Interactive Interfaces for Capturing and Annotating Videos of Human Movement Performance

Kobla Setor Zilevu

Thesis submitted to the Faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Computer Science & Application

Aisling Kelliher
Thanassis Rikakis
Kurt Luther
Margaret Ellis

May 13th 2019
Blacksburg, Virginia

Keywords: Stroke Rehabilitation, Iterative Design Process, Service Design, Patient Experience, Human Movement Performance

Interactive Interfaces for Capturing and Annotating Videos of Human Movement Performance

Kobla Setor Zilevu

Abstract

We propose a methodology for designing robust, low cost semi-automated home rehabilitation systems through gradual integration of computer intelligence with the expertise of clinicians and the active participation of patients. We describe an iterative participatory design process towards the development of a Semi-Automated Rehabilitation At the Home system (SARAH). Our first design cycle concluded with a study with nine stroke survivors. We analyzed the interactions between system, patients, and therapists so as to understand system limitations and explore how therapist knowledge could be leveraged to address these limitations. During a second design cycle we worked with our therapist team to codify their experience into a standardized movement assessment ontology and related rating rubric. We then developed an interface that supported coherent application of this rubric by therapists to rating of videos of rehabilitation training tasks. We used the results of a multi rater study to improve the rating rubric and the interface. We are now scaling the rating and producing a database of rated videos where the ratings are interpretable by both humans and computers. These data sets can be used for developing knowledge constrained machine learning algorithms that can address key limitations of low cost home rehabilitation systems.

Interactive Interfaces for Capturing and Annotating Videos of Human Movement Performance

Kobla Setor Zilevu

General Audience Abstract

In this thesis, I describe the iterative service design process I used in identifying and understanding the needs of diverse stakeholders, the development of technologies to support their mutually beneficial needs, and the evaluation of the end-user experience with these technologies. Over three iterative design cycles, the set of identified end-user customers expanded to include the patient, the supervising therapist, the annotating therapist, and other members of the development team. Multiple versions of interactive movement capture and annotation tools were developed as the needs of these stakeholders were clarified and evolved, and the optimal data forms and structures became evident Interactions between the stakeholders and the developed technologies operating in various environments were evaluated and assessed to help improve and optimize the entire service ecosystem. Results and findings from these three design cycles are being used to direct and shape my ongoing and future doctoral research.

Table of Contents

List of Figures

Preface

The motivation and passion for my thesis stems directly from my father, who had a stroke right before my undergraduate graduation. During that time period, I was contemplating whether to attend graduate school or to go straight to a job in industry. That one moment in my life was very pivotal in my decision to attend graduate school. During my first semester at graduate school at Virginia Tech, I knew I wanted to focus on Human Computer Interaction (HCI,) as well as learn more about stroke. I was fortunate enough to be presented with the opportunity to join the Interactive Neurorehabilitation Lab (INR) which focuses on creating home-based systems for stroke rehabilitation. This has been a great learning opportunity for me, an opportunity I could not have imagined if I did not attend graduate school. One of the greatest joys of my graduate school career is being able to show my father the work I have done, and having him see the different interfaces our team is currently working on. It is truly worthwhile to know that one day, the system we are building will help make stroke rehabilitation affordable to many families.

Introduction

Stroke is one the leading causes of disability in the United States and the most common neurological disorder worldwide. According to the American Heart Association, almost 800,000 people experience stroke in the US each year, with approximately 80% of stroke survivors experiencing muscular weakness or temporary paralysis, which significantly impacts everyday life [1]. Stroke survivors experience challenges in performing everyday activities such as dressing, eating, and basic hygiene routines. Large scale studies highlight the positive impact of long-term therapy in supporting recovery [42], but issues of cost, transport to the clinic, and availability of therapists present considerable challenges, particularly in more rural areas like south-west Virginia. In response, home-based rehabilitation systems are proposed as a potential support solution for patients and therapists alike.

Adaptive home-based rehabilitation can provide evidence-based customization of therapy in the home, together with the increased intensity necessary for better functional outcomes over a shorter duration. While there are a number of challenges to scaling adaptive rehabilitation in the home, the three primary issues are I) replicating the functions of the therapist in the home; ii) motivating the patient to increase adherence and iii) implementing adaptive rehabilitation through low costs systems.

Since the full replication of therapist functions is not feasible [2, 16]researchers and system developers are currently focusing on semiautomated rehabilitation solutions. The approach aims to combine coarse automated real time assessment of rehabilitation exercises with regular remote supervision and adaptation of the training protocol by an expert therapist so as to fit the patient's needs and learning curve. Prior research indicates that even coarse feedback can facilitate active learning and increase [2].

It is important to note that automated assessment of rehabilitation tasks presents challenges that are typical of any complex human performance. From a human performer perspective, performance can be varied as people adopt different strategies in approaching complex sensorimotor learning and performance. From an assessment perspective, in many cases, there is limited agreement between experts regarding standardized and quantitative rubrics for evaluating such performance[28, 43][28, 43]. This includes a lack of consensus among physical therapists regarding the standardized, quantitative evaluation of movement quality components and the influence of such components on overall functional ability [28]. From an automation perspective, challenges arising from variation in performance and lack of consensus on rating performance are compounded by limited availability of standardized training data.

The Interactive Neurorehabilitation Lab at Virginia Tech is developing a low-cost semi-automated rehabilitation system for the home. The current focus of the group is on upper extremity rehabilitation for stroke survivors. The team is building the Semi Automated Rehabilitation At Home (SARAH) system, which combines computational intelligence, therapist expertise, and active patient engagement within a cyber human framework [38] aimed at improving human and cyber intelligence. The SARAH system can be understood as operating within a broad and adaptive ecosystem, which contains diverse stakeholders, technologies, processes, resources, and environments [22]. A service design perspective is ideally suited to conceptualize and operationalize a systematic solution in this complex space. Service design is used to construct a *meta* *design*, where people, institutions, and technologies collaborate together on the co-production of value [22]. Customer (e.g. patients, therapists, research team members, etc.) actions create value for service providers, who in turn provide customers with services. A service design perspective requires an expansion of the user-centered, product focused UX approach. Instead, the service designer focuses on developing product-service systems for multiple stakeholders, who may change roles (from customer to service provider, and vice versa) across a series of intersecting interactions.

In this thesis, I describe the iterative service design process I used in identifying and understanding the needs of diverse stakeholders, the development of technologies to support their mutually beneficial needs, and the evaluation of the end-user experience with these technologies. Over three iterative design cycles, the set of identified end-user customers expanded to include the patient, the supervising therapist, the annotating therapist, and other members of the development team. Multiple versions of interactive movement capture and annotation tools were developed as the needs of these stakeholders were clarified and evolved, and the optimal data forms and structures became evident.

Interactions between the stakeholders and the developed technologies operating in various environments were evaluated and assessed to help improve and optimize the entire service ecosystem. Results and findings from these three design cycles are being used to direct and shape my ongoing and future doctoral research. The figure below introduces the service design blueprint used in this thesis (Figure 1). Details of the stakeholders, relationships, resources, and technologies used in each design cycle will be described in detail in the following chapters. In my concluding remarks, I will return again to this diagram and describe how it will develop and evolve in my doctoral research.

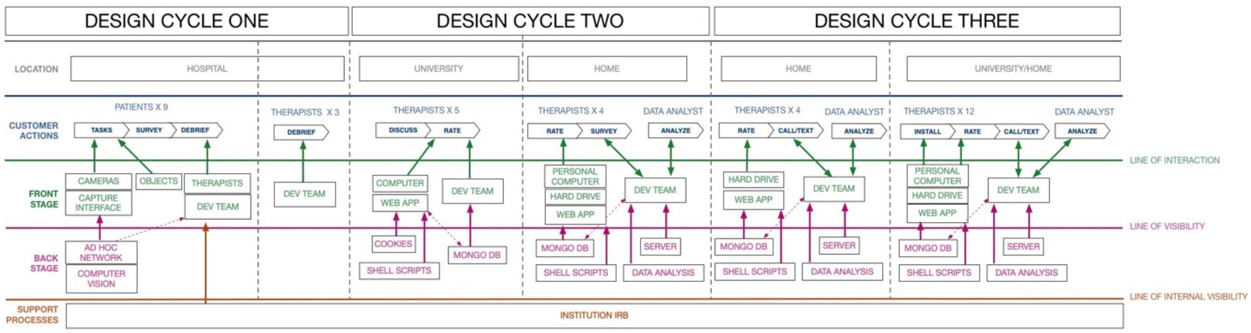

Figure 1. Service Design Blueprint for the SARAH project

1.2: My Role

My primary role in the SARAH project at the INR lab is as a Service Designer, with two sub-roles as a developer and researcher in User Experience (UX) and User Interface (UI). My research goals included: 1) identifying and understanding the needs of the different (and emergent) stakeholders encountered throughout the development process; 2) determining the optimal functionality for the digital support tools required to address stakeholder needs; 3) optimizing the data capture and storage approach for data analysis; 4) capturing high quality data to assist in the development of computer vision algorithms; and 5) assessing the efficiency, usability, and overall stakeholder experience with the SARAH system. To achieve these goals, I used a mixed methods approach involving qualitative and quantitative strategies.

The three design cycles in my thesis can be understood with respect to the *"Integrated design process and people centered research"* diagram depicted in Figure 2, and proposed by service design pioneers, Hugh Dubberly and Shelley Evenson [17]. Although aspects of each of the exploratory, generative, and evaluative strategies are found in each chapter of this document, it is helpful to broadly consider each chapter as follows:

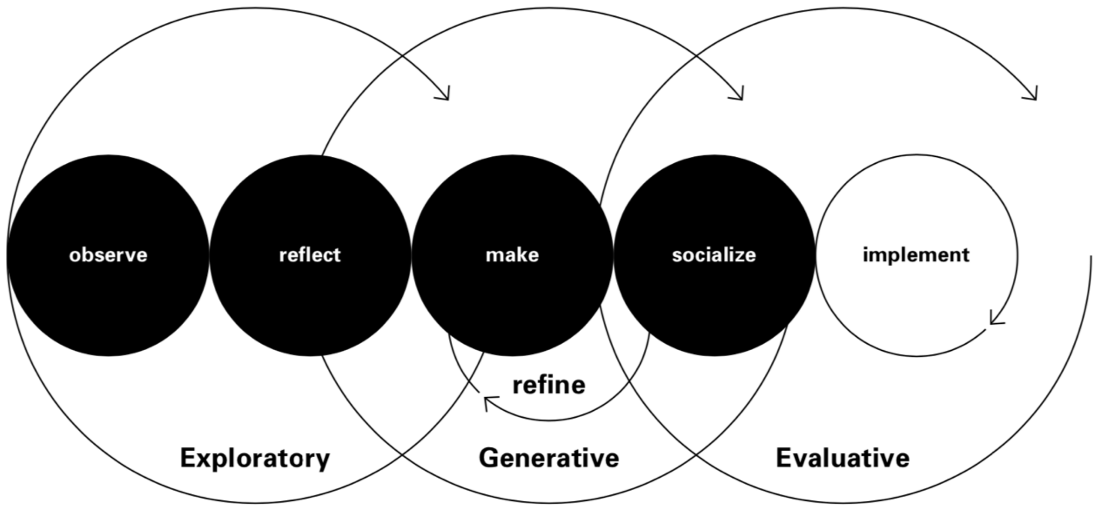

Figure 2. The integrated and iterative design cycles of people centered research (Dubberly & Evenson)

In Chapter 3, I describe the exploratory qualitative work (observations, interviews, and debrief sessions) that helped me to define the multiple project stakeholders, and to scope the functionality of the proposed tools. I also describe the initial generative prototypes of the movement capture and annotation tools and pilot studies with a small set of users. In Chapter 4, I detail how I extended my generative and evaluative output to include more robust tools and a more extensive evaluation period with a group of therapists. In Chapter 5 I introduce the current version of the toolset created for the SARAH system, and describe promising directions for my doctoral studies.

The work in this thesis has been published in two venues to date. I was the first author of a poster on *"A Machine Learning Approach for the Quantitative Assessment of the Upper Extremity Movement in Stroke Survivors*", presented at the first conference on

Progress in Clinical Motor Control: Neurorehabilitation, held at Penn State in July 2018. I was also a co-author on a more extensive paper describing our evaluation studies published at the ACM PETRA ’18 conference, entitled *“Semi-automated home-based therapy for the upper extremity of stroke survivors”.* Both papers are presented in the Appendix.

Chapter 2 - Background and Prior Work

2.1: Stroke Rehabilitation

Stroke is one the leading causes of disability in the United States and the most common neurological disorder worldwide. Every year, over 795,000 people experience stroke in the US, with approximately 80% of stroke survivors experiencing muscular weakness or temporary paralysis, which significantly impacts everyday life [1]. Several large-scale studies demonstrate that recovery can be significantly supported through extended rehabilitation therapy [29, 43] , where the overall goal is to improve long-term functionality to allow the stroke survivor to become independent over time. Traditional stroke rehabilitation therapy of the upper extremity in the clinic is comprised of patients performing movement tasks such as reaching and grasping for an object. These tasks are performed under the supervision of a physical therapist who monitors and assesses the quality of movement over time to provide personalized rehabilitation therapy. The therapist tracks and analyzes each of the activities the patient has performed over a specific time period to generate an overall assessment. The final assessment score is derived by the therapist by analyzing the patient interactions with different objects, environment, and people over a period of time. The observations made by the therapists include the overall assessment, as well as the individualized movement qualities that affect functionality.

2.2: Advancements in Stroke Rehabilitation

Over the last two decades, stroke rehabilitation methods have expanded beyond an instruction-based approach, and towards a more problem-solving framework. There is now a greater focus on the use of patient driven, goal-driven, active problem-solving strategies, and experimentation with training tasks that can be mapped to, and generalized to, multiple activities of everyday relevance. The role of the therapist in developing a treatment plan is to decide which activities of daily living are addressed, what kinds of training may be appropriate, and the overall dosage of the therapy which encompasses frequency and duration.

As the therapy progresses, the therapist needs to change and adapt the protocol based on their observations of the changing performance of the patient [25, 26]. This form of adaptation is used by the therapist to counter issues of fatigue, frustration, or boredom during therapy. The therapist needs to adapt their approach to the different learning styles of each unique patient, as people have varying thresholds for acceptance of challenge, dealing with frustration, and preferences for feedback (some prefer more carrot, less stick). The therapist must help the patient to purposefully map their improvements in training to daily living functions. Finally, the therapist also needs to closely monitor whether the task activities and completion of tasks bring any emotional or physical discomfort to the patient. To address all these challenges and responsibilities, the therapist develops an overall adaptive training plan. Due to the nature and complexity of generating successful adaptive training plans, the most important step for an automated, home-based interactive rehabilitation will be to reproduce a complex adaptive experience in the home without the presence of the therapist while advancing patient self-efficacy, improving adherence, and increasing patient quality of life [39].

2.3: Rehabilitation Systems for Stroke

There are a wide variety of physical, digital, and hybrid systems for stroke rehabilitation including virtual/augmented/mixed reality systems, robotic assist systems, computer vision systems, and tangible interaction systems.

Virtual reality rehabilitation systems often focus on making the rehabilitation process more immersive and appealing for patients. These programmable systems strive to create an interesting and novel environment for stroke survivors to complete their training tasks in a virtual space that is safe and readily customized [12, 34]. However, these VR systems can be disorienting for older users [10]. Relatively low-cost gaming system such as the Microsoft Xbox and the Nintendo Wii have proven effective in improving patient mobility and focus during rehabilitation, although several of these systems still require significant assistances from a therapist [34]. Joystick based systems are another relatively low-cost and customizable technique for lightly supervised therapy.

While detailed tracking of movement through marker-based capture [18] or complex exoskeletons [35-37] is very accurate, it is also costly, complex, and obtrusive. For example, data gloves may not fit those with hand contractures or joint inflammation, especially among stroke survivors, and markers on hands can hinder performance of detailed functional tasks [20, 21]. Furthermore, different low-level features are known to work for different types of movement quality assessment and a combination of different sensors may be needed to acquire the right features. There is also computational and representational incompatibility in the features which makes fusion non-trivial.

2.4: Interactive Therapy Interfaces

As technology and therapy continue to co-align, interactive therapy interfaces are emerging as a mechanism for potentially replicating the instructions and the "voice" of the therapist. Researchers in interactive multimedia environments are beginning to use interactive interfaces to assist therapists in the rehabilitation process [33]. These systems focus on using visualization interfaces with which the patient interacts to complete their assigned rehabilitation tasks. One well known system uses a Microsoft Kinect camera, along with an online virtual environment world to record and assess rehabilitation exercises performed by children with disabilities [3]. This particular application focuses on using a non-invasive approach for assisting children with disabilities in completing their rehabilitation exercises. By creating a 3D multimedia environment, patients can complete their exercises in the form of an interactive game. Each rehabilitation session is controlled by both a menu and a speech user interface [33].

Many interactive interfaces are game based and utilize the RGB and RGBD cameras in creating a customized rehabilitation process for the patient. Creating game-based interfaces can prove to be a motivating and positive experience for patients [20]. A key component in designing any therapeutic interactive interface is ensuring that they are created in a natural and effective way that ultimately promotes end-user learning. By adding approaches such as positive reinforcement, the patients will potentially be more engaged and motivated in continuing to interact with the interface. For several years, the relatively inexpensive Kinect camera tracking system supported multiple promising rehabilitation systems [13, 27, 38] but with manufacturing of the Kinect discontinued, together with other use-case criticisms [40], alternative low-cost movement capture approaches are needed.

2.5: Annotation Tools

The conception of documentation as socially formed and culturally mediated can reasonably be first attributed to the librarian and pioneering information scientist, Suzanne Briet [5]. Her seminal treatise [4] proposed a dynamic interplay between the production and documentation of knowledge, where the creation, combination and use of documents comprise an open and ongoing network of cultural activity. Briet's model allowed and indeed welcomed the future introduction of new technologies for capturing, indexing, sorting, annotating, and presenting documentation. She posited document *use* and *context* on an equal footing with *form*, foreshadowing the approach of contemporary technical archivists engaged in reconciling digital containers, content and contexts [41].

This form of reflective documentation and annotation requires a carefully designed architecture integrating semantics, structure, and syntax. From the Dewey Decimal system, to ISO and Dublin Core, considerable efforts have been made to create comprehensive metadata schemas aimed at facilitating document management, labeling, and retrieval [31]. Insights gained from these large-scale initiatives can inform the development of much smaller domain specific documentation and annotation approaches.

Video annotation tools emerged in the early 90's and included automated annotation of video structure [44] and shot boundary detection [9]. Prior to the advent of shared media platforms, such as Flickr or YouTube, researchers at the MIT Media Lab created shared online video websites [11], where users could upload their personal video collections and then use annotation and moviemaking interfaces to create stories [10], define their movie sequence structure, and co-create movies collaboratively [10].

More recent video analysis and annotation tools adopt a sensemaking approach in collectively determining the quality of both the total video document, as well as particular segments within the document [14]. With the social media explosion in the past decade, a rich variety of community-based assessment and ranking applications have emerged, with the end goal of measuring the perceived quality of contributed content. These applications can be human-centered, end-user centered, designer-centered or hybrids of all three-machine centered, or a hybrid of all four [14]. These applications primarily use either human-based assessment, machine-based assessment or sometimes a combination of the two. For example, with human-based methods, crowds of people voluntarily label data by, for example, writing Yelp reviews or using the thumbs up/down feature for videos on YouTube [30]. In contrast, machine centered approaches use, for example, supervised approaches to determine the 'usefulness' of user-generated content [32], or the similarity between product reviews [23], or machine learning approaches for discovering personally relevant content without user input [6].

2.6: Interactive Neurorehabilitation Lab

Over the past decade, members of the INR Lab have developed systems for supervised interactive rehabilitation in the hospital and clinic [7, 15]. Supervised use in the clinic allowed for expert observers to monitor the therapy, document challenges and areas of improvement, and intervene when necessary when, and if, problems appeared. These earlier systems were marker based and used an extensive array of motion capture cameras [16]. Through a series of studies with stroke survivors, the team were able to identify key kinematic features for evaluating movement quality and model their interrelation and their correlation to standardized clinical measures of functionality [8].

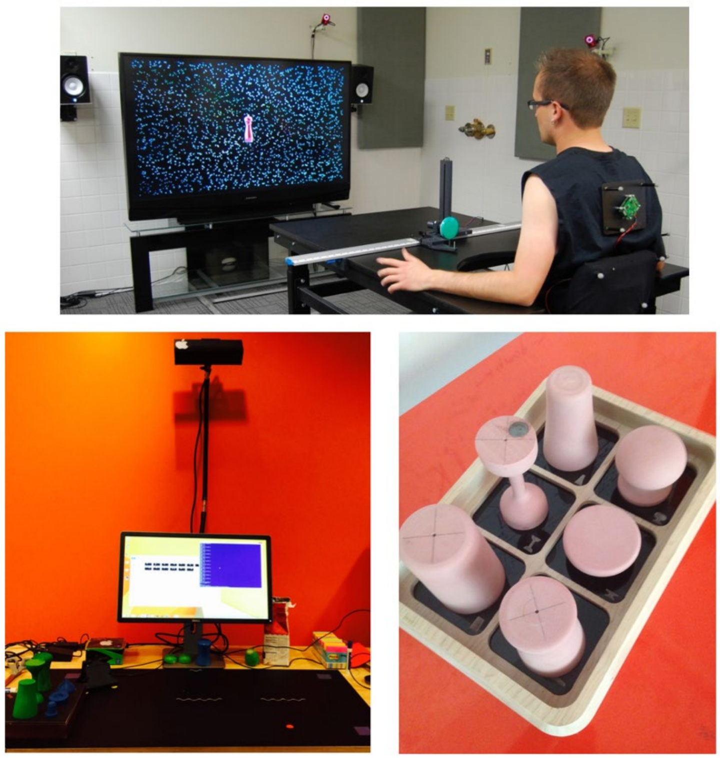

Figure 3. Initial home based prototype developed at ASU (top). Home based system and objects developed at CMU (bottom)

The movement capture and analysis approach with this restricted system combined a significantly reduced marker set, with three motion capture cameras mounted on the screen and smart objects embedded in the table. Therefore, only a reduced set of movement features could be captured and assessed and used to drive automated feedback. However, after testing in the field [2], the variability of tasks and the adaptation options were still too limited for producing generalizable learning. This cycle of testing included offline rating of videos of patient tasks performance by expert therapists in order to further inform the computational assessment of tasks by our system. Although the therapist ratings were limited to a sub-group of the tasks performed, this work showed that an expert constructed movement rating system, used in a consistent manner by trained therapists, could provide good data for training computational agents. The team concluded that successful training at the home needed an even simpler computational rating infrastructure. However, the variability of training tasks needed to be increased, along with their ability to promote generalizable learning that maps to meaningful activities of daily living.

In this thesis, I describe my contributions to the development of the INR lab's SARAH system, including my interactions with multiple stakeholders, iterative development of technology support tools, and evaluation and assessment of the developed system in multiple ecologically valid settings.

Chapter 3 - Design Cycle 1: Data Capture Application

In this chapter, I describe my interactions with primary system stakeholders and the iterative design, development, and evaluation of an early prototype of the SARAH system. This document includes my individual work as a member of a multidisciplinary team, and I will indicate my leadership contributions and that of my colleagues throughout.

This work built initially on some of the core work completed by team members in developing hospital and clinic-based systems as noted in section 2.6. In this instance, our team sought to leverage the extensive experience of the collaborating physical therapists on our team to achieve three key tasks in designing our proposed home-based system and approach. These important stakeholders (and project champions) first identified a limited set of 12 training tasks that scale in complexity (allowing for adaptation to patient progress) and map well to activities of daily living. This produced a constrained set of standardized tasks for automatic rating, and facilitates standardization of training across different clinical settings [39].

The industrial design members in our team (Andrew Gibson and Eric Bottelsen) then designed and created a limited set of modular objects that can be combined in multiple configurations to realize the training tasks at varying rates of patient challenge, as shown in Figure 5. This reduced the number of objects that needed to be recognized and tracked automatically and served to make the home system more compact [24]. Finally, I amended a preliminary interface previously developed at Carnegie Mellon University to create an intuitive digital interface for patients to access the training protocol, control the recording of their training performance, and receive rudimentary feedback on that performance. The design, development, implementation and pilot evaluation of the training system is depicted in Figure 4 below.

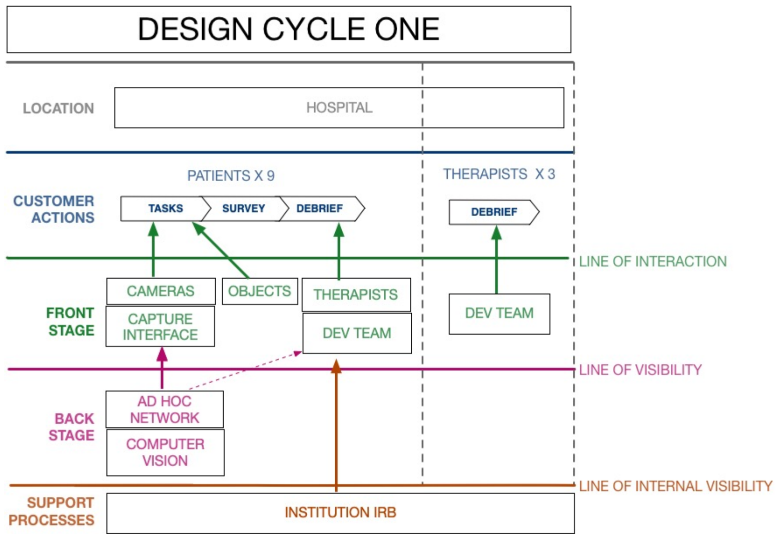

Figure 4. Service blueprint for design cycle one

The primary support process for the work undertaken throughout the three design cycles was the Emory Rehabilitation Hospital IRB approved study (see Appendix). The computer vision system in the first design cycles comprised the Kinect camera and associated custom software (developed initially by Jinwoo Choi). An ad-hoc network was added to the back stage due to network security issues at the hospital (see section 3.3.2 for additional description). The front stage people (those interacting with the primary customer) included the supervising therapists and the observing development team (including me). The front stage technical system involved the SARAH prototype and two video cameras recording observational footage. The patient in this cycle was the initial primary customer, tasked with performing the training, answering some survey questions, and engaging in a brief debrief interview with the therapists and the development team. At the end of this cycle, the therapists migrated roles from being a service provider to becoming a customer in our service ecosystem.

3.1: Data Capture Application

Our initial version of the SARAH system comprised a tablet interface, a Kinect depth camera, a set of therapy objects, and a custom designed tabletop mat. The system focused on training of the upper extremity of stroke survivors who were seated at a table performing reach and touch and transport and/or manipulate tasks. The patient controlled the therapy protocol through the tablet using the unaffected limb(see Figure 5).

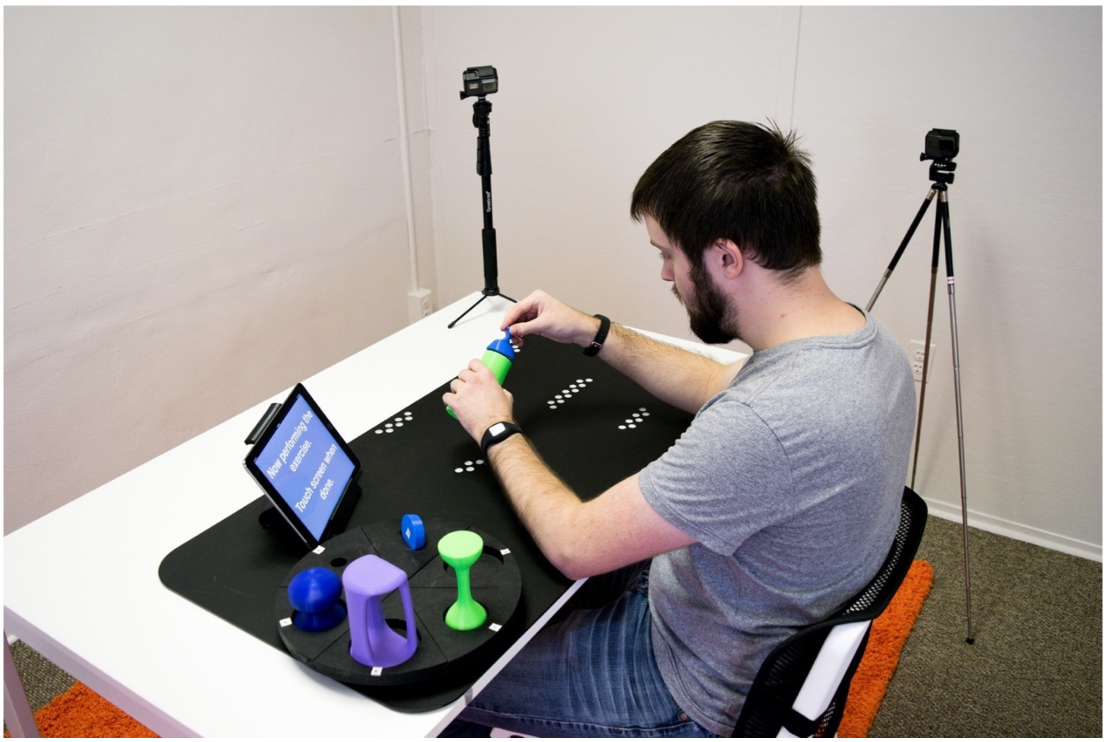

Figure 5. The custom objects, mat, tablet interface, and documenting cameras

We used a co-design process to derive the form of the therapy artifacts (see Figure 6) and their associated activity tasks together with the therapist stakeholders on our team. We leveraged their knowledge and that of the engineers on our team to create objects that were sensible for rehabilitation activities and also optimized for a computer vision system [24]. Objects and tasks were related to functional activities of daily living (ADLs), such as drinking, putting on clothes, and tidying.

Figure 6.The six-initial combinable, modular therapy objects

The system tracked and assessed key mid-level features of standardized upper extremity rehabilitation approaches: time of execution, task completion, misplacement of objects, and compensatory trunk motion. We also incorporated a feedback mechanism in our system to allow both the customer (patient) and the service provider (therapist) to assess and understand how well they were doing in completing each task. The algorithm used to generate this feedback is shown in an equation below and discussed in more depth in Chapter 4. The feedback mechanism was implemented as a way to allow the patient to view their actions as well as see the impact of those actions The system integrated automated ratings across these features into a final 3/2/1 overall rating of task performance, and provided corresponding feedback to the patient through the tablet of "excellent"(3), "good job" (2) , or "nice try" (1). Working with our collaborating therapist team, we then transported our system to Atlanta to test and evaluate this first version of the system and associated training protocol in a two-stage study.

```
if (placScore >= 4 && taskScore >= 4 && avgSC >=4)
{
  //excelent work
  $("#scoreContent").text("Excellent work!");
}

else if (placScore >= 4 && taskScore >= 4 && avgSC < 4)
{
  //good
  $("#scoreContent").text("Good job!");
}

else if (taskScore < 4 )
{
  //nice try!
  $("#scoreContent").text("Nice try!");
}

else if (placScore < 4)
{
  //nice Try, check placement.
  $("#scoreContent").text("Nice try!");
}
```

Figure 7. -Feedback algorithm

3.2: Pilot Study Method

The two-stage study was conducted at the Emory Rehabilitation Hospital in Atlanta, Georgia. Two women and seven men participated in the collection process, two with moderate impairment and seven with mild to moderate impairment. Three of the participants completed three trials with the system over a two-month period, while the remaining six participants completed one trial. We wanted to counteract the novelty effect by having some participants complete several sessions, but we also wanted to have reasonably good coverage across a variety of patient profiles.

The data collection process was approved by the Emory Hospital IRB board [see Appendix] and the participants were compensated for their travel expenses to the hospital and for their time ($50 per session). The participants were informed that this was not a clinical trial or a formal rehabilitation session, but rather an evaluation of an exploratory research therapy tool. Each participant session was video captured through a two- camera setup with one camera placed to the side of the patient capturing the sagittal view (torso and shoulder), while the other was placed on the tabletop capturing arm, wrist and digits during manipulation and transportation activities.

The study sessions were supervised by the consulting rehabilitation medicine expert on our team and a physiotherapist. In addition, one other physiotherapist observed some of the sessions and participated in the debrief meetings. The supervising therapists introduced the purpose of the study and how the system worked to the participants at the beginning of the session. During the session, they helped direct the participants through the set of activities, giving instructions, for example, on when to take a break, or when to move on to another activity. During the hour-long sessions, the participants were asked to complete four repetitions of each of the 12 activity tasks to the best of their ability (again, with the supervising therapist deciding the level of activities/repetitions depending on the ability or fatigue of the participant during the session). Five members of the development team also observed the sessions, three in the first stage with the three repeating participants (I was a member of this team), and two in the second stage with all nine participants. These team members led debrief sessions with each participant at the end of each session. These debrief sessions were recorded and transcribed for subsequent review and analysis.

In addition, the development team held a two-hour meeting at the end of the data collection process with the medical experts and physiotherapists who supervised or observed the sessions. During this session, the therapists gave us important feedback on system and interface improvements including the proposed addition of a training protocol menu option, where the therapist could select and direct the training tasks based on their knowledge and observations of the participating patients. They also discussed how they might systematize their assessment approach and the type of support tool or manual they might need to do this. At this stage in the design cycle, the therapists migrated from a service provider role (where they provided value through feedback to the patient and directions to the development team), to a customer role, where our team needed to support them in making assessment decisions. Qualitative insights gained from this meeting, in conjunction with the extensive notes and documentation collected by the team during the data collection period, were used to help draft a preliminary version of the performance rating rubric described in section 3.6.

3.3 Pilot Study Stage 1 Results

3.3.1 Patient Experience

P1 provided extensive feedback to us during the session itself, and in the debrief period at the end of the session. This participant found considerable fault with the legibility of the interactive interface, in particular with the small font size. All three participants in this initial part of the pilot study commented on this, as they struggled at times to effectively read the different instructions without the help of the therapist or by straining their vision. P1 was also very curious about the feedback score mappings. They wanted to understand the specific reasons as to why they were receiving an excellent work, good job, or nice try assessment. P1 believed that the system was not doing a great job in terms of feedback or reinforcement. While P1 noted that they did not find performing the rehabilitation much fun, they were still very willing to continue the rehabilitation process because they felt it would be help them get better with more practice.

P2 provided interesting feedback about the system throughout their session. As they progressed through the study, they wanted a beginner, intermediate, and expert mode because the exercise difficulties changed as they completed each task. For P2, staying on the easier exercises until they believed they had improved enough to move on to the more complex body movements would have been very helpful. P2 found it very difficult to complete bimanual tasks. However, P2 stated that the exercises helped tremendously with their condition. They noted that each completed task reminded them of how they could utilize their body movements in everyday life. A key feature that P2 enjoyed was the simplicity in our non-time-based interface. P2 did not feel rushed or pressurized to complete any of the tasks, which improved their overall user experience with the system.

P3's feedback was also very useful in helping us refine the interface. As P3 progressed through the study, they did not want to consistently watch the repeated exercise videos. They asked for more concise videos that would instruct them in detail about the specific movements they needed to do to complete the tasks. Though P3 enjoyed interacting with the system, as with P1, they were not happy with the lack of detailed feedback about their performance.

As potential future customers in our service ecosystem, the participants indicated that they saw value in our approach. Through their interactions with the system, they provided value to the development team both through their observed actions and subsequent commentary.

Figure 7. Left and right-handed patients interacting with the SARAH system during the pilot study at Emory Rehabilitation Hospital

3.3.2 System Performance

Based on our initial experiences, it was clear that we had to be more robust in how we installed our system at the hospital. First, we learned that in order for the system to work consistently, we needed to better control and optimize the lighting conditions. We also needed to create our own ad-hoc network, as the hospital firewalls initially prevented us from aligning the communication between the capture interface and the computer vision system. Finally, the object tracking elements of our system did not consistently recognize when objects were moved, which affected the movement assessment. While the supervising therapist was able to give the patient accurate feedback on those occasions, it was clearly something that needed to be addressed. As we work towards building a home-based system, the therapist will not be there to intervene on behalf of the interface. We therefore needed to redefine our algorithm and create a more robust rating scale to ensure that the participants would not lose motivation as they attempted their assigned tasks.

3.4: Data Capture Application Iterative Design

After the completion of the first stage of our study in Atlanta, our team made significant changes to the capture application. One of the first major changes we made was making our interface more focused on the needs of an older population. As the median age for stroke is 63, we needed to ensure that our application was focused on, and better addressed the needs of that age demographic. We made the fonts much bigger, using the standard font and font sizes as described in the user-interface accessibility guidelines published by the US Office of Disease Prevention and Health Promotion [19]. Increasing the font size and changing the font would potentially play a critical role in how each user might understand the task performance instructions. We predicted that by making this change, the therapist would intervene less and allow for the patient to move through the protocol without the need to ask for assistance at the level of instruction.

The second design change we made to the movement capture application was to create and implement a back button. The back button would allow the patients to go back and view the prior instructions. During our initial visit, most of the patients, as they progressed through more difficult tasks, found it challenging to remember what the instructions were after they had clicked through to the next page. We created this button as a method to amplify user agency and to allow users to have more individual control over the therapy progression. Our prediction was that by allowing the user to have the ability to go back and view previous instructions, their feedback on each task would be improved due to the fact that they would the ability to fully understand what each exercise was asking them to complete without feeling rushed.

The third design change we implemented was allowing the patient to practice with the system before any formal assessment began (see Figure 8 below). We created this as a way to alleviate performance anxiety and provide the patients with a “warm-up” space. By allowing them to rehearse with the system, the patient could become more familiar with the interface features and the assessment approach in a relaxed manner.

The fourth design change we implemented was to present an introductory video for the patient to view at the very beginning of the study. This video explained the system components, the study environment setup, the study protocol, and the approach to assessment. We also created a version of the system instructions for left-handed patients (our initial version only depicted right-handed videos and instructions, which was a significant oversight). The fifth change we made to the collection application was the creation of a menu button. This design change was prompted by our growing understanding that the therapist was also an important stakeholder or “customer” in our service design approach. The menu button would serve as a way for the therapist to have more control over the interface and decide which exercises the user should focus on. As patients are all at different stages in their recovery process, the menu button would allow the therapists to hone in on key exercises that they believed the patient needed to focus on. This serves to personalize the experience based on the patient impairment profile and their progress through the tasks.

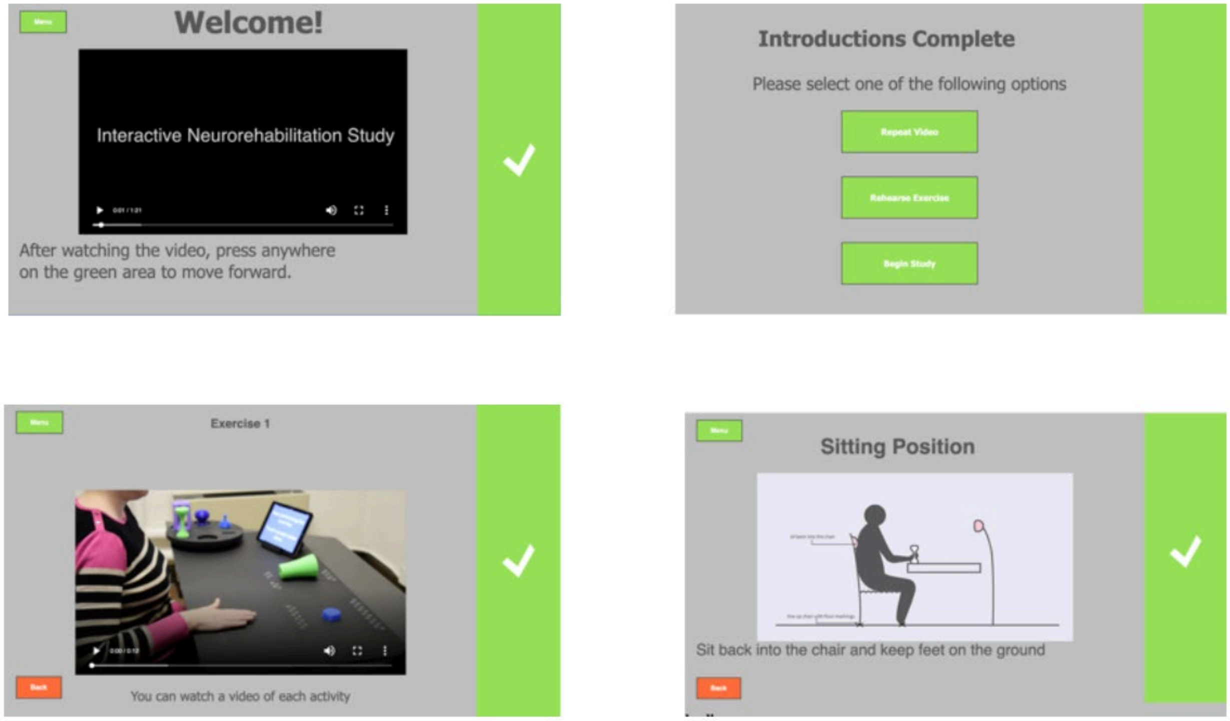

Figure 8. Movement Capture interface including the welcome video, rehearsal options, instruction videos, and helpful tips.

3.5: Pilot Study Stage Two

In the second stage of our user study at Emory, we recruited six new participants, in addition to repeated visits from the three earlier participants. The participants overall offered a rich variety of feedback and critique to the development team regarding additional activities to consider (e.g. more above the shoulder activities, and more bimanual tasks) and additional objects to include (e.g. eating utensils, keys and locks, personal hygiene artifacts). Several of the participants, particularly those with milder physical impairment, noted that while the activities were not physically challenging, they were cognitively challenging.

3.5.1: Patient Experience

Three of the patients did not complete either all four repetitions of an exercise, or did not attempt more than one of the exercise tasks. In some cases, the therapist intervened when the patient could not perform the task but kept trying. In two of the study sessions, the therapist ended the sessions early as the participants were likely not able to complete the remaining exercises, and they were also becoming fatigued. The therapist concluded these sessions gracefully, telling P4 for example *"You did a great job of reaching very hard on a lot of those activities and I think your shoulder is understandably getting a little tired. You've been at it for about an hour and that's good feedback for us too, because, we know that with some patients when they have weakness at their shoulder, then an hour may be too long, or this may be too many repetitions. So I think that's good for all of us to see this, but I also don't want you to strain it anymore."*

The participants had different tolerances for the duration of the training exercises with some of the milder impaired patients (P1, P2) indicating frustration with even the smallest delays in their execution of the task (as compared to what an unaffected limb could complete). Other participants (P3 and P8) persisted in completing the task, even if their ultimate execution time would be considered significantly too long by the therapist. Indeed, P3 noted that they enjoyed having no set task completion time, explaining *"I wasn't being pressured by time. That was good. That was the best part about it really. If there was a time thing, I would have got nervous and stuff probably would have started falling on the floor".*

All nine participants agreed that the training activities reminded them of activities they encountered in their daily lives, with two participants stating that they had done very similar movements earlier in the day. Six of the participants wanted to receive more detailed feedback from the system about their movement performance. Two of the participants were skeptical about the possible accuracy of the feedback, with both candidates expressing their displeasure (e.g. P8*: "Nice try? Bullshit. I did it")* or their ambivalence (e.g. P1: *"Good job. Who knows?*"). The therapist intervened on multiple occasions to assure the patient that their performance was better than the system had assessed. For example, after P6 received several "nice try" assessments after doing the exercise, the therapist carefully stated*: "So, before we move on there. It's saying "nice try" there (points at the interface) and part of what we're trying to align is that you're actually doing a great job of not leaning with your trunk which is exactly what you should be doing. So, it actually should be telling you good job or excellent".*

The participants offered a rich variety of feedback and critique to the development team regarding additional activities to consider (e.g. more above the shoulder activities, and more bimanual tasks) and additional objects to include (e.g. eating utensils, keys and locks, personal hygiene artifacts). Several of the participants, particularly those with milder physical impairment, noted that while the activities were not physically challenging, they were cognitively challenging. P7 stated *"It didn't challenge my strength but it did challenge my pea brain (laughs). Remembering what to do was - the further we got into it, the more I had to think, which is real good for me. It's the deeper I got into it, the more it went from the physical to the brain."*

In summary, the participants found the system and activities potentially useful, particularly from a cognitive challenge perspective. The system needs to be much more adaptive (optimal task duration, order and/or omission of tasks) depending on the impairment profile of the patient and the expectations of the therapist. Where there was clear desire from the patients for more detailed system feedback, this was limited by elements of the system performance as detailed below.

3.5.2: System Performance

Two critical components for assessing functionality in rehabilitation training are the ability to perform the complete training task successfully, and with a duration of execution that is not considered too long to be functional [25, 42].

The system was able to accurately assess duration of the performed task using a simple time/space feature extracted through computer vision (return of affected limb to start position on mat followed by no activity). The accuracy of this approach was confirmed afterward when we manually segmented the sagittal and tabletop captured videos into individual performance tasks and visually inspected them. In addition, expert labeling of this video data using our annotation interface (described in depth in Chapters 4 and 5) also indicated the system accuracy in capturing performance duration. All tasks assessed by the system as performed with a significantly long duration also received low scores from the rating therapists.

The system was not robust in assessing successful task execution. In 30% of the subsequently labeled tasks, the system gave an “excellent” rating and assessed the task execution as complete in contrast to the rating experts who labeled these tasks lower.

Furthermore, 14% of the cases where the system showed the task as incomplete were not corroborated by therapist ratings and annotations. Subsequent inspection of the performance videos of these system-therapist disagreement cases showed a number of repeated errors primarily due to the inability of the system to interpret visual information correctly.

The system assesses task execution by tracking and analyzing objects in relations to the topography of the table mat, and the relations between objects. However, with this approach, the system may not robustly distinguish an incomplete task from object misplacement. For example, on several occasions when the participants placed objects short of their final destination (because they could not reach that far), the system assessed the task as complete with only the object misplaced. The system also did not accurately assess if objects were fully or partially fitted into another object (i.e. whether the screwing of an object into an opening was fully completed). In addition, for this same combinatory task, two participants (P3, P8) executed the task with significant use of their unaffected limb, which was not detected or accurately assessed by the system.

The system was partially successful in tracking significant use of torso compensation during the performance of training tasks. In 92% of cases where the system identified torso compensation issues, the rating therapists also identified torso compensation issues. Even though the precision of the system was high on torso compensation, the precision recall was not. The system only captured 30% of the instances of torso compensation identified later by the team of annotating therapists. Visual inspection of the videos of the missed instances showed that the majority of them involved moderate elements of compensation (i.e. moderate torso leaning or rotation). During the study, we observed that the therapists made multiple movement assessment decisions based on observations of fine movement quality elements (i.e. placement of fingers on objects, raising of shoulder during movement initiation). Furthermore, we also observed that therapists were segmenting the movement into stages, and depending on the stage, they paid attention to different elements. For example, for a reach, grasp and manipulate movement, during the initiation stage, more attention was paid to proximal joints and during the manipulation stage more attention was paid to distal joints. However, although the therapists referred to different movement segments, they never expressed a standardized approach that they were using for this segmentation.

The study indicated that our low-cost prototype system is likely useful for assessing human performance at a coarse level (e.g. task duration and gross torso compensation), but is not, in its current configuration, suitable for providing feedback on fine and complex movement elements (e.g. initial placement of fingers on an object at the end of the reaching stage). Such elements, while deemed critical for assessment by the study therapists, are hard to capture without the use of a high-end capture infrastructure (i.e. markers on digits and objects), which is not feasible for a low-cost home-based system. This confirmed for us that an inexpensive video-based system would need assistance from the humans (patients and therapists) in providing robust and effective therapy at the home.

3.5.3: Discussion with therapists and rehabilitation experts

The debrief sessions with the participating and observing therapy experts also highlighted that accurate movement assessment during rehabilitation requires a significant level of contextual interpretation of the captured features by the therapists themselves. This level of interpretation relies on the extensive experience of the therapists and is hard to automate. The challenge here is also compounded by the well documented issue whereby different therapists have different approaches to assessing movement. This means as there is a no commonly accepted ontology for assessing movement performances [6], therapists may not be able to fully articulate why they made an assessment decision [26].

It became evident from our study that in order to create a semi-automated home rehabilitation system that combines computing and human intelligence, we would first need to work with our therapist team on codifying their expertise into a standardized movement assessment ontology and related rating rubric. Creating this ontology and rubric required the establishment of a common segmentation vocabulary across all tasks and identifying a limited set of movement features deemed critical for assessing performance in each type of movement segment.

To support the needs of the therapists (now customers in our service model) a standardized, consensus method was needed for assessing each of the movement features, and compiling the individual assessments into a rating score for each type of segment. A standardized approach to overall task rating was also needed. The overall task rating needed to address the following: 1) the execution of training exercise (were all the movement segments completed and were they completed in the right order?); 2) timing (was the task completed within a time span that can be considered functional?); and 3) quality of movement (did some/all of the movement segments show movement impairments that need to be addressed?). This rating rubric, would help create consistency in segmentation approaches and segment ratings by experts and would also limit the features needing to be tracked and assessed per movement segment. It would also create consistency in overall task ratings and connect the rating of each movement segment to the overall rating of a training task in an interpretable manner. The resulting hierarchy of coherent ratings along with a methodology for assisting therapists to use this rubric consistently and a user-friendly interface for rating video recordings of patient performances of rehabilitation tasks could create coherent data sets for training of machine learning algorithms. These algorithms, combined with daily input from remote therapists on the adaptation and interpretation fronts, would allow a video based low cost system to robustly and effectively rate performance of rehabilitation exercises at the home.

In the next chapter, I describe how these customer needs were addressed, how I created new technology support tools for the therapists, and how I iteratively developed these tools in response to observations of use and direct customer feedback.

Chapter 4 - Design Cycle 2: Video Annotation Tool

In the second design cycle, depicted in Figure 9 below, I encountered two shifts in stakeholder roles. In the first, the service providing therapists became customers in need of a standardized assessment rubric and an associated interactive support tool. In the second, Heta Patel, the data analyst on our development team, also became a customer who needed data collected and organized in a particular structure and format. From a technology perspective, I created two versions of a video annotation tool, one desktop application, and a subsequent improved online web application. The desktop application was installed on four computers, along with the patient videos for use during an in-situ workshop, while the online application was designed to run on customer's personal computers in their home environment.

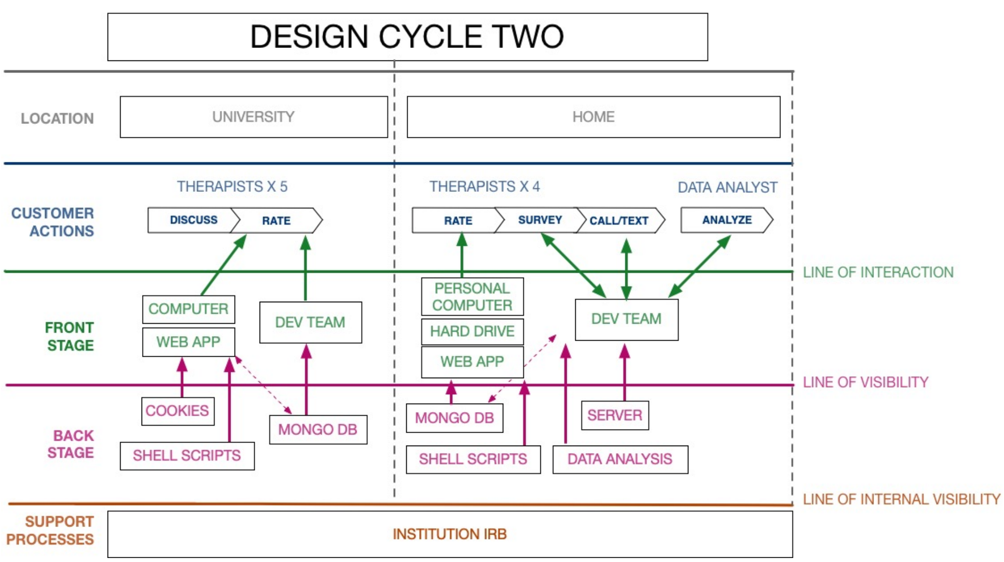

Figure 9. Service blueprint for the second design cycle

I helped facilitate a 3-day workshop at Virginia Tech, where the therapists on our team helped develop the assessment rubric, and where we tried out a preliminary version of the annotation tool. This design cycle was the most intensive of the three design cycles and where I moved extensively and iteratively back and forth through the exploratory, generative, and evaluative stages of people centered design [17] .

4.1: Grading Rubric for Video Annotation Tool

We first leveraged the extensive experience of the therapist's stakeholders on our team to develop a computable rating rubric for assessing videos of human movement. The development of the rubric uses well established motion analysis principles for understanding upper extremity motor control strategies in the context of rehabilitation [1]. The movement rating approach was first drafted by the Dr. Steve Wolf, lead medical researcher on our team, and then thoroughly discussed, edited, and amended over the course of a three-day workshop held at Virginia Tech in March, 2018. Dr. Wolf and four therapists (two physical therapists and two occupational therapists) participated in the workshop. Three of the therapists were from Emory Rehabilitation Hospital, and one therapist was from Radford University.

During the workshop, the therapists developed a common vocabulary of movement segments that can be used in various orders and combinations to segment (and compile) all of the training tasks. In other words, they created a state machine with a limited number of states and transition possibilities that can be used to express all training tasks. This reduces the number of movement segments (states) that need to be recognized, segmented and rated automatically across all training tasks. It also constrains the number of possible movement segment sequences that need to be accounted for.

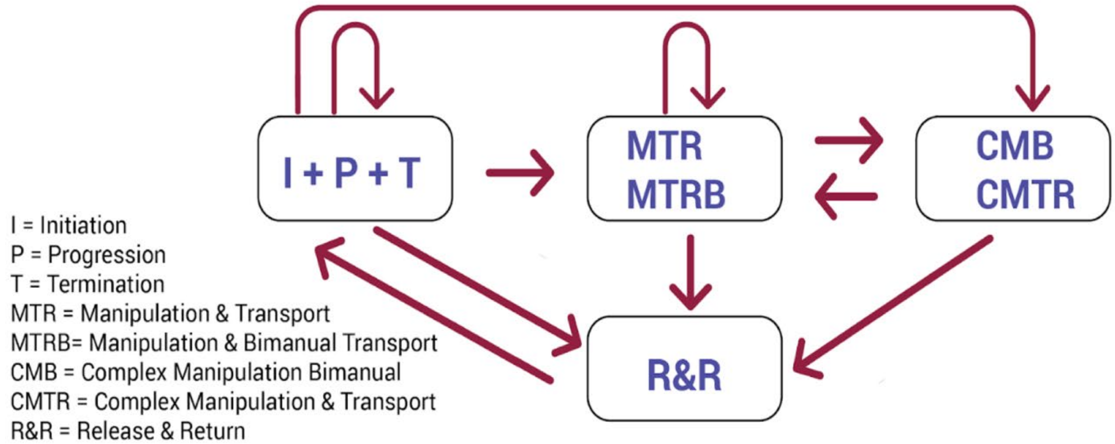

Figure 10.Movement segments and combination possibilities

The rubric establishes four key movement segments that can be combined in different sequences to form all 12 training tasks: Initiation + Progression + Termination (IPT), Manipulate & Transport (M&TR), Complex Manipulation (CM), Release and Return (RR). Figure 11 depicts the possible paths from and to each segment. Assessment and rating of IPT and RR focuses on movement features relating to proximal joints. Assessment and rating of M&TR and CM focuses on movement features relating to distal joints. These segments may also be executed bimanually. A description and evaluation of this rating rubric was recently presented at a motor control in rehabilitation conference [58].

The therapists also developed a consensus limited set of movement features that are critical to assessing performance of each segment. The goal here is to create a standardized, consensus method for assessing each movement feature and compiling individual assessments into a rating score for each type of segment. This creates consistency in segment ratings thus creating coherent scores to train computer vision algorithms in tracking and rating segments automatically. It also limits the features that need to be automatically tracked and assessed per movement segment.

Finally, the therapists developed a standardized, consensus method for assessing and rating the performance of the overall task and the relation of the overall rating score to the rating of the movement segments. The overall task rating needs to address the execution of training task (were all the movement segments completed and were they completed in the right order?), timing (was the task completed within a time span that can be considered functional?) and quality of movement (did some/all the movement segments show movement impairments that need to be addressed?). This approach creates consistency in overall task rating and connects the rating of each movement segment to the overall rating of tasks in an interpretable manner. The resulting hierarchy of coherent ratings facilitates the training of the computer algorithm. An important feature of the therapist's approach was to limit the scoring rubric (for both the movement segments and the overall task) to 3 scores: good performance (3), good performance requiring improvement (2), inadequate performance (1). This approach, (similar to green, yellow, and red used in sports medicine to determine sport participation readiness) keeps the therapist ratings at a high level and makes the rating easier. It also reduces the challenge of reproducing these ratings computationally.

4.2: Video Annotation Tool

In preparation for the 3-day rating workshop, we created a simplified video annotation tool to label the data using the initial rubric draft created by Dr. Wolf. The goal of the tool was to support the therapists at the workshop in assessing and rating the videos as a way to "stress-test" the rubric. In the initial generative design stage, we created sketches and mockups of possible GUI interfaces for the video annotation tool. Working with Yen Troung, we used InDesign to create these initial mockups, as shown in Figure 11. After creating additional low- and high-fidelity mockups of the interface, we began the process of creating both the front-end and back-end of the application.

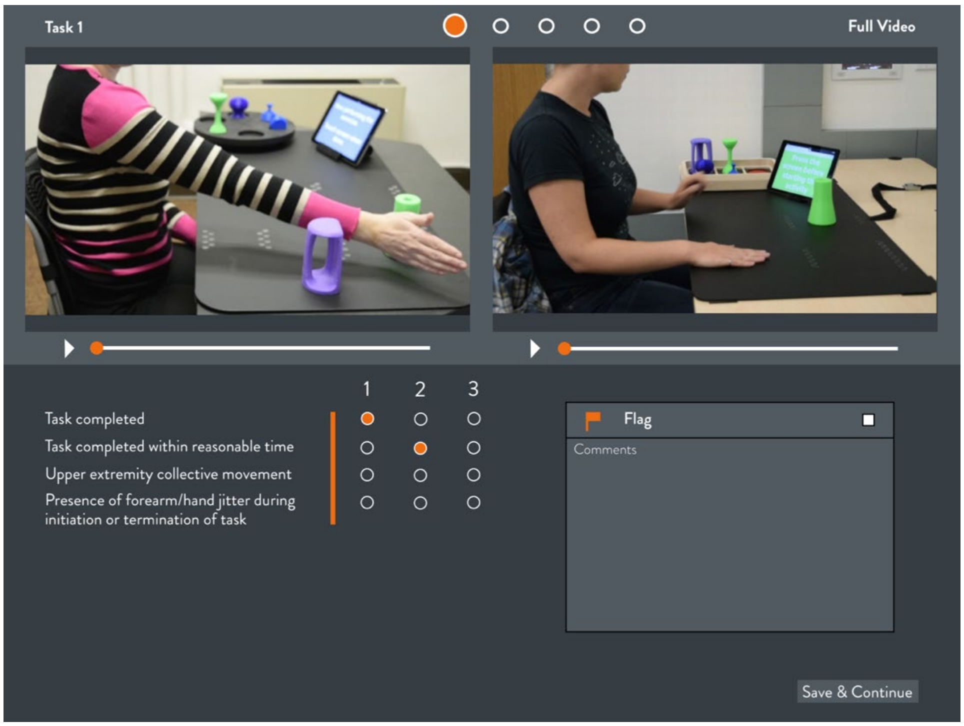

Figure 11. Early video annotation tool GUI mockup

4.3: Front-end development

Our team decided to use a Node JavaScript Framework, along with a MongoDB backend as the optimal technical approach for creating our rating tool. The languages used at this stage of the design process were HTML, CSS, JavaScript, and MongoDB for our server-side logic.

In the first iteration of the video annotation tool, we kept the front-end fairly simple and focused on developing the key functionality to allow the therapists to test out the rubric in rating videos. Figure 12 displays the main features of the desktop application which included displaying two video screens of each patient performing particular tasks from two different angles, a checklist which included the key movement qualities and features the therapists should focus on while evaluating each video, a radio button to allow them to rate each task using a 1/2/3 schema, and a submit button that would store the results in our database.

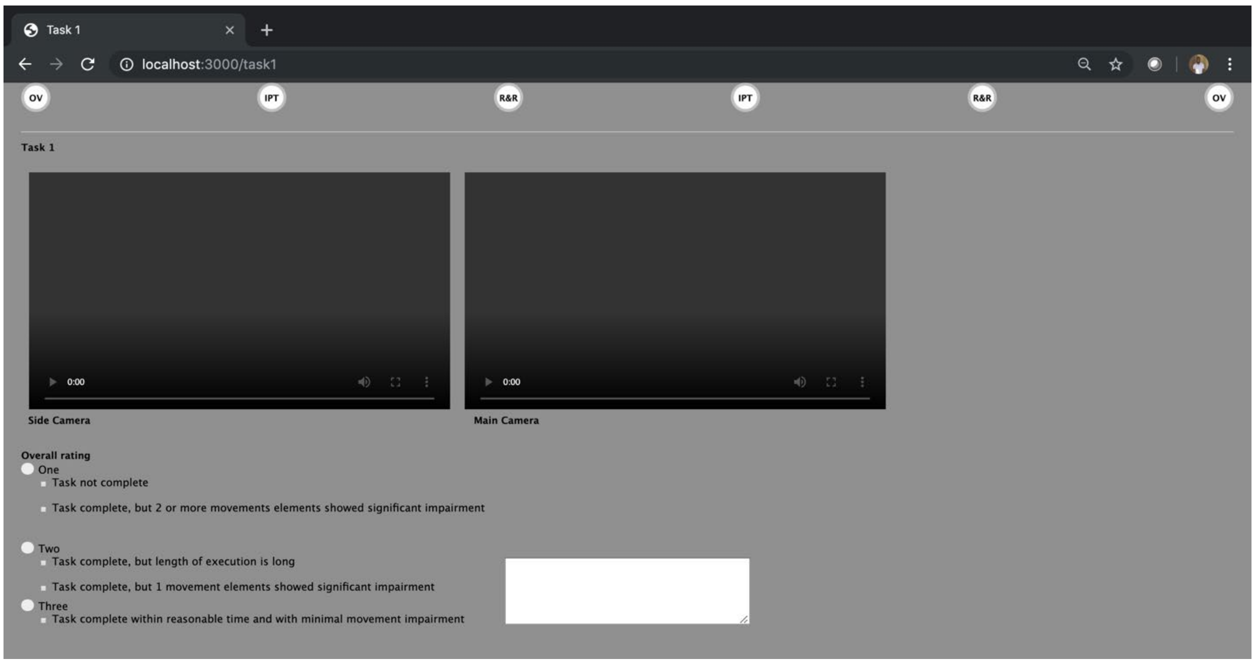

Figure 12. Prototype interface for therapist workshop

4.3: Back-end development

MongoDB was used to provide the back-end of the video annotation tool. MongoDB is a database system that allows different users to store multiple data types. We chose to use MongoDB over SQL Server due to MongoDB's flexibility and accessibility with Node JS. MongoDB provided us with the framework that could successfully keep the data input by the therapists secured through its server encryption.

Security was a major component in designing this application. Due to the nature of our research, we had to ensure that any patient information and the patient videos were kept secure, encrypted and met the HIPPA standards specified in our IRB Protocol (see Appendix). We implemented the backend of the application in three main stages: 1) we created a secure Virginia Tech host server; 2) we created different collection databases for each therapist that they could read/write from, but not modify; and 3) we created a process transferred therapist scores into our system for analysis and interpretation.

We created a host server (inr.cs.vt.edu) to run MongoDB. By hosting our server on the Virginia Tech network, a protective firewall was created in our application, which only allowed authorized users on the Virginia Tech network to access the database collection. We then created and inserted the data into the MongoDB server with a unique read database for each rating therapist. Once the read databases were created, we created write databases for each of the annotating therapists to accurately record and save their scores and comments.

4.4: Video Annotation Tool Workshop

4.4.1. Workshop Method

The goal of the March 2018 workshop at Virginia Tech was to both further clarify and develop the assessment rubric (as described in 4.1), and to assess and rate about 200 videos of patients completing training tasks. Members of our development and design team (Jinwoo Choi, Andrew Gibson, Eric Bottlesen, and Aisling Kelliher) collectively segmented the videos by hand, using the rubric as a segmentation guide. Each of the training tasks included two videos of the entire task attempt from the two different camera angles, and one video angle of each of the task movement segments. The therapist who supervised the capture of the data at Emory worked with us on developing a plan for selecting videos of particular participants and ordering them appropriately. We created an initial group practice session for all rating therapists to work together while viewing and rating videos of two identical exercises performed by two different stroke survivors (one mild/moderate, and one moderate/severe).

We then prepared rating session 1, where the therapists would work by themselves and rate a complete set of exercise videos (exercises 1 – 12) featuring the same two participants. The therapists would look at all the videos sequentially from one participant, before moving to the other participant. In addition, all four rating therapists would be looking at the same data. We also created rating session 2, where the therapists would be assigned two instances of each of the 12 exercises randomly selected from any of the seven participants they had not yet seen. Again, all therapists would be viewing and rating the same video content. We also (very optimistically) created two more rating sessions, where the therapists would be assigned 48 exercise videos, with each exercise video rated by two therapists (to increase our spread across the collected video data).

The four participating therapists at the workshop comprised two occupational and two physical therapists, with Dr. Wolf serving in a consulting role. One of the physical therapists was familiar with our home system and methodology while the other three were not. Each therapist was provided with a computer with the patient videos and rating interface pre-installed (as shown in figure 12), and received training on how to use the rating application.

4.4.2 Workshop Results

Most of the in-situ workshop time ultimately ended up dedicated to considerable discussion about the assessment rubric, with only 18 patient videos rated in the remaining time. Therefore, we had very little results to assess, which was certainly disappointing. However, the therapists gave the development team some pointers about their experience with the annotation interface, and we decided to reorient it to be a more robust online application for remote annotation.

Our tool needed to support therapists in distributed locations in accessing and assessing the videos. From a user experience perspective, the therapists would complete the ratings outside of their regular work hours. Therefore, we needed to support this process by allowing the raters to save their work as they progressed. This would mean that they could do as much rating as their schedule allowed, and all at their own pace. We initially decided to use browser cookies to assist with this process (but updated this approach shortly we began testing the tool – see section 4.6).

We also needed to properly secure the data and the back-end of the application. As per our IRB, we could not put the patient information and/or captured videos on the internet. To address this issue, we secured all the patient videos on an encrypted hard drive that would only be accessible to the receiving therapist. The names of the videos were stored in a database and could only be accessed if the drive was mounted on the therapist's computer. We also chose to place the rating application on the drive as well to give the therapists an easier method of accessing the application. We created "clickable" application buttons so that the therapist could view the application without having to worry about the system configuration on their personal computer.

4.5: Front-end Development

The first step in redesigning the video annotation tool was to adapt the front-end layout and overall tool structure. Again, working in collaboration with undergraduate Computer Science student Yen Troung, we significantly modified the user interface during this design cycle. We wanted to update the look and feel of the tool to be more in line with current human-centered annotation frameworks. We added a login page (Figure 12) to the application to allow for an expanded userbase and to better secure our data.

Figure 11. Video Annotation Tool login

We added an additional video feature that gave the therapists the option to view the instruction task video for the task they were rating (see Figure 13). This was in response to therapist feedback about the possible challenges of coming back to the system (even after a few days) and finding it difficult to remember the appropriate way in which the tasks were supposed to be carried out and assessed.

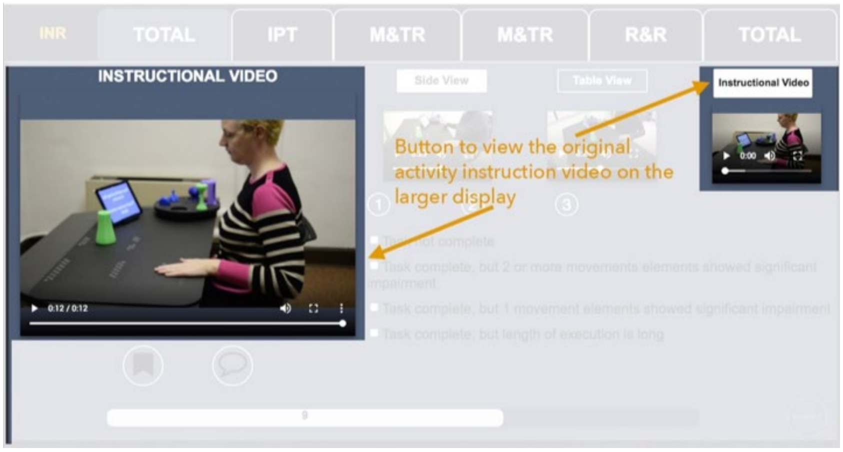

Figure 12. VAT - Instructional Video

From a user experience perspective, we updated the fonts, added color, implemented a progress bar, and refined the overall interface layout to enhance the tool's appeal. Figure 14 depicts the commenting and flagging features implemented to streamline communication between the therapists and the development team, and to draw attention to issues with the rubric or assessment approach. Figure 15 is a screengrab of the interface for annotating the total task video. Here the therapist can assign the video a 1/2/3 and then selected one of the four checkboxes beneath to further annotate their reasoning for the score, using the features assigned by the therapist team during the workshop.

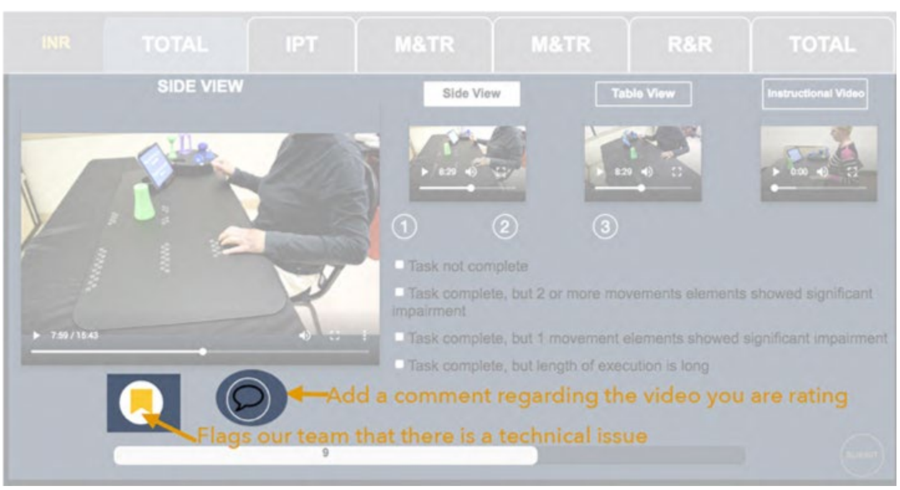

Figure 14. VAT Interface comment box & flag

Figure 13. VAT Interface 3.0

4.6 Back-end Development

The back-end of the application was also completely refurbished. We quickly moved away from using cookies and the therapist's browser history to save state, to instead save their work based on the data from the database. In this iteration of the backend of the application, we moved to a 1:1 mapping of the read and write database vs a 1:2. By moving to a 1:1 mapping, we could easily identify the progress of each therapist and allow our interface to automatically bring each user back to their current state without relying on external services. When we tried the browser cookie approach, we ultimately encountered database errors if the user cleared their browsing history, meaning they would have to begin the rating session again. We also restructured the database in a manner that preserved the authentication of the scores and created a unique database collection for each of the therapists. Finally, we creating a feature that allowed therapists to flag any technical issues through the front-end application during their rating sessions.

4.7 Online Video Annotation Tool Study

4.7.1 Study Method

The four-participating therapists from the in-situ workshop agreed to continue to rate the videos using our refined application. We sent each therapist a personalized hard-drive (with the videos and the amended application) by registered mail. The goal was to have two additional sets of video collections rated by each therapist over a one-month period. The four therapists received instructions on how to log onto the secure online website hosted at Virginia Tech, where they would be presented with an assigned set of task performance videos to label. As noted above, the videos themselves were stored on encrypted external hard drives. The web application would pull the stored videos into the browser for viewing, but only the rating data would be transferred back to our host server (to protect the privacy and security of the patient data). We aimed to check in with each therapist on a weekly basis to see how they were progressing and to address any technical issues.

4.7.2 Study Results

It took considerable effort to get all four therapists up and running with the application. I spent significant amounts of time answering phone-calls and text messages with regard to application installation, mounting hard drives, and occasional data loss. However, despite that, the therapists rated 198 patient exercises over a three-week period.

4.7.3 Study Data Analysis

To analyze the rating data, we first needed to combine the data from the read and write databases into an integrated format. This integrated CSV file needed to include all of the key information such as the patient ID, training task, task repetition numbers, total and segment scores, the checklist information, and any additional comments. I worked closely with Heta Patel in this part of the design cycle as there were inconsistencies initially in how I had set up the database as compared to how she wanted to run her analysis. We were interested in discovering how the raters compared to each other, how they compared to the system results, and how they compared across different patients or indeed, across different tasks.

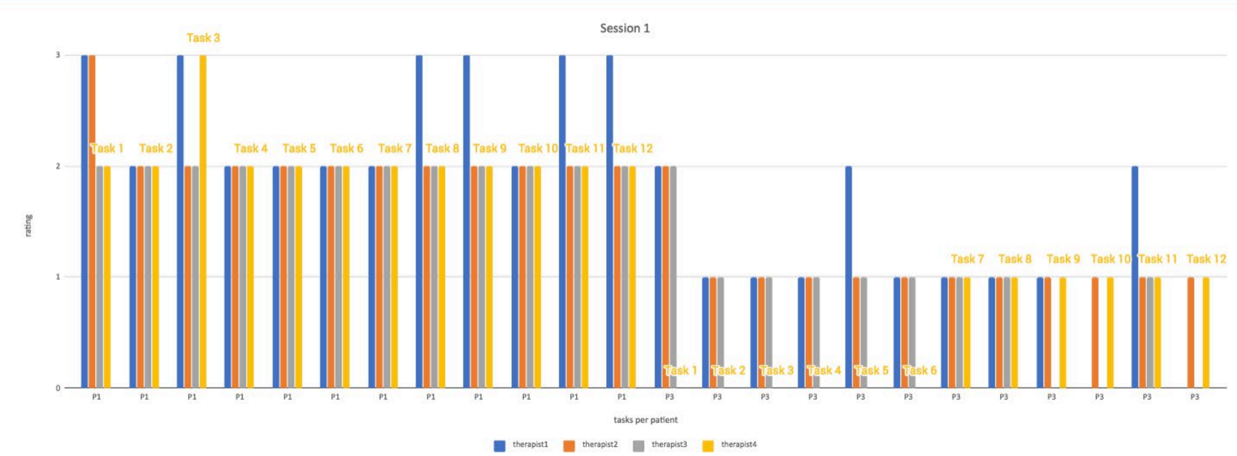

Figure 14. Session 1 Analysis

There was a high inter-rater consistency (80%) when the therapists reviewed a sequence of videos from the same participant (See Figure 17). However, the inter-rater consistency dropped to 55% when the therapists reviewed a sequence of videos featuring different patients (see Figure 18). The inter-rater consistency was high for participants with mild or moderate/severe impairment. However, for participants with moderate impairment, there was much greater variety in rating scores across therapists. There were seven cases where the overall performance rating was evenly split (two therapists gave one score, two therapists gave another score), and two cases where there was even greater variance (three different rating scores per video).

4.7.4 Study Interpretation Survey

We had some intuitions about the consistency levels in the data, but needed the therapists to provide accurate, interpretive context for us. We created a multi-part online survey for the rating therapists and our rehabilitation expert, with sections on rating patient/task order, impact of patient impairment profiles, and interpretation of the split rating scores. The therapists were asked if they expected or were surprised by the drop in inter-rater consistency when rating different patients, and asked for suggestions as to how to optimize the order and presentation of the patient videos. Three of the respondents indicated that this drop was to be expected, although they discussed different possible reasons. T1 noted *"It was difficult and somewhat unnatural as a therapist to break down movement to the degree that we were, especially between the initiation of movement to the progression of movement. It is very difficult to determine when initiation stops and progression begins, so it would not surprise me to have much variability there".*

T5 (the rehab expert on our team) expressed concerns about the challenges for the therapist in rating different patients in succession: "*Once participants are mixed in some random order (or say, all patients doing the same task) the ability to concentrate on the* *totality of movement behaviors that might be unique to any one patient is lost because fundamentally each patient becomes a new set of multiple observations that change from patient to patient and task to task. I would strongly advocate for reviewing as much in one subject as possible rather than mixing across subjects for whatever reason".* T4 offered an alternate solution to increase consistency through extended training: *"My gut tells me that mixing the order may still help generalize rating approaches. I think that we may get better consistency as a result of further clarification and training for rater'.* The suggested requirement of additional training was also repeated by several of the therapists in responses in multiple sections.

In the section on patient impairment profiles, the respondents all agreed that patients with moderate impairment were the most challenging to evaluate, with T1 noting that *"the patients in the middle are always the hardest to judge because their movements patterns are likely more varied. In other words, a person may be considered moderately impaired overall, but will have elements of mild and severe movements intermixed."*

We asked the therapists if there was anything we might need to change in the rubric, or standardize in the rubric instructions to achieve greater inter-rater consistency. Two of the therapists indicated that they weren't sure about changing the rubric, but rather enhanced instructions or longer training sessions might be more beneficial. T4 called for additional clarification in the rubric stating: "*Clarification may remove more of the subjective rating tendencies and give a clearer picture of what exactly is "abnormal" or "non-functional" about the movements. I think the issue is that abnormal and non-functional are not synonymous and individuals can use a wide array of movements to accomplish the same functional task.* T1 also requested additional clarification and proposed a broader, and more philosophical concern: *"Perhaps there needs to be a bigger discussion of the intent of the ratings and software development? The broader the stroke population you try to help with this software, the more difficult it will be to accomplish/rate both quality of movement and completion of tasks with one rating system. In other words, it's difficult to create something appropriate for both mildly and severely impaired patients that will vary so much in their abilities."*

The therapists reaffirmed their disagreement about the split score videos, particularly with regard to the two exercise tasks with ratings of 1, 2, and 3. T1 commented "*I can't see why this would have been scored a 1. In my opinion, movement was completed and with mild-moderate impairments in movement*", while T3 further explained their reasoning for their original low rating *"Certainly cannot be best score...excessive shoulder elevation at initiation of movement, motion is not smooth and wrist position is not appropriate (lacked full pronation at the wrist).* T5 offered a possible reason for the discrepancies stating*: "My guess is that therapists graded the hand position differently. The 3,4,5th digits maintain extension that wouldn't be typical of an unimpaired subject, but his grasp is functional to bring the hourglass to his face. I would guess that raters agree more on the impairments more proximally".*

4.8: Discussion

The interface and rating process helped the therapists assess and improve the rating rubric and reflect on their own assessment process. Focusing on a few key movement elements per segment helped the therapists to track the movement more effectively and to be more consistent with ratings. It is clear that extended rounds of training are required in order to ensure consistency in use of the rubric and rating mechanism. Creating a training set of videos with associated annotations as agreed upon by our therapists and rehabilitation medical experts will be helpful in the process moving forwards.

A key observation from this study concerned the therapists' rating approach while using the system. We had not put any rigorous checks and balances into the system to ensure that all therapists implemented the rubric completely and assessed the data in a similar and consistent fashion. For example, two of the therapists made full use of the checklist features and added extensive comments to explain their rating rationale. However, the other two therapists made only occasional use of the checklist features.

The interface needs to require interpretation of rating scores of segments through selection of the influencing movement elements and interpretation of overall scores through selecting between time, task completion and movement quality as the main reasons for a rating. This approach helps therapists be consistent in their ratings and denote the weights of the elements they used for making a decision. Moving forward, this will result in a database of rated videos where the ratings are interpretable by both humans and computers. This opens the way for addressing key limitations of inexpensive video-based systems for home-based rehabilitation identified in Study 1. The rated videos can be used as training sets for developing knowledge constrained machine learning algorithms that can a) automatically segment videos, b) robustly detect task completion, c) use attention models over pose sequences to assess fine movement quality elements and d) provide robust feedback to the patients in the that helps active learning.

Chapter 5 - Design Cycle 3– Current Video Annotation Tool

For the final design cycle in this thesis project, we redesigned the rating application to include a slightly modified rubric, and a 'stricter' interaction interface. Our research team goal was to increase the pool of therapist raters in order to better integrate our work in the local medical/rehabilitation environment, and to further test the rubric. To this end, we recruited 12 additional occupational and physical therapists from the Radford /Roanoke areas as new annotating 'customers' for our project. Our updated software was designed to be cross platform compatible, and we could fit all the patient videos and the annotating application onto a mini, encrypted thumb drive. I helped facilitate two rating training workshops (one at VTCRI in Roanoke, and one at the INR lab in Blacksburg), where we aimed to install the software, train the therapists, and get them rating more videos at home. Figure 15 below depicts this most recent stage in the service ecosystem.

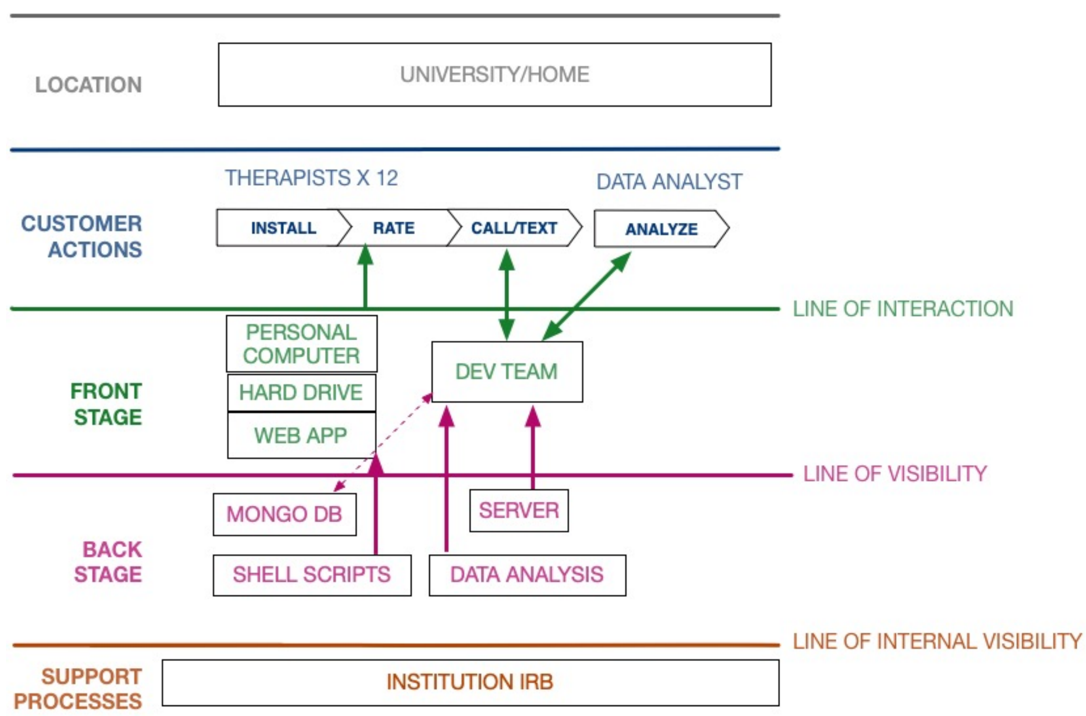

Figure 15. Service blueprint for design cycle three

5.1: SARAH annotation system refinements

Based on therapist rating observation and analysis from the second design cycle, we redefined the checklist used to label the data. The word in the rubric that seemed to give the therapists most difficulty was associating the word "appropriate" with an impaired movement quality as opposed to one that was sufficient. In this design cycle iteration, we reworked the checklist structure in collaboration with Sarah Garrison, the occupational therapist who had joined our team.

We made changes to the tool so that it would require each of the therapists to give a rating, a label, and leave comments to explain the rationale behind their choices. Figure 16 shows the "error" message therapists receive when they choose not to leave a label. The therapists would then be prompted that in order to complete the necessary segments, each total video, as well as each of their segments must be rated and labeled.

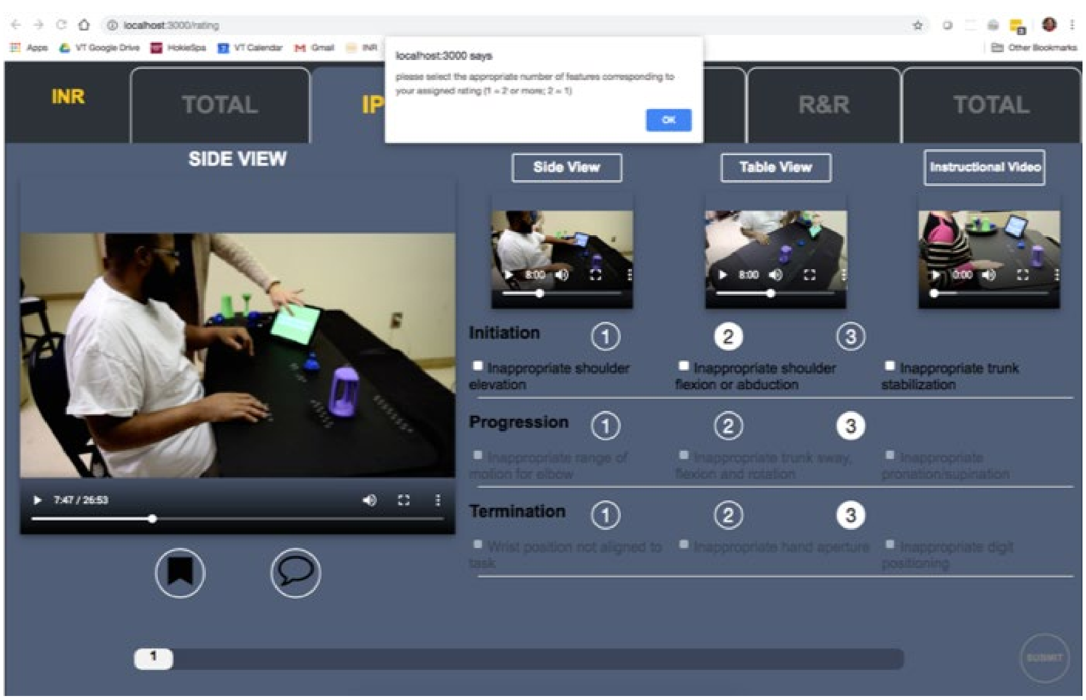

Figure 16. Error message displayed if the therapist did not input a rationale for their score

The most important change that occur within the application was implementing the features that directed the therapists to use the checklist feature on our application. In our first two design cycle, the checklist shown in Figure 16 was optional but after analyzing the data, we decided it was important for the interface to insist on the use of this feature in order to progress through the rating tasks. By making this feature mandatory, our team can better analyze and understand inter-rater correlations.

5.2: Current study

Our team needed to expand the set of potential therapist raters and in particular, we were interested in developing a local team of raters for in-situ work. We recruited twelve physical and occupational therapists from Radford University and Carillion Clinic in Roanoke to participate in our current study. All participants attended an in-person workshop session with our team, where we installed the rating software on their personal computers, instructed them on how to use the interface, and gave each therapist an encrypted thumb drive with the patient videos. We checked in with each therapist on a bi-weekly basis to view how well they were progressing through the rating, and identify and tackle any technical issues that arose.

5.3: Current results

Half of the therapists were able to quickly install the system and began rating videos quickly. However, we encountered challenges with some therapists who had brought relatively "locked down" computers belonging to their medical hospital with them. We also discovered some installation permissions issues, as well as challenges with some therapists who had brought personal computers we were not anticipating using (e.g. chrome books).

Another important issue to note (particularly as the project moves forward) was that very few of the therapists used the implemented flagging/commenting features to notify the therapist or development team that there were issues with the user experience. Instead, they primarily contacted me directly by email or text. While I was able to talk/text them through their issues, this is not particularly sustainable moving forward. For example, many of the therapists encountered the same issues over and over (not being on a secure Wi-Fi which meant that the videos could not load onto their computer based on server communications). Moving forward, we need to begin to add more FAQ features and community support structures.

5.4: Future Doctoral Work

My research development goals moving forward include continued iterative development of the video annotation tool, with particular emphasis on ensuring that the application tool fits the model view controller (MVC) framework. As the cohort of therapists are expected to evolve and grow over time, we will create more steam-lined methods to add or remove annotating therapists. We will also experiment with different hardware approaches, including purchasing customized, inexpensive and encrypted laptops for annotating therapists to individually use for set periods of time. This should help mitigate the need for extended training workshops/installations/copious communications.

Over the longer term, I am interested in seeing our system installed and evaluated in the home. I want to continue collaborating with the patient, therapist, (and next the caregiver) in creating a compelling user experience for the home. I will develop summary interfaces for the therapists and the patients based on their daily activities to help the therapist adapt the therapy protocol, and to help the patient stay motivated over time.

Figure 18 displays the process our overall lab team will be moving towards. As we continue to build our semi-automated system, our goal is to *not* remove the role of the therapist, but rather have the therapist adapt the patient rehabilitation protocol based on the daily, weekly, and monthly summary data and information provided by our system.

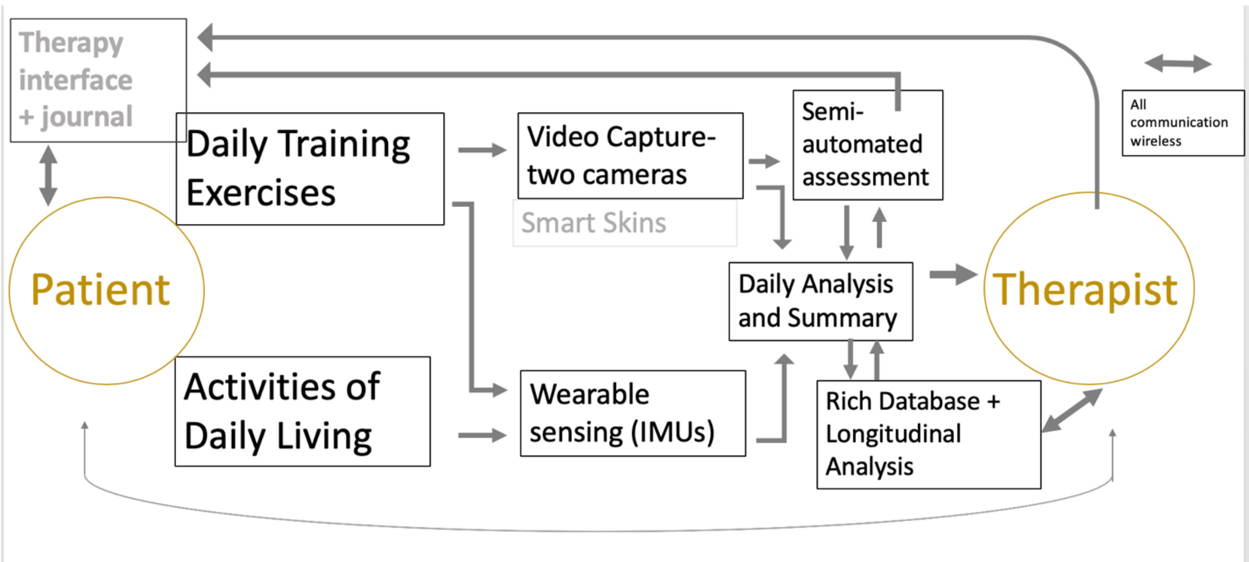

Figure 17. Proposed home based rehabilitation communication system

<u>Chapter 6 - Conclusion</u>

Over the course of an 18-month design and development cycle, the service design model depicted below has evolved and shifted as new stakeholders emerged, new needs were discovered, and new technologies were required.

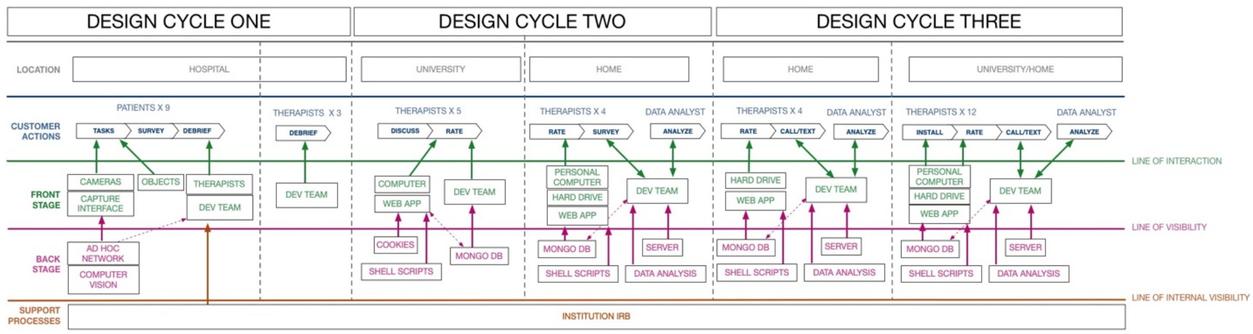

Figure 19. Service Design Blueprint for the SARAH project

I successfully developed and deployed capture and annotation tools for recording and labeling human movement performance within a rehabilitation context. I evaluated these tools in ecologically valid settings and learned many lessons along the way with regard to hardware standardization, software support, and data integration. A service design model provides an adaptive and flexible framework for iterative software development and worked well in this process of customer discovery and needs alignment. I believe our approach will prove successful in generating labeled data of great value not only to our computer vision and machine learning team members, but also to our therapist teams and the rehabilitation community more broadly. I very much look forward to continuing my endeavors in producing functioning, appealing, and rewarding interfaces for home-based rehabilitation and human performance assessment.

Acknowledgements

Aisling Kelliher – Advisor

Thanassis Rikakis – Co-Advisor

Kurt Luther – Committee member

Margaret Ellis – Committee member

Jia-Bin Huang – Computer Vision Lead

Steve Wolf -Emory Rehabilitation Medicine

Sarah Garrison – Occupational Therapist

Heta Patel – Computer Vision Researcher and Data Analysis

Yen Truong – User Experience Designer

Andrew Gibson – Industrial Designer

Eric Bottelsen – Industrial Design

…interactive Neurorehabilitation Team!

Citations

1. al, B.E.e., *Heart Disease and Stroke Statistics – 2017 Update.* A report from the American Heart Association, 2017. **135**(10): p. 146-603.
2. Baran, M., et al., *Design of a home-based adaptive mixed reality rehabilitation system for stroke survivors.* Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 7602-5.
3. Bartoli, L., et al., *Designing and evaluating touchless playful interaction for ASD children*, in *Proceedings of the 2014 conference on Interaction design and children*. 2014, ACM: Aarhus, Denmark.
4. Briet, S. *What Is Documentation?* . 2006. 10.
5. C, M., *Scoring the Work: Documenting Practice and Performance in Variable Media Art*. 2009. p. 59-63.
6. Cafarella, M.B.A.M.E.A.M., *Leveraging noisy lists for social feed ranking.* In Proceedings of the 7th International AAAI Conference on Weblog and Social Media, 2013.
7. Chen, Y., et al., *A low cost, adaptive mixed reality system for home-based stroke rehabilitation.* Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 1827-30.
8. Chen, Y., Duff, M., Lehrer, N., Sundaram, H. Heb, J., Wolf, S., Rikakis, T. , *A Computational Framework for Quantitative Evaluation of Movement during Rehabilitation*, in *AIP Conferernce*. 2011. p. 317-326.
9. Chiu, F.A.A.H.M.-Y., *Feature management for large video databases.* 1993. **Storage and Retrieval for Image and Video Databases**.
10. Davenport, A.K.G., *Everyday Storytelling: supporting the mediated expression of online personal testimony*, in *HCI Applications and Services*. 2007. p. 926-933.
11. Davenport, A.K.J.S.P.P.C.L.G. *Visual interfaces for shareable media*. in *Proceedings of the 10th International Symposium on Electronic Art* 2000. Paris, France.
12. David, L., G. Bouyer, and S. Otmane, *Towards an upper limb self-rehabilitation assistance system after stroke*, in *Proceedings of the Virtual Reality International Conference - Laval Virtual 2017*. 2017, ACM: Laval, France.
13. David, W.O., Celik, *Systematic review of Kinect applications in elderly care and stroke rehabilitation.* Journal of Neuroengineering and Rehabilitation, 2014. **11**: p. 1-24.
14. Diakopoulos., E.M.C.C.N., *A Survey on Assessment and Ranking Methodologies for User-Generated Content on the Web*, in *Computing Surveys (CSUR)*. 2015.
15. Duff, M., et al., *Mixed reality rehabilitation for stroke survivors promotes generalized motor improvements.* Conf Proc IEEE Eng Med Biol Soc, 2010. **2010**: p. 5899-902.
16. Duff, M., et al., *Adaptive mixed reality rehabilitation improves quality of reaching movements more than traditional reaching therapy following stroke.* Neurorehabil Neural Repair, 2013. **27**(4): p. 306-15.
17. Evenson, H.D.S., *Designing for service: Creating an experience advantage.* in Introduction to Service Engineering, 2010.
18. Hanley, B., C. Tucker, and A. Bissas, *Differences between motion capture and video analysis systems in calculating knee angles in elite-standard race walking.* J Sports Sci., 2018. **36**(11): p. 1250 - 1255.
19. https://health.gov/healthliteracyonline/display/section-3-3/.

20. Huang, H., et al., *Novel design of interactive multimodal biofeedback system for neurorehabilitation.* Conf Proc IEEE Eng Med Biol Soc, 2006. **1**: p. 4925-8.
21. Huang, H., et al., *Interactive multimodal biofeedback for task-oriented neural rehabilitation.* Conf Proc IEEE Eng Med Biol Soc, 2005. **3**: p. 2547-50.
22. J, F.J.Z. *Promoting Service Design as a Core Practice in Interaction Design*. in *Proceedings of IASDR13*. 2013.
23. Karahalios, E.G.K., *Understanding deja reviewers*, in *In Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* 2010: New York, NY.
24. Kelliher, A., et al., *Designing Modular Rehabilitation Objects for Interactive Therapy in the Home*, in *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction*. 2019, ACM: Tempe, Arizona, USA.
25. Krakauer, J.W., *Arm function after stroke: from physiology to recovery.* Semin Neurol, 2005. **25**(4): p. 384-95.
26. Krakauer, J.W., *Motor learning: its relevance to stroke recovery and neurorehabilitation.* Curr Opin Neurol, 2006. **19**(1): p. 84-90.
27. Lanzi, A.B.M.P.R.M.P.L., *An integrated low-cost system for at-home rehabilitation*, in *In Virtual Systems and Multimedia (VSMM)*. 2012. p. 553-556.
28. Levin, M.F., J. Kleim, and S. Wolf, *What do motor recovery and compensation mean in patients following stroke?* Neurorehabilitation and neural repair, 2008.
29. Lo, A.C., et al., *Multicenter randomized trial of robot-assisted rehabilitation for chronic stroke: methods and entry characteristics for VA ROBOTICS.* Neurorehabil Neural Repair, 2009. **23**(8): p. 775-83.
30. McAfee, A.G.P. *Incentivizing high-quality user-generated content*. in *In Proceedings of the 20th International Conference on World Wide Web (WWW'11)*. 2011. New York, NY: ACM.
31. P, S.J.R.S., *Metadata Standards Roundup.* IEEE MultiMedia, 2006: p. 84-88.
32. Pedro, S.S.S.C.W.N.J.S. *ow useful are your comments? Analyzing and predicting YouTube comments and comment ratings*. in *Proceedings of the 19th International Conference on World Wide Web (WWW'10)*. 2010. New York, NY.
33. Rahman, M.A., et al., *Multimedia interactive therapy environment for children having physical disabilities*, in *Proceedings of the 3rd ACM conference on International conference on multimedia retrieval*. 2013, ACM: Dallas, Texas, USA.
34. Ram, C., et al., *Design Factors of Virtual Environments for Upper Limb Motor Rehabilitation of Stroke Patients*, in *Proceedings of the 5th Mexican Conference on Human-Computer Interaction*. 2014, ACM: Oaxaca, Mexico, Mexico.
35. Reinkensmeyer, D.J., et al., *How a diverse research ecosystem has generated new rehabilitation technologies: Review of NIDILRR's Rehabilitation Engineering Research Centers.* J Neuroeng Rehabil, 2017. **14**(1): p. 109.
36. Reinkensmeyer, D.J. and M.L. Boninger, *Technologies and combination therapies for enhancing movement training for people with a disability.* J Neuroeng Rehabil, 2012. **9**: p. 17.
37. Reinkensmeyer, D.J., et al., *Computational neurorehabilitation: modeling plasticity and learning to predict recovery.* J Neuroeng Rehabil, 2016. **13**(1): p. 42.

38. Rikakis, T., *Utilizing media arts principles for developing effective interactive neurorehabilitation systems.* Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 1391-4.
39. Rikakis, T., et al., *Semi-automated home-based therapy for the upper extremity of stroke survivors*, in *ACM PErvasive Technologies Related to Assistive Environments (PETRA '18)*. 2018: Corfu.
40. Tang, R.T.X.-D.Y.S.B.J.J.A., *Physio@Home: Exploring Visual Guidance and Feedback Techniques for Physiotherapy Exercises.* , in *Conference on Human Factors in Computing Systems* 2015: New York, NY, USA. p. 4123-4132.
41. Tanner, S., *Managing containers, content and context in digital preservation: Towards a 2020 vision.* The Society for Imaging Science and Technology, 2006.
42. Wolf, S., et al., *Assessing wolf motor function test as outcome measure for research in patients after stroke.* Stroke, 2001. **32**(7): p. 1635-1639.
43. Wolf, S., et al., *Effect of constraint- induced movement therapy on upper extremity function 3 to 9 months after stroke: the excite randomized clinical trial. J.* Jama, 2006. **296**(17): p. 2095-2104.
44. Zhang, H.T., S.Y, *Multimedia Systems* 1995. **2**: p. 256.

Appendix A

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Informed Consent for Participants
in Research Projects Involving Human Subjects

Title of Project: Evaluating interactive systems for semi-automated assessment of movement functionality and movement quality of non impaired subjects

Investigator(s):	Dr. Aisling Kelliher	aislingk@vt.edu/ 540 231-0415
	Dr. Thanassis Rikakis	rikakis@vt.edu / 540 739-0051
	Dr. Jia-bin Huang	jbhuang@vt.edu / 217 55-2771
	Jinwoo Choi	jinchoi@vt.edu / 540 998-1587
	Edward Coe	eoccoe@vt.edu / 540-505-3447

I. Purpose of this Research Project

The purpose of this study is to evaluate the robustness and accuracy of a prototype camera based system designed for conducting stroke rehabilitation exercises in the home. This prototype system uses computer vision techniques, smart objects, and interactive digital feedback to assist stroke survivors in completing rehabilitation activities. This study aims to evaluate the current prototype system with unimpaired, healthy adults. During the study, 20 healthy adults (10 women and 10 men; both left and right handed; no known neurological or neuromuscular disorders; and ranging from 18 – 85 years of age) will be asked to perform sequences of tasks related to functional everyday activity (e.g. picking up a cup). Participants will be asked to perform the tasks at comfortable/normal speed and as accurately as possible. They will also be asked to perform the tasks simulating typical performance errors observed in therapy with stroke survivors (e.g. dropping a cup)**.** The outcomes of the study will be used to improve the system, and will be used to support a series of publication submissions on the design and evaluation of the system. Findings from the study may also be used as part of investigator Jinwoo Choi's dissertation.

II. Procedures

The research will be conducted in one study session located in the Moss Arts Center at Virginia Tech. Should you agree to participate, you will be asked to perform a set of repetitive reaching and grasping tasks with a fixed set of 6 objects while seated at a table. There will be a total of 12 tasks, with each task repeated 4 times (2 correctly, 2 with requested errors), for a total of 48 activities.

At the completion of each set of 4 activities, you will be asked to use the tablet interface on the study table to answer one multiple choice question concerning the relationship of the task to 3 suggested activities of everyday living (e.g. folding a towel). You will then be asked to perform one of the suggested activities using the previous task object (e.g. reach out and pick up the object as if you were folding a towel).

A complete session will take up to 80 minutes, including time for introductions and consent procedures; system setup and calibration; task instruction; task activities; and completion of the multiple choice question at the end of each task.

Each session will be video taped by the system camera located above the study table, and by a secondary video camera place alongside and slightly behind you.

III. Risks

The overall study involves minimal risk to you. One possible risk is that you may become tired or fatigued during the study. You will be allowed to take breaks during the study to help reduce the risk of becoming tired.

IV. Benefits

By participating in this study, you will learn about new advances in computer vision, stroke rehabilitation and human computer interaction. Your participation will also enable you to make a contribution to the development of new home-based stroke rehabilitation systems.

No promise or guarantee of benefits has been made to encourage you to participate.

V. Extent of Anonymity and Confidentiality

You will be assigned and represented by a subject number throughout the study. Your subject number will only be known by the investigators and a sheet linking subject numbers to subject names will be locked in a filing cabinet, accessible only the research team. Video recorded data of your study session will be encrypted and stored on a password protected computer that is also accessible only by members of the research team. Any video images used in publications or dissertation documents will have subject faces completely blurred and unidentifiable.

At no time will the researchers release identifiable results of the study to anyone other than individuals working on the project without your written consent.

The Virginia Tech (VT) Institutional Review Board (IRB) may view the study's data for auditing purposes. The IRB is responsible for the oversight of the protection of human subjects involved in research.

VI. Compensation

You will be compensated with a $20 gift card, pro-rated at $10 for the first half of the study/up to the first 40 minutes, and $10 for the second half of the study/up to the second 40 minutes

VII. Freedom to Withdraw

It is important for you to know that you are free to withdraw from this study at any time without penalty. You are free not to answer any questions that you choose or respond to what is being asked of you without penalty.

Please note that there may be circumstances under which the investigator may determine that a subject should not continue as a subject.

Should you withdraw or otherwise discontinue participation, you will be compensated for the portion of the project completed in accordance with the Compensation section of this document.

VIII. Questions or Concerns

Should you have any questions about this study, you may contact one of the research investigators whose contact information is included at the beginning of this document.

Should you have any questions or concerns about the study's conduct or your rights as a research subject, or need to report a research-related injury or event, you may contact the VT IRB Chair, Dr. David M. Moore at moored@vt.edu or (540) 231-4991.

IX. Subject's Consent

I have read the Consent Form and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent:

____ Date__________
Subject signature

____
Subject printed name

Demographic Information Questions

1) My name is: ______________________________

2) I identify my gender as:
   [ ] Man
   [ ] Woman
   [ ] _________ (please fill in the blank)
   [ ] Prefer not to disclose

3) My age is: _______________________________

4) I am:
   [ ] Right-handed
   [ ] Left- handed

Screening Questions
The screening questions will be administered by email, over the telephone, or in person. Once contact is initiated by a potential subject, the investigator will either next send the email and/or follow the script below to obtain consent to gather screening data. If consent is granted, the investigator will then either send the email and/or follow the script below to determine participant eligibility.

Consent Email:

Hello,

Thank you for your interest in participating in our study. In order to determine your eligibility for this study, we would like to ask you some screening questions. Any information you share with us will be kept completely confidential and is only being gathered to determine eligibility to participate. If you are not eligible to participate, any data you share with us as part of the screening process will be removed from the study and destroyed.

Please let us know by return email if we have your consent to ask you the screening questions.

Many thanks,

Jin-woo Choi
Dr. Aisling Kelliher
Virginia Tech

Consent Script:

Thanks for your interest in participating in our study.

In order to determine your eligibility for this study, we would like to ask you some screening questions. Any information you share with us will be kept completely confidential and is only being gathered to determine eligibility to participate. If you are not eligible to participate, any data you share with us as part of the screening process will be removed from the study and destroyed.

Do we have your verbal consent to now ask you the screening questions?

Thank you.

Screening Email:

Hello again,

Thank you for agreeing to answer the following screening questions. Please answer yes or no after, or underneath, each question.

1) Are you aged between 18 to 85?

2) Do you have any known neurological or neuromuscular disorders?

3) Are you in general good health?

4) Are you proficient in written and spoken English?

5) Are you able to reach for, grasp, and pick up objects using your dominant/writing hand?

After receiving your response, we will determine your eligibility. If you are eligible, we will contact you to set up a time to visit the study location at the Moss Arts Center.

Please let us know also if you have any further questions or concerns,

Many thanks,

Jin-woo Choi
Dr. Aisling Kelliher
Virginia Tech

Screening Script:

Part 1:
Thanks for agreeing to answer the following screening questions. Please answer yes or no after each question.

1) Are you aged between 18 to 85?

2) Do you have any known neurological or neuromuscular disorders?

3) Are you in general good health?

4) Are you proficient in written and spoken English?

5) Are you able to reach for, grasp, and pick up objects using your dominant/writing hand?

Part 2:

If the participant is deemed eligible based on their positive/yes answers, the script will be as follows -
Based on your answers, I believe you are eligible for the study. We will now find a time for you to visit the study location at the Moss Arts Center.

Thanks

If the participant is not deemed eligible based on their positive/yes answers, the script will be as follows -
Based on your answers, I do not believe you are eligible for the study. Moving forward, any records of your answers will now be destroyed. We thank you for your time.

Tablet Interface Questions

1) Which one of the following three activities does the task you just completed remind you of?

The tablet interface will show the participants a randomized selection of three of the options below:

   a. Folding a towel
   b. Lifting a cup
   c. Brushing your teeth
   d. Opening a door with a key
   e. Holding a book
   f. Picking up a fork or spoon
   g. Buttoning clothes
   h. Something else

2) Can you please pick up the object or objects on the table as if you were performing the following task?

   The tablet interface will show the participants one of the three options from the options shown in the randomized selection in question 1.