Objective Assessment of Technical Proficiency in Colonoscopy Training

Matthew S. Holden1, Chang Nancy Wang2, Kyle MacNeil1, Ben Church1, Lawrence Hookey2, Gabor Fichtinger1\*, Tamas Ungi1

1Laboratory for Percutaneous Surgery, School of Computing, Queen’s University, Kingston, ON, Canada

2Gastrointestinal Diseases Research Unit, Department of Medicine, Queen’s University, Kingston, ON, Canada

\*Email: [fichting@queensu.ca](mailto:fichting@queensu.ca)

# Abstract

*Objective*: Manipulation of the colonoscope is a technical challenge for novice clinicians which is best learned in a simulated environment. It involves the coordination of scope tip steering with scope insertion, using a rotated image as reference. The purpose of this work is to develop and validate a system which objectively assesses colonoscopy technical skills proficiency in an arbitrary training environment, allowing novices to assess their technical proficiency prior to real patient encounters.

*Methods*: We implemented a motion tracking setup to objectively analyze and assess the way operators perform colonoscopies, including an analysis of wrist and elbow joint motions. Subsequently, we conducted a validation study to verify whether our motion analysis could discriminate novice colonoscopists from experts. Participants navigated a wooden bench-top model using a standard colonoscope while their motions were tracked.

*Results*: The developed motion tracking setup allowed colonoscopists of varying levels of proficiency to have their technical proficiency assessed, and was able to be operated by a trained non-technical operator. Novice operators had significantly greater median times (101.5s vs. 31.5s) and number of hand movements (62.0 vs. 21.5) than experts. Experts, however, spent significantly more time in extreme ranges of wrist and elbow joint motion than novices.

*Conclusion*: We have developed and implemented a hand motion analysis system for technical skills assessment in colonoscopy. The system is able to discriminate novices from experts based on objective proficiency metrics.

**Keywords:** colonoscopy, simulation-based training, medical education, objective skill assessment

# Introduction

Worldwide, medical education has been undergoing a shift from a time-based model to a competency-based model. The competency-based model allows trainees to progress through their curriculum only once they have achieved specific performance benchmarks. This contrasts the time-based model in which trainees were stuck in courses of fixed length, regardless of their progress or lack thereof. The competency-based model allows trainees who progress faster through a curriculum to advance more quickly and trainees who progress more slowly more time to practice and develop their skills and knowledge. Most importantly, this prevents trainees who have not yet reached competency from moving into a clinical setting with real patients.

For colonoscopy, a competency-based curriculum is particularly important. Studies have reported training correlates significantly with success (e.g. [1]) and reduces complications (e.g. [2]). Such a competency-based curriculum would ensure only well-trained colonoscopists from enter clinical practice, reducing complications and increasing success rate.

For interventional skills training, it is essential to have a simulation environment in which trainees may practice, whether it be a physical, virtual, or mixed environment. These environments are necessary because they allow trainees to learn the intervention’s workflow and practice basic technical skills prior to real patient encounters.

For colonoscopy, several commercial virtual simulators have seen widespread use in practical training scenarios, including the Simbionix GI Mentor (3D Systems Corporation, Littleton, CO), the EndoVR (CAE Healthcare, Sarasota, FL; formerly the Immersion AccuTouch), and the Endo TS-1 (Olympus Corporation, Tokyo, JP). Commercially available physical models include the Koken Colonoscopy Simulator (Koken Co., Tokyo, JP) and the Kyoto Kagaku Colonoscopy Training Model (Kyoto Kagaku Co., Kyoto, JP). Hill et al. provide an overview and evaluation of several common colonoscopy simulators [3], and Preisler et al. have shown good correspondence between performance in virtual and physical environments [4].

Competency-based medical education, however, requires regular monitoring of trainee performance. This way, each trainee’s individual learning curve can be tracked as they progress through the curriculum and the point at which they achieve competency can be quantitatively determined.

Many solutions for proficiency assessment in competency-based medical education are based on structured expert observation, such as global rating scales and procedure-specific checklists. For colonoscopy, Chak et al. [5] designed an assessment method which monitors trainees’ preparedness and technical skills. The Global Assessment of Gastrointestinal Endoscopic Skills (GAGES) has been proposed and validated for assessment of proficiency for both upper endoscopy and colonoscopy [6]. The Simulated Colonoscopy Objective Performance Evaluation (SCOPE) evaluates operators on technical skills such as scope manipulation, tool targeting, loop management, and mucosal inspection [7]. The Assessment of Competency in Endoscopy (ACE) toolbox is designed to assess both cognitive and technical skills [8]. The Gastrointestinal Endoscopy Competency Assessment Tool (GiECAT) was developed based using a Delphi study and validated for the assessment of endoscopy proficiency [9], [10]. These approaches requiring direct expert supervision, however, have several drawbacks. Most notably, constant supervision is time consuming for experts, but there remains inconsistency across experts, even using these structured methods [11].

In accordance with these drawbacks, there has been a recent shift in interest towards the development of methods for quantitatively assessing technical and procedural skill using time and position tracking measurements in numerous medical interventions [12]. Overviews describing the construct validity of many measures of objective proficiency assessment in colonoscopy on several virtual simulators have been provided by Ansell et al. [13], Triantafyllou et al. [14], and Ekkelenkamp et al. [15]. In particular, the majority of studies investigating construct validity for performance metrics in colonoscopy have been for virtual simulators. For the Simbionix GI Mentor, several works have demonstrated the construct validity of total time, efficiency, episodes of view loss, and episodes of patient pain, but note that the difficulty of the clinical scenario has effect on the discriminatory value of these metrics [16], [17]. MacDonald et al. [18] have shown validity for similar metrics including total time, percentage of time with patient pain, scope tip movement, and percentage of diseased region visualized on a predecessor of the CAE EndoVR. Likewise, Haycock et al. [19] demonstrated construct validity for total time, completion rate, use of variable stiffness, and sigmoid looping for the Olympus Endo TS-1. Plooy et al. [20], on the other hand, validated proficiency assessment measures including completion rates, total time and force applied to the colon on a physical model, the Kyoto Kagaku Colonoscope Training Model. Svendsen et al. [21] used hand motion capture on the same model to assess proficiency, and identified that the distance between hands could differentiate novices from experts. Indeed, Telem et al. [22] have demonstrated that performance metrics in simulated colonoscopy translates well to successful colonoscopy in clinical practice.

The objective of this work is to develop a setup for the automatic, objective assessment of colonoscopy technical skills that is not specific to a particular model or simulation environment. To this end, we developed several objective measures of both procedural and ergonomic efficiency, validated their ability to discriminate novices from experts in a particular simulated colonoscopy environment, and integrated the validated metrics into an open-source software for colonoscopy technical skills assessment.

A preliminary version of this work has been reported [23].

# Methods

## System Setup

We designed a motion tracking setup to record and assess hand, wrist, and elbow motions during the intervention. To this end, we attached electromagnetically tracked position and orientation sensors (3D Guidance trakSTAR, Northern Digital Inc., Waterloo, Canada) to operators’ hands, forearms, and biceps and one sensor to the colon model. The PLUS software library ([www.plustoolkit.org](http://www.plustoolkit.org)) [24] was used to acquire and send the tracking data to the 3D Slicer software ([www.slicer.org](http://www.slicer.org)).

We developed a colonoscopy technical skills assessment module based on the Perk Tutor platform ([www.perktutor.org](http://www.perktutor.org)) [25] for image-guided interventions training, within the 3D Slicer environment. We elected to use Perk Tutor, as it allowed us to implement custom performance metrics that are specific to technical skills assessment in colonoscopy. The colonoscopy analysis module performs calibrations, records tracking data and computes the objective metrics of operator performance and efficiency. The module was designed to be usable by non-technical operators.

## Objective Proficiency Metrics

To assess proficiency, we used several motion efficiency metrics already integrated into the Perk Tutor extension: total time of procedure, total path length of hands [25], and the number of discrete hand motions [26]. Additionally, we implemented several objective performance metrics profiling wrist and elbow motions.

To compute wrist and elbow motion, calibration is required. For each wrist, we must determine the axes of flexion/extension motion and abduction/adduction (Fig. 1). For each elbow, we must determine the axes of flexion/extension and supination/pronation (Fig. 1). To determine these axes, the operator rotates their wrist or elbow in each rotational motion. The axes are separately computed over each rotational motion by calculating the eigenvector with the smallest associated eigenvalue of the matrix defined in Equation 1, where is the instantaneous rotation matrix between th and th transform from the hand sensor to the forearm sensor over a total of recorded transforms.

|  |  |
| --- | --- |
|  | Equation |

The axes of rotation can subsequently be used to determine coordinate systems on the hand and on the forearm that are aligned with the wrist or elbow’s axes. The angles of wrist and elbow can thus be determined by the rotation between the aligned coordinate systems on the hand and forearm and the forearm and bicep, respectively (Fig. 2).

For each joint, we measured the number of times the joint entered extreme ranges of motion, as well as the total time spent in extreme ranges of motion. For wrists this was with respect to flexion/extension and adduction/abduction, and for elbow this was with respect to flexion/extension and supination/pronation. A joint was considered to be in an extreme range of motion when the joint’s angle exceeded 20% of the total range of motion for the operator in either direction, following the protocol outlined by Mohankumar et al. [27]. The total ranges of motion were computed from the prior calibration. These metrics are proxies for the total stress and strain exerted on each joint, and assess the ergonomic efficiency of the operator.

Furthermore, we assessed the number of discrete rotational motions for each joint and the cumulative angle through which each joint rotated (the rotational analog of total path length). These metrics are intended to assess the efficiency of joint motions. Finally, we assessed the operators’ coordination by determining the correlation between wrist and elbow rotation for each arm.

## Validation Study

Both the standard efficiency metrics and the joint motion metrics were validated for proficiency assessment in colonoscopy by conducting a study on simulated colonoscopies (Fig. 3). Twenty-two novice and eight expert colonoscopists were recruited to perform simulated colonoscopies. All participants were right-handed. The novice group consisted of medical students with no prior simulated or clinical colonoscopy experience. The expert group consisted of staff gastroenterologists who perform at least 200 colonoscopies per year and have at least five years of experience. Participants performed the simulated colonoscopies on a previously validated wooden bench-top model (Fig. 4). The model is intended for teaching essential colonoscope manipulation techniques to medical trainees with no previous colonoscopy experience, before they move on to more difficult and realistic models. This model has been extensively validated, demonstrating its efficacy as a training tool for low-level trainees [28]. Furthermore, its flexible design allows participants to navigate different training sequences to practice different maneuvers involved in colonoscopy.

First, the electromagnetic sensors were attached securely to the participant’s hands, forearms, and elbows. The participant was subsequently asked to perform the calibration exercises (two motions for each wrist and elbow, totaling eight exercises), and the software module automatically computed the calibrations. Next, the participant was assigned five practice navigation sequences to familiarize themselves with the colonoscope and the wooden bench-top model. Finally, each participant was assigned the same set of eight navigation sequences: four unique sequences each performed twice in random order. Participants were given a maximum of eight minutes to complete each sequence and were stopped if they had not completed the sequence within this time. The order of sequences was the same for all participants. The eight navigation sequences were tracked and analyzed using the colonoscopy software module.

## Statistical Analysis

The difference in the proportion of completed trials between the novice and expert group was tested using Fisher’s Exact test (α=0.05). Proficiency metric data were tested for normality using the Jarque-Bera test and found to be non-normally distributed. Differences between the novice and expert group were tested using the Mann-Whitney U test, with the Bonferroni correction for multiple tests (α=0.0017). The reported effect sizes are computed non-parametrically using Cliff’s ∆, using the interpretation scheme outlined by Romano et al. [29].

# Results

The custom colonoscopy software module was successfully able to record, calibrate, and analyze the tracking data from the simulated colonoscopy procedures. The module is available as an open-source module in the 3D Slicer environment ([www.slicer.org](http://www.slicer.org)), available for use without restriction on Linux, Mac OSX, and Windows. Using the module, the complete calibration can be performed in less than five minutes and the performance metrics are computed automatically. The software was successfully used by several different operators with varying technical backgrounds.

Experts successfully completed the trials within the allotted eight minutes significantly more often than novices (100% vs. 93%, p=0.02). Overall, experts exhibited better scores for all standard performance efficiency metrics, with significantly lower median time, hand path length, and number of discrete hand motions than novices (Fig. 5, Table 1), with large effect size for each performance efficiency metric. This demonstrates validity for the proposed performance efficiency metrics as discriminators of proficiency.

For each of right wrist flexion/extension, right wrist abduction/adduction, right elbow flexion/extension, and right elbow supination/pronation, novices entered extreme ranges of motion significantly more times than experts during a given simulated procedure (Table 1), medium to large effect size. This pattern, however, did not hold for the left side, where only left wrist abduction/adduction was significantly greater for novices. Furthermore, novices spent significantly greater amount of time in extreme ranges of motion than experts for the right side, and for the left side except for left elbow flexion/extension. These results indicate that experts exerted less total stress and strain on their right wrist and elbow joints.

When time was factored out, however, experts appeared significantly less ergonomically efficient for the left side. Experts entered extreme ranges of motion significantly more frequently than novice for left wrist flexion/extension, left wrist adduction/abduction, left elbow flexion/extension, and left elbow supination/pronation. Additionally, experts spent a greater proportion of time in extreme ranges of motion for left wrist flexion/extension and left elbow flexion/extension. For the right side, experts entered extreme ranges of motion at significantly more frequent for only right wrist flexion/extension and effect size was smaller (Δ=0.267).

Novices had more total rotation and rotational motions in both the right wrist and elbow than experts, but there was no difference in the left wrist or elbow (Table 1). Novices exhibited higher joint coordination in the left arm, but joint coordination was not significantly different in the right arm.

# Discussion

Results indicate that experts generally exerted less stress and strain on their joints. When time was factored out, however, experts entered extreme ranges of motion significantly more frequently and spent a greater proportion of time in extreme ranges of motion for the left side. This is likely due to two factors. First, we hypothesize that novices spend more time contemplating how to manipulate the colonoscope correctly, and thus, accumulate time in non-extreme ranges of motion. This is supported by the observation that novices enter extreme ranges of motion a greater number of times during the course of an insertion. We also conjecture that experts use their left side to manage the colonoscope and facilitate insertion, in addition to using the dials.

We used the extreme ranges of motion convention established by Mohankumar et al. [27]. Under this convention, if the joint angle exceeded 20% of the full range of joint motion, it was considered extreme. This threshold, however, was not determined anatomically. For our study, this definition of extreme range of motion is perhaps too strict, as each joint is in an extreme range of motion the majority of the time (Fig. 7). Perhaps a less strict (e.g. 40% of full range of joint motion) or fuzzy definition of this threshold would show an even greater difference between novices and experts.

One difficulty with validating proficiency metrics is determining a gold-standard against which to validate them. Ideally, the gold-standard would be determined by a panel of blinded experts using a previously validated global rating scale or checklist to assess each individual trial. In this study, however, we used experience as a reference for validating the proficiency metrics. While this provides a general depiction of proficiency, it does not provide a full picture. In particular, it cannot account for anomalous performances (i.e. experts performing poorly or novices performing well by happenstance) and “bad habits” developed by experts. We conjecture, however, that the groups were large enough that an individual participant or trial affected by these issues had little impact on our analysis. Any “bad habit” that is pervasive amongst experts must not be of practical importance.

In both groups, participants completed the navigation task fully for the vast majority of cases. While success rate is an important consideration, time and efficiency metrics provide valuable information. In particular, time and efficiency metrics indicate colonoscope manipulation skill and understanding. This is important because greater colonoscope manipulation skill translates to reduced operating times and patient discomfort. Furthermore, when the trainee reaches the “unconscious competence” stage of learning colonoscope manipulation, it allows them to concentrate on other aspects of the procedure such as image interpretation, diagnosis, and patient management.

The proposed system shows promise for use as one component in an experimental colonoscopy training curriculum. This would allow trainees’ improvement to be monitored as they progress through the curriculum, relative to performance benchmarks. Providing this feedback in an objective and automated manner is an important component of the newly evolving competency-based medical education paradigm. This allows trainees to practice their technical skills without requiring continuous expert supervision. As a result, we designed our colonoscopy software module to be easy to use and user-friendly. Several different operators with varying backgrounds were able to use the system.

We did not test, however, if the colonoscopists who participated in this study could use the system without a trained operator. This was done to ensure integrity of the recorded data for validation purposes. In practice, there could be an audience effect causing participants to perform differently with the presence of a technical operator than they would alone [30]. The audience effect in medical education is not well studied, and further study is required to understand how the results of this study translate to a completely self-guided training curriculum.

All software components are free, open-source, and cross-platform. Because the setup uses the PLUS toolkit, it can support a variety of different position trackers. Furthermore, the system is not specific to any particular colonoscopy simulator. It can be used for technical skills assessment equally on any physical simulator or any virtual simulator, as it does not depend on any environment-specific parameters.

The wrist and elbow ranges of motion metrics analyze operators’ ergonomic efficiency. Thus, this training system could be incorporated into a training program which teaches the ergonomic use of the colonoscope, which would contribute to the prevention of workplace injuries.

The ranges of motion analysis is not specific to colonoscopy. Although colonoscopy is one of the highest-risk interventions for injury due to repetitive motion strain, the ergonomic efficiency may be a useful metric in other applications, such as ultrasound-guided needle interventions or laparoscopy. Because the colonoscopy analysis module is built on an open-source, hardware-agnostic platform and the analysis uses no colonoscopy-specific computation, the exact same setup and analysis can be used to compute ergonomic efficiency in other minimally invasive interventions.

This study is limited by the fact that participants performed the procedure on a wooden bench-top model. During the simulated colonoscopies, participants needed not interpret images, identify anatomy, nor manage the patient. It is unclear how performance on the model transfers to proficiency in an operating environment, however, there is evidence in the literature that the technical skills component is indeed transferable [22].

In future work, we suggest that the tip of the colonoscope could be tracked in addition to the hands, forearms, and biceps. This would allow the system to analyze the method in which each operator navigates the colonoscope, which is not possible with sensors exclusively attached to the operators’ hands and arms. Efficiency and accuracy of the path of the colonoscope could have added value as a discriminator of proficiency. These proficiency metrics and others could be readily added to the system as custom performance metrics within Perk Tutor platform. Furthermore, tracking information could be used to provide the operator with various forms of guidance, which may improve the training process for novice operators.

# Conclusion

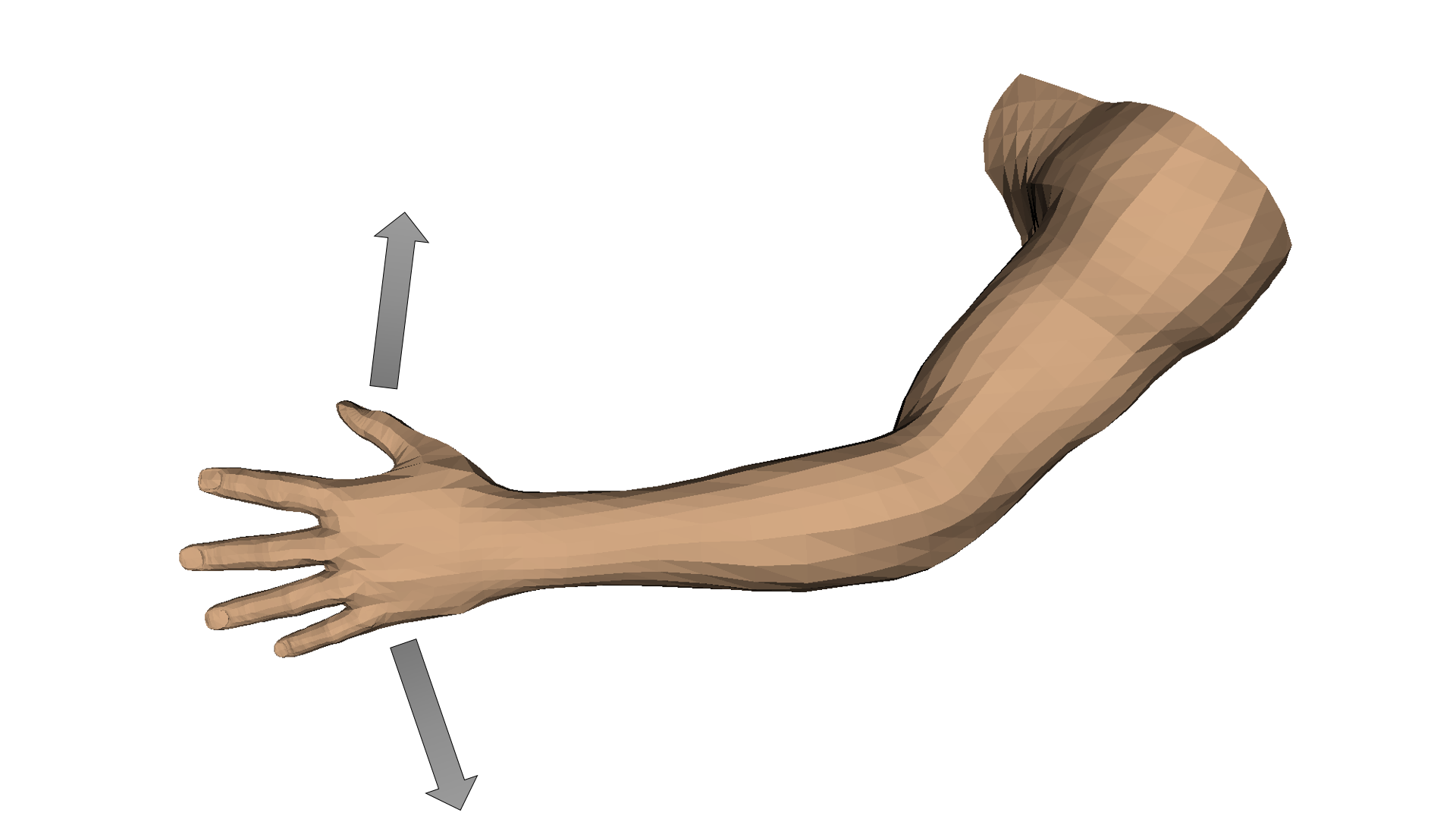
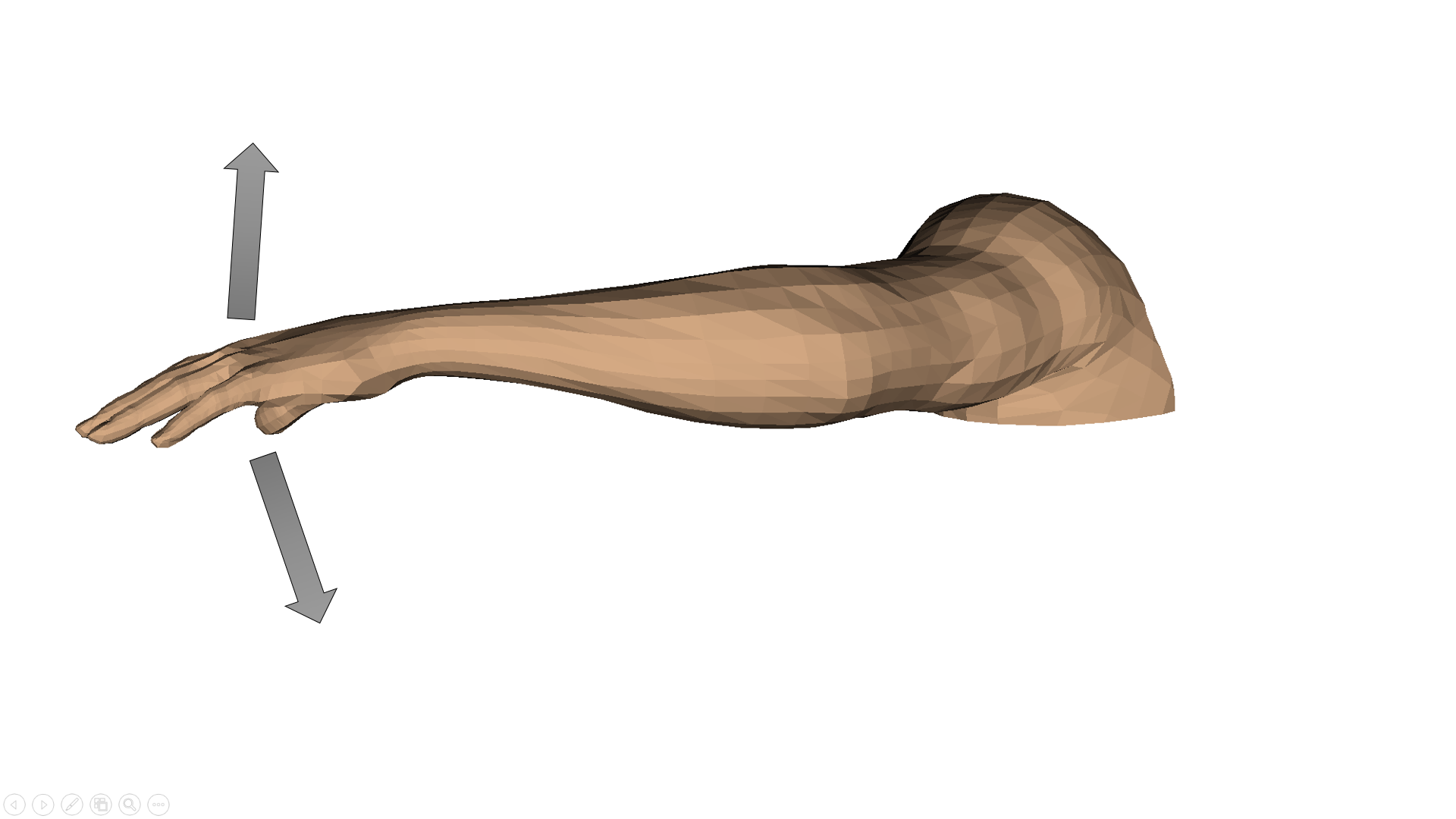
We have implemented a complete hardware and software setup for technical skills assessment in colonoscopy. The system is based on commercially available hardware and open-source, hardware-agnostic software, making it reproducible at other centers under an arbitrary simulation environment. We have shown validity for many of our objective proficiency metrics, especially those assessing overall efficiency and joint rotation efficiency by demonstrating that they can successfully discriminate novices from experts in a simulated colonoscopy study. In the future, we envision that a system using similar automatically computed proficiency metrics for performance assessment could be integrated into a competency-based colonoscopy training curriculum where trainee progress is monitored.

# Acknowledgements

Matthew S. Holden is supported by the Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada. Gabor Fichtinger is supported as a Cancer Care Ontario Research Chair in Cancer Imaging. Financial support was received from the SEAMO Educational Innovation and Research Fund. This work was financially supported as a Collaborative Health Research Project (CHRP #127797), a joint initiative between the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institutes of Health Research (CIHR).

# References

|  |  |
| --- | --- |
| [1] | J. B. Marshall, "Technical proficiency of trainees performing colonoscopy: A learning curve," *Gastrointestinal Endoscopy ,* vol. 42, pp. 287-291, 1995. |
| [2] | M. L. Anderson, T. M. Pasha and J. A. Leighton, "Endoscopic perforation of the colon: lessons from a 10-year study," *Am. J. Gastroenterol.,* vol. 95, pp. 3418-3422, Dec 2000. |
| [3] | A. Hill, M. S. Horswill, A. M. Plooy, M. O. Watson, R. Karamatic, T. A. Basit, G. M. Wallis, S. Riek, R. Burgess-Limerick and D. G. Hewett, "Assessing the realism of colonoscopy simulation: the development of an instrument and systematic comparison of 4 simulators," *Gastrointest. Endosc.,* vol. 75, pp. 631-640, Mar 2012. |
| [4] | L. Preisler, M. B. Svendsen, N. Nerup, L. B. Svendsen and L. Konge, "Simulation-based training for colonoscopy: establishing criteria for competency," *Medicine (Baltimore),* vol. 94, p. e440, Jan 2015. |
| [5] | A. Chak, G. S. Cooper, E. W. Blades, M. Canto and M. V. Sivak, "Prospective assessment of colonoscopic intubation skills in trainees," *Gastrointest. Endosc.,* vol. 44, pp. 54-57, Jul 1996. |
| [6] | M. C. Vassiliou, P. A. Kaneva, B. K. Poulose, B. J. Dunkin, J. M. Marks, R. Sadik, G. Sroka, M. Anvari, K. Thaler, G. L. Adrales, J. W. Hazey, J. R. Lightdale, V. Velanovich, L. L. Swanstrom, J. D. Mellinger and G. M. Fried, "Global Assessment of Gastrointestinal Endoscopic Skills (GAGES): a valid measurement tool for technical skills in flexible endoscopy," *Surg Endosc,* vol. 24, pp. 1834-1841, Aug 2010. |
| [7] | E. M. Ritter, T. C. Cox, K. D. Trinca and J. P. Pearl, "Simulated Colonoscopy Objective Performance Evaluation (SCOPE): a non-computer-based tool for assessment of endoscopic skills," *Surg Endosc,* vol. 27, pp. 4073-4080, Nov 2013. |
| [8] | R. E. Sedlack, W. J. Coyle, K. L. Obstein, M. A. Al-Haddad, G. Bakis, J. A. Christie, R. E. Davila, B. DeGregorio, C. J. DiMaio, B. K. Enestvedt, J. Jorgensen, D. K. Mullady and L. Rajan, "ASGEś assessment of competency in endoscopy evaluation tools for colonoscopy and EGD," *Gastrointest. Endosc.,* vol. 79, pp. 1-7, Jan 2014. |
| [9] | C. M. Walsh, S. C. Ling, N. Khanna, M. A. Cooper, S. C. Grover, G. May, T. D. Walters, L. Rabeneck, R. Reznick and H. Carnahan, "Gastrointestinal Endoscopy Competency Assessment Tool: development of a procedure-specific assessment tool for colonoscopy," *Gastrointest. Endosc.,* vol. 79, pp. 798-807, May 2014. |
| [10] | C. M. Walsh, S. C. Ling, N. Khanna, S. C. Grover, J. J. Yu, M. A. Cooper, E. Yong, G. C. Nguyen, G. May, T. D. Walters, R. Reznick, L. Rabeneck and H. Carnahan, "Gastrointestinal Endoscopy Competency Assessment Tool: reliability and validity evidence," *Gastrointest. Endosc.,* vol. 81, pp. 1417-1424, 2015. |
| [11] | R. Aggarwal, T. P. Grantcharov and A. Darzi, "Framework for systematic training and assessment of technical skills," *J. Am. Coll. Surg.,* vol. 204, pp. 697-705, Apr 2007. |
| [12] | C. E. Reiley, H. C. Lin, D. D. Yuh and G. D. Hager, "Review of methods for objective surgical skill evaluation," *Surgical Endoscopy,* vol. 25, pp. 356-366, 2011. |
| [13] | J. Ansell, J. Mason, N. Warren, P. Donnelly, N. Hawkes, S. Dolwani and J. Torkington, "Systematic review of validity testing in colonoscopy simulation," *Surg Endosc,* vol. 26, pp. 3040-3052, Nov 2012. |
| [14] | K. Triantafyllou, L. D. Lazaridis and G. D. Dimitriadis, "Virtual reality simulators for gastrointestinal endoscopy training," *World J Gastrointest Endosc,* vol. 6, pp. 6-12, Jan 2014. |
| [15] | V. E. Ekkelenkamp, A. D. Koch, A. de Man and R. and E. J. Kuipers, "Training and competence assessment in GI endoscopy: a systematic review," *Gut,* vol. 65, pp. 607-615, Apr 2016. |
| [16] | A. D. Koch, S. N. Buzink, J. Heemskerk, S. M. Botden, R. Veenendaal, J. J. Jakimowicz and E. J. Schoon, "Expert and construct validity of the Simbionix GI Mentor II endoscopy simulator for colonoscopy," *Surg Endosc,* vol. 22, pp. 158-162, Jan 2008. |
| [17] | R. Fayez, L. S. Feldman, P. Kaneva and G. M. Fried, "Testing the construct validity of the Simbionix GI Mentor II virtual reality colonoscopy simulator metrics: module matters," *Surgical Endoscopy,* vol. 24, pp. 1060-1065, 2010. |
| [18] | J. MacDonald, J. Ketchum, R. G. Williams and L. Q. Rogers, "A lay person versus a trained endoscopist: can the preop endoscopy simulator detect a difference?," *Surg Endosc,* vol. 17, pp. 896-898, Jun 2003. |
| [19] | A. V. Haycock, P. Bassett, J. Bladen and S. Thomas-Gibson, "Validation of the second-generation Olympus colonoscopy simulator for skills assessment," *Endoscopy,* vol. 41, pp. 952-958, Nov 2009. |
| [20] | A. M. Plooy, A. Hill, M. S. Horswill, A. S. Cresp, M. O. Watson, S. Y. Ooi, S. Riek, G. M. Wallis, R. Burgess-Limerick and D. G. Hewett, "Construct validation of a physical model colonoscopy simulator," *Gastrointest. Endosc.,* vol. 76, pp. 144-150, Jul 2012. |
| [21] | M. B. Svendsen, L. Preisler, J. G. Hillingsoe, L. B. Svendsen and L. Konge, "Using motion capture to assess colonoscopy experience level," *World J Gastrointest Endosc,* vol. 6, pp. 193-199, May 2014. |
| [22] | D. A. Telem, D. W. Rattner and D. W. Gee, "Endoscopic simulator curriculum improves colonoscopy performance in novice surgical interns as demonstrated in a swine model," *Surgical Endoscopy,* vol. 28, pp. 1494-1499, 2014. |
| [23] | K. MacNeil, C. N. Wang, M. S. Holden, T. Ungi, G. Fichtinger and L. Hookey, "System for objectively evaluating colonoscopy procedural skills using motion analysis," in *14th Annual Imaging Network Ontario Symposium*, 2016. |
| [24] | A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi and G. Fichtinger, "PLUS: Open-Source Toolkit for Ultrasound-Guided Intervention Systems," *IEEE Transactions on Biomedical Engineering,* vol. 61, pp. 2527-2537, Oct 2014. |
| [25] | T. Ungi, D. Sargent, E. Moult, A. Lasso, C. Pinter, R. C. McGraw and G. Fichtinger, "Perk Tutor: an open-source training platform for ultrasound-guided needle insertions," *IEEE Trans Biomed Eng,* vol. 59, pp. 3475-81, Dec 2012. |
| [26] | R. McGraw, T. Chaplin, C. McKaigney, L. Rang, M. Jaeger, D. Redfearn, C. Davison, T. Ungi, M. Holden, C. Yeo, Z. Keri and G. Fichtinger, "Development and Evaluation of a Simulation-based Curriculum for Ultrasound-guided Central Venous Catheterization," *CJEM,* pp. 1-9, May 2016. |
| [27] | D. Mohankumar, H. Garner, K. Ruff, F. C. Ramirez, D. Fleischer, Q. Wu and M. Santello, "Characterization of right wrist posture during simulated colonoscopy: an application of kinematic analysis to the study of endoscopic maneuvers," *Gastrointestinal Endoscopy ,* vol. 79, pp. 480-489, 2014. |
| [28] | C. M. Walsh, S. C. Ling, C. S. Wang and H. Carnahan, "Concurrent versus terminal feedback: it may be better to wait," *Acad Med,* vol. 84, pp. S54--57, Oct 2009. |
| [29] | J. Romano, J. D. Kromrey, J. Coraggio and J. Skowronek, "Appropriate statistics for ordinal level data: Should we really be using t-test and Cohen’s d for evaluating group differences on the NSSE and other surveys," in *annual meeting of the Florida Association of Institutional Research*, 2006. |
| [30] | C. F. Bond and L. J. Titus, "Social facilitation: a meta-analysis of 241 studies," *Psychol Bull,* vol. 94, pp. 265-292, Sep 1983. |



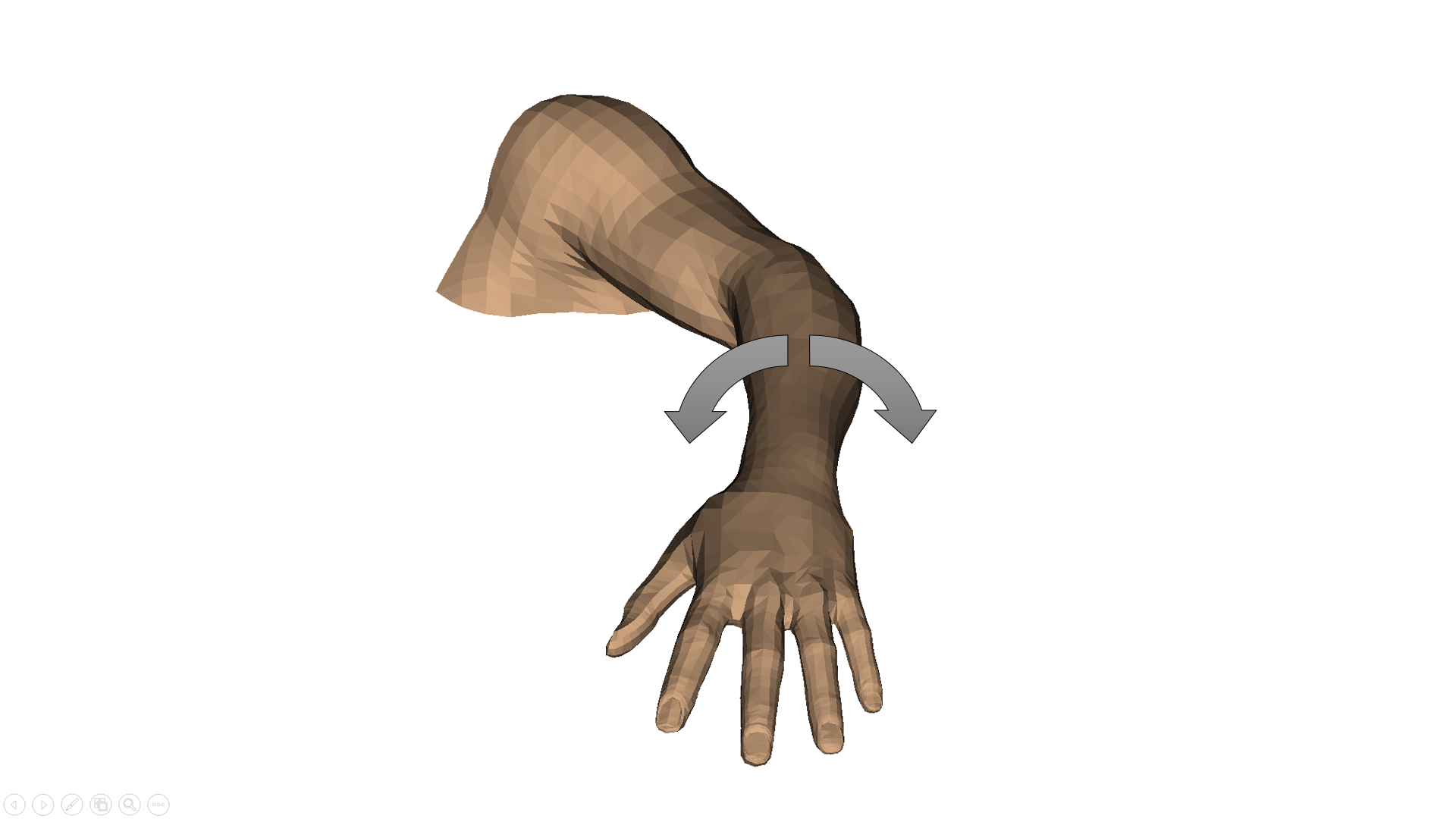
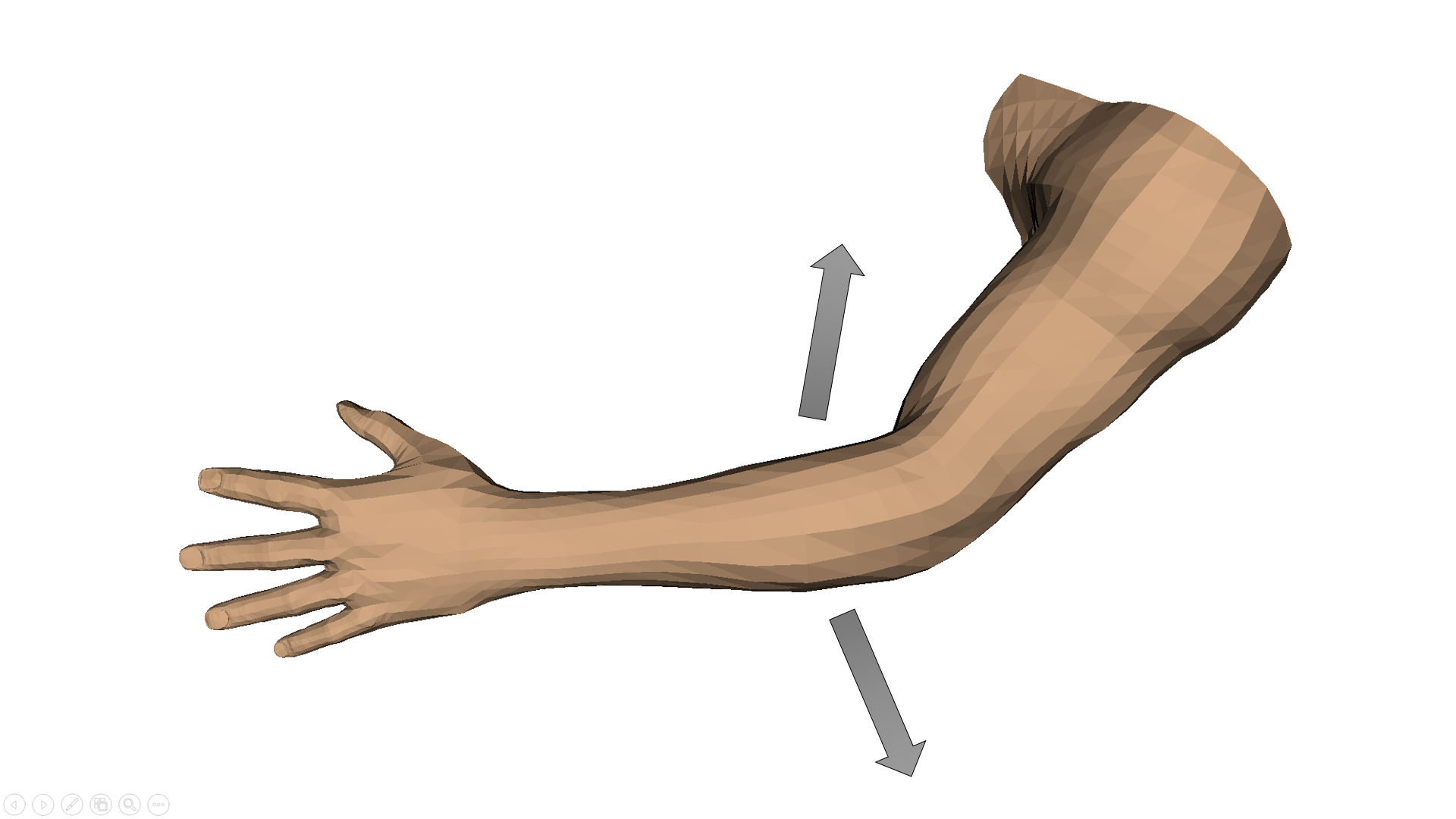


Fig. . Wrist flexion/extension (top-left), wrist abduction/adduction (top-right), elbow flexion/extension (bottom-left), and elbow supination/pronation (bottom-right) for the right side. Motions for the left side are analogous.

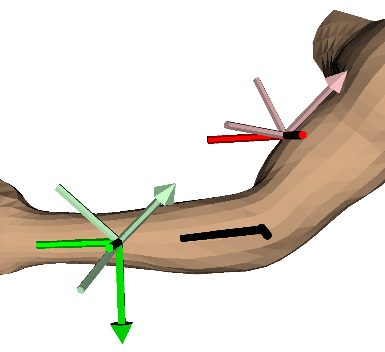
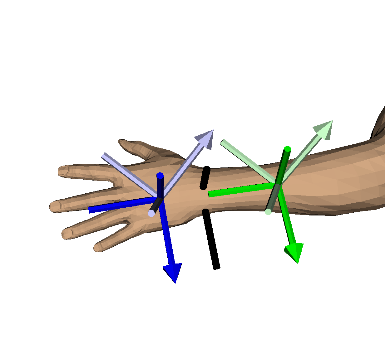


Fig. . Calibration for the wrist (left) and the elbow (right). The pale-colored axes represent the uncalibrated coordinate frames given by the tracking system and the bright-colored axes represent the calibrated coordinate frames. The black vectors represent the axes of rotation for wrist flexion/extension and abduction/adduction and elbow flexion/extension and supination/pronation.

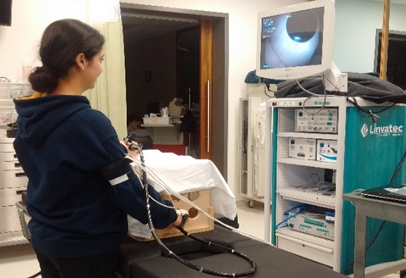
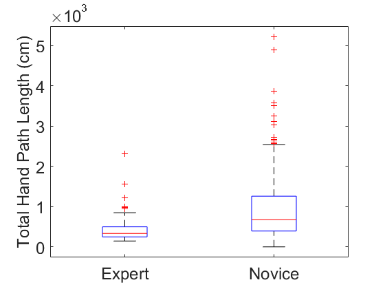
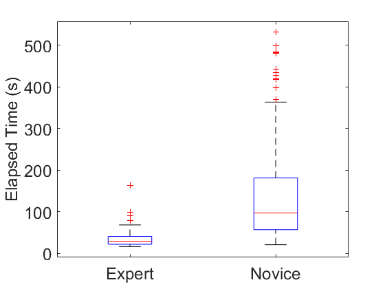


Fig. . A novice participant performing a simulated colonoscopy on the wooden bench-top model. Sensors are strapped to the hand, forearm, and bicep.



Fig. . Close-up view of the wooden bench-top model. The black rings indicate the holes through which the operator must navigate the colonoscope.



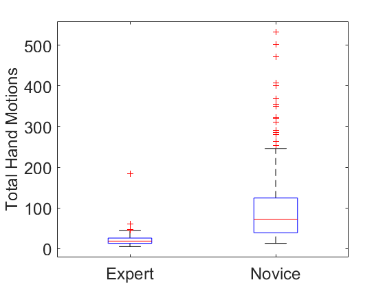
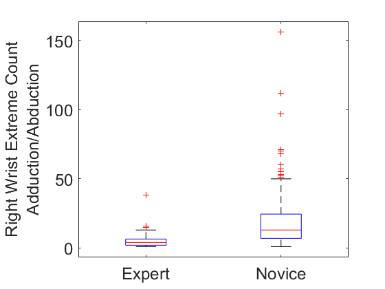
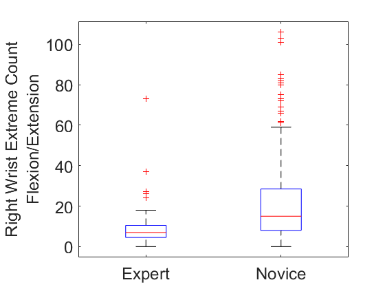


Fig. . Expert and novice total time (top-left), hand path length (top-right), and number of hand motions (bottom) for the recorded simulated colonoscopy sequences performed on the wooden bench-top model.



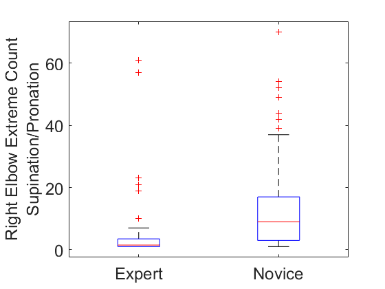
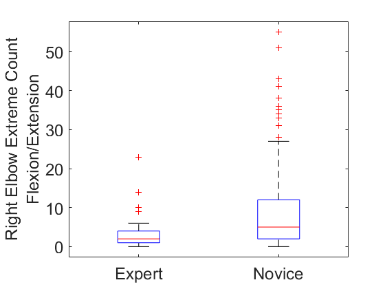
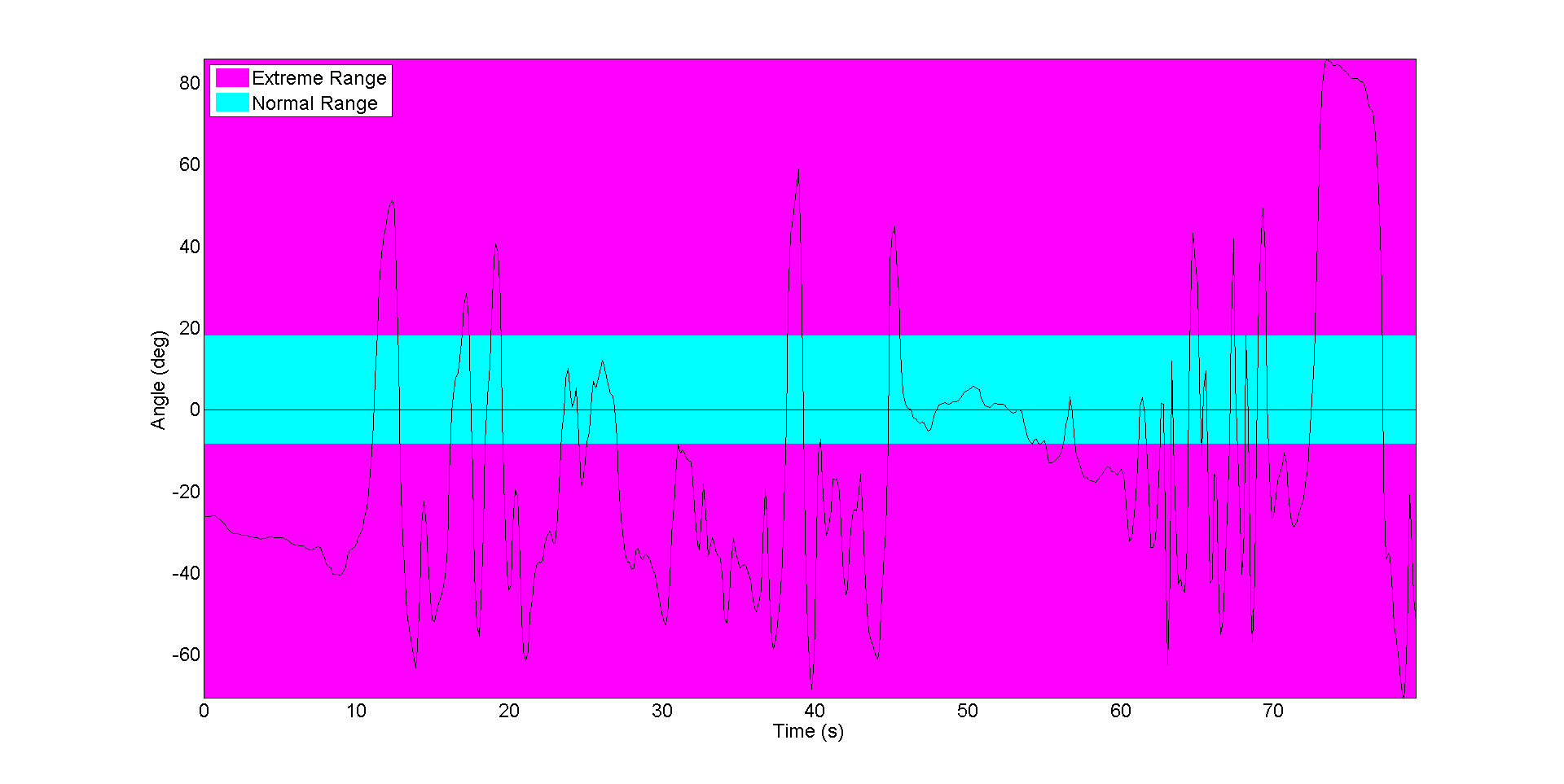


Fig. . Number of times an extreme range of motion is entered for novices and experts for right wrist flexion/extension (top-left), right wrist adduction/abduction (top-right), right elbow flexion/extension (bottom-left) and right elbow supination/pronation (bottom-right).



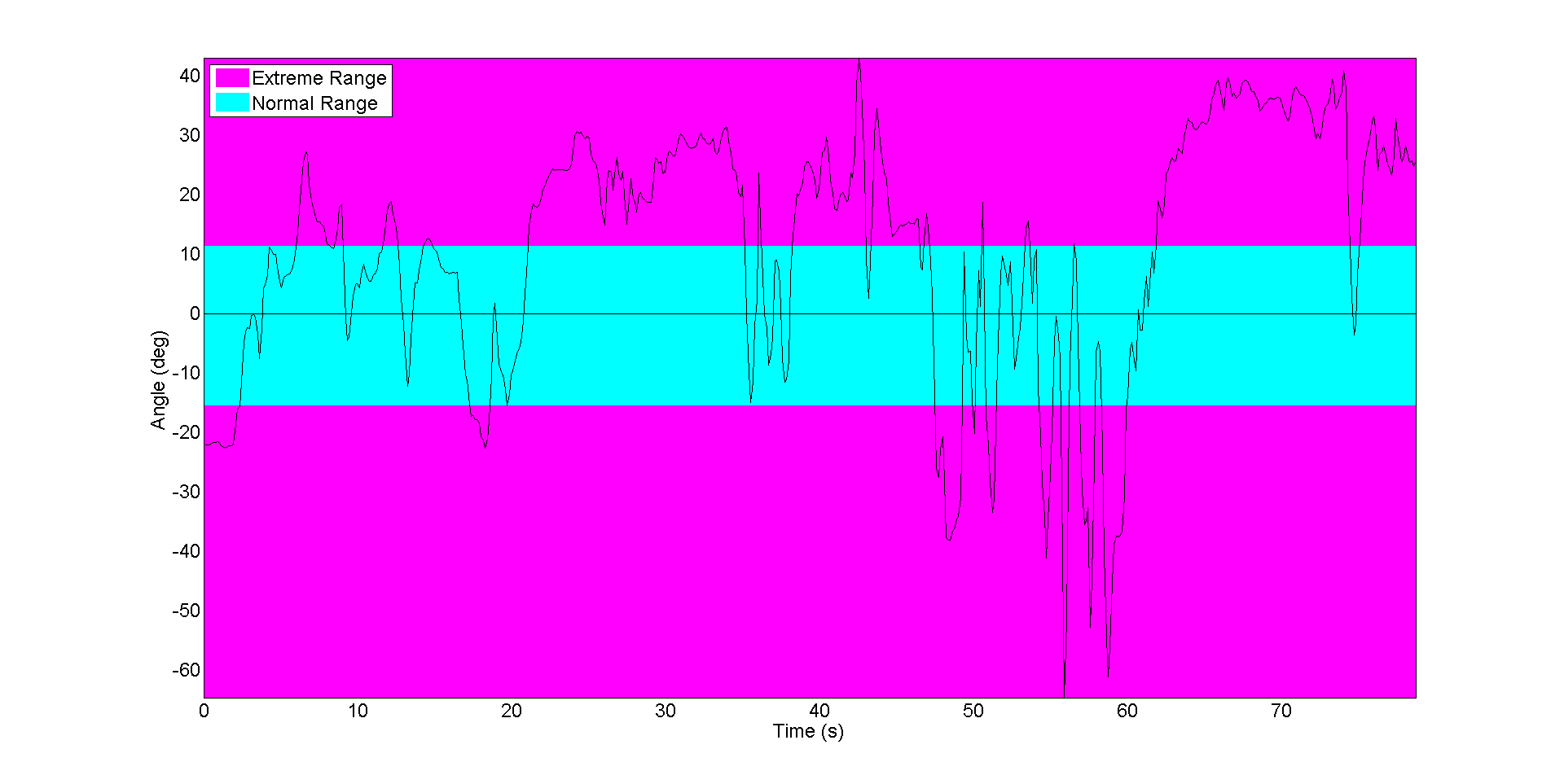


Fig. . Example right elbow supination/pronation angle time series for an expert (top) and a novice (bottom) performing a simulated colonoscopy on the same sequence. Cyan indicates a normal range of motion; pink indicates an extreme range of motion.

Table . Objective proficiency measures for novice and experts groups. Reported values are medians (inter-quartile range). Asterisk indicates a significant difference between the two groups using the Bonferroni corrected alpha value.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Metric | Novice | Expert | p | Δ |
| Elapsed Time (s) | 96.9  (56.9 - 181.0) | 29.0  (21.9 - 40.9) | <0.001\* | 0.792 |
| Total Hand Path Length (cm) | 674.5  (394.1 - 1261.6) | 336.1  (249.7 - 499.0) | <0.001\* | 0.486 |
| Total Hand Motions | 72.5  (39.0 - 124.5) | 18.5  (13.0 - 26.0) | <0.001\* | 0.816 |
| Right Wrist Extreme Count Flexion/Extension | 15.0  (8.0 - 28.5) | 7.0  (4.5 - 10.5) | <0.001\* | 0.517 |
| Right Wrist Extreme Count Adduction/Abduction | 13.0  (7.0 - 24.5) | 4.0  (2.0 - 6.5) | <0.001\* | 0.655 |
| Right Elbow Extreme Count Supination/Pronation | 9.0  (3.0 - 17.0) | 1.5  (1.0 - 3.5) | <0.001\* | 0.487 |
| Right Elbow Extreme Count Flexion/Extension | 5.0  (2.0 - 12.0) | 2.0  (1.0 - 4.0) | <0.001\* | 0.386 |
| Left Wrist Extreme Count Flexion/Extension | 11.0  (5.0 - 22.0) | 11.0  (5.0 - 18.5) | 0.414 | 0.069 |
| Left Wrist Extreme Count Adduction/Abduction | 9.0  (5.0 - 18.5) | 5.0  (2.0 - 8.5) | <0.001\* | 0.382 |
| Left Elbow Extreme Count Flexion/Extension | 1.0  (0.0 - 4.0) | 2.0  (1.0 - 3.0) | 0.041 | 0.169 |
| Left Elbow Extreme Count Supination/Pronation | 1.0  (1.0 - 1.0) | 1.0  (1.0 - 1.0) | 0.224 | 0.05 |
| Right Wrist Extreme Time Flexion/Extension (s) | 35.9  (18.0 – 81.1) | 11.9  (6.7 – 22.0) | <0.001\* | 0.561 |
| Right Wrist Extreme Time Adduction/Abduction (s) | 54.5  (25.6 – 100.2) | 16.2  (8.8 – 26.1) | <0.001\* | 0.676 |
| Right Elbow Extreme Time Supination/Pronation (s) | 82.8  (48.2 – 161.4) | 27.3  (19.9 – 38.1) | <0.001\* | 0.723 |
| Right Elbow Extreme Time Flexion/Extension (s) | 56.1  (25.0 – 115.0) | 12.9  (2.4 – 25.2) | <0.001\* | 0.675 |
| Left Wrist Extreme Time Flexion/Extension (s) | 47.5  (24.5 – 87.5) | 20.3  (15.2 – 30.0) | <0.001\* | 0.488 |
| Left Wrist Extreme Time Adduction/Abduction (s) | 49.8  (25.8 – 97.7) | 13.8  (7.4 – 24.3) | <0.001\* | 0.662 |
| Left Elbow Extreme Time Flexion/Extension (s) | 10.0  (0.0 – 53.2) | 12.7  (6.6 – 19.9) | 0.946 | 0.006 |
| Left Elbow Extreme Time Supination/Pronation (s) | 88.6  (52.7 – 163.8) | 26.7  (20.7 – 39.7) | <0.001\* | 0.776 |
| Right Arm Coordination | 0.907  (0.722 - 0.950) | 0.898  (0.558 - 0.965) | 0.862 | 0.015 |
| Left Arm Coordination | 0.963  (0.872 - 0.985) | 0.807  (0.624 - 0.950) | <0.001\* | 0.416 |
| Right Wrist Total Rotation (rad) | 32.0  (19.6 - 68.0) | 10.7  (7.9 - 16.6) | <0.001\* | 0.708 |
| Right Elbow Total Rotation (rad) | 31.0  (19.4 - 62.8) | 9.7  (6.5 - 17.2) | <0.001\* | 0.667 |
| Left Wrist Total Rotation (rad) | 22.3  (12.0 - 48.4) | 17.2  (10.8 - 26.2) | 0.015 | 0.206 |
| Left Elbow Total Rotation (rad) | 12.2  (6.2 - 25.1) | 8.5  (5.5 – 12.4) | 0.003 | 0.248 |
| Right Wrist Rotational Motions | 6.0  (1.5 - 14.0) | 2.0  (0.5 - 4.0) | <0.001\* | 0.388 |
| Right Elbow Rotational Motions | 6.0  (2.0 - 12.5) | 1.0  (0.0 - 4.0) | <0.001\* | 0.426 |
| Left Wrist Rotational Motions | 4.0  (2.0 - 11.0) | 5.0  (2.5 - 10.0) | 0.869 | 0.014 |
| Left Elbow Rotational Motions | 0.0  (0.0 - 1.0) | 0.0  (0.0 - 2.0) | 0.608 | 0.037 |