DESIGN AND EVALUATION OF A PASSIVE HYDRAULIC SIMULATOR FOR BICEPS SPASTICITY

BY

YINAN PEI

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Mechanical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Advisors:

Professor Elizabeth T. Hsiao-Wecksler

ABSTRACT

This thesis presents the design framework for a passive (unpowered) clinical training simulator using purely mechanical components to help improve the accuracy and reliability of clinical assessment. In the scope of this thesis, a prototype simulator was developed to replicate a common abnormal muscle behavior, biceps spasticity, at different levels of severities. Spasticity is often found in patients with stroke, spinal cord injuries, and other neurological disorders causing abnormal motor activity. The current assessment of spasticity is via in-person evaluation using qualitative clinical scales, and the accuracy and reliability of this method heavily depend on assessors' previous training and clinical experience. However, the current training methods cannot provide students sufficient amount of practice, resulting in lack of proficiency and missing clear understanding of spasticity behavior. The motivation of this project is to develop a self-contained, unpowered simulator to complement the current clinical assessment training. The design process started by characterizing the main behavioral features of the spasticity and selecting the appropriate mechanical design features that provides haptic feedback comparable to a spastic biceps muscle. The prototype was further validated by a two-stage evaluation process. The first part of evaluation involved examining the performance of individual mechanical design features and their combined performance through benchtop experiments. In the second part of evaluation, clinicians were invited to assess the replicated spasticity behavior and to compare the simulation with their previous experience interacting with actual patients. The benchtop performance and clinical feedback help design iteration and provide insights into the future development of the training simulator. Preliminary results suggested the feasibility of using the simulator as a training tool to teach spasticity assessment in a classroom setting.

ACKNOWLEDGEMENTS

I would like to first thank my advisor, Professor Elizabeth Hsiao-Wecksler, who gave me the opportunity to study and work in the Human Dynamics and Controls Lab (HDCL). After getting my undergraduate degree in Civil Engineering, I switched my field to Mechanical Engineering (ME) three years ago. Without much previous experience, finding a ME lab for me was as hard as finding a first job, but I was lucky to find the HDCL and get to know Liz. Her appreciation of diversity in lab culture made me feel comfortable and welcome. I found my skillset was valued and my aspiration was shared by my peers. I would always appreciate her trust and the opportunities she gave me. With her tireless guidance and patient advising, I became a better engineer and researcher. It was her who introduced me to the fascinating world of ME and academic research.

I would like to thank Prof. Randy Ewoldt and Dr. Chris Zallek who provided so many supports in my project. Every meeting with Prof. Ewoldt was inspiring, and his brilliant intuitions and insightful visions always helped me see through problems. Whenever I was stuck in a problem, I often asked myself how Prof. Ewoldt would approach it. Dr. Zallek helped coordinate and arrange the clinical evaluation of the training simulator and many other clinical demo opportunities. His rich clinical experience and passion for technology-driven medical solutions helped me jump out of the role of engineer and see the design problem in more holistic perspective. Without Dr. Zallek, many clinical aspects of this project would not be possible.

I would like to thank past and present HDCL members, Carrie, Alan, Matt, Chenzhang, Cathy, Leo, Nick, Maria, Ayush, and Owen. Carrie, thank you for helping me quickly catch up with the simulator project and teaching me generously on everything from electronics, machining, coding, and design. You are one of the best engineers I ever know. Cathy and Owen, I remember and miss those good time when we were brainstorming creative and fun design ideas for the simulator. Chenzhang, Nick and Ayush, thank you for your countless design and manufacturing suggestions. Maria, thank you for your generous help in statistical analysis and pointing me to all kinds of resources. Leo, it was my pleasure working with you in the same project and I'm glad we made it through so many demos, symposiums, and field trips together with each other's support.

Besides the people I met at UIUC, I would like to thank all my friends, Niantao, Junyi, Zhaochen, Lechen, Gaoran, Yuanxu, Yuqi, Jiwei, Yiming, for their continual friendship and support. Old friends are like stars. Although you don’t see them all the time, you know they are always there for you. Getting to know you all is one of the best things happened to me and this paragraph is for you.

Lastly, I would like to thank my Mom, Dad, Jiayue (Julie), and her parents for their unconditional love. Mom and Dad, you always encourage me to follow my passion and respect my decisions. You value kindness and integrity over personal success. You are the best parents and I’m lucky to be your son. Jiayue, thank you for always believing in what I do and being there for me in the last five years. Without you, my life would become starless.

This project was funded by the Jump Applied Research for Community Health through Engineering and Simulation (ARCHES) program. Special thanks to the following for their assistance with the clinical evaluation and device testing: David Lanier, OTR/L, ATP; Shannon Fraikes, PT; Jianxun Zhou, MD, PhD; Dronacharya Lamichhane, MD; Dyveke Pratt, MD; Prof. Mariana Kersh, PhD; Hafizur Rahman.

TABLE OF CONTENTS

CHPATER 1: INTRODUCTION

1.1 SPASTICITY

1.1.1 Upper Motor Neuron Syndrome

Spasticity affects 12 million people worldwide [1], and it is an important clinical manifestation of upper motor neuron (UMN) syndrome [2–6]. UMN lesion or syndrome is a consequence of neurological damages that affect partial or entire descending motor pathways, such as cerebral palsy, spinal cord injury, stroke, multiple sclerosis, etc. [2–6] (approximately 80% of cerebral palsy and multiple sclerosis patients have varying levels of spasticity [1]). The clinical manifestations of UMN lesion consist of positive and negative phenomena, where positive phenomena are caused by abnormal increase in motor activity and negative ones are due to reduction in motor activity [6]. Spasticity is a typical positive phenomenon of UMN syndrome and other positive features include increased tendon reflex, positive Babinski sign, clonus, spasms, and so on [6]. The disinhibition of spinal reflexes due to interrupted UMN pathways are often considered to contribute to the motor overactivity and give rise to these positive phenomena [7]. On the other hand, negative signs of UMN syndrome include paresis, loss of fine motor control, loss of dexterity, and easy muscle fatigability, etc. [6]. While positive phenomena serve as more important diagnostic signs clinically, the negative ones more tend to lead to functional lose and disability, resulting in discomfort in daily activities and reduced quality of life [5,6].

1.1.2 Mechanism, Behavioral Features, and Quantification of Spasticity

Lance et al. proposed a widely accepted physiological definition of spasticity and described spasticity as "a motor disorder characterized by a velocity dependent increase in tonic muscle reflexes (muscle tone) with exaggerated tendon jerks, resulting from hyperexcitability of the stretch reflex" [8]. This definition highlights the unique speed-dependency of spastic hypertonia (excessive muscle tone), which is often used as an important clinical sign to distinguish spasticity from other types of abnormal muscle behaviors such as rigidity, dystonia, etc.

The occurrence of this hypertonia was found to be associated with hyper-excitable reflex contractions that resist the passive stretch of the affected muscle [9], and these reflex contractions were referred to as tonic stretch reflex (TSR) by Lance and researchers [6,8]. In healthy individuals, the TSRs were present only when stretching their joints at very fast speed (>200°/s), while these reflex contractions did not activate and contribute to the muscle tone when the stretch speed was below a threshold [10,11].

This speed dependency of the hypertonia was attributed to the alteration of the TSR in spastic muscles [6]. It has been found that, in spastic muscles, not only the reflex muscle contractions could be induced at a speed as low as 35°/s [10], but also the magnitude of contraction was positively correlated with the stretch speed [12–14]. There were different hypotheses proposed about what type of alteration was imposed on the TSR by UMN lesions, such as increased reflex gain, lowered reflex threshold, etc. [7,10,14–17]. Although there was still no consensus on how TSR was altered, a common observation was that the UMN lesion resulted in overactive TSR that were sensitive to stretch speed [10,14–17]. In addition, a few other studies pointed out that the TSR in spasticity patients was also muscle length-dependent and became less sensitive when the muscle was in lengthened position [12,18,19].

Except speed-dependent hypertonia, this overactive TSR in spastic muscle is manifested as several other behavioral features often observed in the clinical setting. The most well-known clinical characteristic of spasticity is the "catch-release" behavior or the "clasp-knife" phenomenon in muscle tone [6], which is an important sign for clinical assessment using qualitative scales [20,21]. The "catch" or "spastic catch" refers to a sudden increase in muscle tone at a certain joint position (i.e., catch angle) [6]. In the low end of the clinical scales (i.e., lower spasticity severity) [20,21], the "catch" is usually described to be followed by a quick drop of muscle resistance, called the "release". The presence of catch was considered to be a result of the overactive TSR being triggered by fast stretch speed [6]. After the onset of muscle tone due to stretch reflex contractions, the stretch movement slowed down and fell below the reflex threshold, bringing a stop to the reflex contraction and eventually releasing the hypertonia [6].

As severity increased, clinical scales described the muscle tone as to be considerably increased, but the release behavior is rarely mentioned anymore [20,21], implying that the TSR might not be the only source of increased muscle tone. Several studies suggested that the spastic hypertonia could be attributed to neural and non-neural components [6,7,22]. The neural component of resistance arose due to the hyperexcitability of TSR as discussed previously. The presence of a hyperactive TSR interfered and resisted daily movements and activities [6,7], so patients attempting stretch flexor muscles often encountered involuntary resistance and their joints were confined to a flexed position [7]. As the affected joints/muscles remains in flexed/shorted positions for a prolonged period of time, the non-neural component of hypertonia started to develop as a consequence of soft tissues (i.e., tendons, ligaments, joints, etc.) becoming contracted and less compliant [6,7]. This mixture of neural and non-neural components of hypertonia complicated the clinical diagnosis of spasticity patients and created difficulties in designing therapeutic interventions [6,7]. For example, the release of neural hypertonia might be compensated by the additional resistance due to biomechanical changes, so it would be more difficult for clinicians to identify the release and interfere with their judgements. Generally to distinguish these two components, antispastic medication such as botulinum toxin injection was used first to alleviate the neural component from the overall hypertonia in order to isolate the hypertonia due to soft tissue changes [6,7].

Limited range of motion (ROM) is another common behavioral feature of spasticity [6,7,23–25]. Spasticity resists the lengthening of muscle and triggers the involutory hypertonia, so patients tend to keep the affected muscles in the shortened position, resulting in soft tissue stiffening and contractures [6]. Furthermore, after neurological injuries, certain muscles tend to become immobilized in a shortened position due to paresis, such as lower limb extensors, upper limb flexors, etc. [25]. Although acute care usually involves helping patients stretch the affected muscle for extended periods to restore their full ROM [26], chronologically, immobilization in a shortened position is the main contributor in developing soft tissue contracture [25] and causing limited ROM for chronic spasticity patients [23–25].

Many studies have been undertaken to further quantify these behavioral features, in hope of gaining knowledge of the spasticity mechanism and hence improving the current assessment and treatment approaches [27–46]. Although previous studies varied in protocols, instruments, and subject populations, general characteristics of spasticity behaviors were observed. Several studies demonstrated that spastic muscle resistance intensified with increasing input stretch speed [30,31,41,43,47]. As the spasticity patients became more severe (increasing MAS scores), (a) the abnormal muscle tone became higher [27–29,41,44,45,48], (b) catch angle was encountered earlier in ROM [28,41,44,45,49], and (c) ROM became more limited [27–29,41]. Although a variety of quantitative measures were used to quantify or model the behavior of spasticity in the previous studies, since the definition and understanding of spasticity mechanism is still incomplete, there is still no conclusion on how to measure spasticity most properly [50].

1.1.3 Clinical Evaluation and Training Methods

To effectively design treatment and management plans to spasticity patients, an objective and accurate clinical assessment of a patient's severity is essential. Nevertheless, the instrumented measurements previously discussed are restricted to be used in a well-controlled laboratory setting, but they are not easily applicable in the daily clinical setting [39,51]. Therefore, current clinical assessment relies on qualitative scales such as the Modified Ashworth Scale (MAS) [20] and the Modified Tardieu Scale (MTS) [21]. During an examination, the patient will be instructed to relax and let the clinician passively stretch (i.e., lengthen the muscle when it is not activated) the affected muscle at one or multiple speeds. Based on the resistance felt, the clinician will assign a score to represent the patient severity [20,21]. Although the MTS was considered to be more appropriate for spasticity assessment by some researchers [33,39,50,52], since it takes into account the speed-dependence feature by examining the muscle at multiple speeds [21], clinically the MAS is more commonly used due to its straightforward protocol [52,53]. The remaining thesis will mainly discuss the MAS.

Table 1.1:Modified Ashworth Scale for assessing spasticity [20].

Score	Description
0 (0) [a]	No increase in muscle tone
1 (1)	Slight increase in muscle tone, manifested by a catch and release or by minimal resistance at the end of the range of motion when the affected part(s) is moved in flexion or extension
1+ (2)	Slight increase in muscle tone, manifested by a catch, followed by minimal resistance throughout the reminder (less than half) of the range of motion
2 (3)	More marked increase in muscle tone through most of the range of motion, but affected part is easily moved
3 (4)	Considerable increase in muscle tone, passive movement difficult
4 (5)	Affected part is rigid in flexion or extension

[a] Numbers in parenthesis represents a variant of the Modified Ashworth Scale [42] and this convention was used throughout this thesis.

Table 1.2: Modified Tardieu Scale for assessing spasticity [31].

Score	Description
0	No resistance throughout passive movement
1	Slight resistance throughout, with no clear catch at a precise angle
2	Clear catch at a precise angle followed by release
3	Fatigable clonus (<10 secs) occurring at a precise angle
4	Un-fatigable clonus (>10 secs) occurring at a precise angle
5	Joint immobile

Many researchers have evaluated the reliability of spasticity assessment using the MAS [20,54–60]. Previous reliability studies have reported reliability measures in terms of intra- and inter-rater reliability, where intra-rater reliability was calculated to measure the self-consistency of each rater, and inter-rater reliability was assessed to determine the agreement among all raters [61]. Although these studies examined the reliability of the MAS on accessing spasticity in different muscles, they have suggested a moderate reliability associated with the use of MAS in general [20,54–60].

Common statistics used in those reliability studies included intraclass correlation coefficient (ICC) [54,58,62], percent agreement [55,63], Spearman's correlation coefficient [56,59], Kendall's tau coefficient [20,63], and kappa statistic [57,60,64]. ICC and kappa statistic were used most frequently, but their use should be distinguished by considering the number of repetitions and raters, as well as the ordinal nature of the MAS. Weighed kappa, an extension of the original Cohen's kappa statistic [65] to deal with ordinal data (e.g., the MAS scores), can be only used to estimate intra-rater reliability when there are two repetitions (e.g., test-retest reliability) or inter-rater reliability between two raters [66]. Fleiss's kappa is another extension of the original kappa, and although it can be used for more than two repetitions or raters, it is only appropriate for nominal data [67]. When evaluating the intra- and inter-rater reliability of the MAS with more than two repetitions or raters involved, ICC should be utilized and it is considered mathematical equivalent of the weighted kappa for ordinal data [68]. In addition, Krippendorff's alpha is another alternative statistic in assessing intra- and inter-rater reliabilities for ordinal data given more than two raters or repetitions [69,70]. Based on the above discussion, this thesis used ICC and Krippendorff's alpha to estimate the reliability of clinical study participants in Section 2.4 in Chapter 2.

As previous reliability studies have suggested [20,54–60], the accuracy of clinical scale such as the MAS heavily depends on the assessor's previous training and clinical experience, as the narrative descriptions of qualitative scales are open to interpretation (Table 2.1) [59,63,64,71]. Hence, it is crucial for young healthcare professionals training for spasticity assessment to have a clear understanding of different levels of spasticity and to gain sufficient hands-on assessment experience before entering the job site [71]. But the current training methods, relying on inviting practice patients or asking students to pretend to be the patient for each other, usually result in inconsistent outcomes due to variability in training or lack of proficiency due to limited availability of practice patients [72]. Therefore, to complement current training practices, there is a clear need to develop a spasticity training simulator that can generate realistic and consistent replication of spasticity at different levels and is easily accessible to trainees for frequent practice.

1.2 TRAINING SIMULATORS

1.2.1 Existing Research Simulators

Three research training simulators were found in the literature that were developed to replicate elbow spasticity [35–37,45,48,72–75]. Okamura and colleagues developed a haptic device that replicated elbow spasticity in a child's arm [73]. The device generated a torque to simulate muscle resistance using a servo disc brake controlled by a DC brushed motor. The torque was proportional to the stretch speed used and only began after a certain angular position (similar to catch angle). The angular position was calculated as a function of stretch speed, starting joint position, and severity. However, no detailed information was published regarding their mathematical model and experimental results, so the performance of the device was undetermined.

Komeda and colleagues developed the initial training simulator prototype to replicate spasticity and rigidity (another abnormal muscle behavior often found in Parkinson's diseases patients) in the biceps [72,74]. The simulated spastic muscle tone was generated through a combination of geared DC servomotor and magneto-rheological (MR) fluid viscous brake. However, no mathematical model of spasticity was proposed. The torque profile was recorded when a physical therapist pretended to be a patient and simulated the muscle tone, and the simulator replicated the behavior by playing back the recorded torque. In the subsequent work [35–37], quantitative data were collected from 11 patients and a mathematical model of spastic muscle torque was proposed, which took the stretch speed, catch angle, muscle viscosity, and muscle elasticity into consideration. Experimental results found the simulated torque followed the trend of clinical measures of spasticity patients.

Damiano and colleagues developed the Haptic Elbow Spasticity Simulator to recreate biceps spasticity. The simulator deployed an DC brushless motor with a speed-reduction mechanism [48,75]. To build a mathematical model for spasticity, they developed a portable measurement device and collected quantitative data from four children with cerebral palsy (MAS 1-4). They modeled the spastic joint behavior as a linear mass-spring-damper system and superimposed a catch-release behavior, which consisted of three phases, i.e., pre-catch, catch, and post-catch. The simulated torque profile created by the device matched well with the clinically recorded data in all simulated levels (MAS 1-4). The simulator was further validated by eight clinicians and received a score of realism of 7.63±0.92 out of 10. The percent agreement between clinicians' assessments and setup was 84.4 ±12%. To address the limitation of small patient database, Park and colleagues conducted a large clinical study with 18 patients of varying severities (MAS 1-3) and 22 clinicians, in total 387 trials recorded, and developed a more sophisticated model of spasticity [45]. No further experimental results of the simulator were available.

1.2.2 First-Generation Training Simulator and Limitations

Previous designs using some type of electro-mechanical actuation involved the use of expensive electrical components (i.e., electric motor, sensor, motor driver, hardware controller, etc.) and required an external power supply, limiting their affordability and portability. Although their preliminary results were found promising in recreating realistic spasticity behaviors, to the author's knowledge, none has been deployed outside of academia due to those challenges.

Our group proposed a novel passive (unpowered) training simulator, which utilized a custom viscous hydraulic damper and a linkage system to provide speed- and position-dependent haptic feedback by passively responding to the user's input motion without requiring a power supply [76,77]. The prototype training simulator has an appearance of a human arm with its inertial properties and dimension similar to a 50^{th} percentile Caucasian male's arm. The simulation focused on recreating a speed-dependent increased muscle tone and the catch-release behavior. When the simulator's elbow joint was flexed, the working viscous fluid was forced through orifices on the damper's piston head and generated the speed-dependent resistive force to represent the spastic triceps muscle tone. To replicate different levels of spasticity severity, i.e., MAS 0-4, the piston head was fabricated with five pairs of orifices of different sizes and only one pair was used during operation, while the other four were blocked by a cover plate. The resulting viscosity would increase if a pair of smaller orifice was used. An analytical fluid model was developed to predict the damping force based on input speed and orifice size, and was used to guide the damper design in order to achieve the target viscosity values found in the literature [31,32].

The elbow linkage system converted the rotation of the forearm segment into the linear translation of the damper's piston head. In order to replicate the position-dependent catch-release behavior of spasticity, a Scotch-Yoke linkage with a customized cam and cam follower was used. The rotary-to-linear conversion ratio was governed by the linkage lever arm length, and by this design, the lever arm length was a sinusoidal function of joint angular position. For simplicity and proof of concept, the preliminary catch angle location was assigned to be at the mid-ROM (the total ROM was 82°), so in other words, the linkage lever arm length was the maximum (at the peak of the sine wave) and the user felt the catch (peak resistive torque) at the mid-ROM. After the peak, as the lever arm length gradually decreased in a sinusoidal fashion, the damping force transmitted by the linkage diminished and reached zero at the end of ROM, representing the release behavior.

Bench-top experimental evaluation found that the proposed analytical fluid model of the hydraulic damper matched with experimental results well, and the prototype simulator was capable of replicating general behaviors of spasticity at varying severities (MAS 0-4) [76,77]. However, there were multiple limitations in the first-generation prototype. First, the design was set up to simulate a passive flexion test of the elbow to test spastic triceps muscle. However, we later realized that the spastic biceps muscles were more common and biceps training simulator would have more clinical significance [78,79]. Therefore, the second-generation training simulator was modified to operate in extension to test spastic biceps muscles. Second, the design generated a gradual, sinusoidal catch-release behavior at a fixed catch angle; while the actual behavior observed in the spastic muscle is more abrupt, and the catch angle usually comes earlier in the range of motion as severity of impairment increases [28,41,44,45,49]. Third, the ROM of 82° was too small to accurately represent common ROMs observed in spasticity patients, especially for less severe patients whose ROMs are usually close to healthy individuals. Additionally, the same ROM was applied to all simulated MAS levels; whereas as previously discussed, the ROM tends to decrease with the increase of severity due to soft tissue contracture and muscle stiffening [24,27–29,41,80]. Fourth, from preliminary feedback of three experienced clinicians in the research team, the simulated resistance was found to be too great for low MAS levels (0 and 1) due to the high hydraulic seal friction, and too weak for high MAS levels (2, 3, and 4) compared to their experience with actual spastic muscle tone [81].

1.3 THESIS ORGANIZATION

This thesis presents the design and evaluation of a second-generation training simulator that provides consistent and realistic replications of spasticity to complement the current clinical assessment training. To address the limitations found in the previous prototype, the second-generation design more closely followed the behavioral features of spasticity observed clinically to enhance the fidelity of the replication. The final prototype remained to be a self-contained and portable mechanical design, but it was able to replicate more clinical signs of spasticity and demonstrate how they evolve as the severity changes.

Chapter 1 gives a brief introduction of current understanding in spasticity mechanism and behavioral features and reviews the previous efforts to quantify and characterize its behaviors. This chapter also reviews the common clinical scales used to assess spasticity and previous investigations on their reliability. Moreover, challenges and limitations on current training approach are discussed and the need for a training simulator is motivated. In the end, the chapter reviews the three existing electro-mechanical training simulators developed at other institutions and a novel purely mechanical one developed in our group. Their experimental performances are summarized, and limitations are discussed.

Chapter 2 presents the development and evaluation of the second-generation training simulator. New mechanical design features of the training simulator were developed to mimic the behavioral features of spasticity. Bench-top evaluations suggested mechanical design features performed in accordance with their design intents and the overall replication of spasticity agreed both qualitatively with clinical scales and quantitatively with the clinical measures of actual spastic muscle tone. The prototype simulator also received positive feedback in the clinical evaluation with seven clinicians in terms of its realism and potential to become a training tool in a classroom setting.

Chapter 3 reviews the development of the second-generation training simulator in Chapter 2 and discusses the advantages and disadvantages of the passive design approach comparing to other electro-mechanical designs. This chapter also provides recommendations on future directions of the training simulator development.

CHAPTER 2: PASSIVE HYDRAULIC CLINICAL TRAINING SIMULATOR FOR BICEPS SPASTICITY

2.1 ABSTRACT

This paper presents the framework for developing a passive (unpowered) mechanical training simulator for replication of biceps spasticity to complement current clinical assessment training. The clinical assessment of spasticity relies on qualitative scales such as the Modified Ashworth Scale (MAS), and its reliability and accuracy heavily depend on the assessor's training and proficiency. The training simulator was developed to mimic three main behavioral features of spasticity, i.e., abnormal muscle tone, "catch-release" behavior, and range of motion (ROM) reduction. The simulator can replicate varied levels of spasticity (representing MAS levels 0 to 4) using a unique combination of three adjustable mechanical design features, i.e., resistance level, catch angle, and ROM selectors. Bench-top evaluation was conducted to examine the performance of individual design features, as well as the overall performance. Results qualitatively agreed with the description of the MAS. Peak simulated resistive torque fell within the clinical measures from actual spasticity patients for MAS 1-4 but lower at MAS 0 (0.9, 3.5, 4.2, 6.9, 9.8 Nm for MAS 0-4, respectively). Seven clinicians evaluated the simulator and were asked to identify the simulated MAS level during a blinded assessment and score the realism of the simulated spasticity behavior using a five-point scale (1: too low, 3: about right, 5: too high) during a disclosed assessment. The overall percent agreement between clinicians' judgments and simulator setup was 76% ± 12. The mean realism score throughout MAS 0-4 were 2.84. Preliminary results suggested the simulator's potential to help future students learn and practice the basics of spasticity assessment.

2.2 INTRODUCTION

Spasticity is a common consequence of upper motor neuron (UMN) syndrome and is usually found in patients with neurological disorders (e.g., spinal cord injuries, stroke, multiple sclerosis) that affect descending motor pathways [2–4]. As a positive phenomenon[1] of UMN syndrome, spasticity is characterized by a speed-dependent increased tone (resistance) in the affected muscle when induced by passive movement [2–4,6,8]. When passively stretching the affected muscle during a clinical evaluation, unique behavioral features of spasticity are observed [2,6,24,7,25]. There is a sudden increase in muscle tone, called the "catch", at a certain joint position (catch angle), and the catch angle location may vary with the severity. After the abrupt "catch", a quick drop of muscle resistance, called the "release", and together they are usually referred to as the "catch-release" behavior or the "clasp-knife" phenomenon [2,6,24,7,25]. Furthermore, some patients develop soft tissue contracture and experience limited range of motion (ROM) when their joints are immobilized due to paresis or chronically remain in a flexed position due to involuntary muscle tone [6,7,25].

To effectively design treatment and management plans for spasticity patients, an objective and accurate clinical assessment of spasticity severity is essential. Current clinical evaluation relies on qualitative scales such as the Modified Ashworth Scale (MAS) [20]. During an examination, the patient will be instructed to relax and the clinician passively stretches the affected muscle with increasing speeds, and based on the resistance felt, the clinician will assign a score [20,21]. The accuracy of this method heavily depends on the assessor's previous training and clinical experience, as the narrative descriptions of qualitative scales are open to interpretation [64,71,59,63]. Hence, it is crucial for young healthcare professionals training for spasticity assessment to have a clear understanding of different levels of spasticity and to gain sufficient hands-on assessment experience before entering the job site [71]. Current training methods relying on inviting practice patients or asking students to pretend to be the patient for each other. This practice can result in inconsistent outcomes due to variability in training or lack of proficiency due to limited availability of practice patients [72]. Therefore, to complement current training practices, there is a clear need to develop a spasticity training simulator that can generate realistic and consistent replication of spasticity at different levels and can easily be accessible to trainees for frequent practice.

---

[1] Positive phenomena of UMN syndrome are usually associated with an abnormal increase in motor activity, while negative ones are connected with reduction in motor activity [6].

Several research groups have developed upper-extremity spasticity training simulators [72,35,37,73,74,48,45,75,36,76,77]. Our group previously proposed a completely passive (unpowered) mechanical design that utilized an upper arm segment with a novel hydraulic damper and a Scotch-Yoke linkage system at the elbow [76]. To systematically design the damper performance toward clinical viscosity targets, an analytical fluid model was developed to aid the design process and predict the damping force based on input speed and orifice size. Bench-top experiments found that the proposed analytical model for the hydraulic damper matched with experimental results well and the prototype simulator was capable of replicating general behaviors of spasticity [76,77]. However, there were multiple limitations in this first-generation prototype (Table 2.1, Column 5). First, it could only generate a gradual catch-release behavior (following a sinusoidal resistance vs. angle path) at a fixed catch angle (~50% of ROM); while the actual behavior observed in the spastic muscle is more abrupt and the catch angle usually will be encountered earlier in the ROM as severity of impairment increases [45,28,44,49,41]. Second, the ROM was set to a single value (82°). This value was too small to correctly represent common ROMs observed in patients. Additionally, ROM is known to decrease as severity increases [24,28,41,29,27,80]. Third, from unpublished initial feedback of three experienced clinicians, the simulated resistance was found to be too great for low MAS levels (0 and 1) and too weak for high MAS levels (2, 3, and 4).

To address these limitations of the previous design, we adopted a parallelism design approach which realized the behavioral features of spasticity through appropriate mechanical design features that generated similar kinematic and kinetic responses (Table 2.1). The second-generation training simulator presented in this paper attempts to replicate three main behavioral features of spasticity: a) abnormal muscle tone, b) catch-release behavior, and c) limited ROM (Table 2.1, Column 1). These three features were further quantified by six parameters, i) speed-dependent resistance, ii) severity-dependent resistance, iii) catch duration, iv) catch angle location, v) release duration, and vi) ROM magnitude (Table 2.1, Column 2). These features and parameters were mechanically replicated by the simulator (Table 2.1, Columns 3-4). In the scope of this work, the training simulator was capable of demonstrating all six parameters, and three of them were adjustable, i.e., severity-dependent resistance magnitude, catch angle location, and ROM magnitude to simulate patients with different severities (MAS 0-4) (Table 2.1, Columns 5-6).

Table 2.1: The functional comparison between spasticity behavioral features and training simulator mechanical design features.

Spasticity behavioral feature	Spasticity parameter	Simulator mechanical design feature	Simulator parameter	Gen1 Simulator	Gen2 Simulator
Abnormal muscle tone	Speed-dependent resistance	**Resistance level selector**	Viscous fluid	Constant viscosity for a certain severity	Constant viscosity for a certain severity
	Severity-dependent resistance		Hydraulic damper with selectable viscous effect	MAS 0-4 (too high at low levels and too low at high levels)	MAS 0-4
Catch-release behavior	Catch duration	**Catch angle selector**	Sudden engagement of damper	Gradual (~0.9 second)	Abrupt (0.2-0.3 second)
	Catch angle location		Engagement timing of damper	85°	No catch, immediate (50°), early (80°), late (135°)
	Release duration		Scotch-Yoke linkage system	Gradual (~0.9 second)	Moderate (0.3-0.6 second)
Limited ROM	ROM magnitude	**ROM selector**	Mechanical stop	82°	130°, 115°, 95°, 80°, 60°

2.3 DEVICE DESIGN

2.3.1 Design Overview

The second-generation hydraulic training simulator has the appearance of a human arm and its segment dimensions and inertial properties were based on a 50th percentile Caucasian male's arm [82] (Fig. 2.1). The frame structures and the stand were fabricated from aluminum. Several ergonomic 3D-printed plastic shrouds were installed on the upper arm, elbow, forearm, and wrist of the simulator to ensure all moving mechanical components were properly covered, and meanwhile mimicked the shape of a human arm. The surface of the shrouds was treated with high-friction spray paint to allow users to grip easily and securely during passive stretch trials.

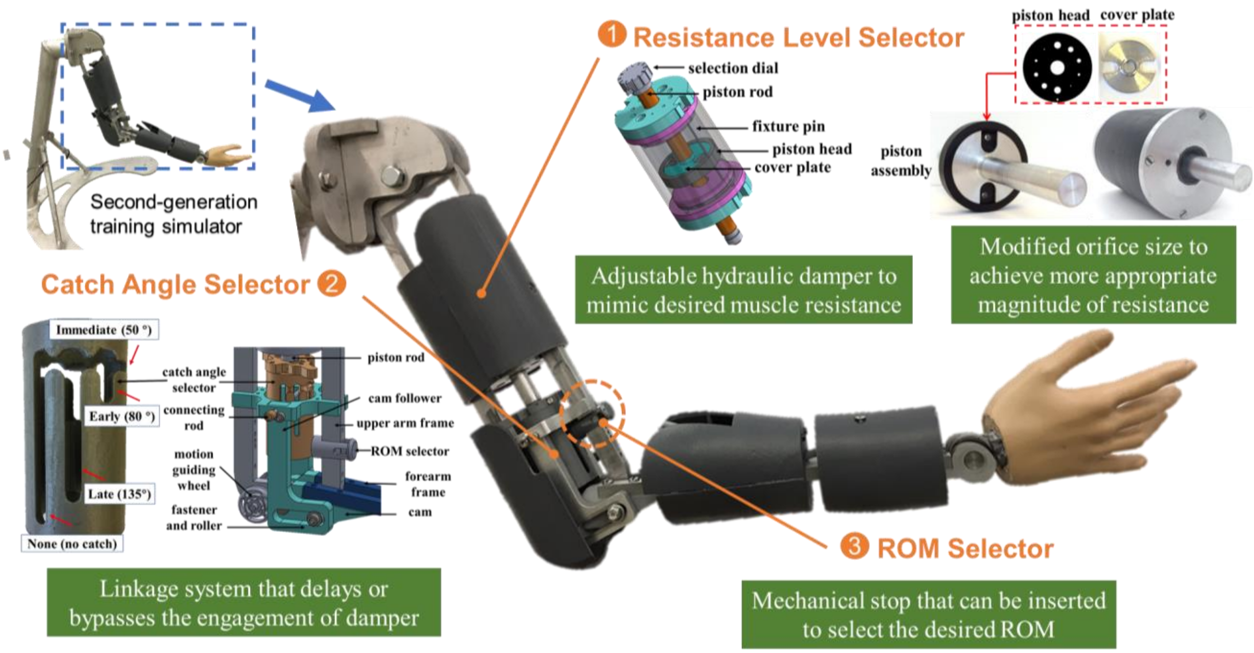


Fig. 2.1: Second-generation passive hydraulic training simulators with three mechanical design features.

2.3.2 Abnormal Muscle Tone

The abnormal muscle resistance was simulated using a custom hydraulic damper with selectable viscous effect. The damper was mounted on the upper arm frame, and as the simulator's elbow was extended, the damper's piston head would move through the viscous fluid and generate a damping force proportional to the stretch speed

(Table 2.1). To simulate muscle resistance to movement for different levels of severity, five pairs of orifices were fabricated on the piston head and only one pair was exposed, while the rest were blocked by a cover plate (Fig. 2.1). By varying the orifice size in operation, different viscous effects could be selected based on the severity of the simulated patient. The damper design was described in detail in [76,77].

In the previous design [76,77], a set of orifice diameters, 1.5, 1.8, 1.9, 2.0, 2.5 mm, were selected to match the clinical target viscosities obtained from [32,31] at MAS levels 0-4, respectively, which were calculated as average viscosity values over several elbow extension-flexion cycles. However, given (unpublished) preliminary clinical feedback from three experienced clinicians of the first-generation simulator and spastic muscle tone reported in other quantification studies [45,44,27], the first-generation viscosity design targets underestimated the resistance level at MAS 2, 3, and 4, while the simulated resistance was too high at MAS 0 and 1 due to the mechanical friction between the piston rod and hydraulic seals.

Due to a lack of clear viscosity design targets, new orifice sizes were iterated from the previous prototype and estimated to be 0.5, 1.0, 1.5, 2.0, and 2.5 mm. Smaller orifice sizes (1.5, 1.0, and 0.5 mm) were used for high MAS levels (2, 3 and 4) to increase the speed-dependency in given severity level. The catch angle selector (discussed below) would bypass the damper completely at MAS 0, while at MAS 1, the catch angle was delayed and would result in a small linkage lever arm length, bringing down the resistance at these two levels.

In addition to mimicking abnormal spastic muscle tone, to provide a more natural joint response and haptic feedback to the user (especially at MAS 0 when the damper was completely bypassed), two linear torsion springs were installed at the simulator's elbow to create an elastic resistance, mimicking the stretching of passive soft tissues. A total effective spring constant (0.00892 Nm/°) was implemented to approximate the value reported for a MAS 0 subject or a healthy individual's elbow joint [27,47] (refer to Appendix D for torsion spring design and testing).

2.3.3 Catch-Release Behavior

During the clinical assessment of spasticity, the assessor will pay close attention to examine the presence of the catch-release behavior and record the location of the catch angle, using these observations to aid their judgement on the patient's severity [6,21], so the realism of the simulated catch-release behavior is crucial to the quality of spasticity replication. We implemented a catch angle selector and an elbow linkage system to replicate this clinical behavioral feature.

The catch angle selector, fixed to the damper piston rod (but can freely rotate relative to the piston rod), was connected to the rest of the linkage system through a metal rod (Fig. 2.1). As the forearm was extended (mimicking biceps stretch), the linkage system would bring the connecting metal rod downward. Since during this “pre-catch” phase, the rod would travel freely in the slot of the catch angle selector, the user would feel minimal resistance. Once the rod reached the end of the slot, if moved further, the linkage would start pulling the damper piston rod. The transition from minimal resistance to the sudden engagement of the damper resulted in an abrupt catch feeling to the user. The catch duration was defined as the time duration between the moments of hitting catch angle and reaching peak resistance, and it was about 0.2-0.3 second in the current design.

Clinically, the locations of the catch angle usually varies with the severity of a patient [45,28,44,49,41]; therefore, four pairs of slots with different lengths were designed into the surface of the 3D printed catch angle selector to create a catch at four different angular locations (immediate, early, late, none) (Fig. 2.1). A longer slot would lead to a later catch angle in the ROM. For each MAS level, the following catch angle was used: MAS 0 – none (no catch), MAS 1 – late (135°), MAS 2 – early (80°), MAS 3 – early (80°), and MAS 4 – immediate (50°) [28,39,40] (refer to Appendix C for more details on how these angles were chosen).

After the catch, the release behavior was created through a Scotch-Yoke linkage system in the elbow. The linkage behaved as a transmission mechanism that converted the rotary motion of the forearm into the linear motion of the damper piston rod for resistance generation. The relationship between the Scotch-Yoke linkage lever arm length (i.e., conversion ratio) and the joint position was described by a sinusoidal function. The phase and magnitude of the sinusoidal function was tuned through two geometric design parameters of the linkage, the slot length on the cam follower and the cam geometry (Fig. 2.1). In the current design, these two parameters were chosen such that the maximum torque would occur at around 90°, which was chosen based on an average catch angle observed in the literature across spasticity patients of different severities [39]. The decrease of lever arm length after 90° would diminish the amount of damping force transmitted to the user and result in the release behavior of spasticity. In the scope of our work, the elbow joint angle was defined relative to upper arm, i.e., the joint angle started from the most flexed position (50°) and ended at full extension (180°). As the design parameters were fixed, the release behavior had a constant position dependency. The release duration was defined as the time duration between the moment of reaching peak resistance and the end of ROM. As the ROM became more limited in higher MAS level, the release behavior was shorter than with the Gen 1 simulator in general, which was about 0.3-0.6 second. Refer to Appendix B for more details.

2.3.4 Range of Motion

During the search of a reasonable set of ROMs, several types of ROM representations were considered, active ROM, passive ROM (pROM), and elbow resting angle. As the simulator was a passive (unpowered) device that cannot be actuated actively (needing external power), only pROM and elbow resting angle needed to be considered. However, no consistent pattern relating pROM with MAS scores were found in the literature [41,27,42], because pROM was achieved by the assessor applying external force to move the patient's joint, and usually the magnitude and direction of this force was not precisely controlled and not consistent among studies. On the other hand, Bhadane et al. reported a negative correlation between elbow resting angle and MAS scores [41]. According to their clinical observations, the ROM design targets were approximated to be 130°, 115°, 95°, 80° and 60° for MAS 0-4, respectively.

2.4 BENCH-TOP EVALUATION

2.4.1 Evaluation Methods and Protocol

Performance was quantified by comparing designed to measured simulated muscle resistive torque and elbow joint angular position and speed for various mechanical design feature settings. Resistive torque was determined by measuring the applied force at the wrist with a load cell (LCM202, OMEGA Engineering; Norwalk, CT) embedded in the forearm shroud with a known lever arm (0.24 m) from the elbow joint center. Joint angular position and speed were determined from one inertial measurement unit (IMU) (MPU9250, InvenSense; San Jose, CA) embedded in each arm segment. The joint stretch speed was calculated as the difference of gyroscope readings (about the axis of rotation) between the two IMUs and the elbow joint angle was calculated from IMU's raw data outputs (linear acceleration, angular velocity, magnetic field intensity) using a Mahoney Filter [83], a sensor fusion algorithm. Sensor readings were collected at 1800 Hz and processed through a microcontroller (Teensy 3.6, PJRC; Portland, OR). All signals were filtered using a 4th order Butterworth filter with a cut-off frequency of 20 Hz.

Prior to each test, the simulator setup for the design features (resistance level, catch angle, and ROM) was adjusted according to the testing condition and then the forearm was reset to the most flexed position (~50°). During the test, the investigator (YP) held the elbow shroud to support the upper arm at ~30° of flexion, while the investigator's other hand held the forearm shroud over the load cell and extended the elbow under a speed prescribed by the testing condition.

Key parameters were extracted from the experimentally measured data, i.e., catch angle, ROM, and peak resistive torque (Fig. 2.2). Catch angle was defined as the joint angular position where the input stretch speed started to drop after reaching its initial peak due to the onset of resistance. ROM was the difference between the starting and ending angular positions. The peak resistive torque was picked as the largest recorded torque. All reported muscle resistance torque values were adjusted to compensate for the inertial effects of the simulator forearm (refer to Appendix D.2 for more details).

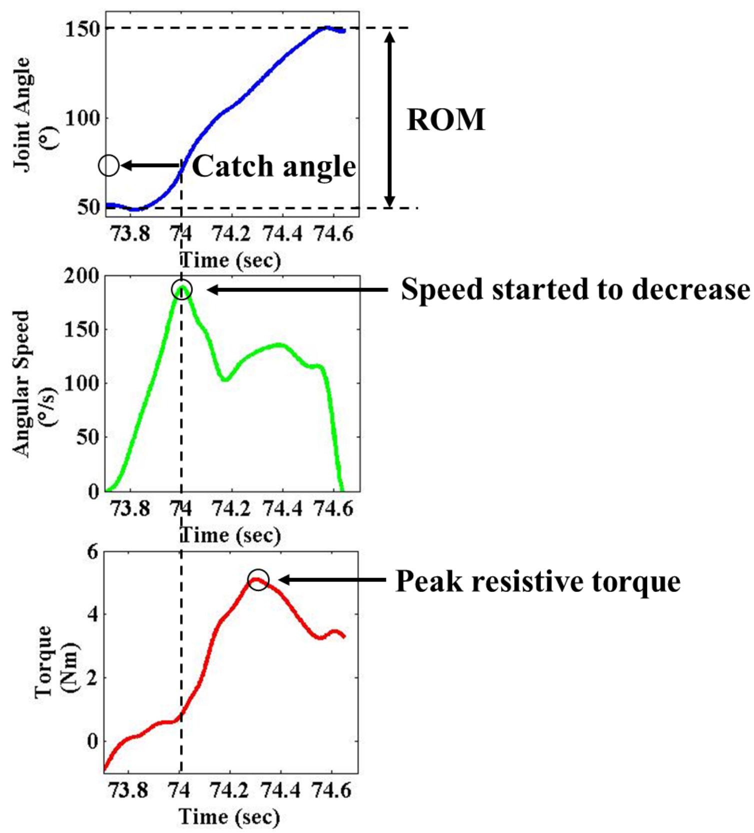


Fig. 2.2: A sample trial of simulated MAS level 2 (Table 2.3) to explain parameter definitions.

The first evaluation was to investigate individually the performance of each mechanical design feature (resistance level, ROM, and catch angle selectors). Three independent tests were designed (Table 2.2) and all testing conditions were repeated for three trials. For the Resistance Level Test, as the magnitude of damping force depended on both piston head orifice size and input stretch speed, each testing condition was performed under three different speeds (slow, medium, and fast) to study the speed-dependency. For the Catch Angle Test and ROM Test, each testing condition was performed under only the fast speed, as these positional features are independent of stretch speed. The stretch speed reported was defined as the average speed throughout the ROM. Speed ranges were slow ($\omega < 20°/s$), fast ($\omega > 80°/s$), and medium (anything in between), chosen similar to actual testing speeds used during clinical assessment [27].

Table 2.2: Three individual mechanical design feature tests, i.e., Resistance Level Test, Catch Angle Test, and ROM Test.

	Testing Condition	Orifice Radius (mm)	Catch Angle	ROM (°)
Resistance Level Test[a]	1	2.5	Early (80°)	95
	2	2	Early (80°)	95
	3	1.5	Early (80°)	95
	4	1.0	Early (80°)	95
	5	0.5	Early (80°)	95
Catch Angle Test	1	0.5	Immediate (50°)	130
	2	0.5	Early (80°)	130
	3	0.5	Late (135°)	130
ROM Test	1	2.5	None	130
	2	2.5	None	115
	3	2.5	None	95
	4	2.5	None	80
	5	2.5	None	60

[a] Each Condition tested at slow, medium, and fast speeds.

For the ROM and Catch Angle Tests, the mean and standard deviation for each measured parameter across all three trials were reported and compared with the design values (Table 2.2). Error Residual was calculated as $(\text{measured value} - \text{design value})$ and Percent Error was calculated as $(\frac{\text{error residual}}{\text{design value}} \times 100\%)$. For Resistance Level

Test, mean and standard deviation of peak resistance torques for each orifice size and average stretch speeds were reported.

The second benchtop evaluation deployed all mechanical design features together to provide a holistic replication of spasticity and each simulated MAS level had its own unique combination of design features (Table 2.3). To examine the realism of each simulated MAS level, the simulator's resistive torque-angle profiles and peak torque values were compared to qualitative descriptions in the MAS Table [20] and quantitative data that were collected from spasticity patients [44,27]. During this evaluation and actual use, only the orifice sizes of 0.5 mm, 1.5 mm, and 2.0 mm were utilized (Table 2.3). For MAS level 0 ("No increase in muscle tone"), no catch should be observed; therefore, the damper was not engaged. MAS level 1 was paired with a late catch angle (135°), which resulted in the damper engaging at a small lever arm length, so a small orifice (0.5 mm) was selected experimentally to generate modest resistance.

Each simulated MAS level was tested under slow, medium, and fast stretch speeds (same speed ranges as defined previously). Fast stretch trials mimicked an actual clinical test using the MAS guideline and the resulting simulator's resistive torque profiles were compared to the narrative MAS descriptions [20]. Medium stretch trials matched the testing speed used in other quantitative studies [44,27], and the recorded peak resistive torque of the simulator was compared to those reported in [44,27]. Including the slow stretch was to give a more systematic investigation of resistive torque's dependency on input speed.

Table 2.3: The combination of mechanical design features to simulate each MAS level.

Simulated MAS Level	Orifice Radius (mm)	Catch Angle	ROM (°)	Abbreviated MAS Description [20]
0	N/A	None	130	No increased resistance
1	0.5	Late (135°)	115	Slight increased resistance, a catch and release at the end of ROM
2	1.5	Early (80°)	95	Slight increase resistance, a catch, followed by minimal resistance throughout the reminder (less than half) of ROM
3	1	Early (80°)	80	More markedly increased resistance through most of ROM
4	0.5	Immediate (50°)	60	Considerably increased resistance, passive movement difficult

2.4.2 Results for Bench-Top Evaluation

Experimental results of ROM and catch angle matched well with design intent (Table 2.4). The experimental average values of catch angle agreed with the design values at 50°, 80°, and 135° with minor differences of 4°-7° (or 4-12% error). Compared to design values, experimental mean values of ROM for MAS 0, 2, 3, and 4 were within ± 9°, but larger at MAS 1.

Table 2.4: Experimental results of Catch Angle Test and ROM Test. The average experimental values over three trials and standard deviations (in parentheses) as well as percent errors were reported.

	Design Value (°)	Measured Value (°)	Error Residual (°)	Percent Error (%)
Catch Angle Test	50 – Immed	57 (0.8)	7	12.3%
	80 – Early	77 (2.1)	-3	-3.9%
	135 – Late	141 (0.8)	6	4.3%
ROM Test	130	132 (1.2)	2	1.5%
	115	132 (0.8)	17	12.9%
	95	104 (1.2)	9	8.7%
	80	84 (1.2)	4	4.8%
	60	59 (1.2)	-1	-1.7%

Experimental torque profiles agreed with the design setup (Fig. 2.3). Across speeds and orifice sizes, the torque profiles started from a pre-catch phase (minimal resistance mainly due to friction and inertial effects), spiked up at the prescribed catch angle ~80° (when the damper kicked in), released as the linkage level arm decreased, and eventually ended at the prescribed ROM of ~95°. As the stretch speed increased, the simulated muscle resistance intensified (higher resistive torque and stronger catch) for all orifice sizes, while under a similar stretch speed, smaller orifice sizes generated higher resistance.

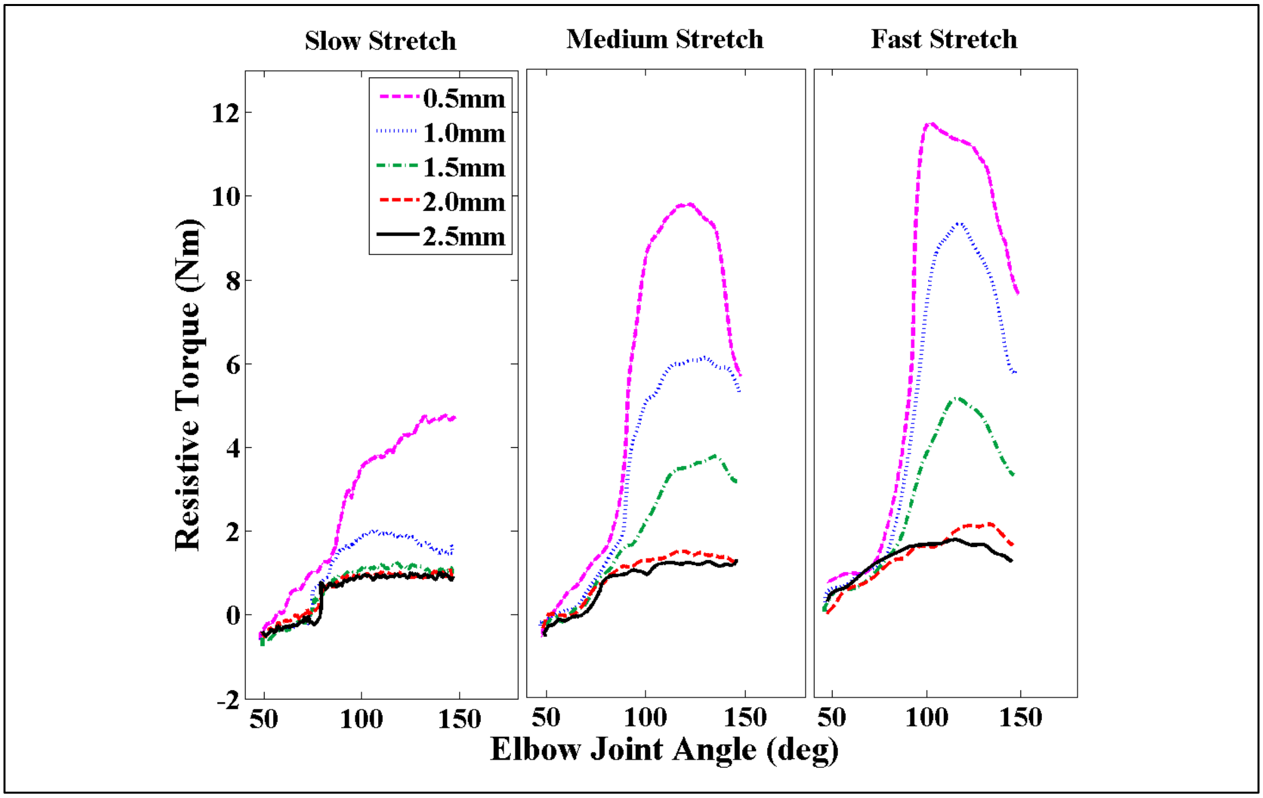


Fig. 2.3: Resistance Level Test results (for a single trial), with ROM at 95° and catch angle at 80°. Resistive torque profiles across five orifice sizes and under three speeds.

The simulated resistive torque profile for each MAS level under fast stretch qualitatively agreed with the MAS (Fig. 2.4). At MAS 0, throughout the ROM, no increased muscle tone was observed. At MAS 1, there was a clear and abrupt "catch and release" behavior towards the end of the ROM. At MAS 2, the resistive torque began to increase around 80° as the catch started, and decreased gradually after about 110° for the reminder of the ROM. At MAS 3, higher resistive torque was created and a ROM of 85° was achieved. Eventually at MAS 4, a substantial increase in resistive torque was observed compared to MAS 3 and the ROM was limited to only about 60°. MAS 5 was not implemented, so no comparison was provided.

The implementation of catch angle and ROM selectors functioned accordingly and matched with design intents in creating varying torque-angle profiles to represent different characteristics of the five MAS levels (Fig. 2.4).

The experimental catch angle locations and ROMs of the simulated MAS levels agreed with individual feature evaluation (Fig. 2.4).

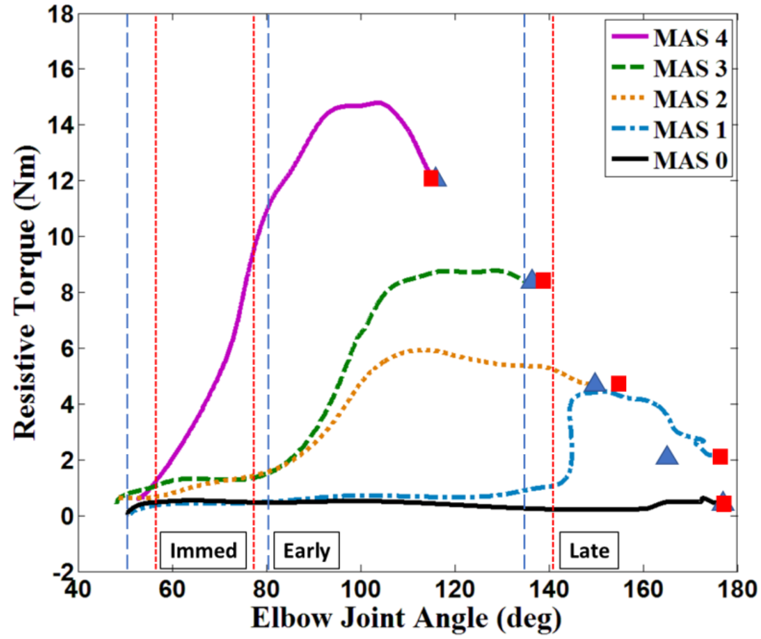


Fig. 2.4: Simulated MAS levels under fast stretch. Desired (blue) and experimental (red) values of catch angle (line) and ROM (symbol).

The experimental peak resistive torques were compared with quantitative results adapted from other two quantification studies [44,27] (Fig. 2.5). Since Nam et al. combined MAS 0 and 1 as low severity group and MAS 3 and 4 as high severity group, a single averaged peak torque value was reported for each severity group with relatively large standard deviations. For MAS 1-4, the simulator's peak resistive torques were comparable with other studies, while for MAS 0 (representing healthy individuals), the experimental result of the simulator was lower than the literature data.

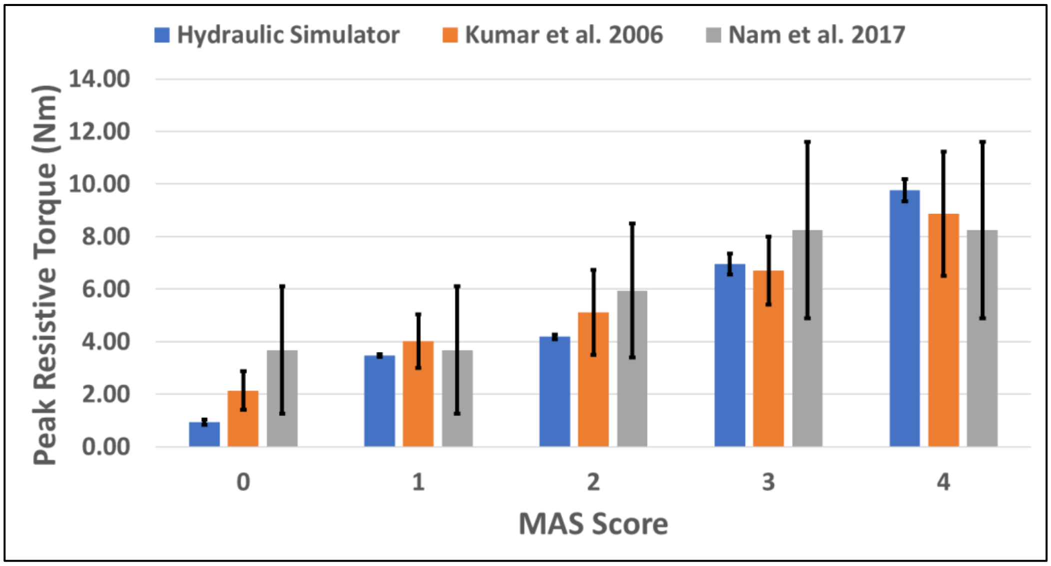


Fig. 2.5: Comparison of simulator's peak resistive torque (mean and standard deviation) with other quantitative studies [44,27]**.**

2.4.3 Discussion for Bench-Top Evaluation

The performance of the three mechanical design features was evaluated. Compared to the first-generation simulator, the resistance level selector's orifice sizes were modified to achieve higher force-generation capacity to match the actual muscle resistance observed in spasticity patients (Fig. 2.3). For catch angle and ROM tests (Table 2.4), angular differences between design values and experimental results might be due to backlash in the linkage system. Relatively loose dimensioning was used to minimize undesired sliding friction between the linkage and upper arm frame to avoid binding. To address both backlash and friction in the system, the rectangular upper arm frame could potentially be replaced by round shafts with linear bearings that provide more precise motion control, higher stability, and lower friction, so that the performance of these two positional features could be further improved.

The experimental torque-angle profiles were in accordance with the descriptions of the MAS levels, in terms of the presence of catch-release behavior, catch angle location, and relative resistance magnitudes between levels (Fig. 2.4). Quantitatively, the experimental peak resistive torques were found comparable to clinical data collected from spasticity patients at MAS 1-4 [44,27], while close to but lower than the clinical data at MAS 0 (Fig. 2.5). At MAS 0 level, the hydraulic damper was completely disengaged, and the simulated muscle resistance was solely created by the torsion springs at the elbow. The spring constant was designed towards numerical values found in the literature, 0.0168 ± 0.0279 Nm/° reported in [27] and 0.015 ± 0.0091 Nm/° reported in [47]. The implemented spring constant of 0.00892 Nm/° was within the range of clinical data but fell at the lower end, which might explain the simulator's lower resistance at MAS 1.

2.5 CLINICAL EVALUATION

2.5.1 Subject Population and Protocol

Seven clinicians were invited to evaluate the fidelity of the second-generation training simulator in replicating various levels of spasticity in the biceps. All participants held medical (Physiatry), physical therapy (PT) or occupational therapy (OT) advanced degrees and had at least 2 years of experience with performing passive stretch tests to assess spasticity using the MAS (Table 2.5). All participating clinicians provided informed consent prior to the test. The test took place in the OSF HealthCare Jump Trading Simulation & Education Center (Peoria, IL). This clinical evaluation study was IRB-approved. Each participating clinician evaluated the simulator individually and was provided an official written description of the MAS [20] for reference. Before the test, the clinician was allowed to practice extending the simulator's elbow when set to the simulated MAS 0. Each clinician performed two tests (Blind Assessment, Disclosed Assessment).

Table 2.5: Summary of participant information.

Subject	Role	Primary Field	Years of Practice	Years of Spasticity Evaluation	Number of Spasticity Patients per Month	Number of Spasticity Patients per Month (using MAS)
1	PT	Neurology	10	9	5	5
2	PT	Cardiac, ICU	2	2	15	15
3	PT	Physical therapy (acute care)	20	20	1	1
4	PT	Physical therapy, neurology	3	2	2-5	2-5
5	OT	Neurology	23	23	40-50	40-50
6	OT	Neurology	6.5	6.5	35	35
7	Physiatrist	Physical medicine rehabilitation	6	8	20	20

The Blind Assessment Test started once the clinician felt comfortable working with the simulator. Before each trial, the simulated MAS level on the simulator was set by the research assistant (YP) and visually blocked from the clinician. The clinician was asked to perform passive stretch tests to assess the simulated biceps spasticity behavior and assign a MAS score for each simulation trial based on prior clinical experience and training. Each simulated MAS level was tested three times, therefore each clinician performed in total 15 trials (5 levels × 3 trials). The order of trials was randomized among all test subjects using a random number generator in MATLAB (R2013b, MathWorks Inc., Natick, MA).

A Disclosed Assessment Test was performed next. In this test, the level being simulated was disclosed, and the clinician was asked to closely examine each simulated MAS level created by the simulator and provide comments/feedback on the realism of replication. The clinician's response was recorded on a questionnaire containing specific questions regarding to the catch-release behavior, resistance magnitude, speed-dependency, and ROM. To quantify their responses, participants were asked to rate each aspect using a five-point scale (1: too little; 3: about right; 5: too much) (the questionnaire is attached in Appendix G).

2.5.2 Data Analysis

For the Blind Assessment Test, percent agreement was obtained per rater and per MAS level. Percent agreement was defined as $\frac{\#\ of\ matched\ trial}{\#\ of\ total\ trial}$ ×100%. A matched trial was defined as a trial when the rater's judgment agreed with the simulator's simulated MAS level. Across all subjects and for each MAS level, average and standard deviation of percent agreement were also reported. For the Disclosed Assessment Test, the raters' realism scores were averaged for each simulation aspect for a given MAS level.

For each clinician (rater), the assessment was recorded for all 15 trials during the Blind Assessment Test. To establish rater reliability [61], the intra-rater reliability and inter-rater reliability were evaluated using both Intraclass correlation coefficient (ICC) [84] and Krippendorff's alpha (K-alpha) [85,70]. Previous studies involving more than two raters or repetitions utilized ICC for intra- and inter-rater reliabilities [86,58], as ICC was considered mathematically equivalent of the weighted kappa for ordinal data [68]. K-alpha is another alternative statistic for assessing intra- and inter-rater reliabilities for ordinal data given more than two raters [70,69].

Each rater was asked to evaluate each simulated MAS level (0-4) for three times (in a randomized order). Intra-rater reliability was calculated to measure the self-consistency of each rater and inter-rater reliability was assessed to determine the agreement among all raters. Both ICC and K-alpha were calculated using R (Version 3.4.3, Package "irr"). ICC estimates and their 95% confidence intervals (CI) were assessed based on a two-way (intra-rater: mixed-effects, inter-rater: random-effects), absolute-agreement, single-measure model [87]. The unit of analysis, model, type of ICC were chosen per guidelines described in [69,88]. ICC values were interpreted based on the following cutoffs: poor reliability for ICC values less than 0.40, fair for values between 0.40 and 0.59, good for values between 0.60 and 0.74, and excellent for values between 0.75 and 1.0 [89]. K-alpha estimates were obtained with weight for ordinal data [90], and the minimum acceptable α value was 0.667; a value larger than 0.8 was considered to be reliable [91].

For the Disclosed Assessment Test, means and standard deviations across the seven raters' scorings for each of the four simulation aspects (catch-release behavior, resistance magnitude, speed-dependency, and ROM) and ensemble across all four aspects were calculated.

2.5.3 Results for Clinical Evaluation

Relatively high percent agreements were found between the MAS scores determined by the raters and training simulator's setup during the Blind Assessment Test (Table 2.6). The average percent agreement among all seven raters was 76 ± 12%. It should be noted that Rater 7 had a substantially lower percent agreement compared to other raters. Documented in the experiment record, Rater 7 consistently scored the simulation one level lower than the simulator's setup except for MAS 0. This observation explained why an "excellent" intra-rater reliability was assigned to Rater 7 (Table 2.8), even though this rater had a low percent agreement. Average percent agreements were higher at MAS 0 and 1 and relatively low at MAS 2-4.

Table 2.6: Percent agreement on MAS scores during Blind Assessment Test. Percent agreements were reported as percentage (Columns 2-8). Average ensemble percent agreement was reported as mean (standard deviation).

MAS Level	Subject ID							Percent Agreement
	1	2	3	4	5	6	7	Mean (SD)
0	66	100	100	100	100	100	100	95 (12)
1	100	100	100	100	100	100	0	86 (35)
2	66	66	100	33	66	100	0	62 (33)
3	66	100	100	33	66	66	66	71 (21)
4	100	100	100	33	100	33	0	67 (40)
Mean (SD)	80 (17)	93 (14)	100 (0)	60 (33)	86 (17)	80 (27)	33 (42)	76 (12)

The intra-rater reliability was obtained in both ICC and K-alpha to determine the reliability of a rater's judgement (Table 2.7). The average intra-rater reliability was considered as "excellent" with an ICC estimate of 0.918, supported by an average K-alpha value of 0.906. Given a wide 95% CI in ICC, Rater 4's reliability was considered as "poor-fair". The K-alpha value (0.652) for this rater raised the same concern, as an K-alpha value below 0.667 was considered as unreliable [91]. Therefore, to cope with the questionable reliability of Rater 4, the inter-rater reliability was reported in both with and without Rater 4. The inter-rater reliability estimates using ICC (95% CI) and K-alpha were calculated to be 0.888 (0.780-0.955) and 0.881 with Rater 4, and 0.899 (0.781-0.961) and 0.893 without Rater 4. Both statistic measures yielded high inter-rater reliability estimates, indicating that raters had a high degree of agreement and the simulated MAS levels were assessed similarly across raters during the Blind Assessment Test. The inclusion of Rater 4 had no significant impact on inter-rater reliability, so Rater 4 remained in the percent agreement analysis that was presented above.

Table 2.7: Intra-rater reliability for each rater reported in ICC and K-alpha.

Subject	ICC	95% CI	Interpretation	K-alpha	Interpretation
1	0.909	0.644-0.989	Good-Excellent	0.903	Reliable
2	0.974	0.891-0.997	Excellent	0.968	Reliable
3[a]	1	/	Excellent	1.000	Reliable
4[b]	0.704	0.23-0.959	Poor-Fair	0.652	Unreliable
5	0.946	0.758-0.994	Excellent	0.936	Reliable
6	0.927	0.686-0.992	Good-Excellent	0.919	Reliable
7	0.965	0.855-0.996	Excellent	0.966	Reliable
Average	0.918	/	Excellent	0.906	Reliable
Standard Deviation	0.092			0.108	

[a] Rater 3 agreed with simulator's setup in all fifteen trials during the Blind Assessment Test, so the percent agreement was 100% and ICC was 1 with no 95% CI calculated.

[b] Considering Rater 4 had a very wide 95% CI, the reliability was interpreted to be "poor" to "fair" by the low end of CI.

From the Disclosed Assessment Test, mean scores for all levels did not deviate much from 3 (Table 2.8). The mean scores were closer to 3 at low MAS levels (0 and 1) and deviated more from 3 at high MAS levels (2,3, and 4). Furthermore, higher variances were also found for high MAS levels. In general, the participants thought the training simulator had potential to be good training tool in teaching the basic spasticity assessment, providing more practice, and helping standardize techniques. Refer to Appendix G for more details.

Table 2.8: Rater similarity score of simulated MAS levels using a five-point scale (1: too little, 3: about right, 5: too much) during Disclosed Assessment Test.

Aspect	MAS 0				MAS 1			
	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM
Mean	2.7	3.0	3.0	3.0	2.9	3.3	2.9	3.0
SD	0.5	0.0	0.0	0.0	0.4	0.7	0.4	0.0
Level Mean	2.9				3.0			
Level SD	0.3				0.5			
Aspect	MAS 2				MAS 3			
	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM
Mean	3.0	2.9	2.9	2.6	2.9	2.9	2.9	2.4
SD	0.0	0.6	0.4	0.9	0.4	0.6	0.4	0.7
Level Mean	2.8				2.8			
Level SD	0.6				0.6			
Aspect	MA 4							
	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM				
Mean	2.4	2.9	2.9	2.4				
SD	0.5	0.4	0.4	0.7				
Level Mean	2.6							
Level SD	0.6							

2.5.4 Discussion for Clinical Evaluation

The clinical evaluators found that the passive hydraulic training simulator mimicked well the general behaviors of the biceps during a passive stretch test for patients typically scored as MAS 0 through 4. An average percent agreement of 76% was obtained during the Blind Assessment Test (Table 2.6). Five of seven raters achieved above 80% percent agreement when blinded to the simulated level. The percent agreements on the simulated MAS 0 and 1 were higher (above 85%), while those on the MAS 2-4 were relatively low (60-80%) (Table 2.6).

Scoring results from the Disclosed Assessment Test agreed with the Blind Assessment Test results (Table 2.6 and 2.8). Mean scores for simulated MAS 0 and 1 were relatively higher (2.9 and 3.0, respectively) in accordance with the higher percent agreements (95% and 86%), suggesting that the spasticity characteristics simulated in these two levels matched with clinicians' expectation. On the other hand, raters had mixed opinion for higher MAS levels (2-4) corresponding to their moderate percent agreements (62%, 71%, and 67%, respectively). In terms of resistance, raters agreed with the designed magnitude of resistance for MAS 1-3, but expected slightly higher resistance at MAS 0 and 4 (average score for resistance at MAS 0 was 2.7 and at level 4 was 2.4). The designed values for speed-dependency and catch-release behavior scored well throughout all five levels with some room for minor fine-tuning.

The average scores for the ROM feature were below 3 and had greater variance at the higher MAS levels. Clinicians' comments and their scoring in the questionnaire provided some explanation to this difference (Table 2.7). For MAS 2-4, PTs and OTs had opposite opinions on the ROM design (large standard deviations were found), where PTs assigned an average score of 2, while average score of 3.2 from OTs. Several participants commented in the questionnaire that, compared to OTs, PTs who frequently work with acute spasticity patients have seen fewer cases of ROM reduction, as this symptom usually develops in spasticity patients whose joints chronically remain in a flexed position as a result of involuntary muscle tone. Therefore, the PT participants in our study tended to consider the ROM reduction as unnecessary, while OT participants, who often interact with long-term patients, agreed with this design value.

2.6 OVERALL DISCUSSION AND FUTURE WORK

This chapter described a framework to develop a clinical training simulator for replicating different levels of biceps spasticity to provide training opportunities for future clinical/healthcare students to practice assess spasticity using the MAS. To systematically replicate spasticity, its behavior was characterized by three main behavioral features, i.e., abnormal muscle tone, catch-release behavior, and limited ROM (Table 2.1). These three features were further quantified through six parameters, i) speed-dependent resistance, ii) severity-dependent resistance, iii) catch duration, iv) catch angle location, v) release duration, and vi) ROM magnitude (Table 2.1). The final simulator prototype was equipped with three mechanical design features (i.e., resistance level, catch angle, ROM selectors) to demonstrate all six parameters and provide adjustability to (ii), (iv), and (vi) for varied patient severity (MAS 0-4). Preliminary bench-top and clinical evaluation results validated the performance of the simulator and suggested the feasibility of using it as a training tool (Table 2.4, 2.6, and 2.8, Fig. 2.3-2.5).

The passive design approach of designing a biceps spasticity training simulator presented in this chapter has advantages and disadvantages. First, benefitting from the self-contained and portable design, our simulator requires no external power to operate and weighs similar to a human arm, so it can be easily transported and deployed in various locations in a training room. Second, the simulator consists of only mechanical components (e.g., linkage, hydraulic device, etc.) without any electrical parts (e.g., sensor, electric motor, hardware controller, etc.), which provides affordability to classrooms.

On the other hand, with lack of real-time sensing and feedback control, our simulator is unable to "playback" actual patient's muscle resistance as other active-controlled simulators are capable but can only replicate a general torque profile of spastic muscle parameterized by discrete values for resistance level, catch angle, and ROM. Our simulator can provide five distinct simulations to replicate MAS 0-4, but it is not reprogrammable. In order to adjust simulation performance, new mechanical components need to be designed, manufactured, tested, and assembled. Additionally, the specific simulator presented in this chapter was designed solely to simulate biceps spasticity. To simulate a different neurological condition such as rigidity, dystonia, etc., a completely new hardware system that operates with a different physical principle needs to be design and little can be inherited from the spasticity simulator prototypes. In contrast, through software control modification, electro-mechanical designs could be reprogrammed to provide a number of variants for each MAS level (similar to real patients) or may even simulate other types of abnormal muscle behaviors without hardware changes.

Clear design targets for spasticity behavior were missing in the literature making it difficult to have a rigorous and deterministic design process. The simulator's ROM, catch angle location, and resistance level were designed iteratively driven by user feedback. Ideally, the development of a training simulator requires (a) establishing a database that documents the kinematic and kinetic data associated with patients at varied severity, and (b) building a mathematical model that relates the kinematic inputs with the spastic muscle kinetic response at varied severities. The current absence of both poses difficulty to systematically design and tune a simulator towards realistic spasticity behaviors. As far as we know, other groups' recent efforts also focus on building a patient database and establish mathematical models for spasticity [37,45]. We are also currently performing a quantification study involving spasticity and rigidity patients such that more quantitative measures could be obtained for the future development and tuning of the simulator [92,93].

2.7 CONCLUSION

This chapter presented the design and validation cycle of a biceps spasticity training simulator. The new training simulator replicated the three main behavioral features of spasticity (abnormal muscle tone, catch-release behavior, and limited ROM) through three mechanical design features (resistance level, catch angle, and ROM selectors). A custom viscous hydraulic damper (resistance level selector) was utilized to generate speed- and severity-dependent abnormal muscle tone; the catch angle selector and elbow linkage created an abrupt catch-release behavior with a variable catch angle whose location depended on the severity of the simulated patient; in addition, the ROM could be adjusted and reduced when simulating more severe patients. The bench-top and clinical evaluation results suggested the feasibility of using this device as a training tool for future clinical/healthcare trainees.

CHAPTER 3: CONCLUSIONS

3.1 REVIEW OF FINDINGS

This thesis presents the design framework for a passive (unpowered) clinical training simulator using purely mechanical components. The design process started by characterizing the main behavioral features of the target disease and selecting the appropriate mechanical design features to provide haptic feedback comparable to the actual disease. The prototype was further validated by a two-stage evaluation process. The first part of evaluation involved examining the performance of individual mechanical design features and their combined performance through benchtop experiments. In the second part of evaluation, clinicians were invited to assess the replicated disease behavior and to compare the simulation with their previous experience interacting with actual patients. The benchtop performance and clinical feedback help design iteration and provide insights into the future development of the training simulator.

Following this design approach, an experimental simulator was specifically developed for replicating different levels of biceps spasticity to provide training opportunities for future clinical/healthcare students to practice assess spasticity using the MAS. The spasticity behavior was parameterized by three main behavioral features, i.e., abnormal muscle tone, catch-release behavior, and limited ROM (Table 2.1). For the purpose of mechanical implementation, these three features were quantified through six parameters, i) speed-dependent resistance, ii) severity-dependent resistance, iii) catch duration, iv) catch angle location, v) release duration, and vi) ROM magnitude (Table 2.1). Three mechanical design features were developed (i.e., resistance level, catch angle, ROM selectors) to demonstrate all six parameters and provide adjustability to (ii), (iv), and (vi) for simulating patients with different severities (MAS 0-4).

Benchtop and clinical evaluations were conducted to validate the performance of the training simulator. Benchtop results demonstrated the adjustability of individual mechanical design features (Table 2.4 and Fig. 2.3). Although experimental results matched with the design intents, positional features (catch angle and ROM selectors) had minor errors due to linkage backlash and instability (Table 2.4). Five unique combinations of these three mechanical design features were used to replicate MAS 0-4 (Table 2.3). The simulator demonstrated its ability to generate five distinct simulations that qualitatively agreed with the MAS descriptions for levels 0-4 (Table 2.3 and Fig. 2.4). The numeric values of peak simulated resistance for each MAS level fell within the range of clinical measures [27,44] (Fig 2.5). In the clinical evaluation, higher percent agreements (95 and 86%, respectively) between clinician's judgements and simulations were found in low MAS levels (0-1) and moderate percent agreements (62, 71 and 67%, respectively) were found in high MAS levels (2-4), where the average percent agreement was 76% (Table 2.6). The lower agreement in high MAS levels were mainly due to two reasons. First, PT participants considered the ROM reduction implemented in the high levels was unnecessary. Second, the simulated resistance at MAS 4 was considered to be low compared to actual spasticity patients. Despite these issues requiring further fine-tuning, the overall performance of the training simulator received positive feedback and were commented to have potential of becoming a training tool in the classroom setting.

3.2 LIMITATIONS AND FUTURE WORK

The passive design of the biceps spasticity training simulator presented in this thesis has its own limitations. The simulator can only replicate a general torque profile of spastic muscle parameterized by resistance level, catch angle, and ROM. The actual spasticity behavior was much more complicated than the simulation. For example, the stretch speed that is required to induce catch may change with patient severity [94], and it has been observed that the catch angle location is not only dependent on severity but also on stretch speed [33,39]. If these more subtle characteristics of spasticity were to be simulated, the mechanical complexity and number of parts would inevitably increase (Fig. 3.1), leading to a decrease of system manufacturability and reliability. As shown in Fig. 3.1, after the turning point, the margin of performance improvement would decrease. In a purely passive design, a higher performance requirement beyond this point would not justify the cost-benefit relationship.

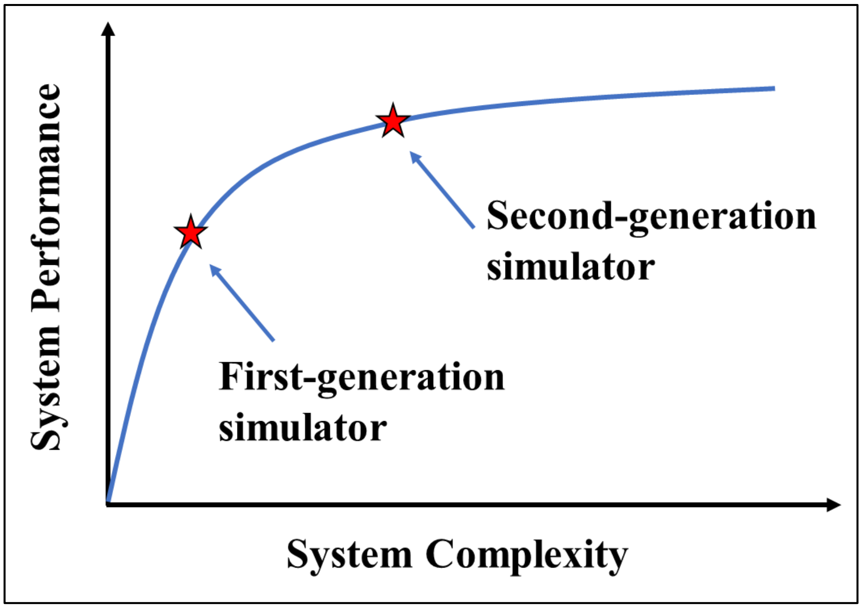


Fig.3.1: System complexity vs. system performance in the passive design approach.

Our simulator can provide five distinct simulations to replicate MAS 0-4 consistently. This capability is valuable for helping beginners who have little experience with spasticity patients to quickly gain understanding of five severity levels and reinforce their understandings through repeated practices. However, the simulator cannot replicate any patient whose severity falling between any two MAS levels or provide any variance around a certain level. This limitation may fail to prepare trainees adapting to the real clinical assessments, because the manifestation of each MAS level is different from patient to patient, and even for the same patient, the symptom may fluctuate as a result of emotion, stress level, etc. To fully train the student from the very beginning to the job site, the simulator needs to have different operation modes at different stages of the program. For example, the beginner mode focuses on standardizing techniques and reinforcing the conceptual understanding, while the expert mode introduces small variances in resistance, catch angle, and ROM to simulate a more realistic patient. This again underscores the need for the simulator to be reprogrammable.

Throughout the development of the spasticity simulator, two main difficulties were encountered. First, healthcare practitioners and researchers lack knowledge of the underlying mechanics of spasticity and have not reached consensus on pathological changes associated with spasticity. Our training simulator was developed based on the clinical observation that spastic muscles exhibited higher viscosity compared to healthy ones [31,32], while other studies did report changes in other muscle properties, such as stiffness [27,47], muscle activation and stretch reflex

[33,50], and so on. Given the current understanding of spasticity is still insufficient, an electro-mechanical approach might be more beneficial, as actively-controlled actuators could be hard-coded to replicate recorded patient's behaviors without fully understanding the physics of behaviors they simulate. On the other hand, the knowledge of pathological mechanics behind spasticity is especially important for the development of a passive training simulator. Without compromising system manufacturability and reliability, a mechanical design cannot generate an equally complicated torque profile as the electro-mechanical ones do, but can only create a generic profile tuned by a limited number of key parameters. Therefore, a designer needs to exploit the spasticity mechanics, identify a few most fundamental and characterizing features, and then choose the appropriate mechanical components that can provide similar physical responses as their biological counterparts.

Second, due to the lack of quantitative data, the simulator's resistance level, catch angle location, and ROM were designed iteratively driven by user feedback, because clear design targets were missing in the literature, making it difficult to have a rigorous and deterministic design process. Ideally, the development of a training simulator requires (a) establishing a database that documents the kinematic and kinetic data associated with patients at varied severity, and (b) building a mathematical model that relates the kinematic inputs with the spastic muscle kinetic response at varied severities. The current absence of both poses difficulty on systematically design and tune a simulator towards realistic spasticity behaviors. As far as we know, other groups' recent efforts also focus on building a patient database and establishing mathematical models for spasticity [37,45]. We are also performing a quantification study involving spasticity and rigidity patients such that more quantitative measures could be obtained for future development and tuning of the simulator [92].

The second-generation design advanced the simulator towards better mimicking typical behaviors observed clinically. Based on the preliminary work on the first-generation prototype, the training simulator incorporated with several design improvements to overcome the previous limitations and more closely mimic the spasticity characterization observed in the literature. The new design featured a more distinct catch-release behavior with a variable catch angle whose location depended on the severity of the simulated patient; in addition, the ROM of 130° was implemented and could be limited for higher severity. A new set of orifice sizes was implemented in the hydraulic damper to generate higher simulated resistance to match the muscle tone of severe spasticity patients more appropriately. Details about design decisions related to these mechanical design features and parameters are included in the attached appendices. Bench-top results of individual design features suggested that the second-generation simulator was capable of improving the fidelity of spasticity replication in terms of variable catch angle, selectable ROM, and more realistic resistance level. Moreover, clinical evaluation results have suggested the feasibility of using this device as a training tool for future clinical/healthcare students.

REFERENCES

[1] “Spasticity – Causes, Symptoms and Treatments” [Online]. Available: http://www.aans.org/Patients/Neurosurgical-Conditions-and-Treatments/Spasticity. [Accessed: 25-Jun-2018].

[2] Mukherjee, A., and Chakravarty, A., 2010, “Spasticity Mechanisms - for the Clinician,” Front. Neurol., p. 149.

[3] Pandyan, A. D., Gregoric, M., Barnes, M. P., Wood, D., Van Wijck, F., Burridge, J., Hermens, H., and Johnson, G. R., 2005, “Spasticity: Clinical Perceptions, Neurological Realities and Meaningful Measurement,” Disabil. Rehabil., **27**(1–2), pp. 2–6.

[4] O’Sullivan, S. B., Schmitz, T. J., and Fulk, G. D., 2014, *Physical Rehabilitation*, F.A. Davis, Philadelphia, USA.

[5] Rekand, T., 2010, “Clinical Assessment and Management of Spasticity: A Review,” Acta Neurol. Scand., **122**(SUPPL. 190), pp. 62–66.

[6] Barnes, Michael P., G. R. J., 2008, *Upper Motor Neurone Syndrome and Spasticity: Clinical Management and Neurophysiology*, Cambridge University Press., New York, USA.

[7] Sheean, G., 2002, “The Pathophysiology of Spasticity,” Eur. J. Neurol., **9**(s1), pp. 3–9.

[8] Lance, JW., Feldman, R.G., Young, R.R. and Koella, W. P. eds, 1980, “Pathophysiology of Spasticity and Clinical Experience with Baclofen,” *Spasticity: Disordered Motor Control. Chicago: Year Book Medical Publishers*, pp. 185–204.

[9] Herman, R., 1970, “The Myotatic Reflex. Clinico-Physiological Aspects of Spasticity and Contracture,” Brain, **93**(2), pp. 273–312.

[10] Thilmann, A. F., Fellows, S. J., and Garms, E., 1991, “The Mechanism of Spastic Muscle Hypertonus. Variation in Reflex Gain over the Time Course of Spasticity,” Brain, **114**(1), pp. 233–244.

[11] Thilmann, A. F., Fellows, S. J., and Ross, H. F., 1991, “Biomechanical Changes at the Ankle Joint after Stroke,” J. Neurol. Neurosurg. Psychiatry, **54**(2), pp. 134–139.

[12] Burke, D., Gillies, J. D., and Lance, J. W., 1970, “The Quadriceps Stretch Reflex in Human Spasticity,” J. Neurol. Neurosurg. Psychiatry, **33**(2), pp. 216–223.

[13] Burke, D., Knowles, L., Andrews, C., and Ashby, P., 1972, “Spasticity, Decerebrate Rigidity and the Clasp-Knife Phenomenon: An Experimental Study in the Cat,” Brain, **95**(1), pp. 31–48.

[14] Burke, D., and Ashby, P., 1972, “Are Spinal ‘presynaptic’ Inhibitory Mechanisms Suppressed in Spasticity?,” J. Neurol. Sci., **15**(3), pp. 321–326.

[15] Powers, R. K., Marder-Meyer, J., and Rymer, W. Z., 1988, “Quantitative Relations between Hypertonia and Stretch Reflex Threshold in Spastic Hemiparesis,” Ann. Neurol., **23**(2), pp. 115–124.

[16] Powers, R. K., Campbell, D. L., and Rymer, W. Z., 1989, “Stretch Reflex Dynamics in Spastic Elbow Flexor Muscles,” Ann. Neurol., **25**(1), pp. 32–42.

[17] Rack, P. M., Ross, H. F., and Thilmann, A. F., 1984, “The Ankle Stretch Reflexes in Normal and Spastic Subjects. The Response to Sinusoidal Movement,” Brain, **107**(2), pp. 637–654.

[18] Li, S., Kamper, D. G., and Rymer, W. Z., 2006, “Effects of Changing Wrist Positions on Finger Flexor Hypertonia in Stroke Survivors,” Muscle Nerve, **33**(2), pp. 183–190.

[19] Meinders, M., Price, R., Lehmann, J. F., and Questad, K. A., 1996, “The Stretch Reflex Response in the

Normal and Spastic Ankle: Effect of Ankle Position," Arch. Phys. Med. Rehabil., **77**(5), pp. 487–492.

[20] Bohannon, R. W., and Smith, M. B., 1987, "Interrater Reliability of a Modified Ashworth Scale of Muscle Spasticity," Phys Ther., **67**(2), pp. 206–207.

[21] Boyd, RN; Graham, H., 1999, "Objective Measurement of Clinical Findings in the Use of Botulinum Toxin Type A for the Management of Children with Cerebral Palsy," Eur. J. Neurol., **6**(Suppl. 4), pp. S23–S35.

[22] Burridge, J. H., Wood, D. E., Hermens, H. J., Voerman, G. E., Johnson, G. R., Wijck, F. Van, Platz, T., Gregoric, M., Hitchcock, R., Pandyan, A. D., Burridge, J. H., Wood, D. E., Hermens, H. J., Voerman, G. E., Johnson, G. R., and Pandyan, A. D., 2005, "Theoretical and Methodological Considerations in the Measurement of Spasticity," Disabil. Rehabil., **27**(1–2), pp. 69–80.

[23] Young, R. R., 1994, "Spasticity: A Review," Neurology, **44**(November), pp. S12–S20.

[24] Thibaut, A., Chatelle, C., Ziegler, E., Bruno, M. A., Laureys, S., and Gosseries, O., 2013, "Spasticity after Stroke: Physiology, Assessment and Treatment," Brain Inj., **27**(10), pp. 1093–1105.

[25] Gracies, J. M., 2005, "Pathophysiology of Spastic Paresis. I: Paresis and Soft Tissue Changes," Muscle and Nerve, **31**(5), pp. 535–551.

[26] Ada, L., and Canning, C., 1993, *Key Issues in Neurological Physiotherapy*, Butterworth Heinemann, Oxford, UK.

[27] Kumar, R. T. S., Pandyan, A. D., and Sharma, A. K., 2006, "Biomechanical Measurement of Post-Stroke Spasticity," Age Ageing, **35**(4), pp. 371–375.

[28] Pandyan, A. D., Price, C. I. M., Rodgers, H., Barnes, M. P., and Johnson, G. R., 2001, "Biomechanical Examination of a Commonly Used Measure of Spasticity," Clin. Biomech., **16**(10), pp. 859–865.

[29] Pandyan, A. D., Price, C. I. M., Barnes, M. P., and Johnson, G. R., 2003, "A Biomechanical Investigation into the Validity of the Modified Ashworth Scale as a Measure of Elbow Spasticity," Clin. Rehabil., **17**(3), pp. 290–294.

[30] Lee, H. M., Huang, Y. Z., Chen, J. J. J., and Hwang, I. S., 2002, "Quantitative Analysis of the Velocity Related Pathophysiology of Spasticity and Rigidity in the Elbow Flexors," J. Neurol. Neurosurg. Psychiatry, **72**(5), pp. 621–629.

[31] Lee, H. M., Chen, J. J. J., Ju, M. S., Lin, C. C. K., and Poon, P. P. W., 2004, "Validation of Portable Muscle Tone Measurement Device for Quantifying Velocity-Dependent Properties in Elbow Spasticity," J. Electromyogr. Kinesiol., **14**(5), pp. 577–589.

[32] Chen, J. J. J., Wu, Y. N., Huang, S. C., Lee, H. M., and Wang, Y. L., 2005, "The Use of a Portable Muscle Tone Measurement Device to Measure the Effects of Botulinum Toxin Type A on Elbow Flexor Spasticity," Arch. Phys. Med. Rehabil., **86**(8), pp. 1655–1660.

[33] Calota, A., and Levin, M. F., 2009, "Tonic Stretch Reflex Threshold as a Measure of Spasticity: Implications for Clinical Practice," Top. Stroke Rehabil., **16**(3), pp. 177–188.

[34] Ewoldt, J. K., Lazzaro, E. C., Roth, E. J., and Suresh, N. L., 2016, "Quantification of a Single Score (1+) in the Modified Ashworth Scale (MAS), a Clinical Assessment of Spasticity," *2016 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, Orlando, USA, pp. 1737–1740.

[35] Ayuni, C. Z. N., Komeda, T., Low, C. Y., and Inoue, K., 2013, "Emulation of Muscle Tone of Upper Limb Spasticity and Rigidity," *In Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference*, IEEE, San Diego, USA, pp. 1590–1593.

[36] Zakaria, N. A. C., Komeda, T., Low, C. Y., Hanapiah, F. A., and Inoue, K., 2015, "Spasticity Mathematical Modelling in Compliance with Modified Ashworth Scale and Modified Tardieu Scales," *ICCAS 2015 - 2015 15th International Conference on Control, Automation and Systems, Proceedings*, IEEE, Busan, Korea, pp.

1893–1897.

[37] Zakaria, N. A. C., Low, C. Y., Hanapiah, F. A., Komeda, T., Inoue, K., Shazidi, M. S., and Hamsan, H. M., 2014, "Evaluation of Upper Limb Spasticity towards the Development of a High Fidelity Part-Task Trainer," *2nd International Conference on System-Integrated Intelligence: Challenges for Product and Production Engineering*, Bremen, Germany, pp. 817–826.

[38] Jonnalagedda, P., Deng, F., Douglas, K., Chukoskie, L., Yip, M., Ng, T. N., Nguyen, T., Skalsky, A., and Garudadri, H., 2016, "An Instrumented Glove for Improving Spasticity Assessment," *2016 IEEE Healthcare Innovation Point-of-Care Technologies Conference (HI-POCT)*, Cancun,Mexico, pp. 167–170.

[39] Wu, Y. N., Ren, Y., Goldsmith, A., Gaebler, D., Liu, S. Q., and Zhang, L. Q., 2010, "Characterization of Spasticity in Cerebral Palsy: Dependence of Catch Angle on Velocity," Dev. Med. Child Neurol., **52**(6), pp. 563–569.

[40] Ansari, N. N., Naghdi, S., Hasson, S., Azarsa, M. H., and Azarnia, S., 2008, "The Modified Tardieu Scale for the Measurement of Elbow Flexor Spasticity in Adult Patients with Hemiplegia," Brain Inj., **22**(13–14), pp. 1007–1012.

[41] Bhadane, M. Y., Gao, F., Francisco, G. E., Zhou, P., and Li, S., 2015, "Correlation of Resting Elbow Angle with Spasticity in Chronic Stroke Survivors," Front. Neurol., **6**, p. 183.

[42] McGibbon, C. A., Sexton, A., Jones, M., and O'Connell, C., 2013, "Elbow Spasticity during Passive Stretch-Reflex: Clinical Evaluation Using a Wearable Sensor System," J. Neuroeng. Rehabil., **10**(1), p. 61.

[43] Lindberg, P. G., Gäverth, J., Islam, M., Fagergren, A., Borg, J., and Forssberg, H., 2011, "Validation of a New Biomechanical Model to Measure Muscle Tone in Spastic Muscles," Neurorehabil. Neural Repair, **25**(7), pp. 617–625.

[44] Nam, H. S., Koh, S., Kim, Y. J., Beom, J., Lee, W. H., Lee, S. U., and Kim, S., 2017, "Biomechanical Reactions of Exoskeleton Neurorehabilitation Robots in Spastic Elbows and Wrists," IEEE Trans. Neural Syst. Rehabil. Eng., **25**(11), pp. 2196–2203.

[45] Park, J. H., Lee, K. J., Yoon, Y. S., Son, E. J., Oh, J. S., Kang, S. H., Kim, H., and Park, H. S., 2017, "Development of Elbow Spasticity Model for Objective Training of Spasticity Assessment of Patients Post Stroke," *2017 IEEE International Conference on Rehabilitation Robotics*, IEEE, London, UK, pp. 146–151.

[46] Lamontagne, A., Malouin, F., Richards, C. L., and Dumas, F., 1997, "Impaired Viscoelastic Behaviour of Spastic Plantarflexors during Passive Stretch at Different Velocities," Clin. Biomech., **12**(7–8), pp. 508–515.

[47] McCrea, P. H., Eng, J. J., and Hodgson, A. J., 2003, "Linear Spring-Damper Model of the Hypertonic Elbow: Reliability and Validity," J. Neurosci. Methods, **128**(1–2), pp. 121–128.

[48] Park, H. S., Kim, J., and Damiano, D. L., 2012, "Development of a Haptic Elbow Spasticity Simulator (HESS) for Improving Accuracy and Reliability of Clinical Assessment of Spasticity," IEEE Trans. Neural Syst. Rehabil. Eng., **20**(3), pp. 361–370.

[49] Lynn, B. O., Erwin, A., Guy, M., Herman, B., Davide, M., Ellen, J., Anne, C., and Kaat, D., 2013, "Comprehensive Quantification of the Spastic Catch in Children with Cerebral Palsy," Res. Dev. Disabil., **34**(1), pp. 386–396.

[50] Calota, A., Feldman, A. G., and Levin, M. F., 2008, "Spasticity Measurement Based on Tonic Stretch Reflex Threshold in Stroke Using a Portable Device," Clin. Neurophysiol., **119**(10), pp. 2329–2337.

[51] Chung, S. G., van Rey, E., Bai, Z., Roth, E. J., and Zhang, L.-Q., 2004, "Biomechanic Changes in Passive Properties of Hemiplegic Ankles with Spastic Hypertonia," Arch. Phys. Med. Rehabil., **85**(10), pp. 1638–1646.

[52] Haugh, A., Pandyan, A., and Johnson, G., 2006, "A Systematic Review of the Tardieu Scale for the Measurement of Spasticity," Disabil. Rehabil., **28**(15), pp. 899–907.

[53] Van Wijck, F. M. J., Pandyan, A. D., Johnson, G. R., and Barnes, M. P., 2001, “Assessing Motor Deficits in Neurological Rehabilitation: Patterns of Instrument Usage,” Neurorehabil. Neural Repair, **15**(1), pp. 23–30.

[54] Yam, W. K. L., and Leung, M. S. M., 2006, “Interrater Reliability of Modified Ashworth Scale and Modified Tardieu Scale in Children with Spastic Cerebral Palsy,” J. Child Neurol., **21**(12), pp. 1031–1035.

[55] Kaya, T., Goksel Karatepe, A., Gunaydin, R., Koc, A., and Altundal Ercan, U., 2011, “Inter-Rater Reliability of the Modified Ashworth Scale and Modified Modified Ashworth Scale in Assessing Poststroke Elbow Flexor Spasticity,” Int. J. Rehabil. Res., **34**(1), pp. 59–64.

[56] Sloan, R. L., Sinclair, E., Thompson, J., Taylor, S., and Pentland, B., 1992, “Inter-Rater Reliability of the Modified Ashworth Scale for Spasticity in Hemiplegic Patients,” Int. J. Rehabil. Res., **15**(2), pp. 158–161.

[57] Mehrholz, J., Wagner, K., Meßiner, D., Grundmann, K., Zange, C., Koch, R., and Pohl, M., 2005, “Reliability of the Modified Tardieu Scale and the Modified Ashworth Scale in Adult Patients with Severe Brain Injury: A Comparison Study,” Clin. Rehabil., **19**(7), pp. 751–759.

[58] Mutlu, A., Livanelioglu, A., and Gunel, M. K., 2008, “Reliability of Ashworth and Modified Ashworth Scales in Children with Spastic Cerebral Palsy,” BMC Musculoskelet. Disord., **9**(1), p. 44.

[59] Allison, S. C., Abraham, L. D., and Petersen, C. L., 1996, “Reliability of the Modified Ashworth Scale in the Assessment of Plantarflexor Muscle Spasticity in Patients with Traumatic Brain Injury,” Int. J. Rehabil. Res., **19**(1), pp. 67–78.

[60] Gregson, J. M., Leathley, M. J., Moore, a P., Smith, T. L., Sharma, a K., and Watkins, C. L., 2000, “Reliability of Measurements of Muscle Tone and Muscle Power in Stroke Patients,” Age Ageing, **29**(3), pp. 223–228.

[61] Portney LG, W. M. P., 2009, *Foundations of Clinical Research: Applications to Practice*, Prentice-Hall, Inc., New Jersey, USA.

[62] Clopton, N., Dutton, J., Featherston, T., Grigsby, A., Mobley, J., and Melvin, J., 2005, “Interrater and Intrarater Reliability of the Modified Ashworth Scale in Children with Hypertonia,” Pediatr. Phys. Ther., **17**(4), pp. 268–274.

[63] Blackburn, M., van Vliet, P., and Mockett, S., 2002, “Reliability of Measurements Obtained With the Modified Ashworth Scale in the Lower Extremities of People With Stroke,” Phys Ther, **82**(1), pp. 25–34.

[64] Pandyan, A., 1999, “A Review of the Properties and Limitations of the Ashworth and Midified Ashworth Scales as Measures of Spasticity,” Clin. Rehabil., **13**, pp. 373–383.

[65] Cohen, J., 1960, “A Coefficient of Agreemnet for Nominal Scales,” Educ. Psychol. Meas., **20**(1), pp. 37–46.

[66] Cohen, J., 1968, “Weighted Kappa: Nominal Scale Agreement Provision for Scaled Disagreement or Partial Credit,” Psychol. Bull., **70**(4), pp. 213–220.

[67] Fleiss, J. L., 1971, “Measuring Nominal Scale Agreement among Many Raters,” Psychol. Bull., **76**(5), pp. 378–382.

[68] Fless, J. L., and Cohen, J., 1973, “The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as a Measure of Reliability,” Educ. Psychil Meas., **33**, pp. 613–619.

[69] Hallgren, K. A., 2012, “Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial,” Tutor. Quant. Methods Psychol., **8**(1), pp. 23–34.

[70] Hayes, A. F., and Krippendorff, K., 2007, “Answering the Call for a Standard Reliability Measure for Coding Data,” Commun. Methods Meas., **1**(1), pp. 77–89.

[71] Biering-Sørensen, F., Nielsen, J. B., and Klinge, K., 2006, “Spasticity-Assessment: A Review,” Spinal Cord, **44**(12), pp. 708–722.

[72] Fujisawa, T., Takagi, M., Takahashi, Y., Inoue, K., Terada, T., Kawakami, Y., and Komeda, T., 2007,

"Basic Research on the Upper Limb Patient Simulator," *2007 IEEE International Conference on Rehabilitation Robotics*, IEEE, Noordwijk, Netherlands, pp. 48–51.

[73] Grow, D. I., Wu, M., Locastro, M. J., Arora, S. K., Bastian, A. J., and Okamura, A. M., 2008, "Haptic Simulation of Elbow Joint Spasticity," *Symposium on Haptics Interfaces for Virtual Environment and Teleoperator Systems 2008 - Proceedings, Haptics*, IEEE, Reno, USA, pp. 475–476.

[74] Takhashi, Y., Komeda, T., Koyama, H., Yamamoto, S.-I., Arimatsu, T., Kawakami, Y., Inoue, K., and Ito, Y., 2011, "Development of an Upper Limb Patient Simulator for Physical Therapy Exercise," *2011 IEEE International Conference on Rehabilitation Robotics*, IEEE, Zurich, Switzerland, pp. 1–4.

[75] Park, H. S., Kim, J., and Damiano, D. L., 2011, "Haptic Recreation of Elbow Spasticity," *2011 IEEE International Conference on Rehabilitation Robotics*, IEEE, Zurich, Switzerland, pp. 1–6.

[76] Liang, J., Ewoldt, R. H., Tippett, S. R., and Hsiao-Wecksler, E. T., 2016, "Design and Modeling of a Passive Hydraulic Device for Muscle Spasticity Simulation," J. Med. Device., **10**(2), p. 20954.

[77] Liang, J., 2016, "Design of a Passive Hydraulic Simulator for Abnormal Muscle Behavior Replication," University of Illinois at Urbana-Champaign.

[78] Gharbaoui, I., Kania, K., and Cole, P., 2016, "Spastic Paralysis of the Elbow and Forearm," Semin. Plast. Surg., **30**(1), pp. 39–44.

[79] Manske, P. R., Langewisch, K. R., Strecker, W. B., and Albrecht, M. M., "Anterior Elbow Release of Spastic Elbow Flexion Deformity in Children with Cerebral Palsy," J. Pediatr. Orthop., **21**(6), pp. 772–777.

[80] O'Dwyer NJ, Ada L, N. P., 1996, "Spasticity and Muscle Contarcture Following a Stroke," Brain, **119**, pp. 1737–1749.

[81] Pei, Y., Liang, J., Ewoldt, R. H., Tippett, S. R., and Hsiao-wecksler, E. T., 2017, "Validation of a Passive Hydraulic Simulator for Spasticity Replication," *ASB Midwest Regional Meeting*, Grand Rapids, MI, Feb 23-24.

[82] Zatsiorsky, V., 2002, *Kinetics of Human Motion*, Human Kinetics.

[83] Mahony, R., Member, S., Hamel, T., and Pflimlin, J.-M., 2008, "Nonlinear Complementary Filters on the Special Orthogonal Group," IEEE Trans. Automat. Contr., **53**(5), pp. 1203–1218.

[84] Bartko, J. J., 1966, "The Intraclass Correlation Coefficient as a Measure of Reliability," Psychol. Rep., **19**(1), pp. 3–11.

[85] Krippendorff, K., 1980, *Reliability*, John Wiley & Sons, Inc.

[86] Fosang, A. L., Galea, M. P., McCoy, A. T., Reddihough, D. S., and Story, I., 2003, "Measures of Muscle and Joint Performance in the Lower Limb of Children with Cerebral Palsy," Dev. Med. Child Neurol., **45**(10), pp. 664–670.

[87] McGraw, K. O., and Wong, S. P., 1996, "Forming Inferences about Some Intraclass Correlation Coefficients," Psychol. Methods, **1**(1), pp. 30–46.

[88] Koo, T. K., and Li, M. Y., 2016, "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research," J. Chiropr. Med., **15**(2), pp. 155–163.

[89] Cicchetti, D. V., 1994, "Guidelines, Criteria, and Rules of Thumb for Evaluating Normed and Standardized Assessment Instruments in Psychology," Psychol. Assess., **6**(4), pp. 284–290.

[90] Krippendorff, K., 2013, *Content Analysis: An Introduction to It's Methodology*, SAGE Publications.

[91] Krippendorff, K., 2004, "Reliability in Content Analysis: Some Common Misconceptions and Recommendations," Hum. Commun. Res., **30**(3), pp. 411–433.

[92] Song, S. Y., Pei, Y., Liang, J., and Hsiao-Wecksler, E. T., 2017, "Design of a Portable Position, Velocity,

and Resistance Meter (PVRM) for Convenient Clinical Evaluation of Spasticity or Rigidity," *ASME. Frontiers in Biomedical Devices, 2017 Design of Medical Devices Conference*, ASME, Minneapolis, USA, p. V001T11A020.

[93] Song, S.Y., Pei, Y., Tippett, S.R., Lamichhane, D., Zallek, C. M. and H.-W., 2018, "Validation of a Wearable Position, Velocity, and Resistance Meter for Assessing Spasticity and Rigidity," *2018 Design of Medical Devices Conference*, American Society of Mechanical Engineers., Minneapolis, USA, p. V001T10A007-V001T10A007.

[94] Tuzson, A. E., Granata, K. P., and Abel, M. F., 2003, "Spastic Velocity Threshold Constrains Functional Performance in Cerebral Palsy," Arch. Phys. Med. Rehabil., **84**(9), pp. 1363–1368.

[95] Liang, J., 2016, "Design of a Passive Hydraulic Simulator for Abnormal Muscle Behavior Replication," University of Illinois at Urbana-Champaign.

[96] Karnopp, D., and Margolis, D. L., 2008, *Engineering Applications of Dynamics*, John Wiley & Sons.

[97] Bennett, D. J., Hollerbach, J. M., Xu, Y., and Hunter, I. W., 1992, "Time-Varying Stiffness of Human Elbow Joint during Cyclic Voluntary Movement," Exp. Brain Res., **88**(2), pp. 433–442.

[98] Corporation, P. H., 2009, *Catalog EPS 5370, Fluid Power Seal Design Guide*.

[99] Parker Hannifin, *Hydraulic Seals*.

[100] Levenberg, K., 1944, "A Method for the Solution of Certain Non-Linear Problems in Least," Q. Appl. Math., **2**, pp. 164–168.

[101] Marquardt, D., 1963, "An Algorithm for Least-Squares Estimation of Nonlinear Paramters," J. Chem. Inf. Model., **11**(2), p. 160.

[102] Bates, D. M., and Watts, D. G., 1988, *Nonlinear Regression Analysis and Its Applications*, Wiley.

[103] Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., and Wagener, T., 2016, "Sensitivity Analysis of Environmental Models: A Systematic Review with Practical Workflow," Environ. Model. Softw., **79**, pp. 214–232.

[104] Giesecke, F. E., 2012, *Technical Drawing with Engineering Graphics*.

[105] Beckwith, T. G. (Thomas G. ., Marangoni, R. D., and Lienhard, J. H., 2007, *Mechanical Measurements*, Pearson/Prentice Hall.

APPENDIX A: SELECTION OF RANGE OF MOTION

For severe spastic patients, their ROMs were commonly reported to be more constrained than healthy individuals. In the simulator design, in order to choose a reasonable set of ROMs for the ROM selector, several types of indication for ROM were considered. There were three choices, active ROM (aROM), passive ROM (pROM) and resting angle. In the literature, neither aROM nor pROM was found to have consistent pattern with MAS scores (Table A.1 and A.2) [41,42].

Table A.1: Columns 1-7 are reproduced from Table 1 in [42] shows the demographic data for 9 subjects. Last two columns are reproduced from Table 3 and 4 in [42] and show MAS scores for subjects in flexor and extensor muscles. MAS level numbering conventions: 0-5 (0-4).

ID	Gender	Age	BW (kg)	D_X*	Side	PROM (°)	*Clinical MAS (Biceps)	*Clinical MAS (Triceps)
1	M	41	76.0	ABI	R	123.31	4 (3)	3 (2)
2	M	73	77.8	MS	L	131.23	0 (0)	0 (0)
3	F	39	74.5	CP	L	128.74	2 (1+)	0 (0)
4	F	67	76.8	MS	L	112.75	4 (3)	1 (1)
5	F	62	60.6	ABI	L	114.66	2 (1+)	2 (1+)
6	F	46	98.7	ABI	R	104.40	3 (2)	3 (2)
7	M	33	84.5	ABI	L	75.08	3 (2)	0 (0)
8	M	28	73.9	ABI	L	95.26	2 (1+)	1 (1)
9	M	40	59.0	SCI	L	121.70	0 (0)	0 (0)

Table A.2: Reproduced from Table 1 in [41], which provides aROMs, pROMs and resting angles for 17 subjects. Full extension is 180°.

ID	Arm	Age	Sex	Weight (lb)	Post stroke	Elbow MAS	Resting angle	Passive ROM	Active ROM
1	Right	57	F	187	65	1+	142	40–180	56–180
2	Right	67	M	272	30	**1+**	**150**	65–169	82–151
3	Left	68	M	176	50	**0**	**168**	52–180	Full
4	Left	59	F	130	16	1	166	55–180	60–180
5	Right	75	M	195	93	1	170	48–180	Full
6	Right	50	M	180	26	0	175	70–180	Full
7	Left	89	M	226	71	**1+**	**150**	60–180	90–170
8	Right	62	M	200	93	**0**	**168**	50–180	50–170
9	Left	70	M	186	74	1+	132	60–152	None
10	Right	37	M	205	29	3	110	60–142	None
11	Right	65	M	135	97	**3**	**120**	84–154	None
12	Left	54	M	185	49	**3**	**120**	30–170	None
13	Right	54	M	230	41	2	136	42–170	None
14	Right	52	M	182	13	2	138	52-280	None
15	Right	76	M	214	12	1	160	55–170	55–160
16	Left	49	F	115	67	2	150	54–180	70–141
17	Right	50	F	144	107	1+	125	56–180	Full

A study relating resting elbow angles with MAS scores was found [1], which might guide us to design the ROM at each MAS level (Fig. A.1 and A.2). This definition of resting angle seems to align with our idea of limiting ROM and the resting angle was found inversely relating to MAS scores in an approximately linear fashion (Eq. (1)). The linear regression equation was rearranged as following and the ROM for each MAS level was calculated accordingly and rounded to the nearest 5° interval tick (Table A.3). In the scope of this thesis, the elbow joint angle is defined relative to upper arm, i.e., the joint angle starts from the most flexed position (50°) and ends at full extension (180°).

$$Resting\ Angle = \frac{10.851 - MAS\ Score}{0.0607} \tag{1}$$

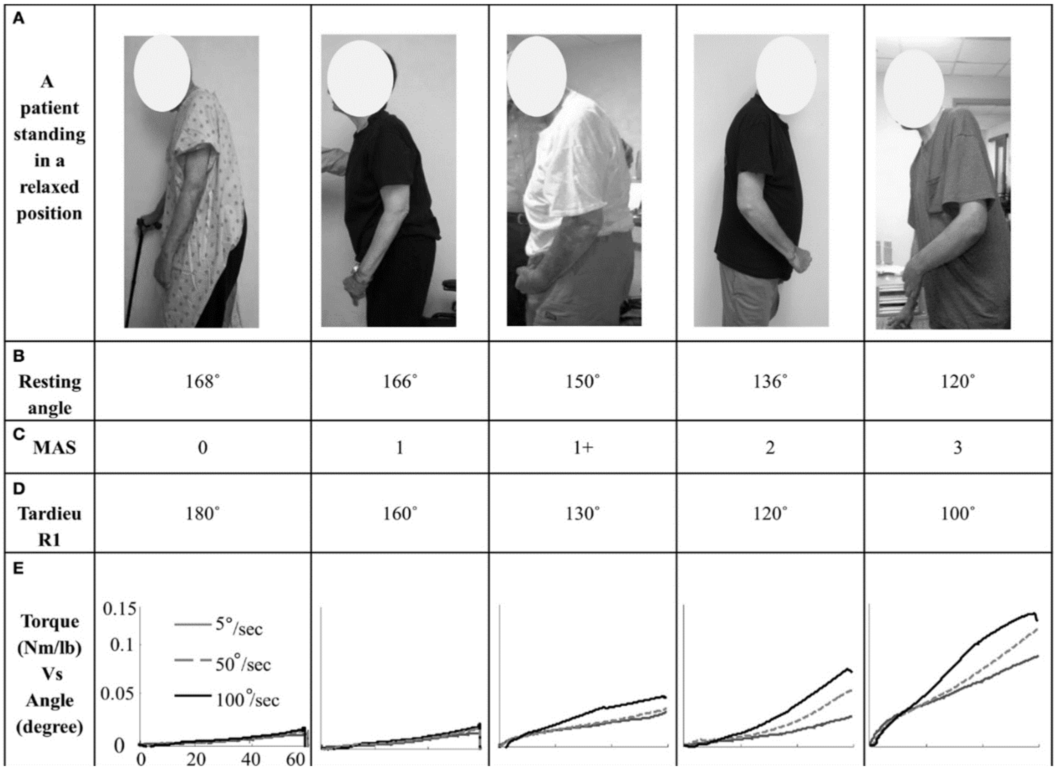


Fig. A.1: Reproduced from Fig.2 in [41] shows the resting elbow joint angle corresponding to each MAS level. Full extension is 180 deg.

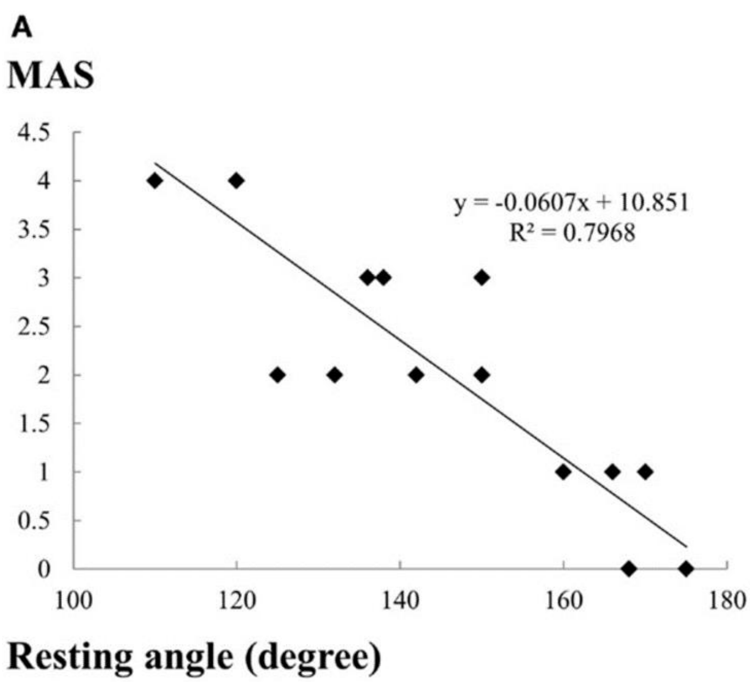


Fig. A.2: Reproduced from Fig. 4A in [41] shows the relationship between MAS level and resting angle.

Table A.3: The summary table of ROMs for MAS score 0-4. The reduction in ROM was calculated using the resting angle equation above (Column 2). ROM was calculated as the difference between the starting and ending elbow angular position.

MAS Score	Elbow Movement Range (°)	ROM (°)
0	50-180	130
1	50-165	115
2	50-145	95
3	50-135	80
4	50-110	60

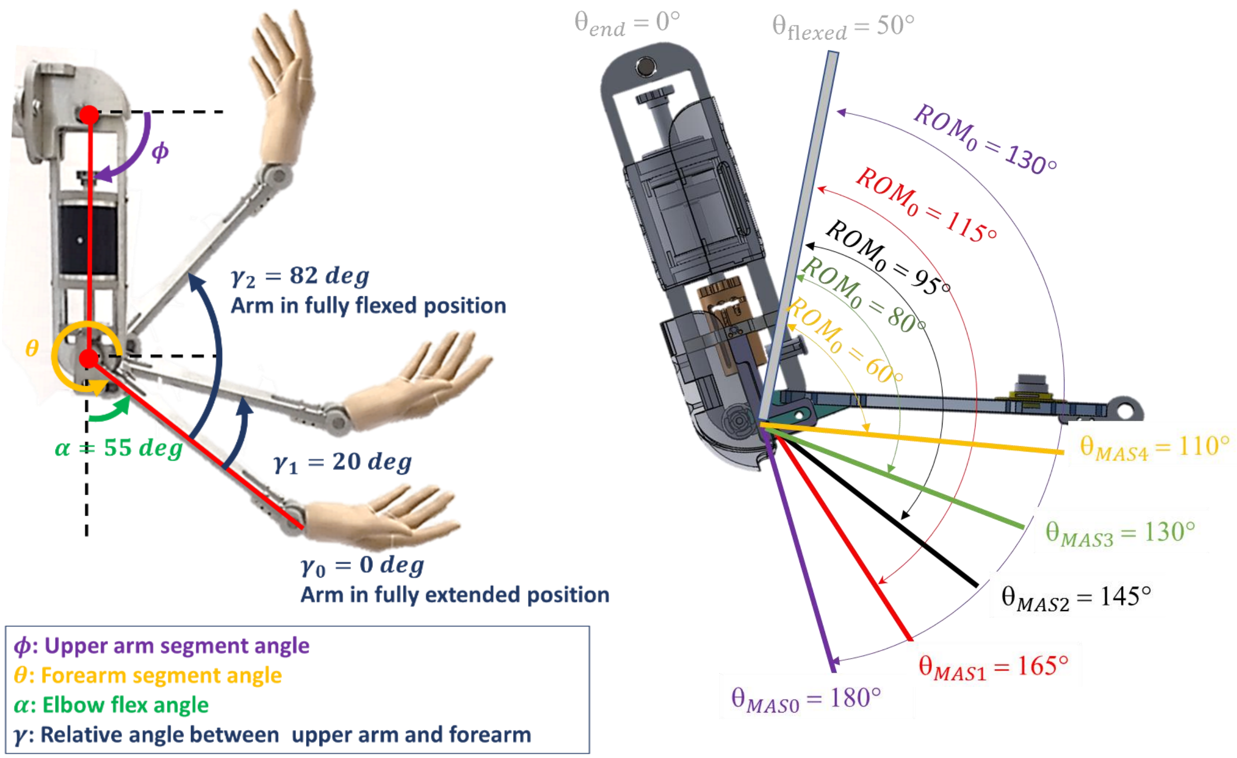


Fig. A.3: (left) The ROM and elbow definition for the first-generation simulator, reproduced from [95]. (right) The definition of elbow angle representation and a demonstration of ROMs for all five MAS levels for the second-generation simulator. The elbow angle convention in this thesis defines the full elbow extension as 180° and the simulator's most flexed position is 50°. All ROMs were defined as the traveling angular distance from the most flexed position.

To assign the proper ROM to the corresponding MAS score, multiple holes were drilled on the side the upper arm frame for the insertion of a ROM selector that functioned as a mechanical stop preventing the forearm from full extension. The kinematic relationship of the Scotch-Yoke linkage mechanism was used to determine the geometric location of each insertion hole on the upper frame (Eq. (2)) [96].

$$x = L_d sin(90° - \theta) \tag{2a}$$

$$\dot{x} = -L_d cos(90° - \theta)\dot{\theta} \quad (2b)$$

θ was the elbow joint angle, and x was linear motion distance of the linkage starting from the position when the elbow angle was 90 ° (forearm perpendicular to upper arm), $\dot{x}$ was linkage's linear motion speed, $\dot{\theta}$ was the elbow angular speed, and L_d was the distance between the pin bearing and the elbow joint center (maximum linkage lever arm). Exact hole locations were summarized in Fig. A.4.

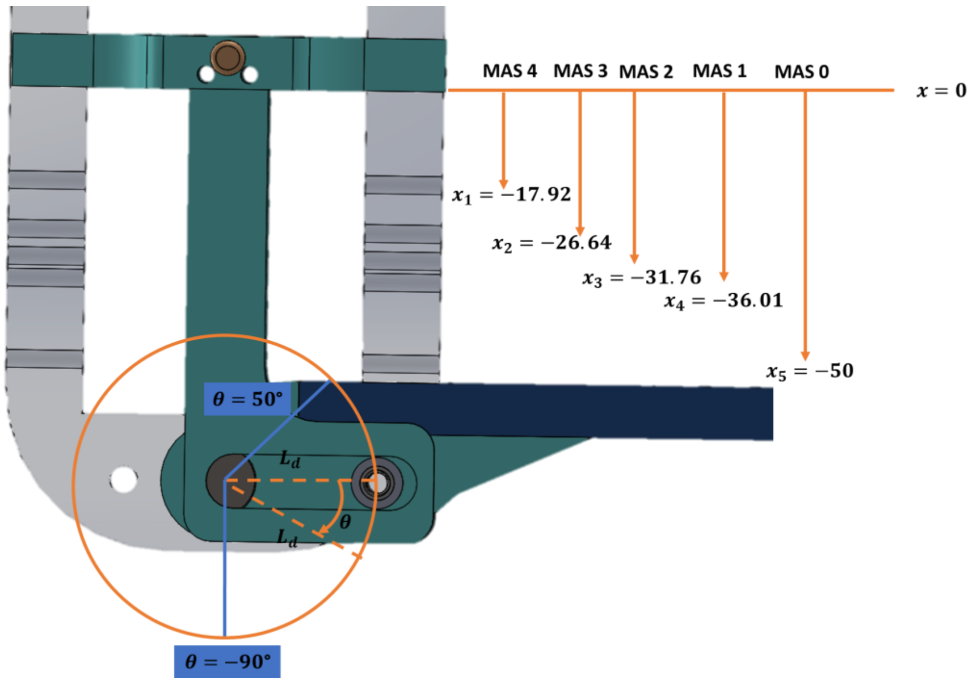


Fig. A.4: The demonstration of frame hole locations to select ROM for MAS levels 0-4. The unit of x_{1-5} is mm.

APPENDIX B: LINKAGE DESIGN

B1 LINKAGE DIMENSIONS

In the first-generation simulator, the linkage system had a ROM of 82° (Fig. B.1). The purpose of this linkage system is to convert the input angular motion of the forearm into a linear push/pull of the damper piston rod. Previously, since the piston rod was fixed to the linkage, the ROM was constrained by the constant cylinder stroke length of 4.2 cm (3 cm + 1.2 cm in Fig. B.1). As the new design decoupled the piston rod and the linkage using an added catch angle selector in the middle, this constraint was removed, giving more flexibility in design.

The new catch angle selector that functioned as a connecting part would travel back and forth in the frame clearance below the hydraulic damper. Its height now became a limiting factor in the place of the stroke length, which was measured to be about 4.8 cm in CAD. Therefore, the dimensions of the linkage needed to be modified to map this distance to a ROM of 130° (from the most flexed position (50°) to full extension (180°)).

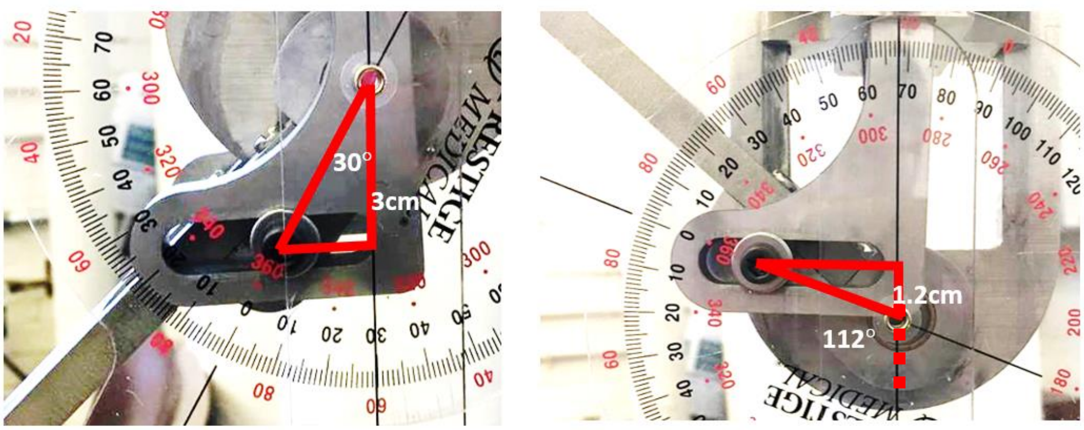


Fig. B.1: The original linkage that has a ROM of 82°. At both ends of the ROM, two parameters were measured: 1) the distance between the elbow joint center and the middle line of the slot was measured; 2) the angle from the vertical direction to the line connecting elbow joint center and pin bearing.

Using the vertical traveling distance of 4.8 cm as a constraint, a simple approach to increase ROM was sketched in Fig. B.2. L_d was defined as the radial distance between the elbow joint center and the pin bearing. Similar to Fig.B.1, two extreme cases were assumed, the most flexed position and full extension.

$$L_d + L_d \cos(50°) = 4.8 \text{ cm}$$

$$L_d = 2.921 \text{ cm} = 29.21 \text{ mm} \approx 29 \text{ mm}$$

From this calculation, the key dimensions of the linkage were determined and were used to guide the rest of the design (Fig. B.3). When the forearm is in full extension, the vertical distance between the elbow joint center and the pin bearing should be 29 mm. Also, when the elbow joint angle is in 90°, the pin bearing travels furthest in the slot (maximum moment arm) and the slot should be at least 29 mm to accommodate the pin's movement.

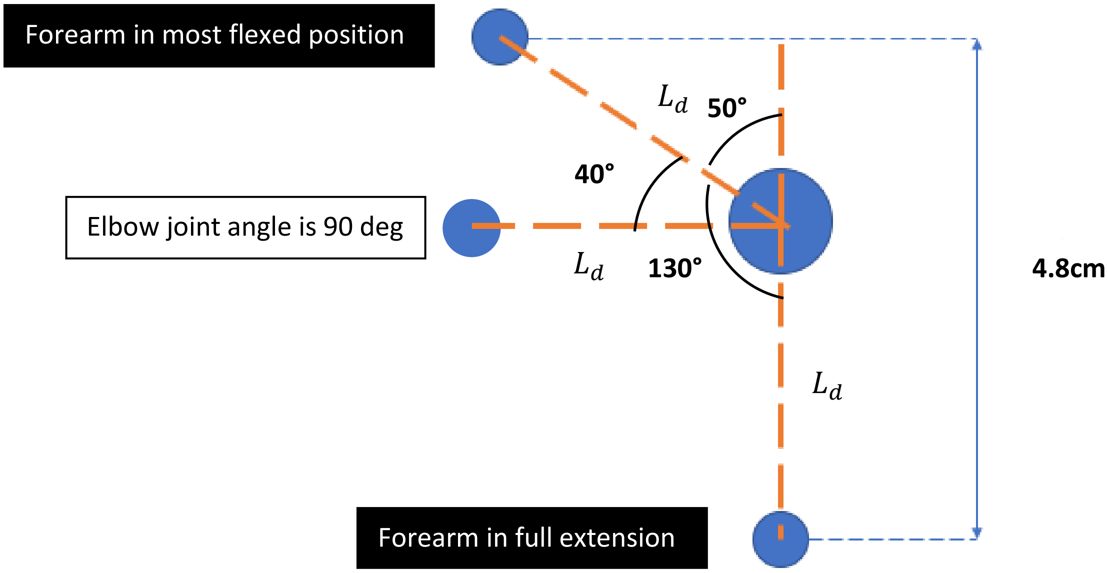


Fig. B.2: A sketch of the mathematical simplification for this design process. Black texts indicate two extreme cases corresponding to two ends of the ROM. The large blue circle represents the elbow joint center, while the small blue circle stands for the pin bearing (each corresponds to an elbow joint angle scenario).

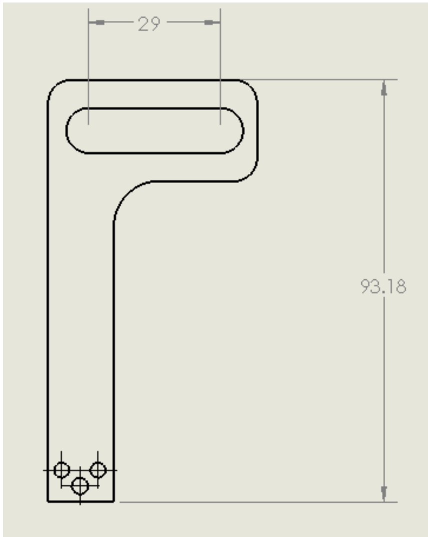


Fig. B.3: The final linkage design drawing. Units are in mm.

B2 Kinematic and Kinetic Analysis for Release Behavior

After the catch, the release behavior was created through a Scotch-Yoke linkage system in the elbow (Fig. B.4). The linkage behaved as a transmission mechanism that converted the user's rotary motion into the linear motion of the damper piston rod for resistance generation. Equation (1) related the output torque felt by the user (T_R), the damping force generated by the hydraulic damper (F_d), and the relationship was tuned through two design parameters, L_d and β (Eq. (2)). L_d defined the maximum possible moment arm length and β determined the location where the maximum output torque would occur.

$$T_R = F_d L_i \tag{1}$$

$$L_i = L_d \sin(\beta + \theta) \tag{2}$$

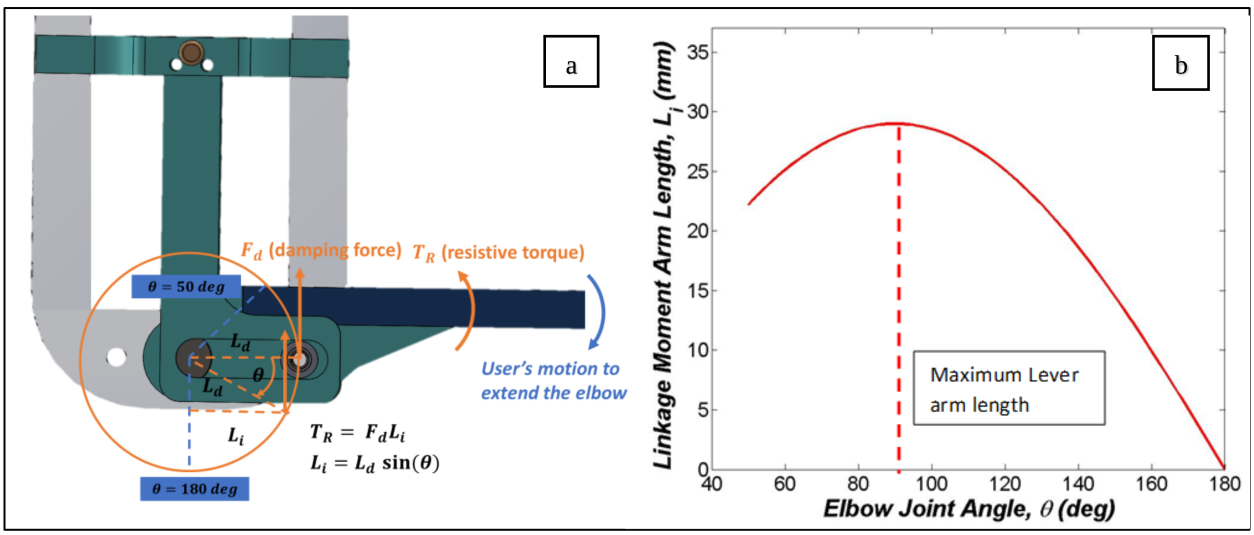


Fig. B.4: Linkage kinematics. (a) Dimensions and forces/torques related to the linkage system (β = 0°). The damping force location varies with elbow angle, θ, (two example locations were shown as two up-pointing orange lines). (b) Simulator linkage lever arm length trajectories. Linkage parameters were chosen to be L_d of 29mm and β of 0°, and the resulting maximum lever arm length occurred at around 90°.

β was the angle formed by the mounting hole on the cam piece, elbow joint center, and the forearm, which was zero in this design as they are all parallel. A nonzero β would shift the torque profile as an angular offset (Eq. (2)). While β and L_d were fixed parameters, the instantaneous linkage lever arm length (L_i) could be calculated at the given elbow angular position (θ) (Eq. (2)). In the current design, L_d was chosen to be 29 mm and β to be zero so that the angular position of the maximum torque would be around 90° (Fig. B.4b), which was chosen based on an average catch angle observed in the literature across spasticity patients of different severities [39]. The decrease of lever arm length after 90° would diminish the damping force and result in the release behavior of spasticity.

APPENDIX C: SELECTION OF CATCH ANGLE

This section will discuss evidence that support the choice of catch angles (length of slots) that determine when the hydraulic damper will kick in and cause the catch feeling. Among studies attempting to quantify spasticity, so far, there is no widely accepted definition of catch angle. For example, some investigators defined the catch angle as the joint angle when clinician's stretch speed decreases below 50°/s due to increased muscle tone [34], some reported the catch angle as the angle where the changing rate of the resistive torque is maximum [39], and some relied on clinician's perceived angle that correspond to a transient increase of resistance [28]

Due to the different definitions and testing protocols, the exact values or ranges of catch angle corresponding to each MAS level of elbow spasticity are still unclear, but some degree of consistency did present in the existing literature. Wu et al. reported that a mean catch angle at around 88°-98° for a subject group of cerebral palsy children with varied severities under different stretch speeds [39]. Pandyan et al. presented a mean catch angle of 137° for patients with MAS level 1 and 88° for those with MAS level 2 [28]. Ansari et al. recruited two raters to evaluate a group of 30 patients using MTS and all patients were accessed to have low MTS scores 0-2, where mean catch angles of 118.57° from Rater 1 and 112.93° from Rater 2 were reported [40].

the linkage mechanism was designed to reach a maximum lever arm length at around 90° and, the catch angle selector could be adjusted to four different angular positions, i.e. immediate, early, late and none (Table C.1), depending on the replication needs. The catch angle was selected to occur at 50°, 90°, 135°, or no catch at all. Based on Eq. (2a) in Appendix A, the slot lengths were derived to be 50 mm, 38.6 mm, 14.5 mm, and 0 mm, respectively.

Table C.1: A summary table for the corresponding catch angle for every MAS level.

MAS Level	Catch Angle Location	Catch Angle Value (°)
0	None	N/A
1	Late	135°
2	Early	90°
3	Early	90°
4	Immediate	50°

APPENDIX D: TORSIONAL SPRING DESIGN AND GRAVITY COMPENSATION

D1 TORSIONAL SPRING DESIGN

Besides viscous resistance provided by the hydraulic damper, other studies [27,47] also reported an elastic component (stiffness) in the muscle resistance and modeled the muscle resistance using a second-order linear system [47,75,97]. Therefore, to provide a more natural joint response, torsion springs were installed at the elbow of the simulator.

In the literature, [27] reported a stiffness of 0.07N/° for MAS 0 subjects and [47] reported a stiffness of $2.08 \times 10^{-4} Nm/°.kg$ for the less affected arm of patients, normalized by body weight. From the design point of view, we needed a stiffness with unit of Nm/°. Therefore, 0.07 N/° was multiplied by simulator's forearm length (0.24 m) and became 0.0168 Nm/°. $2.08 \times 10^{-4} Nm/°.kg$ was multiplied by 73 kg and became 0.015 Nm/°. These two numbers were fairly consistent and used in the selection of torsion springs.

The choice of off-the-shelf torsion spring online was limited and the most applicable one was requested.

Music-Wire Steel Torsion Spring
360 Degree Right-Hand Wound, 0.798" OD

Packs of 1

In stock
$2.19 per pack of 1
9271K133

ADD TO ORDER

Spring Type	Torsion
Material	Music-Wire Steel
Deflection Angle	360°
Wind Direction	Right Hand
OD	0.798"
For Maximum Shaft Diameter	0.516"
Wire Diameter	0.063"
Leg Length	2.000"
Number of Coils	11.50
Spring Length @ Maximum Torque	0.820"
Maximum Torque	5.520 in.-lbs.
RoHS	Compliant

$$k = \frac{d^4 E}{10.8 D N_a}$$

where
k = spring stiffness per turn
d = wire diameter
E = Young's Modulus
D = Mean coil diameter
N_a= number of active coils

(From Shigley's Mechanical Engineering Design)

Since the spring stiffness was not given on the website, k was calculated to be 0.0163 in.-lbs./deg or 0.00184Nm/deg.

In order to increase the spring stiffness, part of coils was cut off per the equation above. With less coils, the spring was made compact enough to fit into the current design.

Fig. D.1: (Left) The most applicable torsion spring found online, from McMaster. (Right) Design equation for torsion spring.

With only approximately four active coils, the stiffness was calculated to be 0.0051Nm/deg and two of these torsion springs in parallel achieved the total stiffness at the elbow to be 0.0101Nm/deg (theoretical value), which was similar to the design target found in the literature.

Fig. D.2: The machined torsion spring with only 4-5 coils.

The shape of the spring was customized so that one end could be embedded in the elbow cam and the other was fixed on the upper arm frame.

Fig. D.3: Modification on the current component to accommodate the added spring.

Table D.1: Summary table comparison of stiffness design target from the literature and the experimental results. Simulator's stiffness value was averaged over five trials (Fig. D.6 displayed a sample trial). All values were reported as mean (standard deviation).

	Stiffness (Nm/°)
Kumar et al.	0.0168 (0.0279)
McCrea et al.	0.015 (0.0091)
Hydraulic Simulator (Analytical)	0.0101
Hydraulic Simulator (Experimental)	0.00892 (0.0008)

The experimental value (0.00892 Nm/°) was close to the analytical design value (0.0101 Nm/°) and was consistent with the muscle stiffness range observed in the actual patients [27,47].

D2 GRAVITY COMPENSATION

We would like to validate the simulated resistance generated by the simulator against other quantification studies in the literature [27,44]. However, direct comparison was not entirely valid because those two studies [27,44] were conducted in a horizontal plane and gravity of the forearm and hand did not play a role in their measurements. On the other hand, the simulator's results were obtained by testing in a vertical plane, so in order to compare with the reported quantitative results, it was necessary to perform gravity compensation on the simulator testing results. In Fig. D.4, instead of $F_{measured}L$, $\tau_{resisance}$ should be used for comparison and was the sum of applied torque and torque due to gravity. This step of considering torque due to gravity was referred to as gravity compensation.

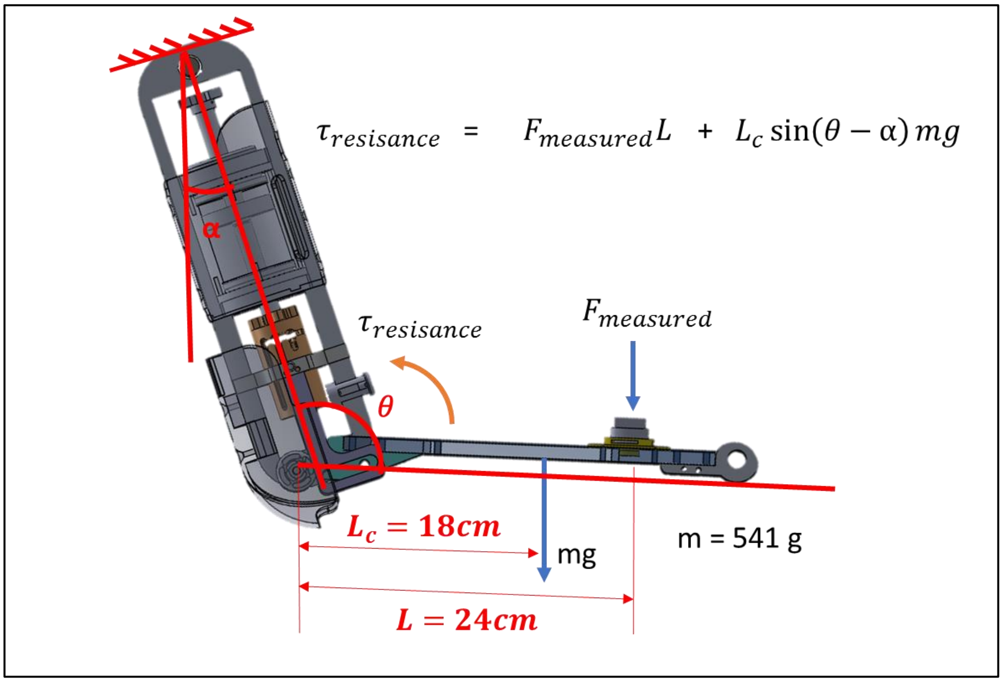


Fig. D.4: The free body diagram and key dimensions of the simulator.

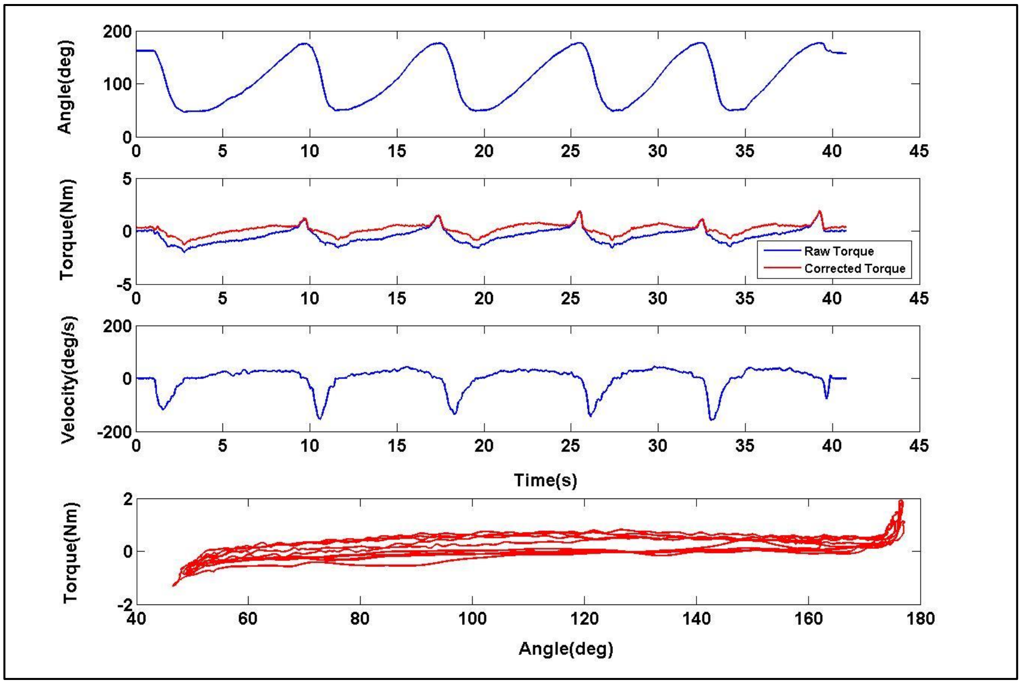


Fig. D.5: Raw testing data of five passive extension stretch. Angle, torque, and angular speed were plotted again time. Torque-time plot included both raw and gravity-corrected torques. Corrected torque was plotted against angle. Corrected torque-angle plot showed a relatively consistent hysteresis behavior.

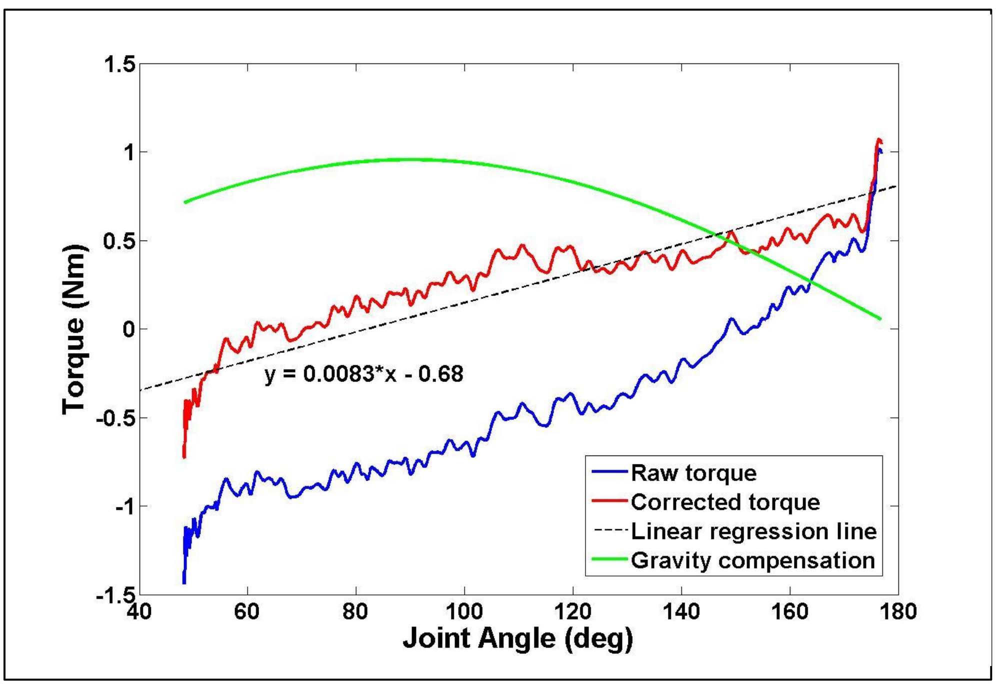


Fig. D.6: A sample data from one extension stretch (MAS 0). Raw torque, corrected torque and the amount of gravity compensation were plotted for demonstration. A linear regression line was plotted for corrected torque curve to find spring stiffness as there was no damping force.

APPENDIX E: HYDRAULIC DAMPER SEAL DESIGN

During the operation and testing of the first-generation simulator, some leakage was found as the piston rod traveled in and out of the hydraulic cylinder. The original custom hydraulic damper was only equipped with one rod seal at each endcap and they could not help the piston rod resist the radial load posed by the linkage. To handle this issue, industry-level hydraulic dampers usually have several layers of seals (piston seal, buffer seal, rod seal and wiper seal) and a guide ring for resisting the radial load upon the piston rod [98,99]. Therefore, to minimize leakage, a new sealing plate containing a guide sleeve bearing and a U-cup seal was fabricated and attached to the bottom endcap (Fig. E.1). This location was chosen because it was more susceptible to leakage due to gravity of the working fluid and its vicinity to the origin of the radial load. The guide sleeve bearing was composed of a stack of three brass bearings to resist any transverse load upon the piston rod. The U-cup seal would support the existing rod seals and function as a wiper seal that prevented the contaminants from entering the cylinder chamber. Finally, an auxiliary rubber O-ring was inserted between the seal plate and the endcap to seal any gap at the interface.

Fig. E.1: The hydraulic damper with the new sealing plate.

APPENDIX F: HYDRAULIC DAMPER TESTING AND ANALYSIS

F1 Hydraulic Damper Design Overview

The speed-dependent muscle resistance was generated by a custom viscous hydraulic damper (i.e., resistance level selector) and its design was described in detail in [76,95]. An analytical fluid model was developed to predict the damper behavior for deterministic design [76,95]

$$F_d = \frac{C_1(24L_{po}R_p^3\mu R_c\pi)/(4R_p(R_c-R_p)^3)}{C_2(\frac{3}{4R_p(R_c-R_p)^3})R_{po}^{C_3 4} + 1}\dot{x} + F_f. \tag{1}$$

Equation (1) related the output damping force (F_d) to the effective radius of the two piston orifices (R_{po}) and the fluid viscosity (μ) at a given input speed, $\dot{x}$. Other terms depended on the damper geometries, where L_{po} was the length of the piston orifice, R_c was the inner radius of the cylinder chamber, R_p was the radius of the piston head. The term, $R_c - R_p$, basically was the gap width between the inner surface of the cylinder and the piston head. F_f was the sliding friction caused by hydraulic seals (experimentally determined). The coefficient in front of $\dot{x}$ essentially represented the linear viscosity of the hydraulic damper (B_L) at a given orifice size (Eq. (2)) and served as a design parameter to characterize the viscosity associated with each MAS level. Eventually, to add model robustness, C_1, C_2 and C_3 were three correction factors to account for any unknown physical, fabrication, assembly, and testing errors and were experimentally determined during the hydraulic damper evaluation.

$$B_L = \frac{C_1(24L_{po}R_p^3\mu R_c\pi)/(4R_p(R_c-R_b)^3)}{C_2(\frac{3}{4R_p(R_c-R_b)^3})R_{po}^{C_3 4} + 1} \tag{2}$$

Table F.1: The summary table for design nominal value for each parameter in the fluid model and the manufacturing uncertainty of each parameter used in the uncertainty analysis.

Symbol	Design Nominal Value	Manufacturing Tolerance
C_1	1	N/A
C_2	1	N/A
C_3	1	N/A
L_{po}	7.0 *mm*	± 0.03 *mm*
R_p	25.4 *mm*	± 0.06 *mm*
R_c	25.7 *mm*	± 0.06 *mm*
$R_{po}{}^{a}$	3.53, 2.83,2.12, 1.41, 0.71 *mm*	± 0.03 *mm*
μ	150 *mPa.s*	± 15 *mPa.s*

[a] Calculated as the effective radius of two equal-size orifices, where the actual orifice radii are 2.5, 2.0, 1.5, 1.0, and 0.5 *mm* for MAS 0-4, respectively.

F2 Hydraulic Damper Evaluation and Results

The purpose of this evaluation was to characterize the dependency of the damping force on the orifice geometry and the input speed. A commercial material testing system (Instron 5967; High Wycombe, United Kingdom) was used. The author used a similar evaluation methodology presented in [95]. The damper was fixed to the testing machine's gripper through a custom bracket (Fig. E.3), and the weight of the damper and the mounting bracket was zeroed out prior to the test so that the reading only reflected the measured damping force. During each test, the piston rod was inserted for 35 mm and displacement vs. force data were collected. Each of the orifice radii (2.5 mm, 2.0 mm, 1.5 mm, 1.0 mm, and 0.5 mm) was subject to under four different speeds (250 mm/min, 500 mm/min, 750 mm/min, and 1000 mm/min (the maximum testing speed allowed)). For each combination (speed and orifice size), the test was repeated three times, and the average damping force as well as standard deviation were reported. In total, 60 trials (4 speeds × 5 orifice sizes × 3 trials) were performed.

Experimental results of the damping force vs. orifice radius under different speeds were fitted with regression lines to investigate the speed dependency of each orifice size. The slope of the regression line was calculated to obtain the experimental values of linear viscosity, B_L, and the residual errors between the experimental data and the regression lines were reported to determine the linearity of the viscosity value for a given orifice size under a variety of speeds. The fluid model prediction (Eq. (1)) were fitted to the experimental results to obtain correction factors (C_1, C_2, and C_3) through a nonlinear curve fitting scheme using Levenberg-Marquardt algorithm [100,101] in MATLAB (R2013b, MathWorks Inc., Natick, MA) (MATLAB Function: lsqcurvefit). These factors were chosen to minimize the error residual between the experimental results and analytical prediction in a least-squared sense. The error residual and percent error between the prediction and experiment were reported for each testing condition. Due to the limited number of experimental combinations of orifice size and input speed, the analytical model was fitted with only 20 experimental data points, so the confidence of these correction factors was also examined. The uncertainties associated with each correction factor were also approximated using the Jacobian matrix and mean squared error obtained in the curve fitting process [102]. The measure of uncertainties was reported in the form of standard deviation and 99% confidence interval of the value of each correction factor.

The correlation of the output damping force (F_d) with respect to input speeds ($\dot{x}$) and orifice radius (R_{po}) was obtained (Fig. F.1). By visual inspection, at a given orifice size, the damping force mostly increased linearly as the speed increased and the linear regression lines well represented the trends of experimental data points. Between the experimental results and regression lines, the mean squared errors (MSE) were 7.36 N, 0.46 N, 0.82 N, 0.32 N, and 0.85 N, respectively for orifice radii from 0.5 mm to 2.5 mm. Experimental values of linear viscosity, B_L, were calculated using the slopes of regression lines (Table F.3).

At a given input speed, when orifice radius varied from 0.5 mm to 2 mm, the output damping force was very sensitive to orifice size (Fig. F.1b). Beyond this range, the damping force approached to the baseline friction on one end (radius > 2 mm), and on the other end (radius < 0.5 mm).It reached a plateau that was determined by the gap width between the piston head and cylinder chamber ($R_c – R_p$). The separation among experimental data points in the Y-direction at each orifice size across different speeds became more prominent as the orifice size decreases and indicated that higher speed-dependencies associated with smaller orifices, which agreed with the steeper slopes in Fig. F.1a as the orifice size decreased.

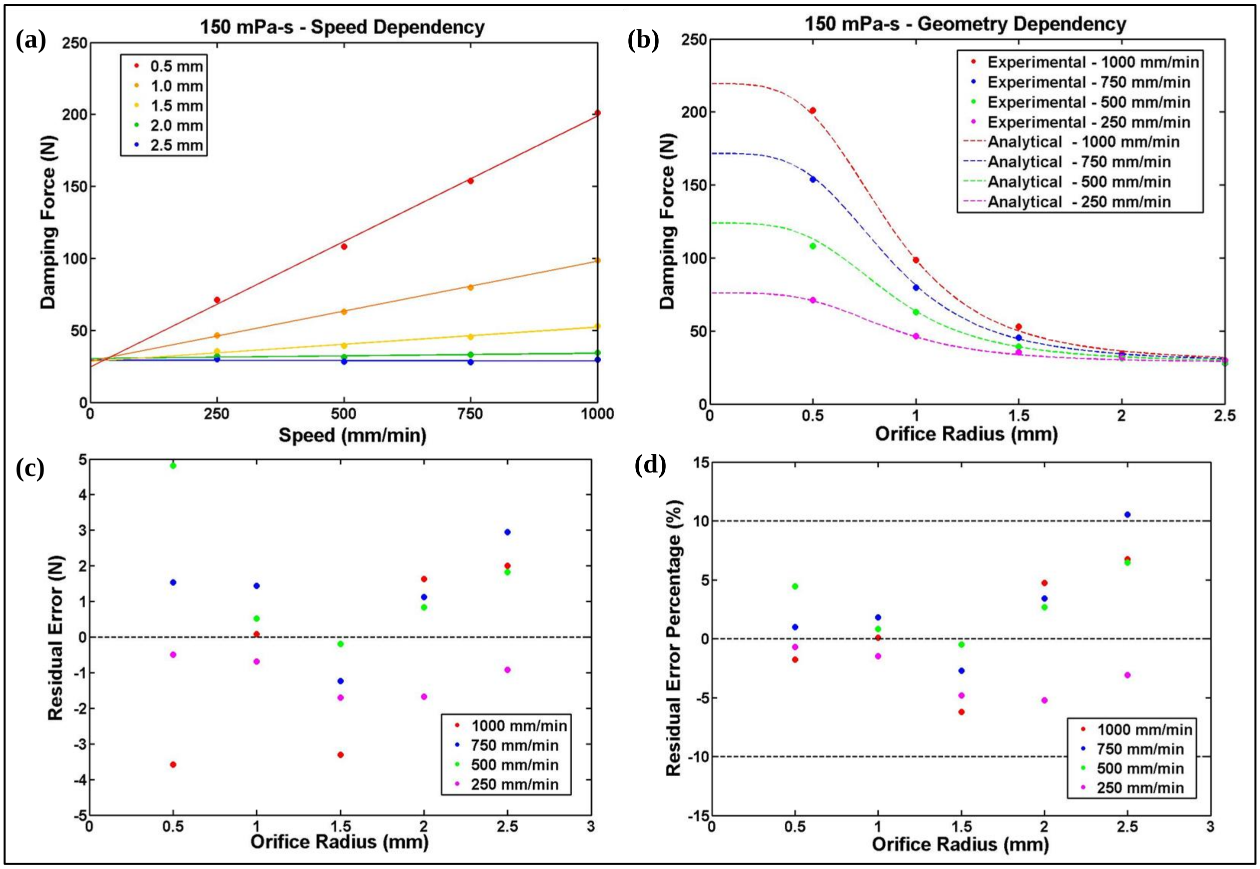


Fig. F.1: (a) The speed dependency of the damping force under different orifice sizes. The experimental data points were plotted as dots and fitted with regression lines. (b) The geometry dependency of the damping force under different input speeds. The experimental data points were plotted as dots and fitted with fluid model predictions with a set of correction factor (C_1= 0.94, C_2 = 0.07, and C_3 = 0.94). (c) The residual error between experimental results and model prediction in (b). (d) The percent error in (b).

The scaled fluid model had a good agreement with the experimental results, where the residual errors were within ± 5 N for all testing conditions and the percent errors were approximately within ±10% (Fig. F.1c and F.1d). The values of correction factors, C_1, C_2, and C_3, and their uncertainties were determined (Table F.2). With the values of correction factors obtained, the analytical linear viscosity (C_1, C_2, and C_3 = 1) and corrected analytical linear viscosity (C_1 = 0.94, C_2 = 0.07, and C_3 = 0.94) were calculated and listed in Table F.2, Column 2-3.

Table F.2: The values of correction factors, C_1, C_2, and C_3, that scaled the analytical fluid model and standard deviations (Column 2) as well as 99% confidence interval (Column 3).

Correction Factor	Value	99% Confidence Interval
C_1	0.94 ± 0.02	0.88-1.01
C_2	0.07 ± 0.09	-0.19-0.33
C_3	0.94 ± 0.05	0.79-1.08

Table F.3: Table compared the model prediction, corrected model prediction, and experimental values of linear viscosity (B_L). Units were converted to N/(cm/s) for easy physical interpretation.

Orifice Radius (mm)	Analytical Linear Viscosity (N/(cm/s))	Corrected Analytical Linear Viscosity (N/(cm/s))	Experimental Linear Viscosity (N/(cm/s))
0.5	95.46	101.64	104.64
1.0	22.62	42.24	41.58
1.5	5.28	13.02	14.22
2.0	1.74	4.80	2.22
2.5	0.72	2.10	-0.36

By fitting the linear regression lines, low MSEs between the experimental data points and regression lines indicated a good linearity between the output damping force and the input speed, and therefore suggested a relatively consistent linear viscosity (B_L) for a given orifice size across speeds (Fig. F.1a). Five different orifice sizes generated five distinct linear viscosities, demonstrating the feasibility of mimicking different levels of muscle resistance with this adjustable damper design.

The damping force was found most sensitive to change of orifice radius between 0.5 and 2 mm (Fig. F.1b). For orifice radius above 2 mm, the damping force generated by the viscous fluid traveling through orifice was minimal and the resistance was dominated by the baseline friction (measured to be 28.3 N). On the other hand, when orifice radius was below 0.5 mm, the amount of flow through the orifice was comparable to that of flow through the gap between piston head and inner surface of the cylinder chamber (gap width: $R_c - R_p$) and therefore the damping force approached to a plateau determined by the gap width. This finding implied that the most effective design space of the orifice radius was between 0.5 mm and 2 mm at current damper dimension. The upper bound of possible B_L (the damping force plateau) was determined by the gap width. If an application requires higher B_L, the gap could be even narrowed down, but its accuracy would be more susceptible to manufacturing uncertainties. On the other hand, the lower bound of possible B_L was at 2.5 mm and the viscous effect became negligible beyond this orifice size.

Three correction factors were obtained through fitting the analytical fluid model to experimental results (Table F.2). Ideally, if all mechanical component dimensions are perfect and the model accounts for all underlying physical processes, these factors should be all equal to 1. However, in our case, C_1 and C_3 were fairly close to 1, while C_2 (0.07) deviated much from 1. Although these three factors were not completely independent, this deviated C_2 value might suggest that terms mainly corrected by C_2 (R_p and R_c) might deviate significantly from the design values. Considering the design value of the gap width ($R_c – R_p$) was only 0.3 mm (Table F.1), the manufacturing error potentially might have a much greater impact on this term. Moreover, it was $(R_c - R_p)^3$ that was corrected by C_2, so the effect of component dimension error would be amplified even more and caused large deviation in C_2. Additionally, although with these determined correction factors, the corrected fluid model matched well with the experimental data (Fig. F.1c and d), this set of correction factors was only applicable to this specific damper prototype used in the experiment and probably be different in another prototype. Along with the values of correction factors, a relatively small standard deviations and narrow confidence intervals were also obtained (Table F.2), indicating that the number of experimental data points was sufficient to characterize the damper behavior. Although more testing conditions (different combinations of input speeds and orifice sizes) could be used to further narrow down the possible range of these correction factors, the current level of certainty was sufficient to draw qualitative conclusions on the potential factors that deviated the model.

F3 Fluid Model Sensitivity Analysis

To investigate the robustness and reliability of the hydraulic damper, a sensitivity analysis was carried out to determine the sensitivity of the resulting linear viscosity with respect to the variation of each parameter in Eq. (2). Considering the design nominal values were known (Table F.1), a local sensitivity analysis that aimed to assess the model performance around a nominal or reference parameter value was used, instead of a global sensitivity analysis that examined the entire design space. A common method for local sensitivity analysis is OAT (One-At-a-Time), which varies one parameter from its nominal value at a time and observing its effect on the model output [103]. Therefore, the partial derivative of B_L was taken with respect to each parameter to calculate the sensitivity

$$s_i(\bar{x}_i) = \frac{\partial B_L}{\partial x_i}|_{\bar{x}_i}. \tag{3}$$

x_i is the i-th parameter, $\bar{x}_i$ is the design nominal value for the i-th parameter (Table F.1), and $s_i(\bar{x}_i)$ is the measure of sensitivity for i-th parameter evaluated at $\bar{x}_i$. Experimentally determined correction factors, C_1, C_2, and C_3 were used in calculation (0.94, 0.07, 0.97, respectively).

The sensitivity of the output linear viscosity with respect to each input parameter were calculated (Table F.4 and analytical expressions are at the end of this section). Sensitivity measures of four geometric parameters $(R_c, R_p, R_{po}, L_{po})$ were grouped into the first category (Table F.4, Column 2-5) and that of the working fluid viscosity (μ) was listed in the second category (Table F.4, Column 6). As two categories had different unit of measurement (Category 1 - $\frac{\Delta N/(cm/s)}{\Delta mm}$, Category 2 - $\frac{\Delta N/(cm/s)}{\Delta \mathrm{m}Pa.s}$), the comparison was only appropriate within a given category. The positive or negative sign for each term indicated the increase/decrease of the linear viscosity if this specific parameter increased.

The sensitivity measures in each category were calculated at all five orifice radii (0.5 – 2.5 mm) and it could be observed that sensitivity measures of all parameters decreased as the orifice radius became larger. At the smallest orifice (radius: 0.5 mm), the output damping force was highly sensitive to errors in dimensions of cylinder chamber inner radius (R_c) and the piston head radius (R_p), while the errors in dimensions of orifice length (L_{po}) and orifice radius (R_{po}) had much less impact. At larger orifices (radii: 1 mm and 1.5 mm), the sensitivity measures of R_c and R_p were quickly decreased but also dominated the other two parameters. At the two largest orifices (radii: 2.0 mm and 2.5 mm), the damping force became equally sensitive to the error in dimension of orifice size (R_{po}), compared to R_c and R_p.

Table F.4: The sensitivity measure of each parameter of the fluid model. Columns 2-5 were geometry sensitivity with a unit of $\frac{\Delta N/(cm/s)}{\Delta mm}$, and Column 6 was working fluid viscosity sensitivity with a unit of $\frac{\Delta N/(cm/s)}{\Delta mPa.s}$.

Orifice Radius (mm)	$\frac{\partial B_L}{\partial L_{po}}$	$\frac{\partial B_L}{\partial R_{po}}$	$\frac{\partial B_L}{\partial R_p}$	$\frac{\partial B_L}{\partial R_c}$	$\frac{\partial B_L}{\partial \mu}$
0.5	15.8	-20.5	1070	-1060	0.735
1	11.0	-67.0	522.0	-512	0.511
1.5	5.02	-43.1	112.0	-108	0.234
2	2.14	-17.3	21.3	-19.0	0.0998
2.5	0.998	-6.98	5.08	-4.00	0.0466

The analytical expressions for sensitivity measures of parameters ($R_c, R_p, R_{po}, L_{po}, \mu$) in the fluid model were presented below (Eq. (3-7)):

$$\frac{\partial B_L}{\partial L_{po}} = \frac{18.8 C_1 R_c R_p^2 \mu}{(R_c - R_p)^3 (\frac{0.75 C_2 R_{po}^{C_3 4}}{R_p (R_c - R_p)^3} + 1)} \tag{4}$$

$$\frac{\partial B_L}{\partial R_{po}} = \frac{56.5 C_1 C_2 C_3 L_{po} R_c R_p R_{po}^{C_3 4 - 1} \mu}{(R_c - R_p)^6 (\frac{0.75 C_2 R_{po}^{C_3 4}}{R_p (R_c - R_p)^3} + 1)^2} \tag{5}$$

$$\frac{\partial B_L}{\partial R_p} = \frac{75.4 C_1 L_{po} R_c R_p^2 \mu (9 C_2 R_{po}^{C_3 4} + 8 R_c^3 R_p + 4 R_p^4 - 12 R_c^2 R_p^2)}{(3 C_2 R_{po}^{C_3 4} + 12 R_c R_p^3 + 4 R_p R_c^3 - 4 R_p^4 - 12 R_c^2 R_p^2)^2} \tag{6}$$

$$\frac{\partial B_L}{\partial R_c} = \frac{75.4 C_1 L_{po} R_p^3 \mu (3 C_2 R_{po}^{C_3 4} - 8 R_c^3 R_p - 4 R_p^4 + 12 R_c^2 R_p^2)}{(3 C_2 R_{po}^{C_3 4} + 12 R_c R_p^3 + 4 R_p R_c^3 - 4 R_p^4 - 12 R_c^2 R_p^2)^2} \tag{7}$$

$$\frac{\partial B_L}{\partial \mu} = \frac{18.8 C_1 L_{po} R_c R_p^2}{(R_c - R_p)^3 (\frac{0.75 C_2 R_{po}^{C_3 4}}{R_p (R_c - R_p)^3} + 1)} \tag{8}$$

F4 Manufacturing Uncertainty Analysis

With the sensitivity measures of input parameters determined in the previous section, how manufacturing uncertainties propagated in the resulting linear viscosity (B_L) was also studied. A common tolerance for machining process such as drilling, milling, and lathing were used in the analysis (Table F.1) [104]. Note that to relate part size with tolerance size, instead of using R_c, R_p, R_{po}, the diameters of these parts were used to select tolerances (for diameter of 0-15 mm: 0.03 mm, diameter of 38-7 1mm: 0.06 mm). Therefore, in calculation, the tolerances associated with R_c, R_p were considered to the half of the values listed in Table F.1, Column 3, while the tolerance of R_{po} was $\frac{\sqrt{2}}{2}$ of the list value, as R_{po} was defined as the effective radius of two orifices. A variance of 10% for the working fluid's viscosity was considered according to the product datasheet.

The uncertainties in these five inputs were propagated to quantify their effects on the model output (B_L) [105] as following,

$$u_{B_L} = \sqrt{\left(u_{x_1} s_1(\bar{x}_1)\right)^2 + \left(u_{x_2} s_2(\bar{x}_2)\right)^2 + \cdots + \left(u_{x_n} s_n(\bar{x}_n)\right)^2} \quad . \tag{9}$$

u_{B_L} was the overall uncertainty in the output linear viscosity and u_{x_i} was the uncertainty in the i-th variable. Furthermore, the contribution from each variable to B_L's uncertainty was also determined to identify the main source of uncertainty by squaring both sides of Eq. (9) and finding the ratio of each term relative to $u_{B_L}^2$.

The overall uncertainty of output linear viscosity (B_L) due to individual uncertainties in the fluid model parameters ($R_c, R_p, R_{po}, L_{po}, \mu$) was propagated at five orifice sizes and in addition the uncertainty contribution of each parameter was also summarized (Table F.5). The overall percent error was high at all orifice sizes (35-55%) but the contribution from the parameters were different. At the small orifice (radius: 0.5 mm and 1.0 mm), the uncertainty in B_L was mainly attributed to the manufacturing error in dimensions of R_p and R_c in accordance with sensitivity analysis results (Table F.4). While the contributions from R_p and R_c was rapidly decreased as the orifice size increased (radius: 2.0 mm and 2.5 mm), the uncertainty due to the manufacturing error of working fluid viscosity became the dominating factor in the overall uncertainty of B_L and other contributions from other parameters became almost negligible when the orifice radius was larger than 1.5 mm.

Table F.5: Manufacturing uncertainty analysis results for five different orifice sizes. Individual model parameter's effect on error of B_L (relative contribution percentage among all five parameters) was reported. Percent error was calculated as the ratio of u_{B_L} and analytical B_L.

Orifice Radius (mm)	Error in B_L due to each model parameter (N/(cm/s))					Total Error u_{B_L} (N/(cm/s))	Analytical Linear Viscosity (N/(cm/s))	Percent Error (%)
	L_{po}	R_{po}	R_p	R_c	μ			
0.5	0.474 (0.010%)	-0.431 (0.010%)	32.1 (48%)	-31.8 (47%)	11.0 (5.6%)	46.5	102.0	46
1.0	0.330 (0.02%)	-1.41 (0.36%)	15.7 (45%)	-15.4 (44.0%)	7.67 (11%)	23.3	42.2	55
1.5	0.151 (0.06%)	-0.905 (2.4%)	3.37 (32.0%)	-3.231 (30%)	3.51 (35%)	5.91	13.0	45
2.0	0.0640 (0.13%)	-0.363 (4.2%)	0.639 (13%)	-0.570 (10%)	1.50 (72%)	1.76	4.80	37
2.5	0.0300 (0.16%)	-0.147 (3.9%)	0.152 (4.2%)	-0.120 (2.6%)	0.699 (89%)	0.741	2.10	35

Results of the manufacturing uncertainty analysis were built upon the sensitivities measures obtained and further identified the sources of uncertainty across orifice sizes (Table F.4). In the overall percent error, the contribution from geometric uncertainties diminished at orifice radii above 1.5 mm and the uncertainty of working fluid viscosity became the major contributor. The overall percent error seemed to converge to approximately 35%. To achieve a reasonable manufacturing cost and a relative robust performance, a design should use high-quality working viscous fluid (viscosity error < 10%) and avoid deploying orifices smaller than 0.5 mm, which would allow regular manufacturing tolerance specifications for R_c and R_p.

It is worth noting that the results of sensitivity and uncertainty analysis only apply to the specific damper used in the testing. When another prototype damper is fabricated, another set of correction factors (C_1, C_2, and C_3) will be experimentally determined, resulting in different sensitivity measures and propagation of the uncertainties. The fluid model still requires additional work to narrow down the range of possible correction factors by more accurately modeling the physical phenomena involved. With large fluctuation in correction factor values, the sensitivity and uncertainty analysis cannot provide consistent insights for design and manufacturing considerations.

APPENDIX G: CLINICAL EVALUATION STUDY MATERIALS AND RESULTS

For Research Staff Only
Date:
Clinician ID:
Research Assistant:

Questionnaire for Clinical Review

First, we would like to know your background as a medical professional, as well as your clinical experience related to spasticity evaluation.

Background information:

Medical Education:
Area of Specialization(s):
Years of Practice:
Years of Spasticity Evaluation:
Number of Spasticity Patients Evaluated per month:

In this part of the test, we would like you to closely examine each simulated MAS level and provide your comments in these aspects:

1) **Magnitude of resistance**
2) **Catch and release behavior**
3) **Speed dependency**
4) **Range of motion**

Please go through each simulated MAS level on the simulator with the research assistant. Based on your previous clinical experience with actual spasticity patients, please select the description closest to your feeling.

Your comments and feedback are highly appreciated and will be crucial for our future development of the simulator.

MAS Level 0

	Too Little		About Right		Too Much
Magnitude of resistance	☐	☐	☐	☐	☐
Catch and release behavior	☐	☐	☐	☐	☐
Speed dependency	☐	☐	☐	☐	☐
Range of motion	☐	☐	☐	☐	☐

Comments:

MAS Level 1

	Too Little		About Right		Too Much
Magnitude of resistance	☐	☐	☐	☐	☐
Catch and release behavior	☐	☐	☐	☐	☐
Speed dependency	☐	☐	☐	☐	☐
Range of motion	☐	☐	☐	☐	☐

Comments:

MAS Level 2

	Too Little		About Right		Too Much
Magnitude of resistance	☐	☐	☐	☐	☐
Catch and release behavior	☐	☐	☐	☐	☐
Speed dependency	☐	☐	☐	☐	☐
Range of motion	☐	☐	☐	☐	☐

Comments:

MAS Level 3

	Too Little		About Right		Too Much
Magnitude of resistance	☐	☐	☐	☐	☐
Catch and release behavior	☐	☐	☐	☐	☐
Speed dependency	☐	☐	☐	☐	☐
Range of motion	☐	☐	☐	☐	☐

Comments:

MAS Level 4

	Too Little		About Right		Too Much
Magnitude of resistance	☐	☐	☐	☐	☐
Catch and release behavior	☐	☐	☐	☐	☐
Speed dependency	☐	☐	☐	☐	☐
Range of motion	☐	☐	☐	☐	☐

Comments:

Do you think this level of realism is sufficient for future medical or healthcare trainees to learn the basics of spasticity assessment across MAS levels 0-4? If no, what improvement(s) do you want to see the most? If yes, how can we make it more suitable as an educational training tool?

You have completed the questionnaire. Thank you very much for your participation!

Table G.1: Complete rater similarity score of simulated MAS levels using a five-point scale (1: too little, 3: about right, 5: too much) during Disclosed Assessment Test.

Aspect / Subject	MAS 0				MAS 1			
	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM
1	3	3	3	3	3	3	3	3
2	2	3	3	3	2	3	3	3
3	3	3	3	3	3	3	3	3
4	3	3	3	3	3	3	2	3
5	3	3	3	3	3	3	3	3
6	3	3	3	3	3	3	3	3
7	2	3	3	3	3	5	3	3
Mean (SD)	2.7 (0.5)	3.0 (0.0)	3.0 (0.0)	3.0 (0.0)	2.9 (0.4)	3.3 (0.7)	2.9 (0.4)	3.0 (0.0)
Level Mean (SD)	2.9 (0.3)				3.0 (0.5)			

Aspect / Subject	MAS 2				MAS 3			
	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM
1	3	3	3	2	3	3	3	2
2	3	2	2	2	3	2	2	2
3	3	3	3	1	3	3	3	1
4	3	3	3	3	3	4	3	3
5	3	2	3	3	3	2	3	3
6	3	4	3	4	2	3	3	3
7	3	3	3	3	3	3	3	3
Mean (SD)	3.0 (0.0)	2.9 (0.6)	2.9 (0.4)	2.6 (0.9)	2.9 (0.4)	2.9 (0.6)	2.9 (0.4)	2.4 (0.7)
Level Mean (SD)	2.8 (0.6)				2.8 (0.6)			

Aspect	MA 4			
	Resistance Magnitude	Catch-Release Behavior	Speed-Dependency	ROM
1	2	3	3	2
2	3	2	2	2
3	3	3	3	1
4	2	3	3	3
5	3	3	3	3
6	2	3	3	3
7	2	3	3	3
Mean (SD)	2.4 (0.5)	2.9 (0.4)	2.9 (0.4)	2.4 (0.7)
Level Mean (SD)	2.6 (0.6)			

Summary and highlights of Clinician 1's comments:

- For MAS 1, the simulator had a good simulation of the minimal resistance at the end of ROM.
- Usually it is difficulty to decipher between MAS 2 and 3. The simulator implemented less available ROM and slightly more resistance at MAS 3 compared to MAS 2, making the assessment easier.
- For MAS 4, the simulator should have a little higher resistance.
- The simulator is a good tool to help student in the classroom gain more experience before
- It will be beneficial to allow more ROM of the elbow joint to simulate more realistic patients.

Summary and highlights of Clinician 2's comments:

- Perfect simulations of MAS 0 and 1.
- For MAS 2 and 3, the catch could be more pronounced.
- After the first one or two stretch, the spastic muscle tone will become lower clinically.
- The simulator offers student to get in more repetitions of practice to feel what spasticity feels like. The fatiguing of an actual patient's arm due to multiple stretches will affect the level of spasticity.

Summary and highlights of Clinician 3's comments:

- For MAS 0, the simulator should have more resistance.
- For MAS 1, the catch-release behavior should occur at the beginning of ROM.
- The simulator is a great device except the catch-release behavior at MAS 1 is too late.

Summary and highlights of Clinician 4's comments:

- For MAS 2, the catch should happen earlier.
- For MAS 3 and 4, more resistance is expected.
- The simulator will be a valuable tool for future students. It is hard for other students to test on one another accurately and convey the resistance, catch-release behavior, and ROM of actual patients. More resistance should be added to MAS 3 and 4.

Summary and highlights of Clinician 5's comments:

- For MAS 0 and 1, slight more resistance should be added to mimic joint friction.
- For MAS 2, resistance felt is appropriate.
- For MAS 3, the catch is too late, which could be confused for MAS 2.
- For MAS 4, there is slight free motion (backlash) at the beginning of ROM and resistance is appropriate.
- For MAS 2-4, the simulator should have full ROM.
- The linkage issue should be fixed. The adjustment of mechanical selectors should be better covered. The resistance at the end of ROM feels very stiff.

Summary and highlights of Clinician 6's comments:

- MAS 1 is realistic.
- For MAS 2, 3, and 4, the simulator should allow full ROM.
- Some degree of reduction of ROM occurs in long-term spasticity patient due to contracture.
- The simulator would help teaching basics of spasticity assessment.

Summary and highlights of Clinician 7's comments:

- The ROM limitations are accurate. ROM tends to decrease and catch angle comes earlier with the increase of MAS scores.
- For MAS 4, the resistance should be higher. In clinical test, clinicians often unable to fully extend the biceps due to muscle tone.
- The simulator will be helpful in providing a more standardized assessment while using the MAS.

APPENDIX H: SUPPLEMENTARY FIGURES AND RESULTS

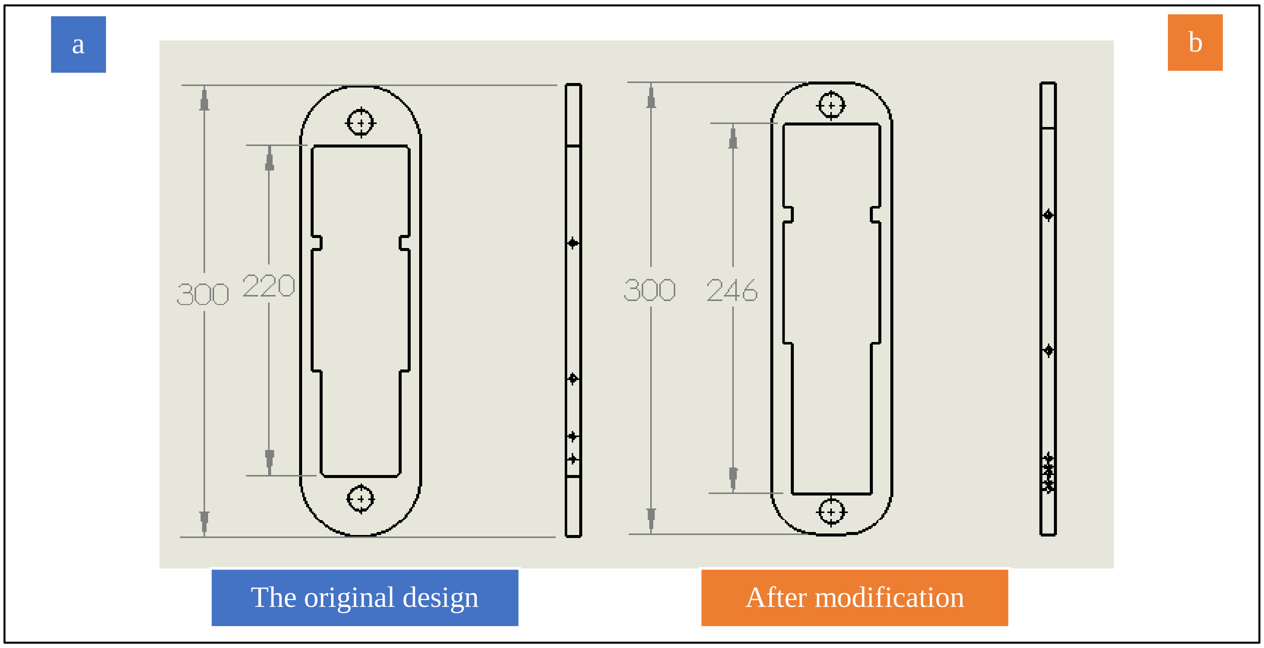
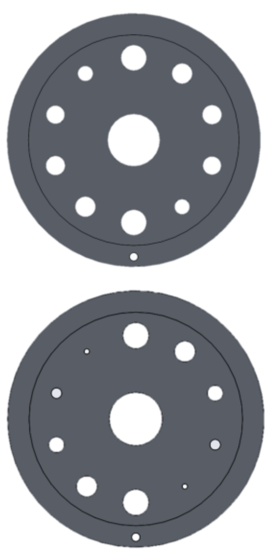


Fig. H.1: The design drawing of the upper arm aluminum frame. (a) The drawing of the first generation. (b) The drawing of the second generation. While the outer dimension remained the same, the inner clearance was increased in the new design.



The First Generation Design

MAS Score	Orifice Radius (mm)
0	2.5
1	2.0
2	1.9
3	1.8
4	1.5

The Second Generation Design

MAS Score	Orifice Radius (mm)
0	2.5
1	2.0
2	1.5
3	1.0
4	0.5

Fig. H.2: The comparison of the orifice design between the first and second generations.

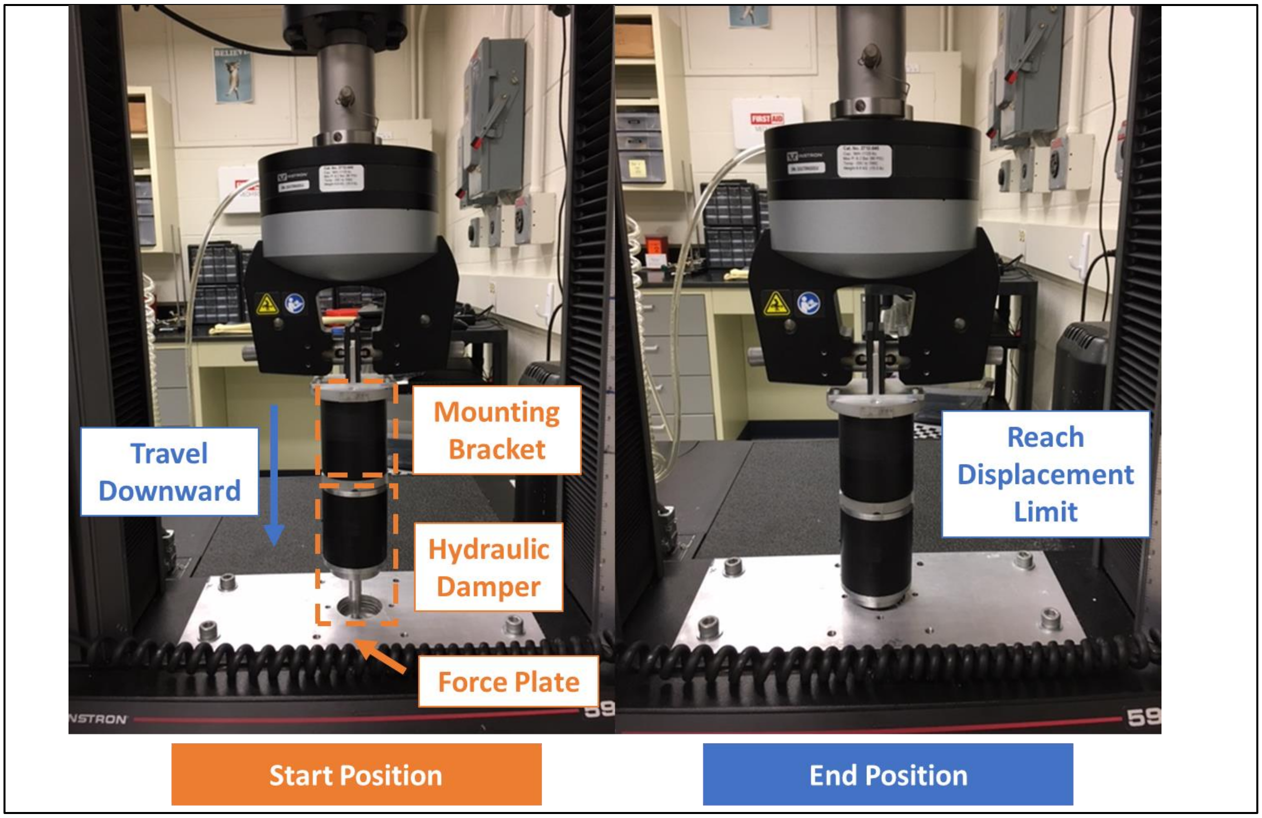


Fig. H.3: The experimental setup for the hydraulic damper evaluation.

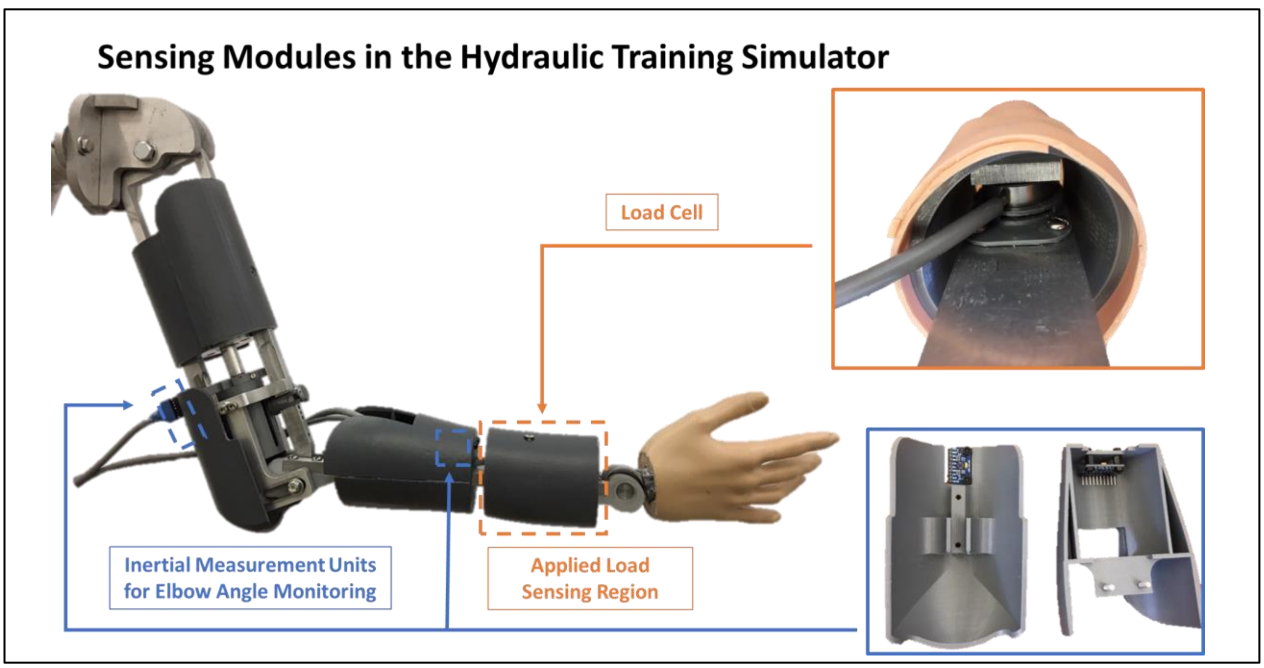


Fig. H.4: Embedded Sensing modules on the hydraulic simulator.

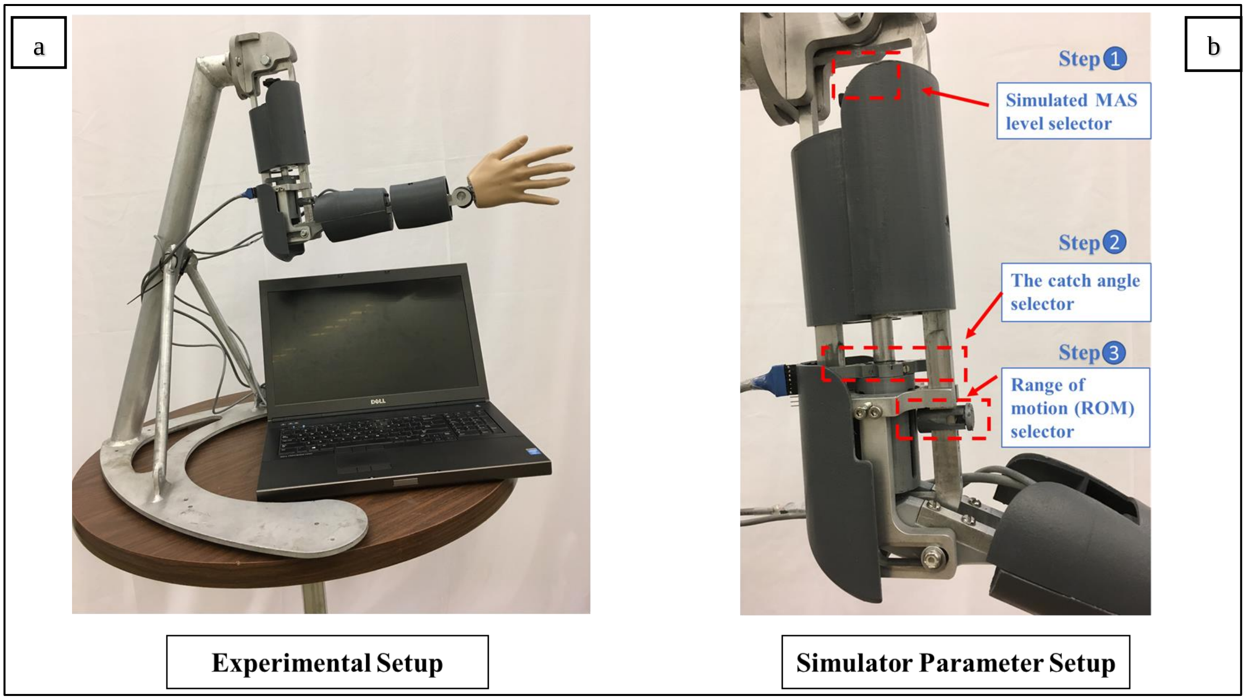


Fig. H.5: The experimental setup to evaluate fully-assembled simulator. (a) Data acquisition setup. (b) Simulator setup procedure prior to each testing condition.

Table H.1: Mean (standard deviation) experimental results of Resistance Level Test under different speeds and orifice sizes.

Orifice Radius (mm)	Stretch Speed	Peak Resistive Torque (Nm)	Average Speed (°/s)
2.5	Slow	1.00 (0.04)	23 (1)
	Medium	1.20 (0.04)	55 (4)
	Fast	1.70 (0.11)	164 (4)
2.0	Slow	1.08 (0.07)	19 (1)
	Medium	1.39 (0.06)	49 (5)
	Fast	2.03 (0.16)	131 (4)
1.5	Slow	1.23 (0.02)	16 (2)
	Medium	3.40 (0.27)	47 (3)
	Fast	5.03 (0.13)	95 (5)
1.0	Slow	2.35 (0.45)	14 (2)
	Medium	5.68 (0.21)	38 (5)
	Fast	8.90 (0.34)	89 (6)
0.5	Slow	4.25 (0.37)	21 (6)
	Medium	9.51 (0.19)	40 (2)
	Fast	12.18 (0.82)	84 (6)