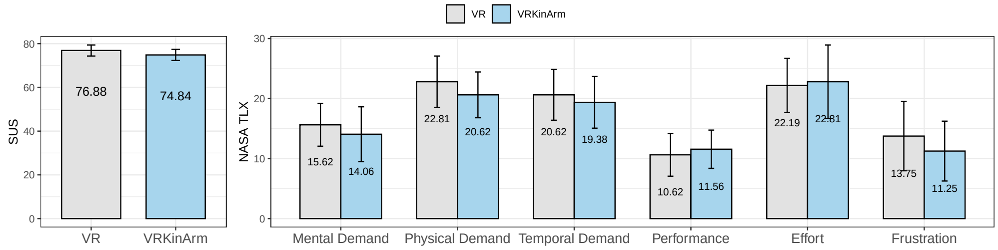

Figure 5: Questionnaire results: (left) system usability scale (SUS) and (right) NASA task load index (TLX).

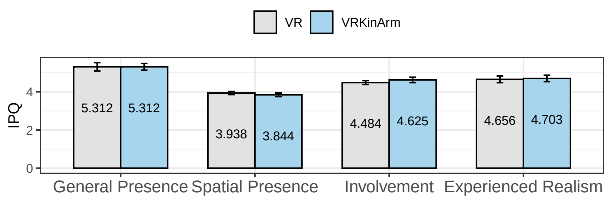

Figure 6: Results of *Igroup* presence questionnaire (IPQ).

on average lower than the *VR KinArm* condition (see Fig. 5). However, the scores are slightly different. We found no significant differences between two *Conditions* for all the task load items. The results show that the participants had similar task loads, and overall scores below 20 are still considered low for both *Conditions*. However, this similarity might be caused by the conditions being similar, or due to the participants not having upper limb impairments.

Presence We assessed the sense of presence in the immersive environment using an IPQ questionnaire (see Fig. 6). There were no significant differences among two *Conditions*. It is noteworthy that the general presence has identical average scores for *VR alone* (M = 5.31, SD = 0.87) and *VR KinArm* (M = 5.31, SD = 0.70) condition. The spatial presence of the *VR alone* condition (M = 3.94, SD = 0.33) is averagely higher than *VR KinArm* (M = 3.84, SD = 0.38), while the involvement (M = 4.48, SD = 0.41) and experienced realism (M = 4.66, SD = 0.70) in the *VR alone* condition are on the average lower than the *VR KinArm*: involvement (M = 4.63, SD = 0.57) and experienced realism (M = 4.70, SD = 0.67). The descriptive results, however, show slight differences between the conditions.

4.3. Qualitative Participant Feedback (RQ3)

In the following, we summarize the participants' feedback during the semi-structured interview.

All participants were positive regarding using VR and end-point robots for upper extremity rehabilitation. Six participants clearly stated that it could be used to engage and motivate patients to exercise, especially for upper extremity therapy, through the immersive experience. Moreover, it benefits patients with limited mobility because the virtual content and surrounding environments can be customized. Two participants with no previous VR experience mentioned that the VR headset was quite heavy for them, and they stated that it could also be an issue for the patients. Thus, using lightweight devices would be more comfortable.

Regarding *Conditions*, four participants stated that using the VR tracker in the *VR alone* condition was easy and portable to hold. Moreover, it is advantageous for *home therapy* with its portability and cost-effectiveness.

Five participants expressed their thoughts about using *VR* and KinArm robot (*VR KinArm*). One advantage is that it is steady and provides haptic feedback. Combining the KinArm robot with VR requires calibration with a VR setup to fit in with the virtual KinArm in the virtual environment. Four participants stated there were tracking issues, especially during the training phase, which required some time to match them with the virtual environment.

5. Discussion

In this section, we summarize the main findings and discuss the implications of the results described in the previous section.

We evaluated our proposed VR and robotics-based upper extremity rehabilitation system in a pilot study using objective user performance metrics, subjective questionnaires, and qualitative feedback. Our hypothesis **H1** can be rejected because there was no significant difference among both factors (*Condition* and *Task*) and their interaction effect. The results indicate that the setups for *VR alone* (using VR tracker) and *VR KinArm* (using KinArm end-point handle) are comparable, thus interchangeable in terms of task completion time. This can mean that the VR representation of the KinArm can produce the same performance results, which can be beneficial to patients since the *VR alone* system is more accessible and home-based than a *VR KinArm* system. However, based on the slight trends in descriptive results, the *VR KinArm* condition is advantageous. Some participants stated that using the *VR KinArm* provides steadier movement and allows them to perform the tasks faster. The results of normalized number of mistakes, however, showed that there was a significant difference between the *Tasks* in the *VR KinArm* condition. Participants made more mistakes performing *Diamond Task* within this *VR KinArm* condition. It is noteworthy that this task was performed faster than others. It could indicate that *VR KinArm* condition provided a steady movement, and the participants attempted to perform it quickly.

No significant difference in the main effect was observed between the *VR alone* and *VR KinArm* conditions for the resistance change. This indicates that both *Conditions* of the study produced comparable outcomes for resistance change measured at the elbow using a wearable sensor. However, there was a significant difference in the *Task* factor, which shows that the participants applied more resistance changes on the *Diamond* task in the *VR alone* condition. This difference in resistance change values between the two *Tasks* underscores how the wearable sensor is sensitive enough to successfully distinguish study *Tasks* from each other. The *Circle Task* has a curved shape which results in lower resistance change values being recorded from participants compared to the *Diamond Task*, which has sharper trajectories and leads to more bending of elbow causing a higher change in resistance. This difference in resistance change may be due to the shape of the task and the type of movement required to complete it.

Descriptive results also show that the resistance change measured during the *Diamond* task is higher in both *Conditions*. Nonetheless, no significant difference in the interaction effect was found.

The use of wearable sensors with *VR* setup enables the accurate measurement of human movements outside the line of sight of VR trackers. These sensors are extremely sensitive, breathable, and comfortable to wear, making them non-invasive and convenient for patients during therapy (Doshi et al., 2022). Using VR therapy with wearable sensors can potentially provide accurate data about a patient's progress and compliance. Challenges with respect to calibration, scalability of sensor manufacturing and data protocols, processing, and protection are being addressed by researchers.

We assessed the subjective measures of the system usability and task load with the standardized questionnaires. For hypothesis **H2**, there were no significant differences between these measures. Trends in descriptive results, the average score of usability for *VR alone* condition is slightly higher than the *VR KinArm*. However, results of the SUS questionnaire reveal both benefits of both *Conditions*, as they are higher than the average (68), which classifies them as rather easy to use. The task load between the *Conditions* was assessed using NASA-TLX questionnaire. The results of overall scores indicate that the task load for both *Conditions* is relatively low. Moreover, we observed no significant effect among the items. For descriptive results, the physical demand was rated higher than other items in the *VR alone* condition, while the effort was rated the highest in the *VR KinArm* condition. The results show the advantages of both *Conditions* regarding the task load.

Concerning hypothesis **H3**, the participants were highly engaged and motivated in completing the tasks in the virtual environment. We used the IPQ questionnaire to assess the sense of presence in VR. It is also noteworthy that nine among 16 participants had no previous VR experience. The results show that experienced realism was rated as the highest among other scales of IPQ for both conditions. It could indicate that the system provides a good experience and realism of the environment. In contrast, the spatial presence was rated as the lowest. One reason might be that the physical KinArm robot was placed in front of the participants, reducing the degrees of freedom. However, they reported that they still had a sense of presence while experiencing the immersive environment. The general presence and involvement, which measures the awareness and attention in the virtual environment, were rated as a good, suitable, and reasonable environment for VR experience. Compared to the work proposed in (Hussain et al., 2019), which used semi-immersive equipment for rehabilitation, our proposed framework offers fully immersive experience using head-mounted display (HMD) headset. Moreover, it can be used for personalized and customized settings according to the needs of each patient. For instance, therapists can adjust the wearable sensor settings to optimize the personalized VR experience throughout the recovery process (Jurioli et al., 2020).

Limitations and Future Directions This preliminary study included non-clinical participants, a fundamental step before including clinical subjects. In future studies, we aim to expand our research to ensure that our findings are applicable in real-world therapeutic contexts. This includes incorporating a wider variety of tasks that test and evaluate the capabilities of the system over the long term with clinical subjects and exploring VR interventions for lower limb rehabilitation (Doshi et al., 2019). Moreover, future research will extend into comparisons of our system with traditional rehabilitation techniques as well as other VR approaches, with additional elbow measurements and personalized calibration with a digital goniometer to provide a more holistic assessment of the framework and participant movement.

6. Conclusion

We have presented a framework for showcasing upper extremity rehabilitation using immersive VR and robotics. We conducted a pilot user study to assess the feasibility and potential usability of this approach while examining user performance and movement using a wearable sleeve sensor made of knitted nanotube fabric. Results of subjective evaluations obtained from self-administered questionnaires, including usability, task load, and presence, were reported. The proposed framework demonstrates the potential advantages of an immersive, multi-sensory approach and provides future avenues for research in developing more cost-effective and personalized upper extremity rehabilitation solutions.

7. Acknowledgments

We wish to acknowledge the support from our sponsors, the National Science Foundation (2222663), and The National Institute of General Medical Sciences of the National Institutes of Health (P20 GM103446 and P20 GM103446E).

We also extend our sincere appreciation to the entire research team and the study participants for their contributions. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

References

Nicholas Aderinto, Gbolahan Olatunji, Muili Opeyemi Abdulbasit, Mariam Edun, Gbolahan Aboderin, and Emmanuel Egbunu. Exploring the efficacy of virtual reality-based rehabilitation in stroke: A narrative review of current evidence. *Annals of Medicine*, 55(2): 2285907, 2023.

Ricardo Alexandre, Octavian Postolache, and Pedro Silva Girão. Physical rehabilitation based on smart wearable and virtual reality serious game. In *IEEE International Instrumentation and Measurement Technology Conference*, pages 1–6, 2019.

Aaron Bangor, Philip Kortum, and James Miller. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3):114–123, 2009.

Lauren Baron, Qile Wang, Sydney Segear, Brian A Cohn, Kangsoo Kim, and Roghayeh Barmaki. Enjoyable physical therapy experience with interactive drawing games in immersive virtual reality. In *Proceedings of the 2021 ACM Symposium on Spatial User Interaction*, pages 1–8, 2021.

Lauren Baron, Vuthea Chheang, Amit Chaudhari, Arooj Liaqat, Aishwarya Chandrasekaran, Yufan Wang, Joshua Cashaback, Erik Thostenson, and Roghayeh Leila Barmaki. Virtual therapy exergame for upper extremity rehabilitation using smart wearable sensors. In *Proceedings of the 8th ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies*, pages 92–101, 2023.

John Brooke. SUS: A quick and dirty usability scale. *Usability evaluation in industry*, 189, 1995.

Jianjun Cui, Shih-Ching Yeh, and Si-Huei Lee. Wearable sensors integrated with virtual reality: a self-guided healthcare system measuring shoulder joint mobility for frozen shoulder. *Journal of Healthcare Engineering*, 2019, 2019.

Sophie Dewil, Shterna Kuptchik, Mingxiao Liu, Sean Sanford, Troy Bradbury, Elena Davis, Amanda Clemente, and Raviraj Nataraj. The cognitive basis for virtual reality rehabilitation of upper-extremity motor function after neurotraumas. *Journal on Multimodal User Interfaces*, 17 (3):105–120, 2023.

Sagar M Doshi, Colleen Murray, Amit Chaudhari, and Erik T Thostenson. Carbon nanotube coated textile sensors with ultrahigh sensitivity for human motion detection. In *2019 IEEE SENSORS*, pages 1–4. IEEE, 2019.

Sagar M Doshi, Colleen Murray, Amit Chaudhari, Dae Han Sung, and Erik T Thostenson. Ultrahigh sensitivity wearable sensors enabled by electrophoretic deposition of carbon nanostructured composites onto everyday fabrics. *Journal of Materials Chemistry C*, 10(5):1617–1624, 2022.

Paul MG Emmelkamp and Katharina Meyerbröker. Virtual reality therapy in mental health. *Annual Review of Clinical Psychology*, 17:495–519, 2021.

Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, 39(2):175–191, 2007.

Antonio Frisoli, Luigi Borelli, Alberto Montagner, Simone Marcheschi, Caterina Procopio, Fabio Salsedo, Massimo Bergamasco, Maria C Carboncini, Martina Tolaini, and Bruno Rossi. Arm rehabilitation with a robotic exoskeleleton in virtual reality. In *2007 IEEE 10th International Conference on Rehabilitation Robotics*, pages 631–642. IEEE, 2007.

Sandra G Hart. NASA-Task Load Index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908, 2006.

Samar M Hatem, Geoffroy Saussez, Margaux Della Faille, Vincent Prist, Xue Zhang, Delphine Dispa, and Yannick Bleyenheuft. Rehabilitation of motor function after stroke: a multiple systematic review focused on techniques to stimulate upper extremity recovery. *Front. in human neuroscience*, 10:442, 2016.

Vanesa Herrera, David Vallejo, José J Castro-Schez, Dorothy N Monekosso, Ana de los Reyes, Carlos Glez-Morcillo, and Javier Albusac. Rehab-immersive: A framework to support the development of virtual reality applications in upper limb rehabilitation. *SoftwareX*, 23:101412, 2023.

Emil Rosenlund Høeg, Begüm Becermen, Jon Ram Bruun-Pedersen, and Stefania Serafin. Co-creating virtual reality applications for motor rehabilitation with physiotherapists. In *Interactivity, Game Creation, Design, Learning, and Innovation*, pages 379–389. Springer, 2020.

Emil Rosenlund Høeg, Jon Ram Bruun-Pedersen, Shannon Cheary, Lars Koreska Andersen, Razvan Paisa, Stefania Serafin, and Belinda Lange. Buddy biking: a user study on social collaboration in a virtual reality exergame for rehabilitation. *Virtual Reality*, 27(1):245–262, 2023.

Vincent S Huang and John W Krakauer. Robotic neurorehabilitation: a computational motor learning perspective. *Journal of neuroengineering and rehabilitation*, 6(1):1–13, 2009.

Netha Hussain, Katharina S Sunnerhagen, and Margit Alt Murphy. End-point kinematics using virtual reality explaining upper limb impairment and activity capacity in stroke. *Journal of neuroengineering and rehabilitation*, 16(1):1–9, 2019.

Lewis A Ingram, Annie A Butler, Matthew A Brodie, Stephen R Lord, and Simon C Gandevia. Quantifying upper limb motor impairment in chronic stroke: A physiological profiling approach. *Journal of Applied Physiology*, 131(3):949–965, 2021.

Jerome Iruthayarajah, Magdalena Mirkowski, MMO Reg, Alice Iliescu, Sarah Caughlin, Joceyln Harris, Sean Dukelow, John Chae, Jayme Knutson, Tom Miller, et al. Upper extremity motor rehabilitation interventions. *Evidence-Based review of stroke rehabilitation (EBRSR). Ontario: Canadian Partnership for stroke Recovery*, 2016.

M Juan, Julen Elexpuru, Paulo Dias, Beatriz Sousa Santos, Paula Amorim, et al. Immersive virtual reality for upper limb rehabilitation: comparing hand and controller interaction. *Virtual Reality*, pages 1–15, 2022.

Mateus Michelin Jurioli, Alexandre Fonseca Brandao, Bárbara Cristina Silva Guedes Martins, Eduardo do Valle Simões, and Cláudeo Fabino Motta Toledo. Wearable device for immersive virtual reality control and application in upper limbs motor rehabilitation. In *Computational Science and Its Applications–ICCSA 2020: 20th International Conference, Cagliari, Italy, July 1–4, 2020, Proceedings, Part VII 20*, pages 741–756. Springer, 2020.

A Donald Keedwell and József Dénes. *Latin squares and their applications*. Elsevier, 2015.

Behdokht Kiafar, Pavan Uttej Ravva, Asif Ahmmed Joy, Salam Daher, and Roghayeh Leila Barmaki. Mena: Multimodal epistemic network analysis for visualizing competencies and emotions. *arXiv preprint arXiv:2504.02794*, 2025.

Gert Kwakkel, Boudewijn J Kollen, and Hermano I Krebs. Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review. *Neurorehabilitation and neural repair*, 22(2):111–121, 2008.

Ho Shing Lo and Sheng Quan Xie. Exoskeleton robots for upper-limb rehabilitation: State of the art and future prospects. *Medical engineering & physics*, 34(3):261–268, 2012.

Laura Marchal-Crespo and David J Reinkensmeyer. Review of control strategies for robotic movement training after neurologic injury. *Journal of neuroengineering and rehabilitation*, 6(1):1–15, 2009.

Pouya Mohammadi, Milad Malekzadeh, Jindrich Kodl, Albert Mukovskiy, Dennis L Wigand, Martin Giese, and Jochen J Steil. Real-time control of whole-body robot motion and trajectory generation for physiotherapeutic juggling in vr. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 270–277. IEEE, 2018.

Omar Mubin, Fady Alnajjar, Nalini Jishtu, Belal Alsinglawi, Abdullah Al Mahmud, et al. Exoskeletons with virtual reality, augmented reality, and gamification for stroke patients' rehabilitation: Systematic review. *JMIR rehabilitation and assistive technologies*, 6(2):e12010, 2019.

Pavan Uttej Ravva, Pinar Kullu, Mohammad Fahim Abrar, and Roghayeh Leila Barmaki. A machine learning approach for predicting upper limb motion intentions with multimodal data in virtual reality. *arXiv preprint arXiv:2405.13023*, 2024.

Samir A Rawashdeh, Ella Reimann, and Timothy L Uhl. Highly-individualized physical therapy instruction beyond the clinic using wearable inertial sensors. *IEEE Access*, 10:14564–14574, 2022.

Ryan E Ross, Emerson Hart, Ewan R Williams, Chris M Gregory, Patrick A Flume, Christina M Mingora, and Michelle L Woodbury. Combined aerobic exercise and virtual reality-based upper extremity rehabilitation intervention for chronic stroke: Feasibility and preliminary effects on physical function and quality of life. *Archives of Rehabilitation Research and Clinical Translation*, 5(1): 100244, 2023.

Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001.

Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–12, 2019.

Sandeep K Subramanian, Mackenzie K Cross, Cole S Hirschhauser, Vineet BK Johnson, and Timothy A Reistetter. Post-stroke upper limb rehabilitation using virtual reality interventions: Do outcome measures assess extent or type of motor improvement? In *2019 International Conference on Virtual Rehabilitation (ICVR)*, pages 1–6. IEEE, 2019.

Barbara G Tabachnick and Linda S Fidell. *Experimental designs using ANOVA*, volume 724. Thomson/Brooks/Cole Belmont, CA, 2007.

Daniela Tarnita, Ionut Daniel Geonea, Doina Pisla, Giuseppe Carbone, Bogdan Gherman, Nicoleta Tohanean, Paul Tucan, Cristian Abrudan, and Danut Nicolae Tarnita. Analysis of dynamic behavior of parreex robot used in upper limb rehabilitation. *Applied Sciences*, 12(15):7907, 2022.

Pinar Tokgöz, Dirk Wähnert, Andreas Elsner, Thomas Schack, Miguel Angel Cienfuegos Tellez, Jens Conrad, Thomas Vordemvenne, and

Christoph Dockweiler. Virtual reality for upper extremity rehabilitation—a prospective pilot study. In *Healthcare*, volume 11, page 1498. MDPI, 2023.

Michael Walker, Thao Phung, Tathagata Chakraborti, Tom Williams, and Daniel Szafir. Virtual, augmented, and mixed reality for human-robot interaction: A survey and virtual design element taxonomy. *J. Hum.-Robot Interact.*, 2022.

Bo Wang, Min Shen, Yan-xue Wang, Zhi-wen He, Shui-qing Chi, and Zhao-hui Yang. Effect of virtual reality on balance and gait ability in patients with parkinson's disease: a systematic review and meta-analysis. *Clinical rehabilitation*, 33(7):1130–1138, 2019.

Anna-Karin Welmer, Lotta Widén Holmqvist, and Disa K Sommerfeld. Limited fine hand use after stroke and its association with other disabilities. *Journal of Rehabilitation Medicine*, 40(8):603–608, 2008.

Brenda K Wiederhold and Giuseppe Riva. Virtual reality therapy: emerging topics and future challenges. *Cyberpsychology, Behavior, and Social Networking*, 22(1):3–6, 2019.

Murphy Wonsick and Taskin Padir. A systematic review of virtual reality interfaces for controlling and interacting with robots. *Applied Sciences*, 10 (24):9051, 2020.

Jingyi Wu, Hao Cheng, Jiaqi Zhang, Shanli Yang, and Sufang Cai. Robot-assisted therapy for upper extremity motor impairment after stroke: a systematic review and meta-analysis. *Physical Therapy*, 101(4):pzab010, 2021.