
Electronic Thesis and Dissertation Repository

11-21-2016 12:00 AM

Performance Factors in Neurosurgical Simulation and Augmented Reality Image Guidance

Ryan Armstrong
The University of Western Ontario

Supervisor
Roy Eagleson
The University of Western Ontario

Graduate Program in Biomedical Engineering
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy
© Ryan Armstrong 2016

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Other Biomedical Engineering and Bioengineering Commons](#)

Recommended Citation

Armstrong, Ryan, "Performance Factors in Neurosurgical Simulation and Augmented Reality Image Guidance" (2016). *Electronic Thesis and Dissertation Repository*. 4288.
<https://ir.lib.uwo.ca/etd/4288>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Virtual reality surgical simulators have seen widespread adoption in an effort to provide safe, cost-effective and realistic practice of surgical skills. However, the majority of these simulators focus on training low-level technical skills, providing only prototypical surgical cases. For many complex procedures, this approach is deficient in representing anatomical variations that present clinically, failing to challenge users' higher-level cognitive skills important for navigation and targeting. Surgical simulators offer the means to not only simulate any case conceivable, but to test novel approaches and examine factors that influence performance. Unfortunately, there is a void in the literature surrounding these questions. This thesis was motivated by the need to expand the role of surgical simulators to provide users with clinically relevant scenarios and evaluate human performance in relation to image-guidance technologies, patient-specific anatomy, and cognitive abilities. To this end, various tools and methodologies were developed to examine cognitive abilities and knowledge, simulate procedures, and guide complex interventions all within a neurosurgical context. The first chapter provides an introduction to the material. The second chapter describes the development and evaluation of a virtual anatomical training and examination tool. The results suggest that learning occurs and that spatial reasoning ability is an important performance predictor, but subordinate to anatomical knowledge. The third chapter outlines development of automation tools to enable efficient simulation studies and data management. In the fourth chapter, subjects perform abstract targeting tasks on ellipsoid targets with and without augmented reality guidance. While the guidance tool improved accuracy, performance with the tool was strongly tied to target depth estimation – an important consideration for implementation and training with similar guidance tools. In the fifth chapter, neurosurgically experienced subjects were recruited to perform simulated ventriculostomies. Results showed anatomical variations influence performance and could impact outcome. Augmented reality guidance showed no marked improvement in performance, but exhibited a mild learning curve, indicating that additional training may be warranted. The final chapter summarizes the work presented. Our results and novel evaluative methodologies lay the groundwork for further investigation into simulators as versatile research tools to explore performance factors in simulated surgical procedures.

Keywords

Surgical Simulation, Neurosurgery, Human Performance, Human-Computer Interaction, Augmented Reality, Image-guidance.

Acknowledgments

At this point, I would like to acknowledge my examiners for taking the time out from their busy schedules to evaluate the rigor of this work.

Co-Authorship Statement

As this thesis is presented in an integrated format consisting of chapters adapted from published works, I wish to acknowledge to work of co-authors.

Chapter 2 co-authors: Leah Plumley, Sandrine de Ribaupierre and Roy Eagleson

Leah Plumley was instrumental in recruiting subjects and assisting in running the user study, which was later expanded upon. In addition, Leah performed an initial analysis that provided a foundation for further investigation. Roy Eagleson and Sandrine de Ribaupierre assisted in conceptualizing and designing the study as well as reviewing the manuscript. I would also like to acknowledge Tim Wilson for the use of his equipment.

Chapter 3 co-authors: Roy Eagleson and Sandrine de Ribaupierre

Roy Eagleson and Sandrine de Ribaupierre assisted in conceptualizing and designing the study as well as reviewing the manuscript.

Chapter 4 co-authors: Trinette Wright, Matt Kramers, Saeed Bakhshmand, Roy Eagleson, Sandrine de Ribaupierre and Aaron Fenster

Trinette Wright assisted in development and testing of the current augmented reality guidance interface. Matt Kramers developed a significant portion of the original interface which has since been redesigned, but provided the conceptual foundation. Saeed Bakhshmand contributed writing to the manuscript. Roy Eagleson, Sandrine de Ribaupierre, and Aaron Fenster assisted in conceptualizing and designing the study, providing equipment, and reviewing the manuscript.

Chapter 5 co-authors: Dayna Noltie, Roy Eagleson, and Sandrine de Ribaupierre

Dayna Noltie assisted in the study conceptualization, recruitment of subjects, and the collection of data. Roy Eagleson and Sandrine de Ribaupierre assisted in conceptualizing and designing the study as well as reviewing the manuscript.

Table of Contents

Preliminary Pages: 196.

Abstract	i
Acknowledgments.....	iii
Co-Authorship Statement.....	iv
Table of Contents	v
List of Tables	x
List of Figures	xi
List of Appendices	xvi
List of Abbreviations	xvii
Chapter 1	1
1 Introduction	1
1.1 Cognition.....	2
1.1.1 Perception	2
1.1.2 Visuospatial Reasoning.....	5
1.1.3 Psychometrics	6
1.1.4 Learning	6
1.1.5 Human Performance	9
1.2 Human-Computer Interfaces.....	12
1.2.1 Input Devices and Techniques	12
1.2.2 Displays, Visualization and Graphics	13
1.3 Augmented Reality	15
1.4 Ventriculostomies	16
1.4.1 Hydrocephalus	16
1.4.2 Hydrocephalus Treatment: Ventriculostomies	18

1.5	Surgical Simulation.....	19
1.5.1	Patient-Specific Simulation	19
1.5.2	Ventriculostomy Simulation	21
1.5.3	The NeuroTouch.....	21
1.5.4	Simulator Validation.....	23
1.6	Objectives	24
1.6.1	Chapter 2: Neuroanatomy Training and Evaluation in a Virtual Environment.....	25
1.6.2	Chapter 3: Automating Simulator Studies	25
1.6.3	Chapter 4: Examining Human Factors in Augmented Reality Guidance .	26
1.6.4	Chapter 5: Augmented Reality for Ventriculostomy Guidance.....	26
1.6.5	Chapter 6: Closing Remarks	27
	Bibliography.....	28
	Chapter 2.....	43
2	Neuroanatomy Training and Evaluation in a Virtual Environment.....	43
2.1	Introduction.....	43
2.2	Background.....	45
2.3	Methodology.....	46
2.3.1	System Design Considerations	46
2.3.2	System Description	47
2.3.3	User Study.....	52
2.4	Results.....	56
2.4.1	Anatomical Training	56
2.4.2	The Influence of Stereoscopy	58
2.5	Discussion.....	59
2.6	Conclusion	61

Bibliography.....	62
Chapter 3.....	66
3 Automating Simulator Studies	66
3.1 Introduction.....	66
3.2 Background.....	67
3.3 Methodology.....	68
3.3.1 Design Considerations	68
3.3.2 Hardware.....	69
3.3.3 Data Management	69
3.3.4 Study Automation	72
3.3.5 Security and Privacy	75
3.4 Results.....	76
3.5 Discussion.....	77
3.6 Conclusion	77
Bibliography.....	79
Chapter 4.....	81
4 Examining Human Factors in Augmented Reality Guidance	81
4.1 Introduction.....	81
4.2 Background.....	82
4.2.1 Augmented Reality Guidance	82
4.2.2 Evaluation of Image-guidance	83
4.3 Methodology.....	84
4.3.1 Augmented Reality Guidance System	84
4.3.2 Evaluation of AR Guidance System	86
4.3.3 Simulation Platform.....	87
4.3.4 Experimental Design.....	90

4.3.5	Analysis.....	91
4.4	Results.....	93
4.4.1	AR System Accuracy.....	93
4.4.2	User Performance.....	93
4.5	Discussion and Future Work.....	96
4.6	Conclusion	97
	Bibliography	98
Chapter 5	103
5	Augmented Reality for Ventriculostomy Guidance.....	103
5.1	Introduction.....	103
5.2	Background.....	104
5.3	Methods.....	107
5.3.1	Simulator Design	107
5.3.2	Experimental Design.....	112
5.4	Results.....	120
5.4.1	The Role of AR in Part-Task Ventriculostomies (Phase I).....	120
5.4.2	Objective Examination of Ventriculostomies (Phase II)	125
5.5	Discussion and Future Work.....	126
5.6	Conclusion	129
	Bibliography	130
Chapter 6	135
6	Closing Remarks	135
6.1	Future Directions	139
	Bibliography.....	142
Appendix A: A Mobile VR Simulation Environment		143
Introduction.....		143

Background	144
Methodology	145
System Description	145
User Study	147
Results	150
Discussion	151
Conclusion	152
Bibliography.....	153
Appendix B: A Pipeline for Creating Patient-Specific ETV Scenarios.....	157
Introduction	157
Methodology	158
Visualization	159
Segmentation.....	163
Conclusion and Future Work	164
Bibliography.....	166
Appendix C: Published Works Reprint Permissions	169
Appendix D: Research Ethics Approval.....	171
Curriculum Vitae	176

List of Tables

Table 1: Study parameters encapsulated within XML. While the table is incomplete, it offers insight into the general organization of XML data for tracking the progress of an ongoing study.....	70
Table 2: Mean values from all users for each modality for the percentage of intersecting trajectories, the deviation in the angle from the ellipsoid orientation, and the relative engagement.	94
Table 3: Mean values for measures examining translational error. Positive X is towards the right side of the mannequin head, positive Y is towards the anterior, and positive Z is towards the superior direction.	95
Table 4: Performance of users for all cases (without AR guidance) divided by difficulty classification with anatomical measures.....	122
Table 5: Performance of users for all cases (with AR guidance) divided by difficulty classification with anatomical measures.....	124
Table 6: The average engagement for residents and novices over each difficulty class.	125

List of Figures

Figure 1: An illustration of binocular vision. The left and right eyes view an object from different perspectives, converging on the point of focus or fixation point.	4
Figure 2: Illustrated learning curves for variable cognitive abilities and task consistency. A) High cognitive abilities and high task consistency. B) Low cognitive abilities and high task consistency. C) High cognitive abilities and low consistency. D) Low cognitive abilities and low consistency.	9
Figure 3: An illustration of a theoretical task of varying difficulty with two subjects. The slope of the line produces the inverse of the <i>IP</i> for each user on this particular task. User <i>B</i> indicates a higher <i>IP</i> compared to user <i>A</i> ; as the <i>ID</i> increases, the task completion time for <i>B</i> increases at a lower rate than it does for <i>A</i>	11
Figure 4: Interaction between a user and human-computer interface.	12
Figure 5: A cone rendered using 4 different approaches: A) wire mesh, B) flat shading with ambient lighting, C) smooth/Phong shading with ambient lighting, and D) smooth/Phong shading with a point light.	14
Figure 6: A view of a mannequin A) without AR and B) with AR depicting internal lateral ventricles.	16
Figure 7: A basic illustration of the brain ventricles. The lateral ventricles are located most superior and are generally symmetric across the midline. CSF flows from the lateral ventricles to the third ventricle through the Foramen of Monro. Each lateral ventricle possesses its own channel, as the lateral ventricles do not communicate under normal circumstances.	17
Figure 8: Comparison of axial brain slices. A) depicts hydrocephalic ventricles (from a CT image) and B) depicts normal ventricles (from an MRI).	18
Figure 9: The NeuroTouch simulation platform.	23
Figure 10: Simple UML diagram of rendering objects.	48

Figure 11: Depiction of rendered spatial cursor (green - indicated by arrow) and stationary markers (grey) placed in space using the Polhemus Patriot™ 51

Figure 12: The system in use. The user on the left controls the spatial input device (A) while examining the screen (C) projected from the dual polarized projectors (B)..... 52

Figure 13: Superior view of the lateral ventricle mesh. A) This region indicates the left anterior horn. B) This region indicates the right anterior horn. 53

Figure 14: The flow of feature targeting. A) The user is presented with a grey matter rendering and instructed to target features absent from the rendering. B) The user positions the cursor at the target location. C) The user selects the position and the anatomy is revealed. D) The user can examine the accuracy of their targeting. 54

Figure 15: Analysis of targeting. The error is taken as the distance between the user's selection (shown as the green sphere) and the closest point on the target area of the anatomical mesh..... 55

Figure 16: Screenshots depicting the flow of anatomical reconstruction. In (a) users must place the ventricular system to the correct location and with the correct orientation. After the ventricles are localized, they remain in place (b) while the user advances to placement of the next anatomical feature – the cerebellum (c)..... 56

Figure 17: Speed and accuracy of the neurosurgical resident. Pointing trial P1 is the set recorded before training and P2 is the set recorded after 3D anatomical display training. The vertical axis represents a measure of speed: the inverse of time ($1/\text{time}$ or s^{-1}). The horizontal axis is a measure of accuracy: the inverse deviation or error ($1/\text{error}$ in mm^{-1}). Points in the lower left corner indicate trials of generally poor performance; both the speed and accuracy were low. The fit curves indicate the *IP* or throughput. The red curve indicates higher overall performance, demonstrating a learning effect. 58

Figure 18: A screen capture of the study setup GUI. The wizard provides a number of facilities to customize a study 71

Figure 19: A sample of data stored in the XML document. Each tag specifies a metric for scoring the specific simulation module of the study. The category attribute is used to distinguish between the preoperative planning approaches. Additional attributes can be added to specify parameters specific to each study.	71
Figure 20: A flow chart of typical user interaction for a study session.	73
Figure 21: A code excerpt from the AutoHotKey script. This excerpt closes the open Slicer scene, activates the open Blender window, runs the script, and then opens the two rendered images to display the feedback of the user's targeting.	74
Figure 22: User interaction depicting automation as a black box.	75
Figure 23: The image-based cuboid marker has 4 trackable sides, allowing for a flexible workspace. As our platform's mannequin head is not fully supine, the apparatus is affixed to the mannequin head.	85
Figure 24: Left: The AR system overlays the rendered spheres onto the target points. In this instance, the mannequin head has not yet been covered for subject trials, allowing edges of the recesses to be seen. Right: While not used in this study, the mannequin head possesses a ball joint at the location of Kocher's point for use in additional simulation scenarios.	87
Figure 25: Results of NeuroTouch tool-mannequin registration process. Each point in the process corresponds to a physical recess on the mannequin head.	89
Figure 26: A) Screenshot taken from the augmented reality tool during the planning stage. B) Screenshot taken from Slicer 3D during the slice-based planning stage. C) Screenshot from feedback stage showing right sagittal view. The user's trajectory is rendered in green. D) Screenshot from feedback stage showing anterior coronal view.	90
Figure 27: A user guides their trajectory using the AR interface.	91
Figure 28: A) Measures for a trajectory that intersect the ellipsoid. B) Measures for targets that do not intersect the ellipsoid.	93

Figure 29: A screenshot of the slice-based interface. Unlike conventional radiographs, the right side of the image corresponds to the right side of the head.	108
Figure 30: An illustration of the image-based marker cuboid, the positional landmark at the nasion, as well as patient-derived lateral ventricles.....	110
Figure 31: The NeuroTouch used in conjunction with the AR guidance tool. The ventricles can be seen overlaid on the mannequin head in addition to a green mesh showing alignment between the virtual head and the physical model.	111
Figure 32: Two renderings providing feedback. A view from the right side (A) and from the front (B) are provided during experiments. The green line represents the user's trajectory. The right anterior horn of the lateral ventricles can be seen highlighted red – this roughly indicates the target area.....	112
Figure 33: Sagittal cross section of mannequin head with all ventricle cases overlaid.	113
Figure 34: Measures taken along an axial slice to determine Evan's ratio.	114
Figure 35: Summary of user study.....	116
Figure 36: In green, the trajectory pointed by the user, in blue, the ideal trajectory. A) Index of accuracy (IA) of 1 with a perfect trajectory. Note that the blue trajectory is occluded by the user's trajectory which is perfectly coincident. B) IA of 2 with a small angle change in one of the planes. C) IA of 2 with a small modification of angle in both planes, resulting in the trajectory coming out of the ventricle. D) IA of 6 with a poor trajectory.....	118
Figure 37: An illustration of the accuracy metrics. Engagement (a) is the length of the trajectory intersecting the anterior horn. The projected area (b) is the area from projecting the anterior horn onto the viewing plane from the burr-hole. The relative sagittal (c) and coronal (d) angulations are the angles the trajectory makes with the Foramen of Monro's respective path.....	120
Figure 38: Mean time and IA for each user without AR guidance. The users are numbered and their class of experience is indicated. For both time and IA, a lower score is an indicator of higher performance.....	121

Figure 39: Mean time and IA for each user with AR guidance. The users are numbered and their class of experience is indicated.	123
Figure 40: The mobile system depicted in its entirety (mobile device obstructed by headset).	147
Figure 41: A screenshot of the mobile simulator in use.	150
Figure 42: Pipeline for creation of patient-specific content from raw medical data.	159
Figure 43: Rendering of skull from CT with encased lateral ventricles and grey matter. The central blue circle represents user input.....	160
Figure 44: Rendering workflow. 1 – Images and label map are co-registered within the scene. 2 – Rays are cast into the scene for sampling. 3 – The label at each sample directs the algorithm to a transfer function and associated image. 4 – the result is composited to the viewing plane.....	162

List of Appendices

Appendix A: A Mobile VR Simulation Environment	143
Appendix B: A Pipeline for Creating Patient-Specific ETV Scenarios.....	157
Appendix C: Published Works Reprint Permissions	169
Appendix D: Research Ethics Approval	16971

List of Abbreviations

- AR** – Augmented Reality
- CI** – Confidence Interval
- CL** – Cognitive Load
- CLT** – Cognitive Load Theory
- CT** – Computed Tomography
- CSF** – Cerebrospinal Fluid
- DVR** – Direct Volume Rendering
- ETV** – Endoscopic Third Ventriculostomy
- EVD** – External Ventricular Drain
- GLSL** – OpenGL Shader Language
- GUI** – Graphical User Interface
- HCI** – Human-Computer Interaction
- ICU** – Intensive Care Unit
- ID** – Index of Difficulty
- IA** – Index of Accuracy
- IP** – Index of Performance
- K-S** – Kolmogorov-Smirnov
- MRI** – Magnetic Resonance Imaging
- OpenGL** – Open Graphics Library
- OR** – Operating Room
- RT** – Response Time
- VR** – Virtual Reality

Chapter 1

1 Introduction

According to the World Health Organization [1], approximately 3-20% of surgical procedures performed in industrialized countries result in complications. Incidence of preventable error varies by procedure, but error rates are generally higher for complex procedures [2]. Given that the majority of neurosurgical complications are of a technical nature, approximately 80% could likely be reduced or altogether prevented by mitigating surgical error [3]. For example, complex endoscopic surgeries — particularly those that exhibit a steep learning curve due to their highly technical nature — can have higher rates of technical errors, especially among novice surgeons [4,5]. While technical skill training is becoming increasingly valuable for many surgical disciplines, duty hour constraints on residents may impact their ability to sufficiently practice within the traditional apprenticeship training model [6].

The limitations of traditional residency programs to sufficiently train new physicians on increasingly technically demanding surgical technologies and techniques calls for the re-development of surgical training methods. Concurrent innovations in simulation hardware and software techniques allow for the creation of new approaches to surgical training. Virtual simulation technologies increasingly allow for the augmentation the classical training regime of apprenticeship in the operating theatre with virtual environment based surgical simulators. There is growing consensus that simulation technology will have a major impact on the clinic by reducing risk for error, reducing costs, improving operating room efficiency, and ultimately increasing the overall flexibility of technical skills training [7,8]. Indeed, research indicates that acquired skills from simulation can transfer to actual operations and potentially improve patient safety [9–12]. Surgical simulation is gaining popularity in many surgical fields, particularly those with complex technical skill requirements. Despite the promise of surgical simulation and increased interest and application, there continue to be significant limitations in this field. Unfortunately, most simulation platforms lack thorough validation and implement rudimentary performance metrics. Beyond training, modern

simulators provide capabilities to test novel techniques and evaluate factors that impact surgical performance to inform future developments, but this field remains largely unexplored.

Ahead of addressing these shortcomings, the remaining sections of this chapter will provide the contextual background information. As surgical skill and training reside in the psychological domain, *cognition* is a key area of focus as it pertains to the learning and performance of complex tasks. The design and implementation of *human-computer interfaces* are essential to provide a conduit between the surgeon and the *surgical simulator* or *image-guidance* tool. These techniques are all examined as they apply to *ventriculostomies* – common neurosurgical procedures that largely rely on the surgeon’s *visuospatial ability*.

1.1 Cognition

A theorization of cognition is essential to the study of complex surgical skills. Cognition is approached here as a constellation of factors which include: perception, memory, attention, reasoning, problem solving, decision making, and knowledge. Taking into account cognitive processes is a fundamental consideration in surgical skills training and image-guidance. This section examines the components of cognition most often considered when analyzing surgical interfaces: perception, visuospatial reasoning, and learning. Psychometrics are explored as a means to investigate and quantify cognitive processes.

1.1.1 Perception

In a surgical context, a physician is exposed to a wide array of stimuli. For example, a surgeon encounters and needs to make sense of the view of the surgical field, the physical sensation of the surgical tools and their interaction with the tissue, and sounds of the operating room. This low level information is filtered and processed by cognitive mechanisms and combines high level information such as a surgeon’s knowledge of anatomy and procedures in order to inform the decision making process. When examining image-guidance technologies and graphical interfaces, we are primarily concerned with how the user perceives visual stimuli.

Visual perception has been modeled as simultaneous processing of top-down and bottom-up processes [13]. Bottom-up processes are driven by the low level sensory processing of visual information from abstract entities into groups and objects. Top-down processing relies on prior knowledge in order to use the information stream from the bottom-up process. These two processes occur simultaneously in order to recognize discrete visual stimuli and process their role in a given context. Reliance on top-down processing for understanding the structure of objects and bottom-up processing for parsing the necessary visual cues are essential in depth perception, which remains an ongoing challenge in simulator and guidance systems.

Depth perception is the visual ability that allows perception of the world in three dimensions. Depth estimation of an object from visual stimuli is an essential ability for performing rudimentary motor tasks such as prehension — the action of reaching and grasping an object, which is an essential surgical ability. Processing depth cues involves the interaction of numerous aspects of the visual system. Understanding of the types and processing of depth cues is essential for exploiting them in graphical interfaces to convey depth virtually where there may be no physical component. There are two categories of depth cues: oculomotor (relying on the oculomotor responses of the eyes) and visual (further classified into monocular and binocular) [14]. Within these categories, psychologists generally differentiate between sixteen unique depth cues [15].

Oculomotor cues require an object's depth and its perceived depth to align for maximal depth estimation. While this may seem like a given, visual displays most often create a disparity by providing a rendering at one location (the screen), but rendering the image in such a way that the object appears to be at a further distance from the observer. Monocular cues involve properties such as shading and relative size of stimuli to determine depth. The majority of these are often trivially exploited in graphical renderings and have been employed by artists for hundreds of years [16]. In addition to static cues, dynamic virtual environments allow users to make use of motion cues such as motion parallax (the relative motion of objects as a function of their distance from the observer), which offer significant benefits to depth perception in the absence of binocular cues [17,18]. Stereopsis, known as binocular cues, involves the disparity between what is

seen by the right and left eyes (such as different angles). In order to provide binocular cues, a display system must provide separate images to each eye, simulating what each eye would see in a physical manifestation of the virtual scene. The presence of binocular cues in complex psychomotor tasks (such as prehension) is often beneficial to the performance of that task [19]. These cues play a strong role on our depth perception within a short distance of the observer [20]. Surgeons rely on this perceptual ability in order to effectively predict the depth of target anatomical features. An important note is that a small proportion of the population has no sense of stereopsis. Stereopsis is illustrated in Figure 1.

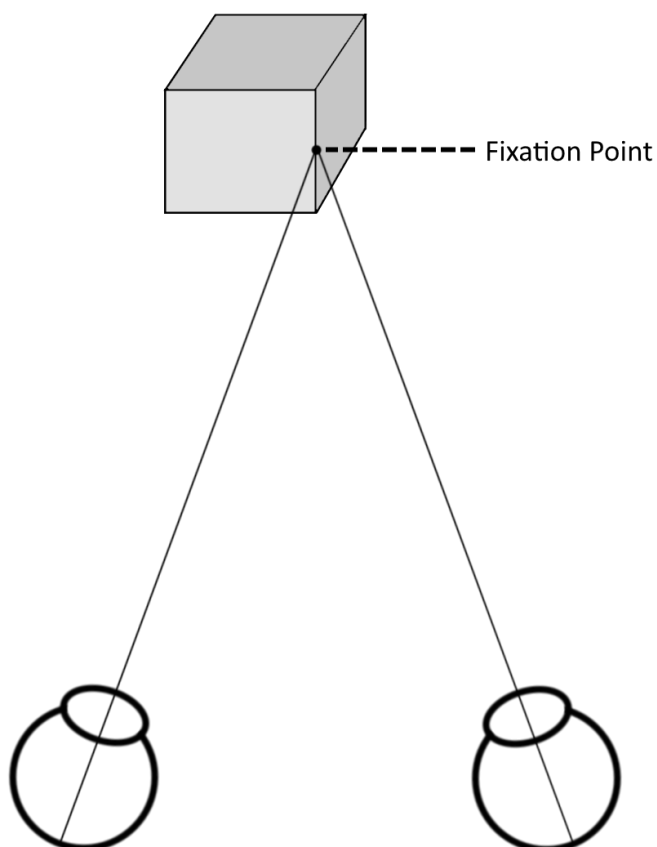


Figure 1: An illustration of binocular vision. The left and right eyes view an object from different perspectives, converging on the point of focus or fixation point.

1.1.2 Visuospatial Reasoning

Visuospatial reasoning is a fundamental psychological construct that is responsible for the ability to manipulate objects in multiple dimensions. Visuospatial representations and transformations can be approached from the bottom-up or the top-down; that is, we must consider the way that elementary processes provide the foundation for more complex reasoning as well as the way that prior knowledge and complex reasoning influence visuospatial processing [21]. This highlights the importance of a surgeon having both a sufficient aptitude for visuospatial abilities and a strong understanding of the spatial relationships inherent to the relevant anatomy. Indeed, the notion of examining surgeon visuospatial reasoning aptitude has a long history.

Visuospatial ability is a construct that is considered measurable and for which numerous tests have been developed. Shepard and Metzler [22] were among the first to propose a formal test for spatial reasoning ability with their mental rotations test. Many variations have since been developed and the test remains in use for a number of applications. The mental rotations test creates moderately complex 3D shapes using a series of connected cubes. Users are prompted with one such shape and then must correctly choose the same shape in a different orientation from a group of similarly constructed objects. Some consider there to be numerous subordinate types of visuospatial reasoning for which additional tests have been conceived [23]. Visuospatial reasoning is one of few cognitive abilities where strong sex differences have been observed [23, 24]. Age-related decline in visuospatial abilities have also been observed [25]. Naturally, visuospatial ability has been examined in numerous surgical contexts. Visuospatial abilities have been seen to play a role in the learning phase of complex skills, although the contribution of these abilities diminishes as experience is gained. The role of visuospatial abilities in learning have been examined in laparoscopic and endoscopic surgery – procedures with constrained workspaces and restricted vision of internal anatomy – both of which necessitates increased reliance on visuospatial reasoning abilities [26,27,28]. However, some studies have shown that high visuospatial reasoning continues to influence performance as expertise is gained [29]. A possible

explanation for this divergence is varying task consistency. This is explored further in in the *Learning* section.

1.1.3 Psychometrics

Psychometrics is a field concerned with the quantitative measurement of mental capacities and processes. Psychometric tests include measurements of both relatively innate abilities and acquired skills. 'Innate' abilities include fairly stable, albeit life course mediated, attributes like memory and general intelligence. In the present case, specific skills, such as surgical proficiency can also be quantifiably measured. The essential components of a robust psychometric test are its reliability and validity. Briefly, reliability is concerned with how consistently the measure performs whereas validity concerns the metric actually measuring what is intended to be measured. Ensuring validity may seem trivial by choosing appropriate measures however, almost counter-intuitively, this is often a difficult task, particularly when determining how to demonstrate validity

Developing metrics to measure surgical performance requires the ability to collect the appropriate data from the simulator or operating environment, an identification of the link between the metric and patient outcome and then thorough validation to ensure that the metric is, indeed, appropriate. Generally, a surgical procedure must first be decomposed into part-tasks where clear geometric goals can be established and measured. This is generally accomplished in consult with a surgical team.

1.1.4 Learning

Learning is the process where – in the context of surgery – new knowledge, skills or behaviors are acquired. The distinction is often made between ability and skill; abilities are thought to be innate qualities, such as intelligence quotient (a popular construct that has received an abundance of contemporary criticism), whereas skills are learned behaviors. However, it is not universally accepted that abilities exist that cannot be trained. In support of innate abilities (at very least as a concept by definition) many rationalize observed baseline changes in testing as reflecting a skill trained through the use of an imperfect test – one that does not solely measure the intended construct. The

question of whether training on a single skill affects other cognitive functions is a long-standing research area [30].

The acquisition of surgical skills is often approached theoretically via cognitive learning theory (CLT) [31]. CLT views cognition as an information processing system that can store information; the theory conceptualizes learning as the process where perceptual systems intake new information which is subsequently processed in working memory and ultimately stored in long term memory. This information can then be retrieved from long term memory for use in working memory when needed. This learning framework led to the development of cognitive load theory. The principal tenet of the theory is the acceptance of human limitations on information processing; the speed at which we can perceive and process stimuli is limited and our working memory has a relatively small capacity. When these systems are overloaded with stimuli presenting novel information, they become less effective. This phenomenon is known as cognitive overload and can severely impact the learning process as well as the performance of complex tasks [32]. Ultimately, when considering CLT in the design of simulators and guidance tools, extraneous information should be minimized, interfaces should be as intuitive as possible, and complex novel tasks should not overlap during the acquisition stage of learning [33]. This view aligns with Fitts' conceptualization of motor learning where he envisioned skill acquisition as consisting of three phases: the cognitive phase, the associative phase and finally the autonomous stage [34]. The cognitive phase builds the requisite knowledge to understand the task that must be completed. The associative phase involves refinements to movement and the development of consistency from trial to trial. Finally, the autonomous stage is characterized by fluid movements which require minimal conscious attention. In the early stages of learning, there is an increased demand on cognition before those skills ultimately become learned and autonomous.

In addition to CLT considerations, numerous factors have been observed to influence the learning of complex skills. The development of expertise in a given skill requires both deliberate and sustained practice [35]. When we examine surgical skills, we find that they rely on a combination of both conceptual and procedural knowledge. That is, the surgeon must recognize and understand the role of conceptual knowledge,

including vital anatomy, and then apply this through the use of procedural motor skills to perform the operation. It is thought that these domains of knowledge develop iteratively together and may be mediated by the presentation of the problem (how the learning phase of the task is presented to the user) [36]. The presentation of the problem may also impact a trainee's motivation – a factor that influences the learning curve [37].

Contextualization, personalization, and freedom in the approach of the problem have been observed to improve motivation and engagement, ultimately benefiting the learning process [38]. Thorndike's Law of Effect experiment demonstrated that quantitative feedback can facilitate learning [30]. To this end, mere repetition may not be sufficient for improvement at a given skill.

A common research challenge is the demarcation of tasks; many problems inherently possess variability, allowing for the possibility of presenting novel information between trials. It is not always trivial to determine what constitutes variability within a task and what constitutes a separate task altogether. In the domain of surgery, we generally consider different procedures analogous to different tasks and anatomical variability as a form of variability within the task, although many surgical procedures can be decomposed into common part-tasks. When examining task performance as a function of task consistency and general cognitive abilities, training on consistent tasks generally decreases performance variability as well as the influence of cognitive abilities, whereas inconsistencies in the presentation of the task will bound the performance by cognitive abilities [39- 41]. It follows that – while between-subject variability generally decreases on tasks with training – higher task complexity and inconsistency maintains the inter-individual gap [42]. This effect is explained by the raised demands on controlled processing that inconsistency imposes; performance of inconsistent tasks is relegated to the cognitive stage of learning [43, 44]. Figure 2 depicts a general illustration of learning curves for variable cognitive abilities and task consistency on a hypothetical technical skill.

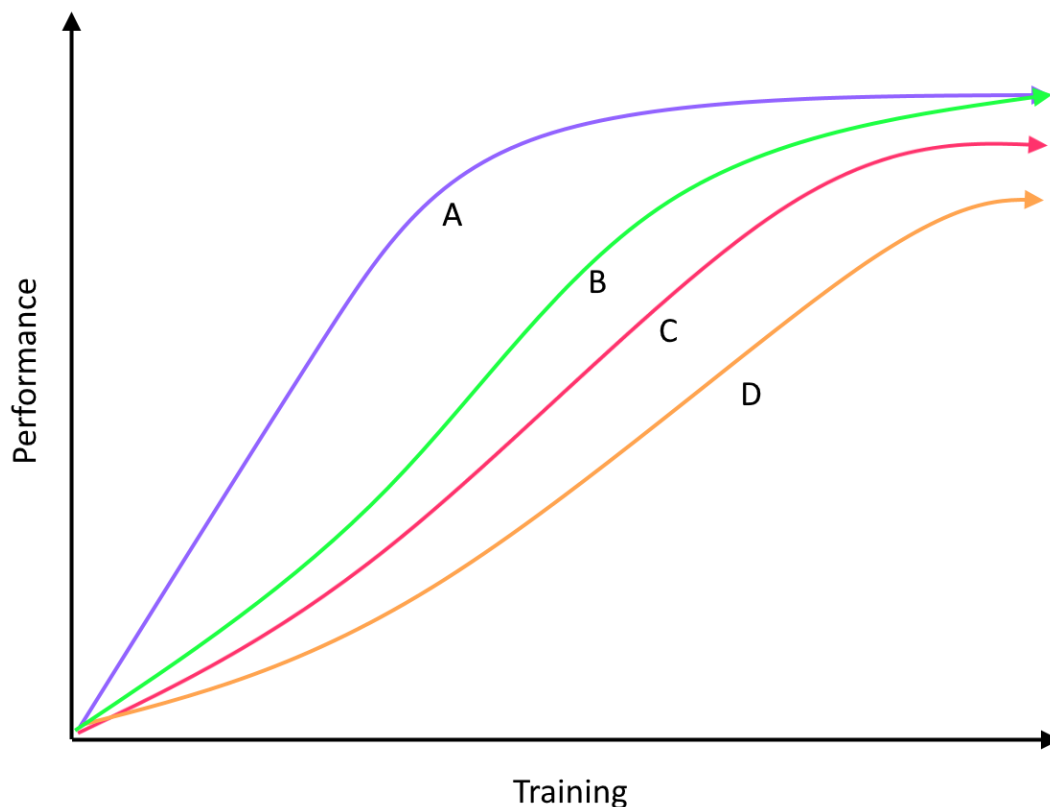


Figure 2: Illustrated learning curves for variable cognitive abilities and task consistency. A) High cognitive abilities and high task consistency. B) Low cognitive abilities and high task consistency. C) High cognitive abilities and low consistency. D) Low cognitive abilities and low consistency.

1.1.5 Human Performance

The study of human performance on simple and complex tasks is in many ways an extension of psychometrics. Studies of human performance aim to quantify performance with the intent of optimization in relation to controllable factors, such as competency in a surgical procedure. Evaluation of human performance is generally concerned with psychomotor tasks requiring the manipulation of physical objects. As this work is primarily concerned with the evaluation of such tasks with the intent of optimizing human-computer interfaces for surgical simulation and image-guidance, Fitts's law and methodological paradigm provide the foundation of this research.

Fitts's originally explored the relationship between speed and accuracy in simple psychomotor tasks involving variations in task difficulty [45]. For a simple one-dimensional targeting task, Fitts proposed the following formulation, adapting an information processing model to the study of human performance:

$$ID = -\log_2 \left(\frac{2D}{W} \right) \quad (1)$$

This equation solves for ID – the *index of difficulty* as a measure of bits. Additionally, W represents the width of a given target and D represents the distance to that target. The equation was applied to quantify the difficulty in a target selection task, but can be extended to additional applications. The *index of difficulty* is thought to be conceptually analogous to Shannon's information theory, which models the information capacity of communication [46]. In Fitts' formulation, D is comparable to the signal and W comparable to noise. In this conceptualization, the ID is representative of human signal-to-noise for a given task. In addition to the ID , Fitts formulated a measure of human performance:

$$IP = \left(\frac{ID}{MT} \right) \quad (2)$$

The IP represents the *index of performance* and is the ratio of the ID and the *movement time* (MT) for the given task. Hence, the units for the IP are bits per seconds, which represent a rate of information transmission.

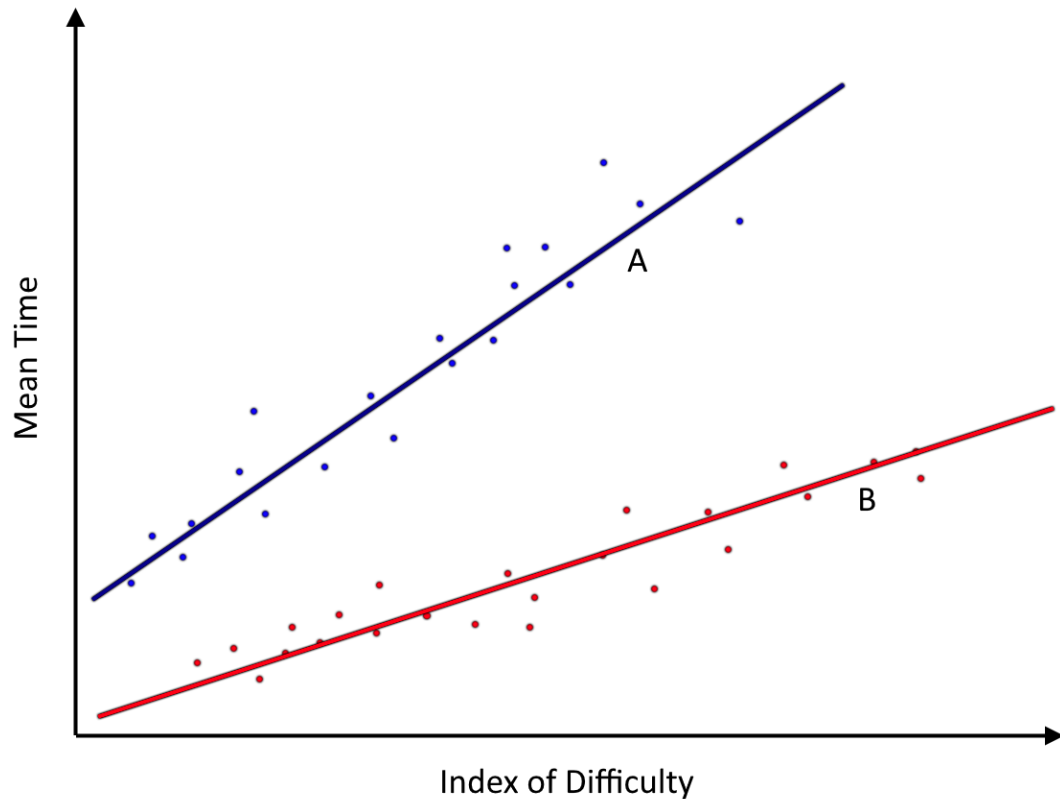


Figure 3: An illustration of a theoretical task of varying difficulty with two subjects. The slope of the line produces the inverse of the *IP* for each user on this particular task. User *B* indicates a higher *IP* compared to user *A*; as the *ID* increases, the task completion time for *B* increases at a lower rate than it does for *A*.

Contemporary investigations of human performance demonstrate numerous extensions and adjustments to Fitts's methodology [47–51], but a full discussion of these formulations is beyond the scope of this introduction.). In surgical applications Fitts' law may not be feasible or desirable for quantifying surgical proficiency; however, it can provide foundation for quantifying performance in tasks that can be decomposed into simple motor skills. While surgical tasks can theoretically be decomposed into an arrangement of simple target selection tasks, it's not always practical to do so. Furthermore, the reliance on task completion time as an evaluative measure may obscure results, given that in many surgical tasks, there is no need for task time optimization. One area where Fitts's methodology has made a significant impact is the design of

human-computer interfaces. Human-computer interfaces are often designed with motor movement simplicity as a driving design consideration, allowing for the simple decomposition of tasks and quantification of performance as an information processing paradigm.

1.2 Human-Computer Interfaces

Human-computer interfaces are at the core of modern surgical simulation and image-guidance techniques. HCIs encompass not only the means by which we send information to computer systems (such as traditional mouse and keyboard devices), but also the ways in which computer systems send information to us (such as the display on a computer monitor). This is depicted in Figure 4.

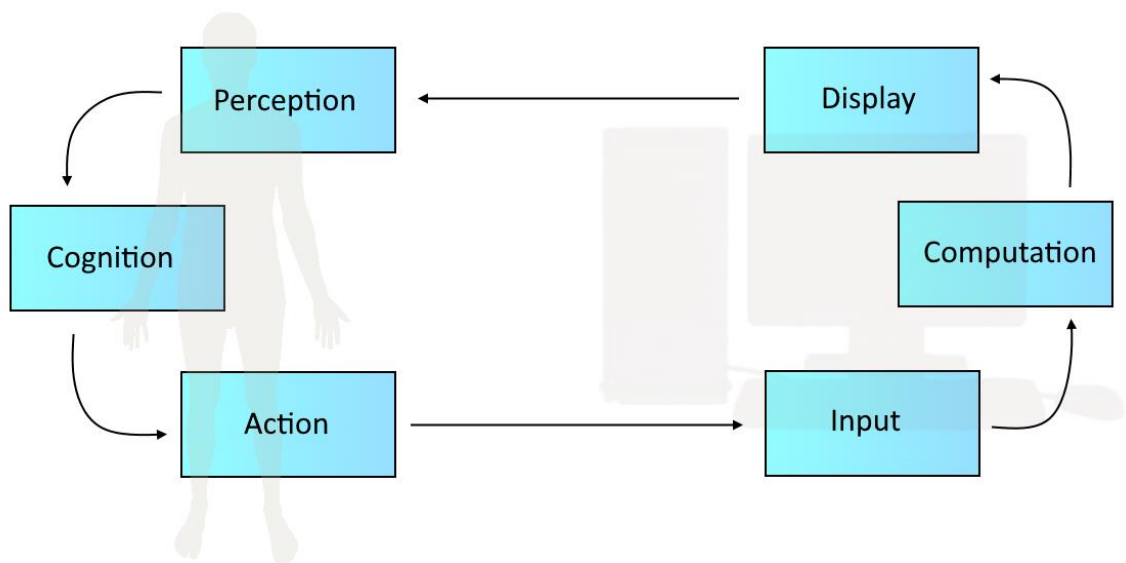


Figure 4: Interaction between a user and human-computer interface.

1.2.1 Input Devices and Techniques

Traditional computer input takes the form of a mouse and keyboard – tools of substantial utility for common use, however these devices are not suitable the simulation of surgical tasks requiring complex motor skills. Traditional interfaces also present numerous shortcomings when interacting with complex, three-dimensional medical imaging. They do not offer input in the same spatial domain as the data and are difficult to interact with

in some contexts (such as when hands-free input is desired). Various devices have been developed to address these shortcomings. These range from rudimentary tools to track position and orientation of an object using optical, mechanical, or electromagnetic tracking to systems that parse camera data to recognize specific gestures and actions [52-55]. A deficiency in many spatially-based input devices is the unidirectional constraint of tactility; users may apply force to the input device, but the input device cannot apply force back, resulting in a loss of tactile feedback that would otherwise be present in a physical simulation and the actual procedure. Haptic devices address this concern by simulating force on the input device. Implementation of haptics in surgical simulation has shown enhanced learning for some procedures [56–59], but haptic tools often drastically reduce the allowable range of motion and diversity of tasks that can be performed. Thus, it is not yet suitable for all simulation environments.

1.2.2 Displays, Visualization and Graphics

Fundamental to HCIs are computer displays that enable the visual perception of system variables. The most widely employed display system is the computer monitor, which allows depiction of two dimensional images of varying resolution and colour. Among the shortcomings of this traditional display is the limitation of monocular depth cues. There is growing evidence that stereoscopic displays can provide benefits to a number of medical domains [60]. Numerous three-dimensional display systems have been developed to address this concern, but few have seen any widespread use [61]. In recent years, virtual reality headsets have shown promising developments as their popularity in the consumer domain has fueled cheaper, more robust, and more accessible implementations. However, VR headsets are still plagued by the lack of environmental awareness, limited positional tracking, and the introduction of ‘VR sickness’ [62].

Visualization is the ongoing study of how to represent data in a way that best lends itself to accomplishing a specific task. When considering medical imaging, common objectives are to reach a diagnostic conclusion or to understand the spatial relationships of the underlying anatomy. Graphics (computer graphics more precisely) involve the algorithmic transformation of computer data into a format presentable as images on the display device. The traditional approach to the display of medical imaging

is to assign a graphical value to each data point of an image. Such a transfer function allows mapping a CT image of Hounsfield units in various orientations directly to grayscale values as a two dimensional cross section of the medical volume. In addition to conventional two dimensional mappings, anatomy derived from medical imaging can be rendered in ways to convey inherent spatial properties. By segmenting the features of interest, anatomy can be represented by meshes constructed from a manifold of triangles using an algorithm such as marching cubes [63]. However, this approach is time consuming to prepare and preserves only surface features, which may be insufficient for some applications. Direct volume rendering is an alternative approach that uses real-time geometric transformations and transfer functions to map a volumetric image onto a two dimensional plane [64–66]. There is compounding evidence indicating that three-dimensional representations of data better convey complex anatomy, facilitating learning [67, 68].

When displaying complex 3D anatomy, the primary intent of the graphics system is often to convey the spatial relationships inherent to the rendered object. This requires the visualization of edges, curvature and often transparency. Simple modifications to the shading system can have a significant impact on a viewer's ability to discern contours [69]. Figure 5 illustrates 4 different rendering techniques.

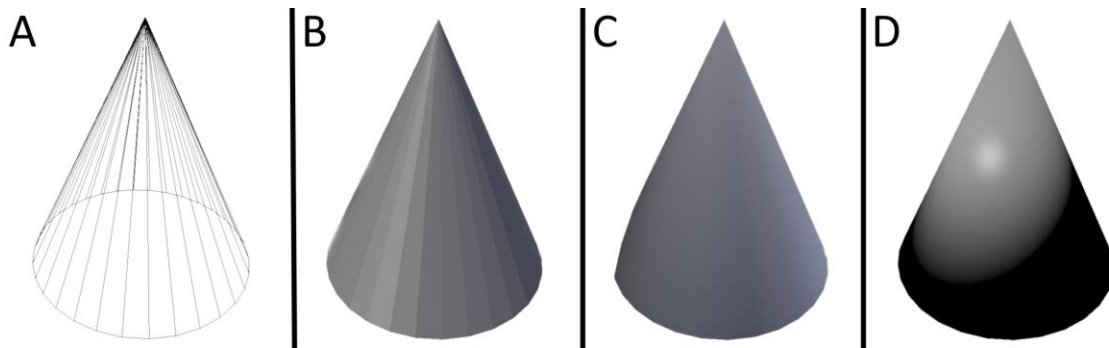


Figure 5: A cone rendered using 4 different approaches: A) wire mesh, B) flat shading with ambient lighting, C) smooth/Phong shading with ambient lighting, and D) smooth/Phong shading with a point light.

1.3 Augmented Reality

Augmented reality (AR) in the simplest sense is the juxtaposition of virtual and physical information. This can take many forms, but the focus of this introduction is towards visual information. In the context of surgery, AR is generally applied to augment the surgical setting with information not typically available to the surgeon at the time, or information presented in a more intuitive form. AR can generally be decomposed into two separate challenges: tracking and visualization.

Tracking for AR applications involves gathering information about the physical and virtual scenes to allow for registration between the two. When implementing real-time systems, the computational demands of tracking algorithms will impact the feasibility of their use. There are numerous methods for aligning physical and virtual scenes. Common tracking modalities include image-based, optical, and electromagnetic. Generally, these tracking modalities are used to select physical landmarks that correspond to known locations within the virtual scene. The type of computing platform as well as the use requirements for the system will determine the tracking chosen, although image-based tracking is gaining popularity due to the development of efficient algorithms that track predesigned markers as well as general features [70]. Hybrid approaches using multiple tracking modalities have also found clinical applications [71].

Once alignment of the physical and virtual scenes has been satisfied, information from each scene can be combined to create an augmented view. In most applications, the physical scene is used as the visual foundation with virtual information superimposed. There are various displays that support this paradigm. A relatively unobtrusive approach involves the use of a projector displaying information directly onto the physical scene (such as the patient) [72]. There are many challenges with this approach – such as the unpredictability of surface topology and rigidity of projector location – and it is generally considered to lack versatility. Head-mounted and hand-held displays are alternative approaches that either implement video see-through using the device's camera or optical see-through using transparent displays (as in Microsoft's HoloLens) [73,70]. Figure 6 provides a comparison between a non-AR and AR view of a mannequin head.

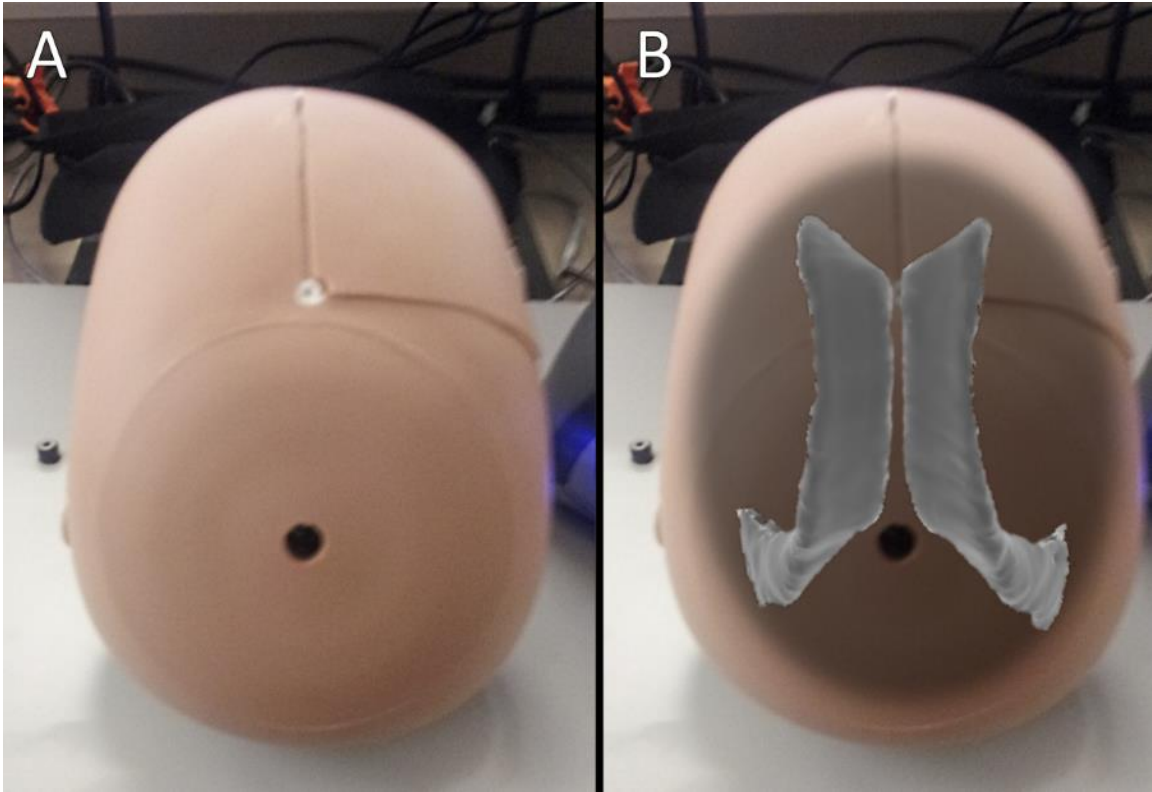


Figure 6: A view of a mannequin A) without AR and B) with AR depicting internal lateral ventricles.

1.4 Ventriculostomies

1.4.1 Hydrocephalus

Hydrocephalus is a condition characterized by an accumulation of cerebrospinal fluid in the brain due to excess fluid production (communicating hydrocephalus) or blockage (obstructive hydrocephalus), resulting in enlargement of the ventricular system [74]. The ventricles hold cerebrospinal fluid and produce it within choroid plexus. A rudimentary illustration of ventricle anatomy is presented in Figure 7. While there are four separately identified ventricles, discussion largely focuses on the two lateral ventricles and the third ventricle, as these features are of primary interest to the procedures examined in this thesis (and generally to the treatment of hydrocephalus).

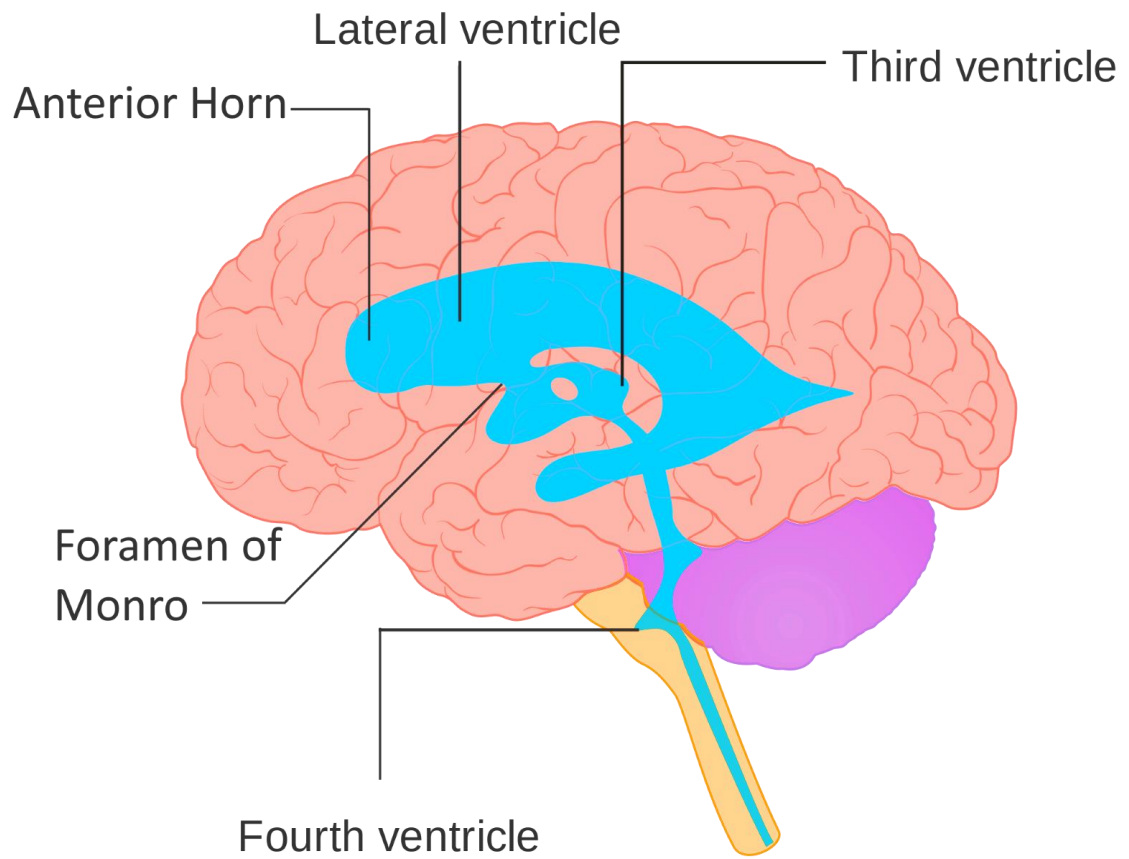


Figure 7: A basic illustration of the brain ventricles. The lateral ventricles are located most superior and are generally symmetric across the midline. CSF flows from the lateral ventricles to the third ventricle through the Foramen of Monro. Each lateral ventricle possesses its own channel, as the lateral ventricles do not communicate under normal circumstances.

Hydrocephalus results in an increase in intracranial pressure, which can lead to a number of neurological disorders or even death. While a large proportion of cases are congenital – largely affecting children – hydrocephalus can also occur secondary to existing pathologies or injury. Due to the nature of hydrocephalus, there is significant anatomical variation seen between patients [75,76]. See Figure 8 for a depiction of hydrocephalus.

The Hydrocephalus Association estimates that hydrocephalus occurs in 2 out of every 1,000 births [77]. Over a 10 year-long study conducted at the Sick Kids Hospital in Toronto, 839 patients had a total of 1183 hydrocephalus-related procedures throughout the duration of the study [78].

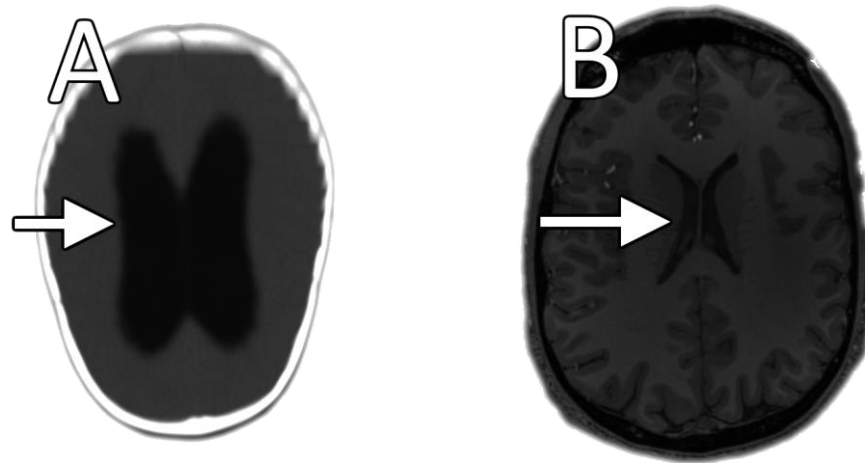


Figure 8: Comparison of axial brain slices. A) depicts hydrocephalic ventricles (from a CT image) and B) depicts normal ventricles (from an MRI).

1.4.2 Hydrocephalus Treatment: Ventriculostomies

There are two primary treatments indicated for chronic hydrocephalus: cerebral shunts and the endoscopic third ventriculostomy (ETV). Cerebral shunts act as cerebrospinal drainage systems, allowing fluid to flow from ventricular catheters in the brain to separate body cavities. While cerebral shunts have historically been the treatment of choice and have benefited from numerous technical advances, long-term complications are still common - often a result of infection or shunt malfunction [79]. The complication rate of shunts, coupled with an increased understanding of the third ventricle anatomy and improved surgical tools has led to widespread adoption of the ETV for obstructive hydrocephalus [80]. Several advantages of the ETV over mechanical shunts have been

reported in the literature [81]. For acute hydrocephalus, external ventricular drains (EVD) allow for the drainage of cerebrospinal fluid external to the head.

Although all ventriculostomies are procedurally unique, decomposition into part-tasks reveals common components. A fundamental skill of this class of procedures involves the targeting of the ipsilateral anterior horn of the ventricular system. This psychomotor skill requires the analysis of preoperative imaging combined with extensive anatomical knowledge to build a mental representation of the target. Following the planning stage, the surgeon must localize a burr-hole location on the skull and advance a trajectory into the anterior horn while blind to internal anatomy. Although this process has long been thought to be relatively safe, recent reports are shedding light on the high incidence of error that in some cases results in adverse events [82–87]. Anatomic variation has been identified as a common source of complications. [88,83,89].

1.5 Surgical Simulation

Training for surgical interventions—particularly complex endoscopic procedures—requires significant mastery of low-level technical skills in addition to high-level knowledge and cognition. Complex surgical procedures exhibit distinct learning curves which have direct impacts on patient outcomes [90]. Surgical simulation arose as a field to address a number of limitations to the current surgical training models. By providing a safe and versatile virtual training environment, surgical simulators have the potential to reduce the learning curve effect on patients, reduce the required time to skill mastery, and provide financial benefits to training programs and hospitals.

1.5.1 Patient-Specific Simulation

While surgical simulators are improving in quality and are increasingly adopted into educational settings, there are still numerous limitations which reduce their utility. Clinical demands require surgeons to operate on numerous cases, each of which may present unique conditions and challenges. However, the majority of current-generation simulators often provide only a fixed set of generic scenarios with prototypical anatomical features, failing to represent the variation seen in practice [91]. While direct clinical practice may always be a necessary component of neurosurgical residents'

curriculum, the limitation imposed by generic scenarios further separates simulations from reality. Patient-specific simulation provides the means to overcome this limitation by allowing for a multitude of practice scenarios to be created to represent actual clinical cases. By providing clinicians with the tools to generate such content preoperatively, the possibility to perform a rehearsal surgery on patient-specific data is also made a reality for both trainees and experienced surgeons.

Recent advances in simulation and image processing permit the feasibility of patient-specific imaging data to be incorporated into simulations [92]. While efforts in creating patient-specific simulations are gaining in momentum, major developments are still necessary as many simulators lack this functionality and most existing implementations are still prototypes [93]. Although often involving similar workflows, implementations are very domain-specific for both content creation and simulation, necessitating unique development for separate surgical procedures.

To date, there has been limited use of patient-specific information in surgical simulations. Willaert et. al. [93] outlines some recent advances in patient-specific simulation [93]. The most significant developments in the field have focused on simulation of carotid artery stenting, often relying on preoperative CT and MR angiography images to reconstruct the patient's anatomical features within the simulator [92, 94–100]. Additional implementations have been developed for brain tumour resection [101], cervical hip fracture surgery [102], sinus surgery [103], and more. Patient-specific simulation has been proposed for the ETV procedure, but projects have not come to fruition [104-105].

Although a number of patient-specific simulations exist, very few are used clinically or even for training. This is partly due to the lack of validation – an often difficult and time-consuming task. For training purposes, it is sufficient to demonstrate that scenarios represent possible cases, but scenarios for clinical rehearsal and planning must demonstrate either likeness to the patient, improved clinical outcomes, or superior operational performance. Typically, face validity is obtained through questionnaires attaining experts' subjective opinions to examine likeness to real patients. While such an

approach lacks objectivity, it has been the approach predominantly taken in the field. At very least it provides the opportunity to identify grossly unrealistic cases. Examining patient outcomes is a significant undertaking that has yet to be thoroughly examined. Finally, superior operational performance can be examined and linked to patient-specific rehearsal providing intra-procedure performance metrics can be obtained. These often take the form of subjective analyses by expert surgeons. Simulators with demonstrated construct validity can also be able to evaluate user performance.

1.5.2 Ventriculostomy Simulation

There are a number of existing simulators for various forms of ventriculostomies. The ImmersiveTouch (developed by the aptly named ImmersiveTouch company) is perhaps the most established simulator for ventriculostomies as it allows for patient-specific scenarios and offers the greatest depth of evaluation. The ImmersiveTouch has been used to examine catheter depth, location in the ventricular system, and distance from the Foramen of Monro [106]. While these measures are useful, they aren't fully descriptive of the approach and may prove insufficient when examining unique approaches resulting from unique anatomical variation. The Dextroscope VR platform has been used for simulating and planning ventriculostomy procedures, but it exists as more of an exploratory VR visualization than a fully functional simulator (it offers no pressure sensation – haptics) [107]. A mixed-reality simulator established by Hooten et al. [108] examined numerous clinically functional measures such as damage to eloquent structures but used only distance from the Foramen of Monro as the standard for tip localization. Additionally, a number of physical simulators have made a recent resurgence, although there has not been a significant effort to expand the traditional repertoire of performance metrics [109,110]. A VR-based neurosurgical simulator, the NeuroTouch contains a module for users to select a burr-hole location and trajectory to target the anterior horn in addition to a haptically enabled ETV simulation [111].

1.5.3 The NeuroTouch

The NeuroTouch is a commercial surgical simulation platform developed by the National Research Council of Canada [101]. Alternative neuroendoscopy simulators have been

implemented and proposed [105,104], including physical model-based simulators [112], but the NeuroTouch excels in its faithful representation of ventriculostomy procedures, overall fidelity, and versatility. Physical models in particular are limited in their ability to easily modify scenarios, making it difficult to implement patient-specific cases.

Initially developed for cranial microneurosurgery, the NeuroTouch contains a module for simulation of the ETV and EVD procedures [111]. The platform provides a phantom skull equipped with a Geomatic® Touch™ Haptic Device. The device provides the same range of motion available to a surgeon in the clinic. The simulator provides high-fidelity, real-time rendering of relevant anatomy as well as realistic deformation and tool-tissue interaction. The NeuroTouch is pictured in Figure 9.

The simulator offers some metrics for user performance, but is ultimately limited in the versatility of its metrics. Additionally, the NeuroTouch does not provide patient-specific simulations or anatomic variability by default, instead necessitating training on a single generic scenario.

The simulation offers only a single patient case and the evaluation of burr-hole location and catheter trajectory is based on comparison to a *gold standard* set by a neurosurgical expert that, although it will produce a functioning drain without damage to eloquent tissue, does not fully represent the full range of functional and safe drain placements. Scoring is based only on deviation from this ideal rather than objective measures that relate to the functionality of the drain, the error margin of the placement, and the avoidance of critical structures. Utilizing this evaluative methodology, it is possible for two dissimilar trajectories to be scored identically; a trajectory deviating a particular distance along one axis may still result in a safe and functioning drain, whereas an opposing deviation may miss the target and damage critical brain structures.

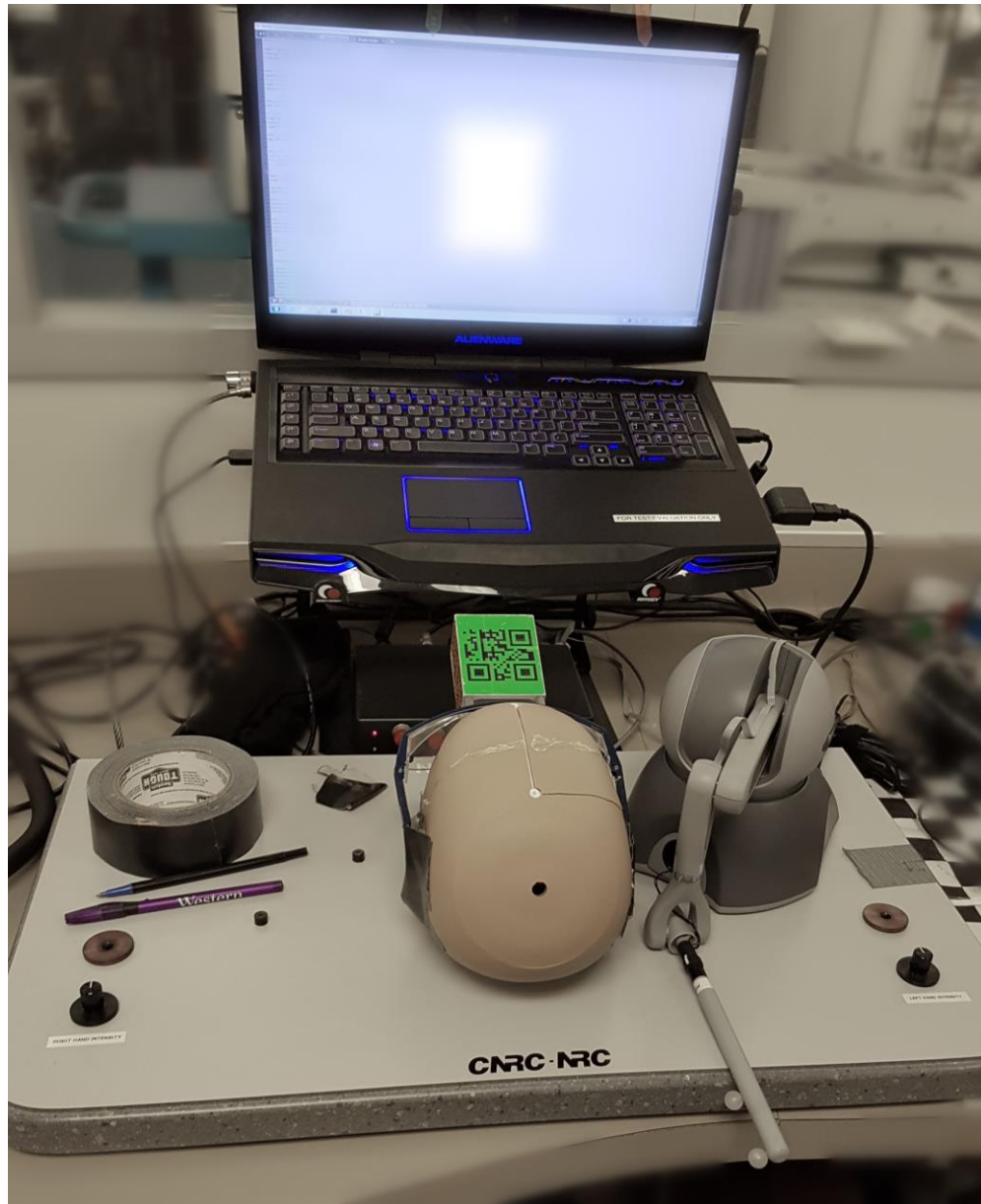


Figure 9: The NeuroTouch simulation platform.

1.5.4 Simulator Validation

Simulator validation is considered to be the process of determining that the implemented metrics measure the intended construct and that the simulator environment accurately depicts the surgical procedure. While efforts have been made towards consensus guidelines and standardization of simulator validation, validation studies with rigorous experimental methodology are in the minority [113]. Gallagher et al. [114]

identified a number of validation criteria that are most widely employed: face validity, content validity, construct validity, concurrent validity, discriminate validity, and predictive validity.

Face validity concerns whether the measures of a test appear to experts to measure what they intend to. Face validity is not considered a strong test for validity and often takes the form of questionnaires aimed at experts.

Content validity examines the content of the test items to ensure that all relevant components of the measured construct are contained within. Evaluation of content validity generally involves consultation with experts, which may take the form of a check list containing decomposed elements of the overall procedure.

Construct validity is the extent to which the test metrics measure construct that they intend to quantify. The evaluation of construct validity is contentious in the literature. Gallagher et al. distinguish between construct, concurrent and discriminate validity, although concurrent and discriminate validity are fundamentally indirect assessments of construct validity. In this text, the construct validity will instead be partitioned into convergent and discriminant construct validity [115]. Convergent validity identifies whether alternative measures and correlates of the construct correlate with the test in question, whereas discriminant validity identifies that the test does not correlate with measures that should not be related.

Predictive validity examines whether the metrics of a certain test correlate with and predict scores on the actual performance of the task. Predictive validity is essential for determining clinical worth of a simulator, but necessitates complex clinical trials that may present ethical conflicts.

1.6 Objectives

The overall objective of this work is to examine the factors that influence surgical performance in simulator environments. This includes examining the role of cognitive abilities, anatomical knowledge, interface design, patient-specific anatomical variations, and augmented reality image-guidance. By understanding these human performance

factors, clinical practice and technological developments can be influenced through the identification and mitigation of shortcomings that lead to human error.

1.6.1 Chapter 2: Neuroanatomy Training and Evaluation in a Virtual Environment

Neuroanatomy is one of the most challenging sections of anatomy to learn, partially related to the intricate relation of multiple 3D structures. As part of the medical student curriculum, it is usually taught in 2D using illustrations and plastinated brain sections. Since the number of hours devoted to anatomy has dropped in the curriculum, the dissection of brain is considered too time-consuming (in addition to the increased costs). In this chapter, we develop a novel stereoscopic VR tool with spatial input for neuroanatomy and analyze the role of innate spatial ability of novices in learning some basic structures and placing them back in a 3D volumetric brain. We investigated the hypothesis that specific spatial training and visualization tools would impact the evaluation and learning of complex spatial relationships in anatomy. Two tasks are performed after a short training session: the first one is to localize the ventricular tip as would be required during a temporal lobectomy, and the second task requires that the subject ‘reconstruct’ 3D anatomical structures within the context of our 3D brain model. We report our findings on the performance scores obtained from a population of subjects of differing backgrounds and spatial abilities.

1.6.2 Chapter 3: Automating Simulator Studies

While there is a plethora of promising new research directions made available by surgical simulation, recruiting eligible study participants remains a challenge. Not only do surgeons and residents work under substantial time constraints, but for many specializations the population size is small and sparse. This necessitates the employment of efficient user studies to maximize the collection of suitable data. Unfortunately, the majority of surgical simulation platforms are closed source and many are inflexible with regards to incorporation of customized modules and scenarios. These issues compound the difficulty of running a simulator study that handles data gathering while providing a streamlined series of tasks at a reasonable pace to subjects. In this chapter, we addressed

these issues by creating an automated framework for running user studies with custom modules on a rigid, closed-source surgical simulator.

1.6.3 Chapter 4: Examining Human Factors in Augmented Reality Guidance

Numerous augmented reality image-guidance tools have been evaluated under specific clinical criteria, but there is a lack of investigation into the broad effect on targeting ability and perception. In this chapter, we evaluated performance of 18 subjects on a targeting task modeling ventriculostomy trajectory planning. Users targeted ellipsoids within a mannequin head using both an augmented reality interface and a traditional slice-based interface for planning. We hypothesized that users would see improved performance using AR guidance. Users were significantly more accurate by several measures using augmented reality guidance, but were seen to have significant targeting bias; depth was underestimated by users with low targeting success. Our results further demonstrate the need for superior depth cues in augmented reality implementations while providing a framework for objective evaluation of augmented reality interfaces.

1.6.4 Chapter 5: Augmented Reality for Ventriculostomy Guidance

The placement of an external ventricular drain is one of the most commonly performed neurosurgical procedures, and consequently, is an essential skill to be mastered for neurosurgical trainees. In this chapter, we describe the development of a simulation environment to train residents on the acquisition of these skills before attempting the placement on live patients. In addition, the environment provides a safe test-bed for the evaluation of novel techniques and tools as well as for the examination of the factors that influence performance, particularly in regard to anatomical variations that occur clinically. A module for patient-specific simulation of free-hand ventriculostomy and integrated an augmented reality (AR) image-guidance system was developed. Patient-specific cases were included to represent a progression of burr-hole selection and ventricle targeting tasks. Seven residents and one expert neurosurgeon completed a number of targeting tasks with and without AR guidance for evaluation using an

extension of Fitts's methodology. A strong correlation between task accuracy and residency experience ($r^2 = 0.87$) was observed. Expert subjective classification of difficulty was found to be a better predictor of the challenge of a case than ventricle volume or Evan's ratio. AR guidance showed slight performance improvement, but not significant. Evidence indicates that the platform can discern experience and may be useful in training, as well as examining the use of novel tools and other factors that influence user performance.

1.6.5 Chapter 6: Closing Remarks

Chapter 6 will conclude with closing remarks and a discussion of future research directions.

Bibliography

1. Organization, W.H., others: WHO guidelines for safe surgery 2009: Safe Surgery Saves Lives. Geneva: World Health Organization. (2009).
2. Flum, D.R., Koepsell, T., Heagerty, P., Sinanan, M., Dellinger, E.P.: Common bile duct injury during laparoscopic cholecystectomy and the use of intraoperative cholangiography: adverse outcome or preventable error? *Archives of Surgery*. 136, 1287 (2001).
3. Stone, S., Bernstein, M.: Prospective error recording in surgery: an analysis of 1108 elective neurosurgical cases. *Neurosurgery*. 60, 1075–1082 (2007).
4. Schroeder, H.W.S., Niendorf, W.R., Gaab, M.R.: Complications of endoscopic third ventriculostomy. *Journal of neurosurgery*. 96, 1032–1040 (2002).
5. Teo, C., Rahman, S., Boop, F.A., Cherny, B.: Complications of endoscopic neurosurgery. *Child's Nervous System*. 12, 248–253 (1996).
6. Woodrow, S.I., Segouin, C., Armbruster, J., Hamstra, S.J., Hodges, B.: Duty hours reforms in the United States, France, and Canada: is it time to refocus our attention on education? *Academic Medicine*. 81, 1045–1051 (2006).
7. Babineau, T.J., Becker, J., Gibbons, G., Sentovich, S., Hess, D., Robertson, S., Stone, M.: The “cost” of operative training for surgical residents. *Archives of Surgery*. 139, 366 (2004).
8. Reznick, R.K., MacRae, H.: Teaching surgical skills-changes in the wind. *The New England journal of medicine*. 355, 2664 (2006).
9. Chaer, R.A., DeRubertis, B.G., Lin, S.C., Bush, H.L., Karwowski, J.K., Birk, D., Morrissey, N.J., Faries, P.L., McKinsey, J.F., Kent, K.C.: Simulation improves

- resident performance in catheter-based intervention: results of a randomized, controlled study. *Annals of surgery*. 244, 343 (2006).
10. Aggarwal, R., Ward, J., Balasundaram, I., Sains, P., Athanasiou, T., Darzi, A.: Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. *Annals of surgery*. 246, 771–779 (2007).
 11. Grantcharov, T.P., Kristiansen, V., Bendix, J., Bardram, L., Rosenberg, J., Funch-Jensen, P.: Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *British Journal of Surgery*. 91, 146–150 (2004).
 12. Seymour, N.E.: VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. *World journal of surgery*. 32, 182–188 (2008).
 13. Kinchla, R., Wolfe, J.: The order of visual processing: “Top-down,” “bottom-up,” or “middle-out.” *Perception & psychophysics*. 25, 225–231 (1979).
 14. Reichelt, S., Häussler, R., Fütterer, G., Leister, N.: Depth cues in human visual perception and their realization in 3D displays. *Proc. SPIE*. p. 76900B (2010).
 15. Cutting, J.: Vishton, PM (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. W. Epstein. 69–117 (6)AD.
 16. Crozier, W.R., Chapman, A.J.: *Cognitive processes in the perception of art*. Elsevier (1984).
 17. Dey, A., Cunningham, A., Sandor, C.: Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments. *Proceedings of the 17th ACM Symposium on Virtual Reality Software and*

- Technology. pp. 211–218 (2010).
18. Jones, J.A., Swan II, J.E., Singh, G., Kolstad, E., Ellis, S.R.: The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. Proceedings of the 5th symposium on Applied perception in graphics and visualization. pp. 9–14 (2008).
 19. Servos, P., Goodale, M.A., Jakobson, L.S.: The role of binocular vision in prehension: a kinematic analysis. *Vision Research*. 32, 1513–1521 (1992).
 20. Nagata, S.: How to reinforce perception of depth in single 2D pictures,-comparative study on various depth cues. Proceedings of the SID. pp. 239–247 (1984).
 21. Tversky, B.: Visuospatial reasoning. *The Cambridge handbook of thinking and reasoning*. 209–240 (2005).
 22. Shepard, R., Metzler, J.: Mental rotation of three-dimensional objects. *Science (New York, NY)*. 171, 701 (1971).
 23. Halpern, D.F.: Sex differences in cognitive abilities. Psychology press (2013).
 24. Peters, M.: Sex differences and the factor of time in solving Vandenberg and Kuse mental rotation problems. *Brain and cognition*. 57, 176–184 (2005).
 25. Kirasic, K.C.: Age differences in adults' spatial abilities, learning environmental layout, and wayfinding behavior. *Spatial Cognition and Computation*. 2, 117–134 (2000).
 26. Keehner, M.M., Tendick, F., Meng, M.V., Anwar, H.P., Hegarty, M., Stoller, M.L., Duh, Q.Y.: Spatial ability, experience, and skill in laparoscopic surgery. *The*

- American journal of surgery. 188, 71–75 (2004).
27. Ritter, E.M., McClusky III, D.A., Gallagher, A.G., Enochsson, L., Smith, C.D.: Perceptual, visuospatial, and psychomotor abilities correlate with duration of training required on a virtual-reality flexible endoscopy simulator. *The American journal of surgery*. 192, 379–384 (2006).
 28. Wanzel, K.R., Hamstra, S.J., Anastakis, D.J., Matsumoto, E.D., Cusimano, M.D.: Effect of visual-spatial ability on learning of spatially-complex surgical skills. *The Lancet*. 359, 230–231 (2002).
 29. Keehner, M., Lipka, Y., Montello, D.R., Tendick, F., Hegarty, M.: Learning a spatial skill for surgery: How the contributions of abilities change with practice. *Applied cognitive psychology*. 20, 487–503 (2006).
 30. Woodworth, R.S., Thorndike, E.: The influence of improvement in one mental function upon the efficiency of other functions.(I). *Psychological review*. 8, 247 (1901).
 31. Baron, R.A., Byrne, D.E.: *Social psychology: Understanding human interaction*. Allyn & Bacon (1984).
 32. Kirschner, P.A.: Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and instruction*. 12, 1–10 (2002).
 33. Sweller, J.: Cognitive load during problem solving: Effects on learning. *Cognitive science*. 12, 257–285 (1988).
 34. Fitts, P.M.: Perceptual-Motor Skill Learning 1. Categories of human learning. 243–285 (1964).

35. Ericsson, K.A., Krampe, R.T., Tesch-Römer, C.: The role of deliberate practice in the acquisition of expert performance. *Psychological review*. 100, 363 (1993).
36. Rittle-Johnson, B., Siegler, R.S., Alibali, M.W.: Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of educational psychology*. 93, 346 (2001).
37. Kanfer, R.: Task-specific motivation: An integrative approach to issues of measurement, mechanisms, processes, and determinants. *Journal of social and clinical psychology*. 5, 237–264 (1987).
38. Cordova, D.I., Lepper, M.R.: Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of educational psychology*. 88, 715 (1996).
39. Ackerman, P.L.: Determinants of individual differences during skill acquisition: cognitive abilities and information processing. *Journal of experimental psychology: General*. 117, 288 (1988).
40. Ackerman, P.L.: Individual differences in information processing: An investigation of intellectual abilities and task performance during practice. *Intelligence*. 10, 101–139 (1986).
41. Schneider, W., Shiffrin, R.M.: Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review*. 84, 1 (1977).
42. Ackerman, P.L.: Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological bulletin*. 102, 3 (1987).

43. Schneider, W., Fisk, A.D.: Degree of consistent training: Improvements in search performance and automatic process development. *Perception & Psychophysics*. 31, 160–168 (1982).
44. Fisk, A.D., Schneider, W.: Category and word search: generalizing search principles to complex processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 9, 177 (1983).
45. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*. 47, 381 (1954).
46. Shannon, C.E.: A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*. 5, 3–55 (2001).
47. Stoelen, M.F., Akin, D.L.: Assessment of Fitts' Law for Quantifying Combined Rotational and Translational Movements. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 52, 63–77 (2010).
48. MacKenzie, I.S.: Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*. 7, 91–139 (1992).
49. Guiard, Y., Beaudouin-Lafon, M.: Fitts' law 50 years later: Applications and contributions from human-computer interaction. *International Journal of Human-Computer Studies*. (2004).
50. Murata, A., Iwase, H.: Extending Fitts' law to a three-dimensional pointing task. *Human movement science*. 20, 791–805 (2001).
51. MacKenzie, I.S., Buxton, W.: Extending Fitts' law to two-dimensional tasks. *Proceedings of the SIGCHI conference on Human factors in computing systems*. pp.

- 219–226 (1992).
52. Hinckley, K., Pausch, R., Goble, J.C., Kassell, N.F., others: A three-dimensional user interface for neurosurgical visualization. *Proc. of the SPIE Conference on Medical Imaging*. pp. 126–136 (1994).
 53. LaViola, J.J., Keefe, D.F.: 3D spatial interaction: applications for art, design, and science. *ACM SIGGRAPH 2011 Courses*. p. 1 (2011).
 54. Caggianese, G., Gallo, L., Neroni, P.: Design and Preliminary Evaluation of Free-Hand Travel Techniques for Wearable Immersive Virtual Reality Systems with Egocentric Sensing. *Augmented and Virtual Reality*. pp. 399–408. Springer (2015).
 55. Robison, R.A., Liu, C.Y., Apuzzo, M.L.: Man, mind, and machine: the past and future of virtual reality simulation in neurologic surgery. *World neurosurgery*. 76, 419–430 (2011).
 56. Yiasemidou, M., Glassman, D., Vasas, P., Badiani, S., Patel, B.: Faster simulated laparoscopic cholecystectomy with haptic feedback technology. *Open Access Surg*. 4, 39–44 (2011).
 57. Shamsunder, S., Manivannan, M., others: Haptic guided laparoscopy simulation improves learning curve. *Studies in health technology and informatics*. 132, 454 (2008).
 58. Morris, D., Tan, H., Barbagli, F., Chang, T., Salisbury, K.: Haptic feedback enhances force skill learning. *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint*. pp. 21–26 (2007).

59. Van der Meijden, O., Schijven, M.: The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review. *Surgical endoscopy*. 23, 1180–1190 (2009).
60. Van Beurden, M.H.P.H., van Hoey, G., Hatzakis, H., Ijsselsteijn, W.A.: Stereoscopic displays in medical domains: A review of perception and performance effects. *IS\&T/SPIE Electronic Imaging*. p. 72400A–72400A (2009).
61. Holliman, N.: 3D display systems. to appear. 0–7503 (2005).
62. McGill, M., Boland, D., Murray-Smith, R., Brewster, S.: *A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays*. (2015).
63. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3D surface construction algorithm. *ACM Siggraph Computer Graphics*. pp. 163–169 (1987).
64. Engel, K., Hadwiger, M., Kniss, J., Rezk-Salama, C., Weiskopf, D.: Real-time volume graphics. *Eurographics Association*. 112–114 (2006).
65. Zhang, Q., Peters, T.M., Eagleson, R.: *Medical Image Volumetric Visualization: Algorithms, Pipelines, and Surgical Applications*. *Medical Image Processing*. 291–317 (2011).
66. Zhang, Q., Eagleson, R., Peters, T.M.: Volume visualization: A technical overview with a focus on medical applications. *Journal of Digital Imaging*. 24, 640–664 (2011).
67. Estevez, M.E., Lindgren, K.A., Bergethon, P.R.: A novel three-dimensional tool for teaching human neuroanatomy. *Anatomical sciences education*. 3, 309–317 (2010).

68. Nicholson, D.T., Chalk, C., Funnell, W.R.J., Daniel, S.J.: Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. *Medical Education*. 40, 1081–1087 (2006).
69. Abhari, K., Baxter, J.S., Khan, A.R., Peters, T.M., Ribaupierre, S.D., Eagleson, R.: Visual enhancement of MR angiography images to facilitate planning of arteriovenous malformation interventions. *ACM Transactions on Applied Perception (TAP)*. 12, 4 (2015).
70. Sielhorst, T., Feuerstein, M., Navab, N.: Advanced medical displays: A literature review of augmented reality. *Display Technology, Journal of*. 4, 451–467 (2008).
71. Fischer, J., Eichler, M., Bartz, D., Straßer, W.: Model-based Hybrid Tracking for Medical Augmented Reality. *EGVE*. pp. 71–80 (2006).
72. Tardif, J.-P., Roy, S., Meunier, J.: Projector-based augmented reality in surgery without calibration. *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*. pp. 548–551 (2003).
73. Sauer, F., Khamene, A., Bascle, B., Rubino, G.J.: A head-mounted display system for augmented reality image guidance: Towards clinical evaluation for imri-guided neurosurgery. *Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 707–716 (2001).
74. Scarff, J.E.: Treatment of hydrocephalus: an historical and critical review of methods and results. *Journal of Neurology, Neurosurgery & Psychiatry*. 26, 1–26 (1963).

75. Yadav, Y.R., Parihar, V., Pande, S., Namdev, H., Agarwal, M.: Endoscopic third ventriculostomy. *Journal of Neurosciences in Rural Practice*. 3, 163 (2012).
76. Morota, N., Watabe, T., Inukai, T., Hongo, K., Nakagawa, H.: Anatomical variants in the floor of the third ventricle; implications for endoscopic third ventriculostomy. *Journal of Neurology, Neurosurgery & Psychiatry*. 69, 531–534 (2000).
77. Tator, C., Bray, G., Morin, D.: The CBANCH Report-The Burden of Neurological Diseases, Disorders, and Injuries in Canada. *The Canadian Journal of Neurological Sciences*. 34, 268–269 (2007).
78. Tuli, S., Drake, J., Lawless, J., Wigg, M., Lamberti-Pasculli, M.: Risk factors for repeated cerebrospinal shunt failures in pediatric patients with hydrocephalus. *Journal of neurosurgery*. 92, 31–38 (2000).
79. Jabaudon, D., Charest, D., Porchet, F.: Pathogenesis and diagnostic pitfalls of ventricular diverticula: case report and review of the literature. *Neurosurgery*. 52, 209–212 (2003).
80. MISSING:brockmeyer2004techniques: MISSING:brockmeyer2004techniques. (2016).
81. Scarff, J.E.: Evaluation of treatment of hydrocephalus Results of third ventriculostomy and endoscopic cauterization of choroid plexuses compared with mechanical shunts. *Archives of Neurology*. 14, 382–391 (1966).
82. Bouras, T., Sgouros, S.: Complications of endoscopic third ventriculostomy. *Journal of Neurosurgery: Pediatrics*. 7, 643–649 (2011).
83. Kakarla, U.K., Chang, S.W., Theodore, N., Spetzler, R.F., Kim, L.J.: Safety and accuracy of bedside external ventricular drain placement. *Neurosurgery*. 63,

ONS162–ONS167 (2008).

84. Lollis, S.S., Roberts, D.W.: Robotic catheter ventriculostomy: feasibility, efficacy, and implications. *Journal of Neurosurgery*. 108, 269–74 (2008).
85. Z. Roitberg, N.K.M.A.T.H.F.C.J.A.B.: Bedside external ventricular drain placement for the treatment of acute hydrocephalus. *British journal of neurosurgery*. 15, 324–327 (2001).
86. Saladino, A., White, J.B., Wijdicks, E.F., Lanzino, G.: Malplacement of ventricular catheters by neurosurgeons: a single institution experience. *Neurocritical care*. 10, 248–252 (2009).
87. Schödel, P., Proescholdt, M., Ullrich, O.-W., Brawanski, A., Schebesch, K.-M.: An outcome analysis of two different procedures of burr-hole trephine and external ventricular drainage in acute hydrocephalus. *Journal of Clinical Neuroscience*. (2011).
88. Maniker, A.H., Vaynman, A.Y., Karimi, R.J., Sabit, A.O., Holland, B.: Hemorrhagic complications of external ventricular drainage. *Neurosurgery*. 59, ONS–419 (2006).
89. Hsieh, C.-T., Chen, G.-J., Ma, H.-I., Chang, C.-F., Cheng, C.-M., Su, Y.-H., Ju, D.-T., Hsia, C.-C., Chen, Y.-H., Wu, H.-Y., Liu, M.-Y.: The misplacement of external ventricular drain by freehand method in emergent neurosurgery. *Acta Neurol Belg*. 111, 22–8 (2011).
90. Udaya Kumar, M., Gill, I.S., others: Learning curve in human laparoscopic surgery. *Current urology reports*. 7, 120–124 (2006).

91. Soler, L., Marescaux, J.: Patient-specific surgical simulation. *World journal of surgery*. 32, 208–212 (2008).
92. Willaert, W., Aggarwal, R., Bicknell, C., Hamady, M., Darzi, A., Vermassen, F., Cheshire, N.: Patient-specific simulation in carotid artery stenting. *Journal of Vascular Surgery*. 52, 1700–1705 (2010).
93. Willaert, W.I.M., Aggarwal, R., Van Herzeele, I., Cheshire, N.J., Vermassen, F.E.: Recent Advancements in Medical Simulation: Patient-Specific Virtual Reality Simulation. *World journal of surgery*. 1–10 (2012).
94. Willaert, W., Aggarwal, R., Harvey, K., Cochennec, F., Nestel, D., Darzi, A., Vermassen, F., Cheshire, N.: Efficient implementation of patient-specific simulated rehearsal for the carotid artery stenting procedure: part-task rehearsal. *European Journal of Vascular and Endovascular Surgery*. 42, 158–166 (2011).
95. Roguin, A., Beyar, R.: Real case virtual reality training prior to carotid artery stenting. *Catheterization and Cardiovascular Interventions*. 75, 279–282 (2010).
96. Willaert, W.I.M., Aggarwal, R., Daruwalla, F., Van Herzeele, I., Darzi, A.W., Vermassen, F.E., Cheshire, N.J., others: Simulated Procedure Rehearsal Is More Effective Than a Preoperative Generic Warm-Up for Endovascular Procedures. *Annals of surgery*. 255, 1184 (2012).
97. Cates, C.U., Patel, A.D., Nicholson, W.J.: Use of virtual reality simulation for mission rehearsal for carotid stenting. *JAMA: The Journal of the American Medical Association*. 297, 265–266 (2007).
98. Willaert, W., Aggarwal, R., Nestel, D., Gaines, P., Vermassen, F., Darzi, A., Cheshire, N.: Patient-specific simulation for endovascular procedures: qualitative evaluation of the development process. *The International Journal of Medical*

- Robotics and Computer Assisted Surgery. 6, 202–210 (2010).
99. Hislop, S., Hedrick, J., Singh, M., Rhodes, J., Gillespie, D., Johansson, M., Illig, K.: Simulation case rehearsals for carotid artery stenting. *European Journal of Vascular and Endovascular Surgery*. 38, 750–754 (2009).
 100. Gallagher, A.G., Cates, C.U.: Virtual reality training for the operating room and cardiac catheterisation laboratory. *The Lancet*. 364, 1538–1540 (2004).
 101. Delorme, S., Laroche, D., DiRaddo, R., Del Maestro, R.F.: NeuroTouch: A Physics-Based Virtual Simulator for Cranial Microneurosurgery Training. *Neurosurgery*. 71, 32–42 (2012).
 102. Pettersson, J., Knutsson, H., Nordqvist, P., Borga, M.: A hip surgery simulator based on patient specific models generated by automatic segmentation. *Studies in health technology and informatics*. 119, 431 (2005).
 103. Parikh, S.S., Chan, S., Agrawal, S.K., Hwang, P.H., Salisbury, C.M., Rafii, B.Y., Varma, G., Salisbury, K.J., Blevins, N.H.: Integration of patient-specific paranasal sinus computed tomographic data into a virtual surgical environment. *American journal of rhinology & allergy*. 23, 442–447 (2009).
 104. Brown, N., Natsupakpong, S., Johannsen, S., Manjila, S., Cai, Q., Liberatore, V., Cohen, A.R., Cavusoglu, M.C.: Virtual environment-based training simulator for endoscopic third ventriculostomy. *Studies in health technology and informatics*. 119, 73 (2005).
 105. Sierra, R., Dimaio, S., Wada, J., Hata, N., Székely, G., Kikinis, R., Jolesz, F., others: Patient specific simulation and navigation of ventriculoscopic interventions. *Studies in health technology and informatics*. 125, 433 (2006).

106. Lemole Jr, G.M., Banerjee, P.P., Luciano, C., Neckrysh, S., Charbel, F.T.: Virtual reality in neurosurgical education: part-task ventriculostomy simulation with dynamic visual and haptic feedback. *Neurosurgery*. 61, 142–149 (2007).
107. Lee, C.K., Tay, L.L., Ng, W.H., Ng, I., Ang, B.T.: Optimization of ventricular catheter placement via posterior approaches: a virtual reality simulation study. *Surg Neurol*. 70, 274–7; discussion 277–8 (2008).
108. Hooten, K.G., Lister, J.R., Lombard, G., Lizdas, D.E., Lampotang, S., Rajon, D.A., Bova, F., Murad, G.J.: Mixed Reality Ventriculostomy Simulation: Experience in Neurosurgical Residency. *Neurosurgery*. 10, 576–581 (2014).
109. Ryan, J.R., Chen, T., Nakaji, P., Frakes, D.H., Gonzalez, L.F.: Ventriculostomy Simulation Using Patient-specific ventricular anatomy, 3D Printing, and Hydrogel Casting. *World neurosurgery*. (2015).
110. Coelho, G., Zymberg, S., Lyra, M., Zanon, N., Warf, B.: New anatomical simulator for pediatric neuroendoscopic practice. *Child’s Nervous System*. 31, 213–219 (2015).
111. Jiang, D., Hovdebo, J., Cabral, A., Mora, V., Delorme, S.: Endoscopic third ventriculostomy on a microneurosurgery simulator. *SIMULATION*. (2013).
112. Filho, F.V.G., Coelho, G., Cavalheiro, S., Lyra, M., Zymberg, S.T.: Quality assessment of a new surgical simulator for neuroendoscopic training. *Neurosurgical Focus*. 30, E17 (2011).
113. Carter, F., Schijven, M., Aggarwal, R., Grantcharov, T., Francis, N., Hanna, G., Jakimowicz, J.: Consensus guidelines for validation of virtual reality surgical simulators. *Surgical Endoscopy and Other Interventional Techniques*. 19, 1523–

1532 (2005).

114. Gallagher, A., Ritter, E., Satava, R.: Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surgical endoscopy*. 17, 1525–1529 (2003).
115. William, M., Donnelly, J.P.: Research methods knowledge base. Retrieved on June 10, 2014. (2006).

Chapter 2

2 Neuroanatomy Training and Evaluation in a Virtual Environment

This chapter is adapted from ‘A Software System for Evaluation and Training of Spatial Reasoning and Neuroanatomical Knowledge in a Virtual Environment’[1], and ‘Spatial Ability and Training in Virtual Neuroanatomy’ [2].

My contribution to this chapter involved (i) gathering system requirements, (ii) designing and implementing the software, (iii) testing and analyzing the system, (iv) designing the user study, (v) analyzing data, and (vi) writing the manuscripts.

2.1 Introduction

When we consider the design of interactive virtual reality (VR) systems for neurosurgical training, it is essential to develop evaluation methodologies based on objective metrics; quantification of learning allows for the comparative evaluation of learning scenarios. The situation is the same in the design of surgical simulators where training can be used to increase performance in basic and complex skills. Performance can be quantified using the objective metrics recorded directly to the simulation platform. In general surgical training, however, there is a hierarchy of knowledge and information that must be learned, ranging from the low-level psychomotor skills, to the task and sub-task skills that involve decision-theoretic performance coupled with movements such as targeting, navigation, and rehearsed skills such as cutting and suturing. The range of knowledge extends upwards to higher levels of knowledge abstraction: decision- making, situational assessment, spatial reasoning, diagnostics, and general surgical knowledge. Ultimately, the performance at each of these levels will need to be evaluated using a framework of compatible metrics. In this study, we explore a range of neuroanatomical training and knowledge evaluation that includes basic interaction skills, spatial reasoning and pattern recognition, and a form of spatial reasoning and anatomical knowledge involving the spatial relations between anatomical structures in the brain. Traditionally, the process of learning neuroanatomy occurs largely in a two-dimensional (2D) manner utilizing tools

such as textbooks, x-rays, plastinated slices, as well as medical imaging such as MRI and CT. In addition, time spent in dissection labs where interactive three-dimensional (3D) anatomy of the brain could be visualized has decreased in most institutions due to the cost and upkeep requirements of such implements [3]. Students must therefore be able to transform their understanding of 2D anatomical structures into a 3D context in order to use that knowledge in a practical sense [4]. There are two alternative approaches to facilitate the way by which this 2D to 3D translation can be successful: an individual may have innate high spatial ability, or they may be trained enough to have developed the skills required within a 3D environment. These same concepts can be applied to learning in a 3D environment; both high spatial ability and sufficient training could be expected to lead to better performance in a 3D task with high demand on spatial knowledge and reasoning.

The purpose of this study was two-fold. The first objective was to develop an application that provides stereoscopic visualization of anatomical brain structures while allowing for real time rotation and translation of the object in 3D and supporting 3 degrees-of-freedom (DOF) input for selection of points in the 3D virtual workspace. The software was to be used to test the user's knowledge of brain anatomy in 3D, and their ability to localize features in the context of a 3D surface rendering of the brain.

The second objective was to analyze human performance on the system with a particular interest in the correlation between spatial reasoning abilities and novice performance positioning anatomical structures within the brain. Based on current literature where sufficient training can attenuate the effects of spatial abilities on overall performance in some technical skills [4], we hypothesize that certain patterns and trends can be expected [5,6]. For complete novices, we expect their scores of being able to place a structure back in a 3D brain (in the correct position and orientation) to be correlated with their spatial abilities, while that difference would be minimal in comparison in subjects with previous or superior anatomical knowledge. In addition to anatomical knowledge, experience in VR environments may also facilitate performance, as is seen in the performance improvements in surgical simulators realized by training on video games where anatomical and procedural information is absent [7-9]. It follows that there must be

some point at which training becomes the better predictor of performance in a 3D environment over innate spatial ability [10]. In addition, we are investigating whether the form of presentation of the material (2D vs. 3D) will affect the understanding of spatial relationships of the trainee. We hypothesize that training emphasizing spatial relationships will show improved performance in examinations that rely on that spatial knowledge. In addition, we expect stereoscopic visualization to improve targeting by enhancing depth perception.

2.2 Background

Widely-employed traditional anatomical learning techniques such as 2D diagrams and cadaver labs present a number of limitations [11]. Two-dimensional representations, in particular, are not able to convey the full complex spatial relationships of 3D anatomical structures. Virtual reality teaching tools aim to rectify these limitations by providing controlled, portable, versatile and relatively inexpensive environments where complex spatial relationships inherent to neuroanatomy can be easily viewed and interacted with. It is widely believed that virtual models can enhance anatomy education [12], but it has been shown that not all visualizations are equally effective; optimizing the design of such visualizations is an active area of research [13,14]. One common issue seen in many computer visualizations is an increase in cognitive load resulting from non-intuitive or complex interfaces between the user and the software [15,16].

A difficult aspect in introducing new anatomical teaching tools is evaluating their effect on student knowledge and performance. Often, such evaluations are based on subjective measurements that rely on reports of user experience [17,18,19]. Part of the issue stems from the fact that there is no standard for testing anatomy knowledge. There has been an even smaller effort put in to evaluating the understanding of complex spatial anatomical relationships, which studies often evaluate through two-dimensional testing [67].

2.3 Methodology

2.3.1 System Design Considerations

The primary goal of the system was to provide discernable anatomical high-density surface renderings with maximal depth cues in real-time, while allowing the user to input spatial locations using 3-DOF input. The system was designed to evaluate user success in either localizing a point at a desired location in space within the brain, or placing a neuroanatomical structure in the correct location based on their anatomical knowledge. All points and structures were to be placed within a surface rendering of the brain's grey matter, which was rendered with appropriate transparency as to not impede viewing other structures and 3D cursors, but to also be entirely visible to the user. In addition, the system provided a mode to allow evaluators to setup evaluation sessions for the user. The overall requirement of the software package was to be predictive of user spatial skills and neuroanatomical knowledge as well as provide the appropriate practice and feedback to improve such skills and knowledge within a neuroanatomical context.

Abstraction and possibility for future extension were essential requirements in development. As such, an object oriented approach was adopted in order to facilitate these needs. The design allowed for simple management of geometric models as well as rendering routines. The application is easily extensible, allowing for swappable runtime modules that define the tasks and geometries involved.

The initial intent was to develop a cross-platform tool allowing portability. Software tools and packages used in the project were chosen with this in mind. Due to certain constraints on cross-platform tools, however, the software required an operating system specific implementation. Since the design is largely abstracted from the implementation and the majority of tools used are cross-platform, minimal changes are required in order to port the software to other systems, especially due to the encapsulation of platform-specific code.

2.3.2 System Description

The software was developed in C++ using Microsoft's Visual Studio 2010® targeting Windows 7®. The OpenGL graphics library was used to obtain real-time high-performance mesh rendering. The target machine is based on an Intel Extreme® i7-980X processor at 3.33GHz with 12MB Cache, 24GB Triple Channel RAM (1333Mhz DDR3), with a pair of dedicated graphics processors: dual GTX 470.

The brain model and corresponding neuroanatomy used were acquired from the Visible Human data set and reconstructed into segmented models manually from axial cryosections [20]. A triangle reduction was performed on the model prior to use with the software to obtain the ideal triangle density within the model for real time display. A trial and error approach was used to determine the optimal triangle density at which the rendering would run smoothly in real-time (minimum of 60 frames-per-second). The data was stored on disk in a standard format that separates the vertices, vertex normals and faces (triangle indices) into three distinct files, allowing selective loading at initialization if required. The following anatomical features were modeled and packaged with the system: the grey matter, the hippocampus, the cerebellum, the ventricles, the thalamus and the fornix.

The rendering system was designed to encapsulate OpenGL functionality into a modular system. All scene objects extend the abstract *Entity* base class, allowing all entities to be stored within Standard Template Library containers within the instanced *Renderer* class that controls the windowing and rendering operations. Each frame, a Standard Template Library Vector of *Entity* is iterated, drawing each model in succession. Models can also be referenced by index to modify transformation matrices. In this way, models must be pre-sorted before load-time to ensure proper drawing order depending on the transparency effects used. Each model was rendered in OpenGL using a *Model* object (extending *Entity*), containing the normalized vertices, normal vectors and face declaration, or indices, for each triangle. Each model was loaded from disk on program initialization into a *Model* object and then stored in global video memory as a static OpenGL Vertex Buffer Object before being added to the *Entity* Vector. The class relationships are illustrated in Figure 10.

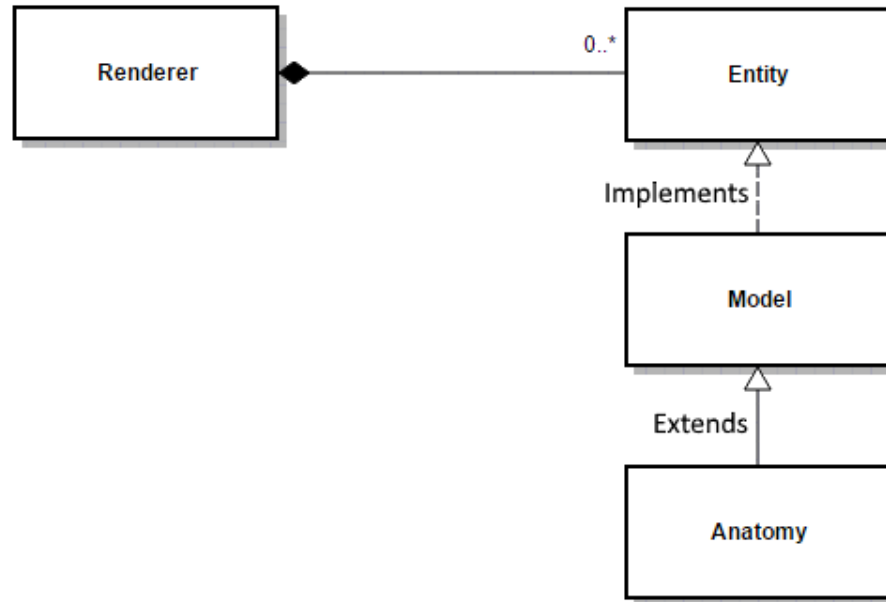


Figure 10: Simple UML diagram of rendering objects.

The challenge with the models was to use appropriate translucency and shading cues to create sufficient depth cues to enable the user to recognize the orientation of the brain that was presented and, given that information, to be able to determine the location of the cued or absent anatomical features in question within the 3D scene. The same challenge applied to rendering additional geometry used for user reconstruction of the brain anatomy. The transparency was made adjustable at runtime in order to accommodate each user's preference. Various monocular depth cues were introduced as well in order to facilitate perception of the scene. Perspective projection was used, although the effect was minimal on the brain model because of the effects of relative positioning. Depth shading was also introduced, although the distance between surface sides within the brain is minimal. A standard lighting model was used, centered at the viewpoint location to produce both ambient light and diffuse reflections on the model, aiding in the perception of structure. There are no inherent motion cues of the grey matter rendering as the brain remains static within trials and rotation between trials is switched instantly to new viewpoints.

The lighting and rendering utilized GLSL shaders. For the brain rendering, a single vertex and fragment shader was used, each with a single pass. All the lighting calculations occur in the vertex shader using a simplified Gouraud shading method with the addition of transparency. The role of the fragment shader was to simply receive and output the interpolated values. The main vertex shader computations occur in calculating the diffuse lighting.

Stereoscopic projection was implemented in order to improve depth perception within the scene, thereby facilitating the targeting tasks. The task of pinpointing a location within the scene mimics the task of prehension, which involves reaching and grabbing objects in space. Stereoscopic vision has been attributed to an increase in prehension performance [21], making it a desired addition to this visualization.

For the stereo display, a passive projection system was used, making use of polarized lenses with dual projectors and the corresponding polarized glasses for the user. The toed-in method for achieving stereo pairs was used for simplicity. Vertical parallax causing increased eye strain is a common issue with the chosen method, but was negligible in our case as trials were relatively quick, and there were no renderings in the extremities of the window where distortion is most likely to occur.

Since the target machine was running on commercial GPUs that lack quad-buffer stereo support, a single back-buffer was used for rendering each eye scene. This caused the displays to update asynchronously, but the effect was minimally visible to the user. In order to implement this type of display, the graphics engine was built directly on top of the native Windows 32 API. Each window was created with the same window class, but required distinct device contexts. Each device context was assigned to the same OpenGL-compatible pixel format. Each device context could be given a designated rendering context, but for simplicity, a single rendering context was used. This way, geometry and shader resources wouldn't have to be shared between contexts, but would exist only in a single context. This rendering context is then passed between device contexts to update each display every frame.

There are two interfaces to the software. The first is involved in setting up experiments and requires the keyboard and mouse. This allows the examiner to setup the display correctly, choose brain orientations, and run through the examination. The second input is used by the examiner to select points of reference and later controlled by the examinee during evaluation. This interface makes use of the Polhemus Patriot™ tracking device to obtain spatial input. The device is a magnetically tracked wand that is able to measure and input 6 DOF real-time, although we are currently only interested in position, not orientation. The Patriot™ has a 60Hz update rate with a latency of 17ms and a static position accuracy of 1.5mm RMS, which makes the device suitable for accurate real-time position tracking in our case. The magnetic sensor corresponding to the tracker will ideally be kept close to the user as resolution of the device decreases significantly with range. There is a single button on the wand that acts as a signal to query the device's coordinates and record them in order to target a location within the model or scene. The position can also be queried from the software, which allows the position to be updated periodically during normal movement.

The tracking device controls the position of a marker within the visualization, allowing the user to position this marker to visually target desired locations. When the user is tasked at reconstructing the inner anatomy of the brain, the marker is replaced with the substructure required for the task. This 3D cursor can be set to any desired arbitrary model loaded into the scene. The tracker is controlled via a custom Tracker class setup using the Polhemus Developer Interface API. Each frame, the rendering method first requests that the Tracker methods update its member coordinates. These coordinates can then be called from the *Renderer* instance to draw the cursor, but when direct rendering of the coordinates was used, hand tremors were visible in the display of the rendered marker, decreasing accuracy in localization. In order to smooth the movement of the marker within the scene, the updated location is filtered with a moving average algorithm, with larger weight given to most recent positions. This technique effectively smoothed the movement of the user-controlled cursor.

In the display, the cursor was represented as a small opaque sphere when point localization was required (see Figure 11). After the user identified the location within the

model that they wish to target, it is important to allow them to accurately pinpoint that location using the tracking device as input and the 3D cursor (sphere) as a reference. As such, depth cues were not only presented for the brain model, but for the cursor as well, which was lit by the same view-centered lighting model. The cursor moved linearly with the tracking device movements in real-time, facilitating a natural pointing experience for the user. In effect, motion based depth cues were introduced, such as motion parallax, helping the user to determine relative depth by providing input. The size gradient effect is also apparent when the user moves the marker into and out of the screen. Another important effect was introduced by the opaque marker interacting with the semi-transparent brain model. The user was able to identify when the marker was within the bounds of the brain by the partial-occlusion effect, acting as an effective cue in the target localization task [22]. The sphere geometry was loaded into global video memory at initialization, similar to the brain model. A separate shader was used to render the sphere opaque and with a Phong-based algorithm to provide precise shading for the relatively small object.

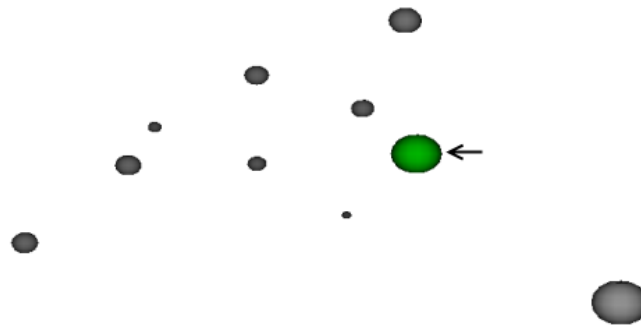


Figure 11: Depiction of rendered spatial cursor (green - indicated by arrow) and stationary markers (grey) placed in space using the Polhemus Patriot™

When the user was tasked with reconstructing the brain anatomy, the spherical cursor was rendered with the substructure chosen for placement within the brain. The same rendering techniques were applied to these structures. Each time a structure was placed within the brain, however, the anatomy became persistent within the model as a

reference and was rendered transparently. The model controlled by the tracking device was then advanced to the next model, representing another anatomical substructure of the brain for the user to place using it as a cursor. The system is shown in its entirety in Figure 12.

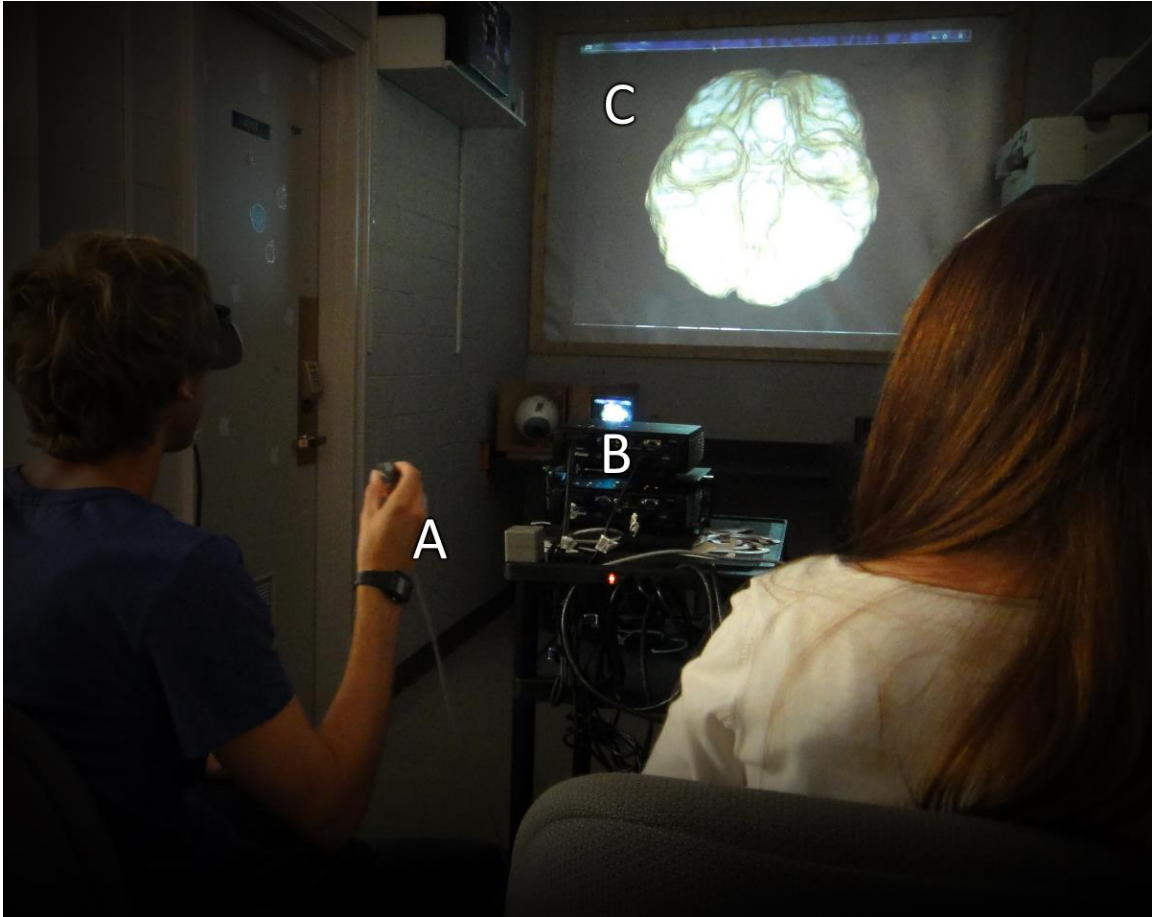


Figure 12: The system in use. The user on the left controls the spatial input device (A) while examining the screen (C) projected from the dual polarized projectors (B).

2.3.3 User Study

2.3.3.1 Spatial Abilities and User Performance

A total of 12 subjects were recruited for various anatomical tasks. Participants began by first completing the Santa Barbara Solids Test [23] to evaluate their spatial abilities. They were then stratified to either a 2D or 3D model from which they studied basic neuroanatomy. In each group, they had 10 minutes to learn 5 brain structures.

The 2D model was presented as a PowerPoint presentation consisting of labeled images including the lobes of the brain, major sulci, gyri and fissures, internal brain structures, the brain stem, and the cerebellum. The 3D model was viewed in a video format that was 1.5 minutes in length. The subject could rewind, stop or fast-forward the video and play it multiple times. The rotating image of the brain revealed the internal structures as in Miller's principles of syncretion [24]. The assembly of internal structures was differentiated using unique colour coding, and was deconstructed and reconstructed in various sequences throughout the video. The same structures were displayed in 2D and 3D.

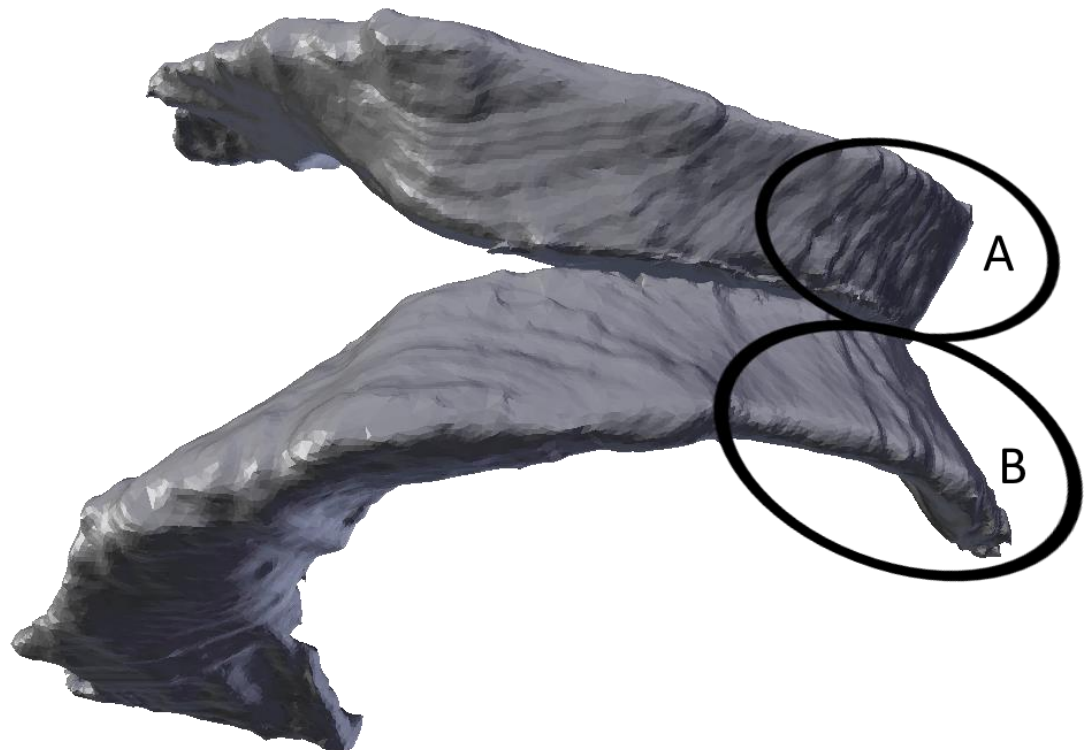


Figure 13: Superior view of the lateral ventricle mesh. A) This region indicates the left anterior horn. B) This region indicates the right anterior horn.

Following the review of neuroanatomy, participants completed two separate tasks using the stereoscopic display and software developed for this study. In the first task (anatomical targeting task), subjects were required to localize a point on the anterior

horn of either the right or left lateral ventricle; either location would be considered a correct response (Figure 13 depicts the rough boundaries of the anterior horns). In order to choose a location within the 3D brain, participants pressed a button on the magnetically tracked 3D wand whose movement corresponded with the spherical cursor in the scene.

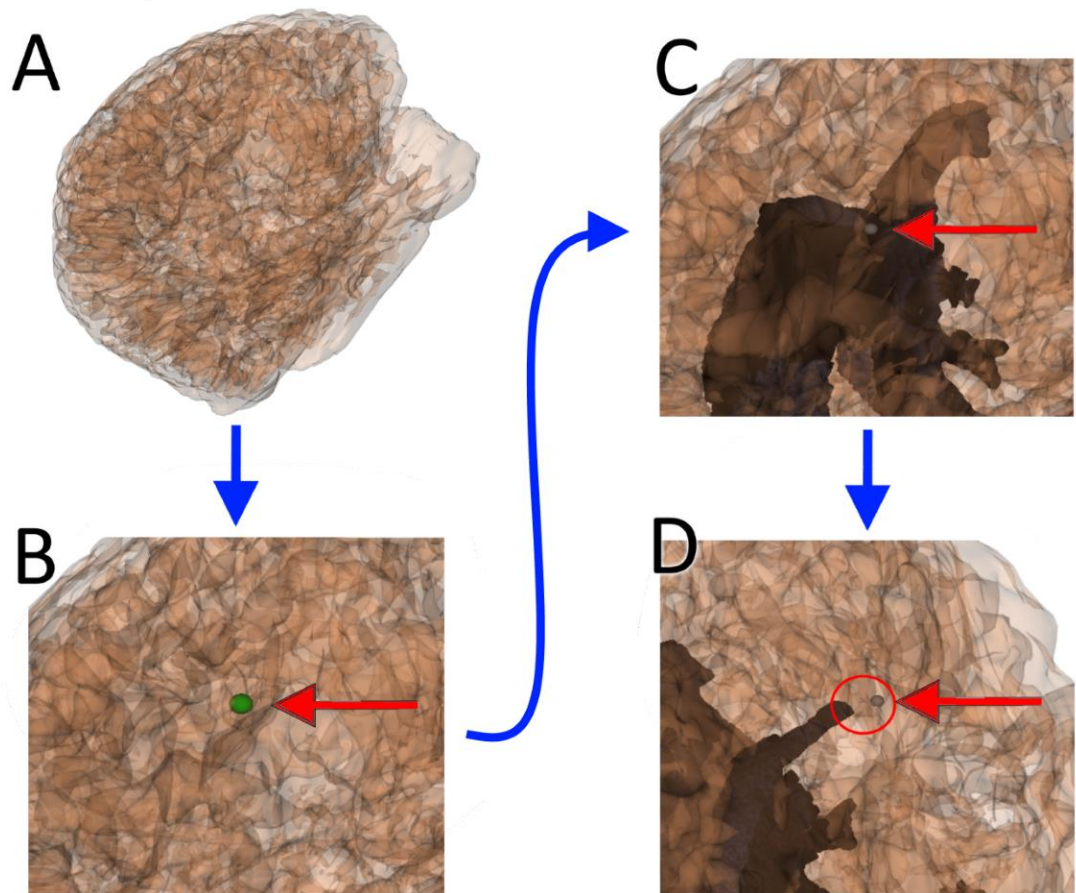


Figure 14: The flow of feature targeting. A) The user is presented with a grey matter rendering and instructed to target features absent from the rendering. B) The user positions the cursor at the target location. C) The user selects the position and the anatomy is revealed. D) The user can examine the accuracy of their targeting.

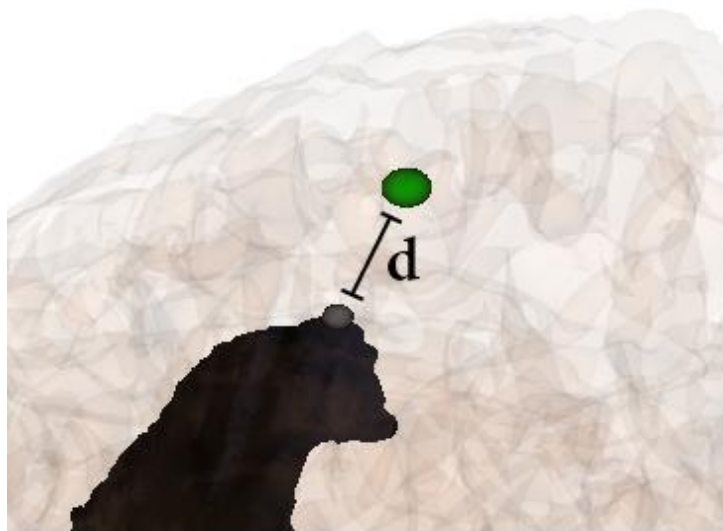


Figure 15: Analysis of targeting. The error is taken as the distance between the user's selection (shown as the green sphere) and the closest point on the target area of the anatomical mesh.

Each user performed this task 37 times using a different orientation of the brain each time. The first 10 attempts were considered as training and not included in the analysis. The flow of the task is pictured in Figure 14 and the analysis is illustrated in Figure 15. The second task was a novel test of anatomical knowledge (we refer to it as the reconstruction task). Although in principle it was a targeting task (involving prehension), the test involved rebuilding shapes in a 3D space, specifically, the 3D tracker controls the position of an anatomical feature and the user must place that feature at the correct location within the brain. The task was a test of anatomical knowledge – not just the recall of the shape of the anatomical structures, but also the positioning of those structures within the anatomical context. Participants were required to place 5 structures within the brain: the ventricles, the cerebellum, the fornix, the hippocampus, and the thalamus. In this task, the tracked wand corresponded with the movement of an individual structure, and the structure was placed at the chosen location by pressing the button on the 3D mouse. The participant had to scale the structure as well as insert it in the right location (Figure 16 illustrates this entire process). For both tasks the system recorded the placement coordinates as well as the time taken to complete the placement.

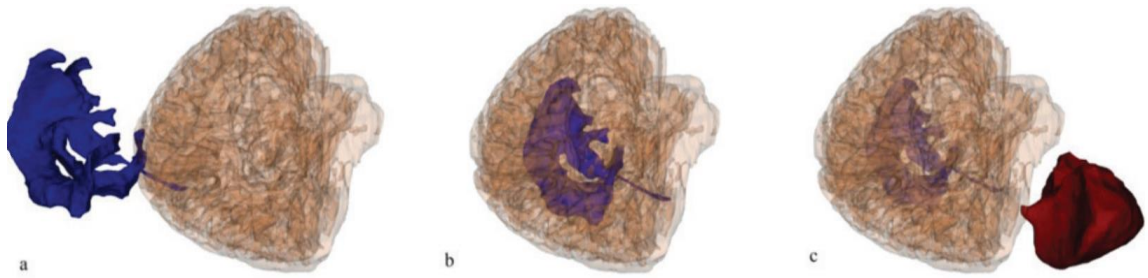


Figure 16: Screenshots depicting the flow of anatomical reconstruction. In (a) users must place the ventricular system to the correct location and with the correct orientation. After the ventricles are localized, they remain in place (b) while the user advances to placement of the next anatomical feature – the cerebellum (c).

2.3.3.2 The Role of Stereoscopy in Anatomy Reconstruction

The second user study examined the role of stereoscopic rendering for aiding in reconstruction of various brain structures using spatial input. A total of 8 subjects were recruited to perform 5 whole-brain reconstructions. Results examined translational error as well as task time. The subjects were not trained on anatomy, but had sufficient anatomical knowledge of the relevant structures. The users were stratified into the viewing order (3D vs. 2D first).

2.4 Results

2.4.1 Anatomical Training

From the 12 participants recruited, one subject was excluded from the analysis because of incomplete data. 10 participants were medically novices (6 women and 4 men with no surgical experience) and one participant was a neurosurgery resident. Spatial scores were separated by gender and were not found to be non-normal through a Kolmogorov-Smirnov test ($p > 0.15$). There was a significant difference between genders in spatial ability scores with men scoring higher on average ($p < 0.05$) using a two tailed t-test, as is seen in the literature [6]. The distance from the ideal target was recorded for each trial, for the pointing task and the object assembly task, as well as the time taken to achieve each trial.

Performance was calculated using a variation of Fitts's law, considering the speed and accuracy trade-off, but without explicit display of the tolerance for the pointing task (in other words, the 'W' parameter was not displayed or recorded. Instead, explicit error scores were taken from the center of the target). The following equation originally proposed by MacKenzie [25] was employed:

$$MT = a + b \log_2 \left(\frac{A}{W} + 1 \right) \quad (3)$$

where MT is the mean time, W is the user's distance away from the correct location or placement, and A was given a value of 3.0. This value was chosen because it is the average estimated distance that the user must move the cursor to the target each time. It is assumed that the distance remains constant throughout the experiment. The ID is contained within the term $\log_2(A/W + 1)$ and the IP is represented by $1/b$.

We did not find any significant correlation between the spatial ability test and the performance. When analyzing just accuracy by separating the subjects into poor and good spatial ability (each distribution was not found to deviate significantly from normality using a Kolmogorov-Smirnov test and obtaining each $p > 0.15$) there was a trend for subjects with lower spatial abilities to have a lower accuracy ($p = 0.09$ performed with a two tailed t-test). We did not find any differences in performance as a result of different training modalities. One interesting finding is that the resident performed well compared to the group, but had a low score at the spatial ability test. This subject was tested twice; once before training, then again after training. The subject's performance improved according to a two tailed Mann Whitney U Test (the data set was found to deviate significantly from a normal distribution using a Kolmogorov-Smirnov goodness-of-fit test with $p < 0.05$) comparing each session ($p < 0.01$ and $U=369$). The results are shown in Figure 17.

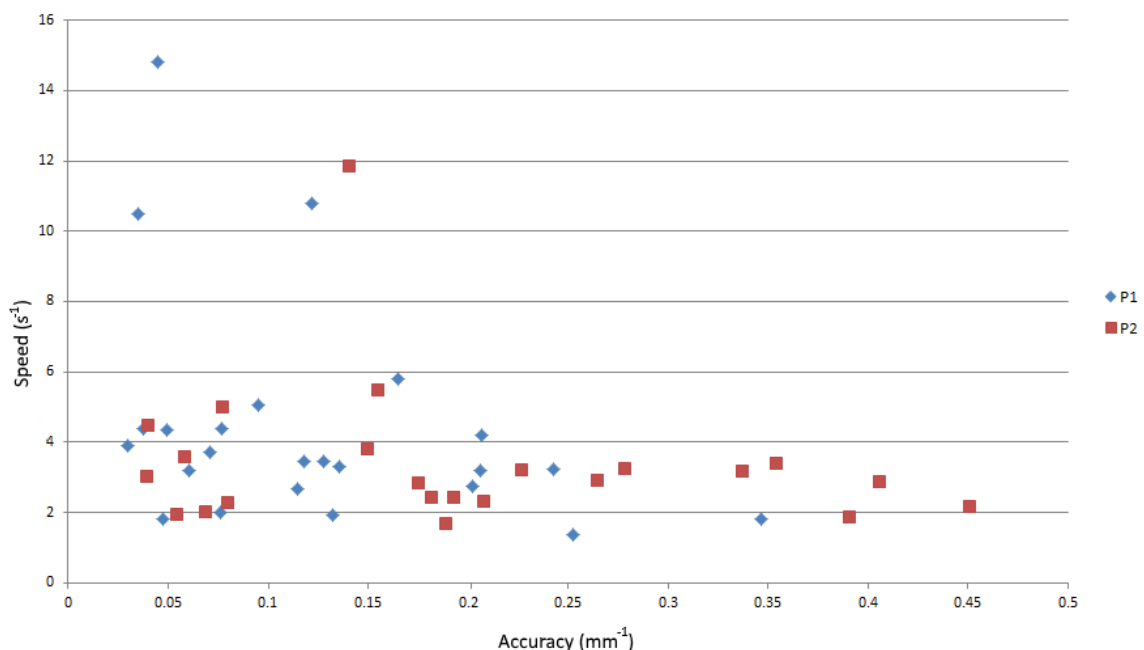


Figure 17: Speed and accuracy of the neurosurgical resident. Pointing trial P1 is the set recorded before training and P2 is the set recorded after 3D anatomical display training. The vertical axis represents a measure of speed: the inverse of time (1/time or s^{-1}). The horizontal axis is a measure of accuracy: the inverse deviation or error (1/error in mm^{-1}). Points in the lower left corner indicate trials of generally poor performance; both the speed and accuracy were low. The fit curves indicate the IP or throughput. The red curve indicates higher overall performance, demonstrating a learning effect. While statistical outliers are present, their use in analysis is justified; if a user felt that they misplaced a target accidentally, they were given the opportunity to replace the marker. As brain rotations were presented randomly, it's quite likely that several initializations of the task presented with opportunistic positioning through pseudo-random variations.

2.4.2 The Influence of Stereoscopy

No difference was seen between the accuracy of the stereoscopic and non-stereoscopic rendering modes. Users did, however, perform the task faster (time samples were found to be non-normal using a K-S goodness of fit test with $p < 0.05$) when presented with stereoscopic rendering ($p < 0.05$ and $U=3458$ using a two tailed Mann-Whitney Test).

2.5 Discussion

We presented a software tool for the evaluation and training of spatial skills. The software is unique in that it offers both virtual training and evaluation of spatial reasoning and spatial knowledge within a neuroanatomical context. Since the initial developments of this software tool, it has since been ported to modern VR and interaction technology. In its current state, the program is implemented on an Android platform making use of the Google Cardboard

Our user study focused on novices because of previous results using 3D models and looking at spatial ability, which had shown that novices had a greater variability depending on their mental rotation scores [26].

When measuring the performance of a task, no matter what level of abstraction, there are only “task time” and “error rate;” or, in other words – speed and accuracy. It is not only striking that there are only two objective measures, but in fact they are inseparably intertwined; because of the speed-accuracy trade-off at any level of the abstraction hierarchy, if a subject has the goal of behaving more rapidly, they will be less accurate and if they have the goal of behaving more accurately, they will require more time. This is the trade-off explored in the seminal work by Paul Fitts’s [27] and provides the foundation for the analysis presented here.

One of the constraints on Fitts’s methodology is that it is explicitly a task in which the accuracy is controlled and the time is measured. Accordingly, we consider the inverse – just as it would have been plausible for Fitts to have set up a paradigm in which the granularity of the time was the controlled variable, and the error was the measured parameter. We can consider how the ‘Index of performance’ methodology would have had to be reformulated in accordance with the following targeting paradigm. First, the subject indicates that they are ready for a new targeting task. The target appears, and then the subject moves towards the target. The subject clicks the mouse button when they decide that they have acquired the target. The screen is cleared, and the subject is then

ready for the new trial. The log files from the datasets are recorded, and the analysis proceeds without bias from the recorded data. In other words, even though the ‘time’ and ‘distance error’ measures are the recorded data stored in the columns, the analysis is based on ‘speed’ and ‘accuracy’, but these are merely extracted as reciprocals from the raw data files. Therefore, ‘speed’ and ‘accuracy’ are constructed as the reciprocal of the time and error measures taken from the raw data files.

We did not find any significant correlation between spatial ability and performance in this test – only a weak tendency to perform better with a better spatial ability. We had chosen the Santa Barbara Solid Test since it is compatible with the need to visualize objects in cross-section and imagine them in 3D. However, this might not have been the best correlate to the 3D perceptual task. We are therefore planning to rerun those subjects using a modified mental rotation test (Sheppard and Metzler [28]) in the future. Interestingly the neurosurgical resident received the lowest score on the Santa Barbara Solids Test yet performed among the top scores in each task, which would imply that her previous knowledge of neuroanatomy did, in fact, play a larger role in her results as opposed to her spatial abilities.

Stereoscopy did not improve accuracy as expected, but did improve the users time when reconstructing anatomy. This lack of improved accuracy may be a result of reliance on cues other than stereopsis. For example, the relative scale of each anatomical feature as it moves in the depth direction may have been the positional cue that users most relied on.

Future plans include testing and evaluating more novices and residents on the current system, repeat spatial ability using a modified Shepard and Metlzer in addition to the Santa Barabara Solid Test, as well as expanding the software capabilities to allow for free rotation of the static/reference internal structures by the user, which could increase its efficacy as a training tool.

2.6 Conclusion

Our analysis allows us to rank the performance of each subject in terms of their errors for the pointing task. There is a correlation between scores on the anatomical pointing task, and the anatomical 'structural reconstruction' task. The performance increase seen from stereoscopic rendering is likely attributable to the improved presence and depth perception. A follow-up study is planned in order to decompose these factors. Some caveats remain; the quantitative improvement in the subject's pointing task performance seems to partially facilitate the task for subjects with low Mental Rotation scores. Currently, the number of subjects in this study is too low to power a test for difference in 2D vs. 3D training effects for these subjects with low MRT scores.

Bibliography

1. Armstrong, R., de Ribaupierre, S., Eagleson, R.: A software system for evaluation and training of spatial reasoning and neuroanatomical knowledge in a virtual environment. *Comput Methods Programs Biomed.* 114, 29–37 (2014).
2. Plumley, L., Armstrong, R., De Ribaupierre, S., Eagleson, R.: Spatial ability and training in virtual neuroanatomy. *Stud Health Technol Inform.* 184, 324–9 (2013).
3. Collins, T.J., Given, R.L., Hulsebosch, C.E., Miller, B.T.: Status of gross anatomy in the US and Canada: Dilemma for the 21st century. *Clinical anatomy.* 7, 275–296 (1994).
4. Hegarty, M., Keehner, M., Cohen, C., Montello, D.R., Lippa, Y.: The role of spatial cognition in medicine: Applications for selecting and training professionals. *Applied spatial cognition.* 285–315 (2007).
5. Keehner, M., Lippa, Y., Montello, D.R., Tendick, F., Hegarty, M.: Learning a spatial skill for surgery: How the contributions of abilities change with practice. *Applied cognitive psychology.* 20, 487–503 (2006).
6. Guillot, A., Champely, S., Batier, C., Thiriet, P., Collet, C.: Relationship between spatial abilities, mental rotation and functional anatomy learning. *Advances in Health Sciences Education.* 12, 491–507 (2007).
7. Schlickum, M.K., Hedman, L., Enochsson, L., Kjellin, A., Felländer-Tsai, L.: Systematic video game training in surgical novices improves performance in virtual reality endoscopic surgical simulators: a prospective randomized study. *World journal of surgery.* 33, 2360–2367 (2009).

8. Bahsoun, A., Malik, M., Elhage, O., Dasgupta, P.: Do video games increase aptitude on surgical simulators? *International Journal of Surgery*. 9, 568 (2011).
9. Lynch, J., Aughwane, P., Hammond, T.M.: Video games and surgical ability: a literature review. *Journal of Surgical Education*. 67, 184–189 (2010).
10. Keehner, M.M., Tendick, F., Meng, M.V., Anwar, H.P., Hegarty, M., Stoller, M.L., Duh, Q.Y.: Spatial ability, experience, and skill in laparoscopic surgery. *The American journal of surgery*. 188, 71–75 (2004).
11. Tendick, F., Downes, M., Goktekin, T., Cavusoglu, M.C., Feygin, D., Wu, X., Eyal, R., Hegarty, M., Way, L.W.: A virtual environment testbed for training laparoscopic surgical skills. *Presence: Teleoperators & Virtual Environments*. 9, 236–255 (2000).
12. Satava, R.M., Jones, S.B.: Current and future applications of virtual reality for medicine. *Proceedings of the IEEE*. 86, 484–489 (1998).
13. Garg, A., Norman, G.R., Spero, L., Maheshwari, P.: Do virtual computer models hinder anatomy learning?. *Academic Medicine*. 74, S87–9 (1999).
14. Garg, A.X., Norman, G., Spero, L.: How medical students learn spatial anatomy. *The Lancet*. 357, 363–364 (2001).
15. Sweller, J., van Merriënboer, J.G., Paas, F.W.C.: Cognitive Architecture and Instructional Design. *Educational Psychology Review*. 10, 251–296 (1998).
16. Levinson, A.J., Weaver, B., Garside, S., McGinn, H., Norman, G.R.: Virtual reality and brain anatomy: a randomised trial of e-learning instructional designs. *Medical education*. 41, 495–501 (2007).

17. Henn, J.S., Lemole Jr, G.M., Ferreira, M.A.T., Gonzalez, L.F., Schornak, M., Preul, M.C., Spetzler, R.F.: Interactive stereoscopic virtual reality: a new tool for neurosurgical education. *Journal of neurosurgery*. 96, 144–149 (2002).
18. Familiari, G., Relucenti, M., Heyn, R., Petruzzello, L., Familiari, P., D'Andrea, G., Mastronardi, L., Ferrante, L.: The use of Brain-suite teaching tools improves learning of neuroanatomy. *Italian Journal of Anatomy and Embryology*. 115, 65 (2010).
19. Estevez, M.E., Lindgren, K.A., Bergethon, P.R.: A novel three-dimensional tool for teaching human neuroanatomy. *Anatomical sciences education*. 3, 309–317 (2010).
20. Spitzer, V., Ackerman, M.J., Scherzinger, A.L., Whitlock, D.: The visible human male: a technical report. *Journal of the American Medical Informatics Association*. 3, 118–130 (1996).
21. Servos, P., Goodale, M.A., Jakobson, L.S.: The role of binocular vision in prehension: a kinematic analysis. *Vision Research*. 32, 1513–1521 (1992).
22. Zhai, S., Buxton, W., Milgram, P.: The partial-occlusion effect: Utilizing semitransparency in 3D human-computer interaction. *ACM Transactions on Computer-Human Interaction (TOCHI)*. 3, 254–284 (1996).
23. Cohen, C.A., Hegarty, M.: Sources of difficulty in imagining cross sections of 3D objects. *Proceedings of the Twenty-Ninth Annual Conference of the Cognitive Science Society*. pp. 179–184 (2007).
24. Miller, R.: Approaches to learning spatial relationships in gross anatomy: perspective from wider principles of learning. *Clinical Anatomy*. 13, 439–443 (2000).

25. MacKenzie, I.S.: Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*. 7, 91–139 (1992).
26. Brewer, D.N., Wilson, T.D., Eagleson, R., de Ribaupierre, S., others: Evaluation of neuroanatomical training using a 3D visual reality model. *MMVR*. pp. 85–91 (2012).
27. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*. 47, 381 (1954).
28. Shepard, R., Metzler, J.: Mental rotation of three-dimensional objects. *Science (New York, NY)*. 171, 701 (1971).

Chapter 3

3 Automating Simulator Studies

This chapter is adapted from ‘Automation of a Closed Source Surgical Simulator: Making User Studies Feasible’ submitted to the American Journal of Biomedical Engineering.

My contribution to this chapter involved (i) gathering system requirements, (ii) designing and implementing the software, (iii) testing and analyzing the system, and (iv) writing the manuscript.

3.1 Introduction

The advancement and adoption of surgical simulation in clinical and educational contexts offers a wealth of fertile research directions within the domain. As an important precursor to clinical deployment, surgical simulators must be validated to ensure that sufficient skill transfer occurs without the training of poor habits. The standard of validation is even higher when using patient-specific data to create surgical rehearsal scenarios; in training, a misrepresentation of a patient is still valuable if it portrays a realistic scenario, but a misrepresentation in rehearsal would be the equivalent of providing the surgeon with preoperative imaging from a separate clinical case. In addition to training and rehearsal, versatile simulator platforms can be adapted to examine performance using novel approaches and techniques, such as those that utilize image-guidance technology [1].

All of the aforementioned research applications require the recruitment and participation of subjects; the development and evaluation of performance metrics crucial to the validation of simulator technology fall within the domain of human performance psychology. This compounds the increasing complexity of simulation studies. Not only must the simulation platform be robust and perform consistently, but in order to run studies, the simulated tasks must progress at a reasonable pace and not require the user to wait for a minute or more between trials. This is especially true when recruiting subjects

from surgical specialties where only a small sample is available for participation and they are under considerable time constraints. Indeed, many studies in this field suffer from a low sample size [2].

When working with commercial surgical simulators and proprietary equipment, it can be difficult to create customized solutions where novel research questions can be investigated through efficient user studies. In this chapter, we present our approach for automating tasks around a closed source commercial surgical simulator using popular open source tools. We added functionality to simulate our desired tasks and automated the process that connects the planning, task performance, and the feedback stages. The automation not only makes user studies with a suitable number of tasks feasible, it also frees the supervising researcher from performing repetitive tasks during each session.

3.2 Background

Surgical simulators require thorough validation prior to clinical deployment; in general, they must demonstrate that they accurately model the surgical task in question, that the implemented metrics measure the desired construct (the specific surgical skill), and that they provide a skill transfer specific to the procedure [3]. While some components may be validated to objective criteria, such as a check list or expert interview for content validity [4], the examination of human performance on these systems is necessary for determining most measures of validity [3], necessitating user studies.

These studies are notoriously difficult to conduct as the target population of highly specialized surgeons for each simulated procedure is relatively small and the subjects are often under substantial time constraints [5][6]. As such, maximizing the efficiency of simulator studies is essential to gather sufficient data and is an important focus as this field grows; a Medline query of Pubmed containing the terms ‘surgical’, ‘simulation’, and ‘validation’ returned no results before the year 1990, but made up 4.27% of indexed publications in the last 5 years [7].

An extensive literature search yielded no relevant results for extending and automating surgical simulators with the intent of creating a research platform for use in

user studies. In addition, we could not find any articles detailing methodology to automate around existing closed source applications with the intent of running user studies using software techniques.

3.3 Methodology

3.3.1 Design Considerations

Our simulation needs were well-defined; we required a platform that allowed users to perform targeting tasks on a mannequin head following an initial planning stage using an image-guidance tool or a traditional slice-based environment of a patient-specific case. From the perspective of the user, the software acts as a black box. Integration with patient-specific imaging data to create novel simulation scenarios was required. The NeuroTouch surgical simulator was chosen as the base platform as it contains modules for numerous neurosurgical procedures using a Windows platform, a mechanically tracked haptic tool, and a mannequin head [8]. Specifically, the NeuroTouch contains a module for performing simple targeting on the mannequin head using the tool, which records the position on the head along with the orientation.

Unfortunately, the NeuroTouch offers only limited facilities for custom modules using novel anatomy and does not provide features for planning on imaging data. We required custom anatomical (and non-anatomical) meshes as virtual targets registered within the mannequin head. In addition, we required a planning stage that allowed users to explore volumetric images depicting the mannequin head and the internal target and a feedback stage that rendered the meshes along with the user's trajectory after each stage. Further, the NeuroTouch does not offer advanced automation and data management capabilities.

For each task, the user's trajectory and task time were required for data collection. This data is written to disk by the NeuroTouch after each simulation, but is overwritten following each simulation session. The trajectory is required in the coordinate space of the mannequin head, which the tool is registered to by default using the NeuroTouch. With numerous users completing numerous tasks for different studies, we required a

comprehensive data management system to track study, task, and user information. There was the additional requirement that the data is parsable across various languages and platforms to allow for extensibility and analysis.

Finally, efficiency between tasks on the simulator was prioritized. Initial implementations of our system required user intervention between tasks in order to record data (written to a text file on disk) and advance to the next task, which was time consuming and precluded running a large number of scenarios. Minimal between-task time maximizes the number of trials that can be run within a given time slot.

3.3.2 Hardware

The hardware used primarily belongs to the NeuroTouch system; the system is equipped with a tabletop simulation bench holding a mannequin head and a Phantom Omni Haptic tool with a foot pedal below the bench. The computer used is an Alienware M17 laptop with dual GTX 470M graphics cards. The hardware requirements will vary depending on the simulation module. Additional hardware includes any tools used for image-guidance, but these can run as standalone systems without interfering with the automation. For our initial implementation, we utilized a Samsung Galaxy Tab S tablet and an image-based marker affixed to an eye glass frame on the mannequin for augmented reality tracking. The original implementation of the image-guidance tool is discussed in Kramers, et al. [9] and expanded on in the next chapter.

3.3.3 Data Management

The data management system keeps track of the details of each study as well as the data obtained from subjects. XML was chosen for database management for its ease of use, scalability, data access speed and reliability [10]. In addition to storing subject data, study parameters are defined in a master XML document, allowing ad hoc modifications to study protocols and simulation cases. These documents are human-readable, allowing users to easily make changes and additions beyond what is provided by the interface at the time. XML also offers benefits over traditional relational databases for storage of clinical data and has been implemented in a number of clinical management settings [11].

All text data is stored in XML format in a single document according to a schema that structures data hierarchically. When creating a new study, a GUI (see Figure 18) guides researchers through entering the parameters for the study. These parameters are described in

Table 1: Study parameters encapsulated within XML. While the table is incomplete, it offers insight into the general organization of XML data for tracking the progress of an ongoing study.

Study Parameter	XML Implementation	Example
Study Name	Single String Element	<studyname>Example Study</...>
Start Date	Single String Element	<startdate>02/04/2016</...>
Status of Ethical Approval	Single Boolean Element	<ethicsapproval>>true</...>
Current Subjects	Single Decimal Element	<currentsubject>12</...>
Expected Subjects	Single Decimal Element	<expectedsubjects>22</...>
Type of Simulation	Single String Element	<simulationtype>neuro</...>
Study Groups	Parent Element to Study Stratum Elements	<studygroups><stratum id="1">...</...>
Study Stratum	Container Element with ID	<stratum id="1"><order>...<users>...</...>
Stratum Order	Single String Element	<order>1;3;5;7;9;2;4;6;8</...>
Stratum Users	Single String Element	<users>1,3,5,7</...>
Subject Results	Container Element with ID	<results userid="4"><metric1>...</...>
Subject Psychometrics	Container Element with ID	<pmetrics userid="4"><metric1>...</...>
Analysis Results	Container Element with ID	<analysis userid="4"><result1>...</...>

The schema is extensible and additional elements need only be accounted for in the parsers. A sample of data analysis stored in the XML is shown in Figure 19.

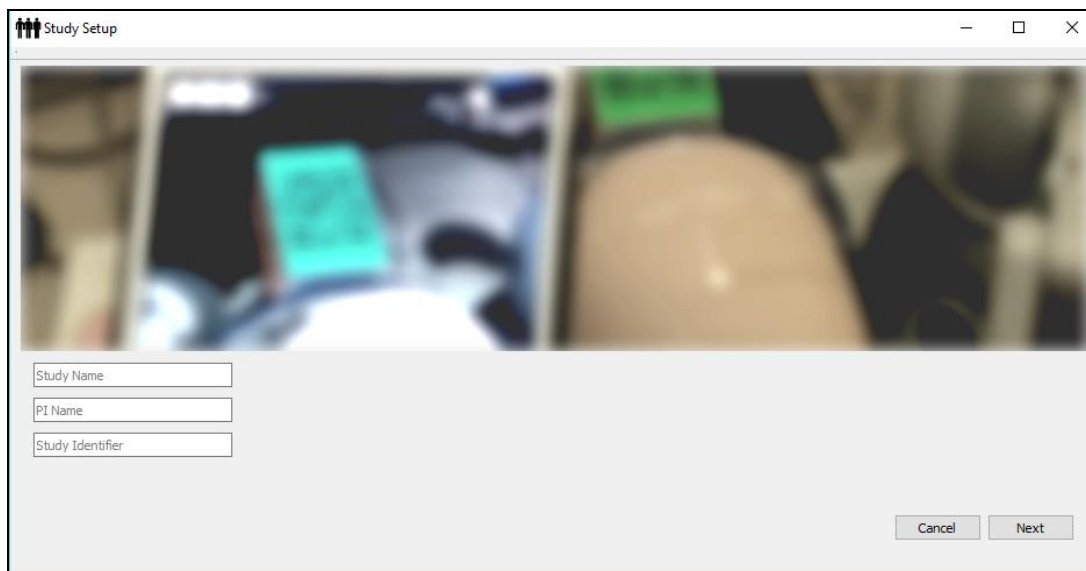


Figure 18: A screen capture of the study setup GUI. The wizard provides a number of facilities to customize a study

```

<analysis>
  <engagement category="AR">2, 3, 4, 5, 2</engagement>
  <engagement category="slice-based">2, 3, 4, 5, 2</engagement>
  <relativecoronalangulation category="AR">1.64, 2.32, 2.2, 2.1, 4.1</relativecoronalangulation>
  <relativecoronalangulation category="slice-based">1.64, 2.32, 2.2, 2.1, 4.1</relativecoronalangulation>
  <relativesagittalangulation category="AR">1.64, 2.32, 2.2, 2.1, 4.1</relativesagittalangulation>
  <relativesagittalangulation category="slice-based">1.64, 2.32, 2.2, 2.1, 4.1</relativesagittalangulation>
  <errormargin category="AR">1.64, 2.32, 2.2, 2.1, 4.1</errormargin>
  <errormargin category="slice-based">1.64, 2.32, 2.2, 2.1, 4.1</errormargin>
  <distancefromforamen category="AR">1.64, 2.32, 2.2, 2.1, 4.1</distancefromforamen>
  <distancefromforamen category="slice-based">1.64, 2.32, 2.2, 2.1, 4.1</distancefromforamen>
  <expertscore category="AR">2, 3, 6, 5, 1</expertscore>
  <expertscore category="slice-based">2, 3, 6, 5, 1</expertscore>
</analysis>

```

Figure 19: A sample of data stored in the XML document. Each tag specifies a metric for scoring the specific simulation module of the study. The category attribute is used to distinguish between the preoperative planning approaches. Additional attributes can be added to specify parameters specific to each study.

When a user begins participation in a study, the study XML is loaded through the GUI to allow manual input of user data (such as psychometrics or user experience with a particular procedure). When performing tasks on the simulator, the data is automatically recorded to the XML file following each scenario. In our experience, the XML document is suitable for every step of the study with the exception of spreadsheet-style analysis.

These types of analysis programs often include facilities for parsing XML directly into spreadsheets (assuming the data is not massively hierarchical).

3.3.4 Study Automation

Due to the closed source nature of the NeuroTouch, additional simulator functionality was added using external programs. Slicer 3D was used for the planning stage of each scenario (each scenario was encapsulated in a unique Slicer MRML scene), allowing users to slice through volumetric images pertaining to each scenario they are tasked with completing. If an image-guidance system is used, this complements or replaces the Slicer image exploration. Following planning, the task is performed using the haptic tool of the NeuroTouch. In our study of ventriculostomy approaches, the tool is used to select a location on the mannequin and a trajectory into the interior anatomy. Finally, the feedback of the user's task (in this case the trajectory into the mannequin head) is rendered using blender and displayed from the desired viewpoints as images. A Blender python script handles the data reading from the NeuroTouch and reads a log file from the automation to determine the current simulation scenario. The flow of the process is modeled in Figure 20.

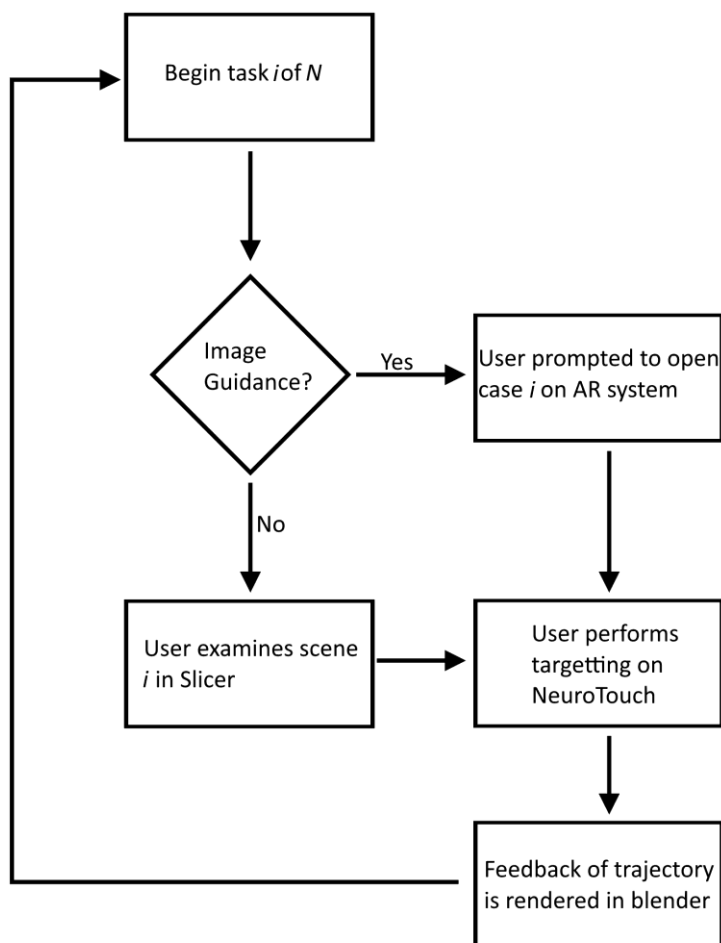


Figure 20: A flow chart of typical user interaction for a study session.

Without automation, each simulation scenario requires the researcher to do the following: open the appropriate Slicer case, open the NeuroTouch module, copy the data saved to text by the NeuroTouch following completion of the task, open the appropriate Blender case, and render and open the feedback images. To improve efficiency, these tasks were automated with AutoHotKey (version 1.1.20.01) – an operating system level scripting and automation tool. While there are alternatives to AutoHotKey, the program was chosen for its ease of use. Drawbacks to implementation include volatility in mouse operations and placement, but these features were avoided in this project. To the authors' knowledge, AutoHotKey has not been utilized in any simulator or psychological study, although it has been implemented in an application to improve radiologic reading

efficiency [12]. The tool uses its own scripting language. The script reads the study parameters directly from the master XML file. For each scenario specified in the chosen order in the XML, the script loads the correct Slicer scene and NeuroTouch module, saves the output data from the NeuroTouch directly into the XML file, and then uses the Blender python script to render and open the images for visual feedback. The only intervention required by the researchers or the users is to press the spacebar one time at the end of each scenario to advance to the next one. Sample code of the automation script is shown in Figure 5.

```

94 WinClose 3D Slicer
95 IfWinExist, 3D Slicer
96 {
97     WinKill 3D Slicer
98 }
99 WinActivate Blender
100 Sleep 100
101 WinMaximize Blender
102 Sleep 200
103
104 Click 650,1035
105 WinMinimize Blender
106 Sleep 3000
107
108
109 temp := caseCount - 1
110 Run, open %renderDir%%temp%\Camera1.png
111 WinWaitActive, Camera1
112 Send #{Left}
113 Run, open %renderDir%%temp%\Camera2.png
114 WinWaitActive, Camera2
115 Send #{Right}
116 state = 1

```

Figure 21: A code excerpt from the AutoHotKey script. This excerpt closes the open Slicer scene, activates the open Blender window, runs the script, and then opens the two rendered images to display the feedback of the user's targeting.

Figure 22 illustrates user interaction with the automation system as a black box.

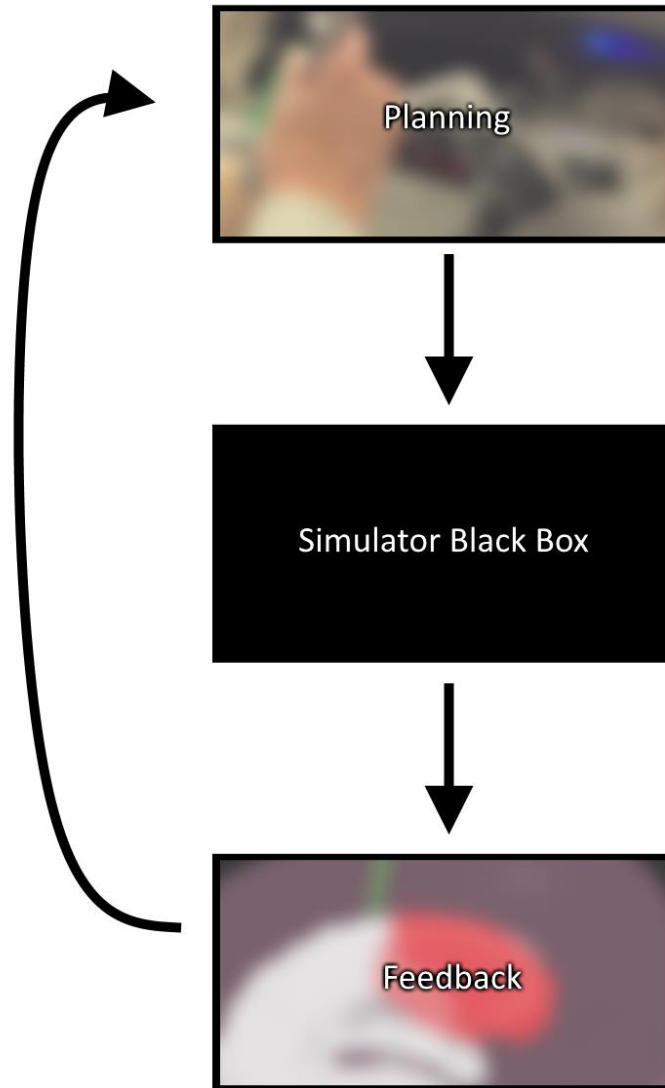


Figure 22: User interaction depicting automation as a black box.

3.3.5 Security and Privacy

Considerations for security and anonymity are essential for research platforms that utilize patient data and information from research subjects. As all data remains on the research computer from its conception to analysis, data is stored locally in an encrypted partition. In addition, users require credentials to log in to the operating system. These considerations alleviate many security concerns by delegating to the underlying operating system. Network connectivity is left to the experimenter's discretion, but is not required

for operation, alleviating the possibility of a network attack. Functionality for subjects performing repeat sessions is not implemented, as all user information is anonymized on entry. However, a user identification value could easily be added to the schema to allow for user tracking for long term studies.

3.4 Results

The system has been successfully deployed for multiple user studies. In a sample of 30 runs of between-user task time for the automation system compared to the same sample from manual operation, there was a significant reduction in time from $32.44 \pm 5.12s$ to $6.69 \pm 1.27s$ using a two-tailed T-test ($p < 0.05$). A Kolmogorov-Smirnov test did not find the data distinguishable from a normal distribution ($p > 0.15$). If a user's average task time is 30 seconds in duration (common for simple part-task simulations), 57 tasks could be completed in 30 minutes manually compared to 98 tasks with automation.

This is essential for gathering sufficient data from populations with limited numbers and availability, although application of the system to larger studies would save a greater amount of time. The system is largely resistant to user error as simulation participants are exposed only to a limited user interface.

In addition to more efficient user trials, the runtime organization of data into the XML hierarchy allows for seamless parsing into various analysis packages following study completion. Most spreadsheet-based software packages contain functionality for automatic import of XML formatted documents.

A weakness of the implementation is its adaptability to novel environments. While the data management system is trivially extensible, the automation environment must be modified for any changes to the simulation environment that cannot be made within the study setup in the XML document.

3.5 Discussion

There were two primary lessons learned throughout the development and deployment process. The first is that user error will occur where possible. This is not a particularly novel observation, but initial versions of our software relied on multiple ‘hotkeys’ and as a result, keys were used in the incorrect order. Two remedies were implemented; the spacebar was the only key enabled for users which allowed progression of scenarios and a mechanism was implemented to save the current system state for reloading if an error occurred. The other lesson learned was in the handling of data. While maximal automation is ideal, there always seemed to be a specific need to make manual modifications. This helped inform the choice of XML as the data document standard; the documents are human-readable and easy to modify by non-developers. Complementary XML schemas can then be used to validate whether the user’s changes maintain internal validity and adhere to the format.

Although the NeuroTouch is a proprietary system, the automation and data management system presented in this chapter can be adapted to other platforms with minor (major in some cases) modifications. To allow for adoption and extension, all technical work is archived on the following public GitHub repository:
<https://github.com/ryarmst/NeuroTouch-Automation>.

The intent of future development is to extend the system to allow for easy modification and adaptation to novel study designs and simulators. The initial setup wizard will be developed to allow for a wide array of possible study designs to be represented within the XML document. The automation script will read the XML and change the behavior based on these parameters. Researchers will be able to specify their own programs and which commands must be run at each step of the study without the need to make scripting modifications.

3.6 Conclusion

This chapter presented the conceptualization, design, development, and evaluation of an automation and data management system for surgical simulation research. The system

improves the workflow of simulator studies, enabling feasibility when working within time constraints. This application lays the foundation for work in the following chapters.

Bibliography

1. Ungi, T., Yeo, C.T., Paweena, U., McGraw, R.C., Fichtinger, G., others: Augmented reality needle guidance improves facet joint injection training. *SPIE Medical Imaging*. p. 79642E–79642E (2011).
2. Haque, S., Srinivasan, S.: A meta-analysis of the training effectiveness of virtual reality surgical simulators. *Information Technology in Biomedicine, IEEE Transactions on*. 10, 51–58 (2006).
3. Schout, B., Hendriks, A., Scheele, F., Bemelmans, B., Scherpbier, A.: Validation and implementation of surgical simulators: a critical review of present, past, and future. *Surgical endoscopy*. 24, 536–546 (2010).
4. Schijven, M., Jakimowicz, J.: Construct validity. *Surgical Endoscopy and Other Interventional Techniques*. 17, 803–810 (2003).
5. Haase, J., Boisen, E., others: Neurosurgical training: more hours needed or a new learning culture? *Surgical neurology*. 72, 89 (2009).
6. Woodrow, S.I., Segouin, C., Armbruster, J., Hamstra, S.J., Hodges, B.: Duty hours reforms in the United States, France, and Canada: is it time to refocus our attention on education? *Academic Medicine*. 81, 1045–1051 (2006).
7. Corlan, A.D.: Medline trend: automated yearly statistics of PubMed results for any query, URL:<http://dan.corlan.net/medline-trend.html>.
8. Delorme, S., Laroche, D., DiRaddo, R., Del Maestro, R.F.: NeuroTouch: A Physics-Based Virtual Simulator for Cranial Microneurosurgery Training. *Neurosurgery*. 71, ons32–ons42 (2012).

9. Kramers, M., Armstrong, R., Bakhshmand, S.M., Fenster, A., de Ribaupierre, S., Eagleson, R.: A Mobile Augmented Reality Application for Image Guidance of Neurosurgical Interventions. *American Journal of Biomedical Engineering*. 3, 169–174 (2013).
10. Vakali, A., Catania, B., Maddalena, A.: XML data stores: emerging practices. *IEEE Internet Computing*. 9, 62–69 (2005).
11. Lee, K.K.-Y., Tang, W.-C., Choi, K.-S.: Alternatives to relational database: comparison of NoSQL and XML approaches for clinical data storage. *Computer methods and programs in biomedicine*. 110, 99–109 (2013).
12. Lee, Y.H.: Efficient radiologic reading environment by using an open-source macro program as connection software. *European journal of radiology*. 81, 100–103 (2012).

Chapter 4

4 Examining Human Factors in Augmented Reality Guidance

This chapter is adapted from ‘Augmented reality for Neurosurgical Guidance: An Objective Comparison of Planning Interface Modalities’ [1], ‘Evaluation of a Mobile Augmented Reality Application for Image-guidance of Neurosurgical Interventions’ [2], and ‘A Mobile Augmented Reality Application for Image-guidance of Neurosurgical Interventions’ [3].

My contribution to this chapter involved (i) developing the AR guidance system, (ii) designing and conducting experiments, (iii) analyzing the data, and (iv) writing the manuscript.

4.1 Introduction

Augmented reality (AR) technologies are gaining traction in the medical domain as tools for image-guidance [4-8], preoperative planning [9-13]. Although translation of these technologies from the research stage to clinical settings has been slow, numerous technologies have been demonstrated to be effective for improving surgical proficiency compared to traditional techniques [4,6,7,13]. While these platform-specific validation studies are providing converging evidence of clinical worth, there is a need for further evaluation of the specific ways that these augmented perceptual cues affect performance in terms of speed and accuracy.

In this chapter, we describe the evaluation of a mobile AR interface in comparison to a traditional preoperative planning approach using 2D views sampled from orthogonal slices from a volumetric image modality. This chapter will examine low-level targeting performance as a complement to high-level metrics involving patient outcomes and procedure-specific tasks. To examine these characteristics in a controlled environment, we specified a number of abstract tasks involving the targeting of ellipsoids of various sizes and shapes, displayed using a commercial neurosurgical simulator. The tasks model ventriculostomies - common neurosurgical procedures that require precise targeting of

subsurface anatomy that is occluded to the surgeon. By providing a non-anatomical context, medical knowledge was not a prerequisite for performance of the tasks. As such, the tasks were entirely visuospatial in nature.

A total of 18 subjects were recruited to participate in the user study. Each user completed 30 tasks with the AR and 30 tasks with a traditional slice-based interface. Numerous measures of accuracy were derived in order to fully characterize user performance in both scenarios.

4.2 Background

4.2.1 Augmented Reality Guidance

Tracking of an object's position is the main aspect which influences accuracy of a system and determines the level of interference with the medical workflow. The majority of AR systems take advantage of Head Mounted Displays (HMD), hand-held or fixed displays to show computer generated scenes to the user. Visualization techniques are used to incorporate preoperative medical images during the intervention as 3D objects rendered in real-time. Considering these three methods of AR, we can classify some recent work and indicate their advantages and deficiencies. As an early prototype, in 1968 Sutherland et al. [14] developed a mechanical tracking system for their HMD 3Ddisplay. It was realized by attaching a mechanical linkage to the HMD which measured head position by computing axial displacement of the joints of a passive robotic arm. In similar work in 1992, Bajura et al. [15] replaced the mechanical linkage with electromagnetic sensors to determine the pose of a HMD and an ultrasound probe. Measuring position remotely (by magnetic or optical tracking systems) leads to a significant improvement in terms usability of the system, since they give more freedom to move the HMD within the operation site. Shamir et al. [7] used magnetic trackers and point based registration to align images, risk surfaces and segmented models to physical head models. In further work, Shamir et al. utilized the ability to track multiple objects simultaneously using optical tracking to develop an AR probe that incorporated a camera attached to a reference plate [16]. The output of the system was an augmented video image of the therapeutic site with relevant superimposed graphical content rendered according to the

position of the probe. HMDs interfere with medical work flow and may restrict a surgeon's natural movement. As a result, hand held displays and cameras (AR probe as an example) are becoming more feasible within the operating room. DEX-ray [17] is a miniaturized version of a hand held probe with an integrated video camera. Naturally, in [16] and [17], displays are fixed at some point in operation room and a surgeon is required to switch between the real scene and the displayed AR scene, and thus increases the system's complexity. Mischkowski et al. found that camera and display units could be combined, as demonstrated by their X-Scope [18]. X-Scope could be used for detection of bony segments in real-time and results were displayed on a hand held LCD. This configuration resembles current mobile devices. Infrared optical tracking became feasible by attaching reflective frames to portable display devices. These mobile devices enable surgeons to inspect patients from different points of view. In contrast, optical trackers impose limitations on this procedure, due to their limited workspace, necessity of attaching multiple reflectors to objects of interest and line of sight issues. An alternative method for tracking utilizes image-based tracking algorithms, which require adequate speed and accuracy. Fisher et al. developed a hybrid tracking scheme to calculate final estimation of the pose in an AR framework for a neurosurgical application [19]. Two streams of sensory data (Infrared and vision based tracking results) are combined by a pose estimation algorithm based on RANSAC (RANDOM SAMPLE CONSENSUS) which is an iterative parameter estimation algorithm. Simulations of this work were performed on an artificially textured cube as a reference model. For an overview of image techniques and technologies, see Peters 2006 [20].

4.2.2 Evaluation of Image-guidance

Evaluation of the accuracy of image-guidance technologies for targeting tasks can be decomposed into two separate problems: evaluation of system accuracy and evaluation of user performance. System accuracy can be evaluated with the use of ground truth models, such as phantoms or known targets that stay fixed within both the physical and virtual spaces. This is largely an evaluation of registration accuracy between the virtual space and the physical space. User performance relies on the registration accuracy, but also takes into account perceptual discrepancies. In this chapter, we examine both facets of

guidance accuracy using our mobile AR guidance tool. Note that this discussion does not encompass additional factors that influence the overall suitability of an image-guidance modality such as the usability or ease of clinical integration.

4.3 Methodology

4.3.1 Augmented Reality Guidance System

The guidance system was constructed using an Android-based mobile tablet and a pair of modified safety glasses affixed to a cuboid frame for image-based marker adhesion. The software was developed using the Unity game development framework making use of the Vuforia™ toolkit (developed by PTC®) to provide image-based marker tracking required for aligning the virtual and physical scenes. Unique image-based markers were applied to each surface of the affixed cuboid, allowing for disambiguation of all possible viewing angles. While Vuforia™ is a closed-source augmented reality toolkit, it offers tools to optimize the robustness of image-based markers for use in tracking, acting as a black box. Generally, images with a large number of asymmetric, high contrast points and corners are preferred. What the toolkit does not account for is the distance at which the marker must be tracked. For example, a marker image can be optimized for a large number of trackable points, but it would perform poorly at a long distance where those points cannot be resolved by the device camera. Our implementation made use of Vuforia's™ cuboid primitive to facilitate tracking for a wide workspace. The image markers were then optimized for the application through an iterative trial-and-error approach. The final tracker setup can be seen in Figure 23.

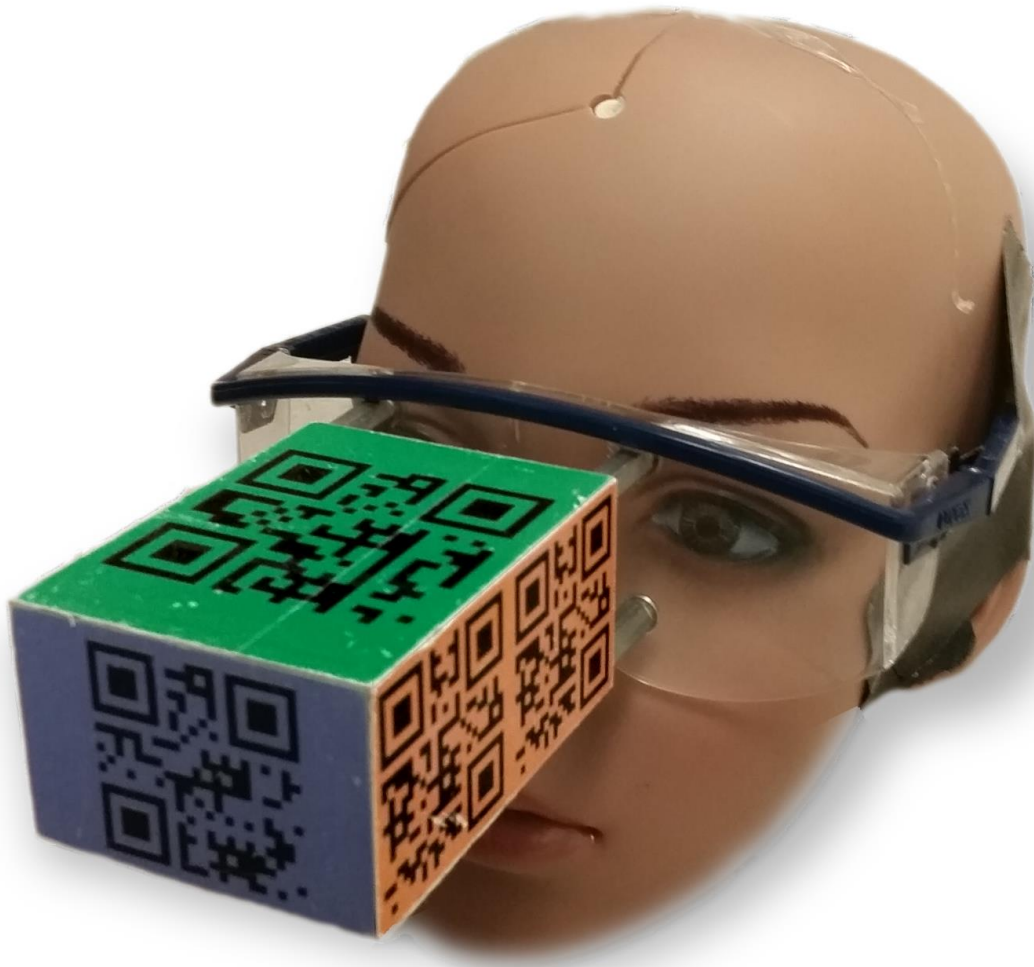


Figure 23: The image-based cuboid marker has 4 trackable sides, allowing for a flexible workspace. As our platform's mannequin head is not fully supine, the apparatus is affixed to the mannequin head.

We used a consumer-grade tablet computer as it provided a mobile and cost-effective all-in-one platform for computation, tracking, and display. We deployed the application for a Samsung Galaxy Tab S™, but the cross-platform nature allows for deployment onto other Android™ and iOS™ devices. Tracking utilizes the tablet's built-in camera to track the image-based markers.

4.3.2 Evaluation of AR Guidance System

To evaluate the accuracy of the system, we performed a study to quantify accuracy of the system applied to an environment of similar scale to a clinical implementation. Accuracy was assessed by having users target known reference points on a mannequin head in 3D space. These points are spaced equidistant from Kocher's point – an anatomical landmark that is often the chosen burr-hole location for a number of ventriculostomy procedures. The image-based tracker was fixed onto the mannequin head and the AR scene on the mobile device was setup to overlay spheres indicating the location of the target reference points. As the points are visible on the mannequin head as recesses, a pen was used to mark the target points on writable film covering the mannequin head, which occluded these points from the user's perspective. The points could then be measured against the ground truth (surface distance) using a vernier caliper to determine the accuracy. Since only accuracy was being evaluated, no constraints on time were imposed. In total, 5 subjects were asked to perform 8 targeting tasks on each of the 4 reference points (each target was selected twice for a task). The mean accuracy and the variance are of interest to quantify the performance characteristics. Figure 24 provides a reference for the evaluation system.

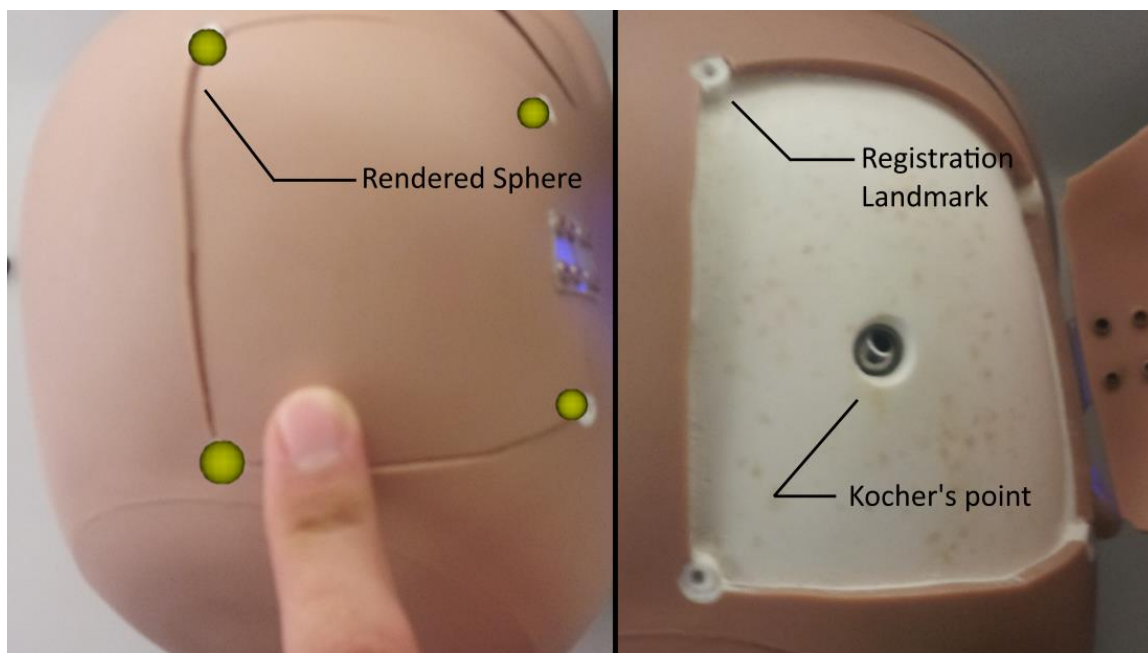


Figure 24: Left: The AR system overlays the rendered spheres onto the target points. In this instance, the mannequin head has not yet been covered for subject trials, allowing edges of the recesses to be seen. Right: While not used in this study, the mannequin head possesses a ball joint at the location of Kocher's point for use in additional simulation scenarios.

4.3.3 Simulation Platform

An existing commercial surgical simulation platform was used as the foundation for simulated tasks. The NeuroTouch is a neurosurgical simulator that offers a number of modules for various procedures [21]. The simulator consists of a computer platform with a physical mannequin head and a haptic device that offers mechanical tracking and mimics the feeling of the tool colliding with tissue. While other comprehensive neurosurgical simulators exist – namely the ImmersiveTouch [22] – the physical mannequin head is necessitated to provide the appropriate test-bed for our augmented reality interface. For the study, the image marker glasses were affixed to the mannequin. The physical and virtual spaces were registered using a graphical mesh of the mannequin head to align the two scenes identically using each ear helix and the nasion as anatomical landmarks. To provide visual feedback for registration accuracy as well as an additional

depth cue, the graphical mannequin head mesh was overlaid on the physical head in the renderings with adjustable transparency.

Our user task is constructed around the NeuroTouch's burr-hole selection task, which is typically used to test a surgeon's ability to localize a catheter tip within the ventricles for a range of ventriculostomy procedures. Ventriculostomies generally require a precise trajectory into the anterior horn of the lateral ventricles. Rather than target virtual anatomical features within the mannequin head, we constructed our module with a number of virtual ellipsoid targets of various shape, size and position within the mannequin head. This approach models the perceptual challenge of traditional ventriculostomies; surgeons must target a relatively small anatomical feature without sight of the feature inside the head and along a specific trajectory to avoid damaging eloquent tissue. Ellipsoids are well suited for 3D targeting tasks as they are a basic geometrical shape that can be used to specify a target which has a location as well as an orientation. The long axis allows us to specify the unique trajectory vector. Where the orientation of a sphere is indistinguishable, an ellipsoid's orientation is uniquely described by the vector of its longest axis (assuming that a longest axis exists on a given ellipsoid). As such, our targeting tasks were created in such a way as to permit a single ideal trajectory in each case. In total, 30 ellipsoids were created and positioned in the head in a pseudorandom approach. One of the limitations on ellipsoid position and orientation was imposed by the mechanical haptic device that acts as the input for trajectory selection; since the tool has a constrained workspace, all the ellipsoids were positioned in a way that they can be targeted by the device.

As tool and mannequin position and registration can drift between sessions, registration was performed at the beginning of each study session. Registration is performed using the NeuroTouch's built-in module which involves the sampling of tip points at 4 discrete landmarks on the mannequin head. The calculated registration error varied from session to session, but never exceeded a mean of 2mm translational registration error (TRE). This registration error applies to the registration between the tool and the physical mannequin head. Results from a registration task as visualized by the system are depicted in Figure 25.

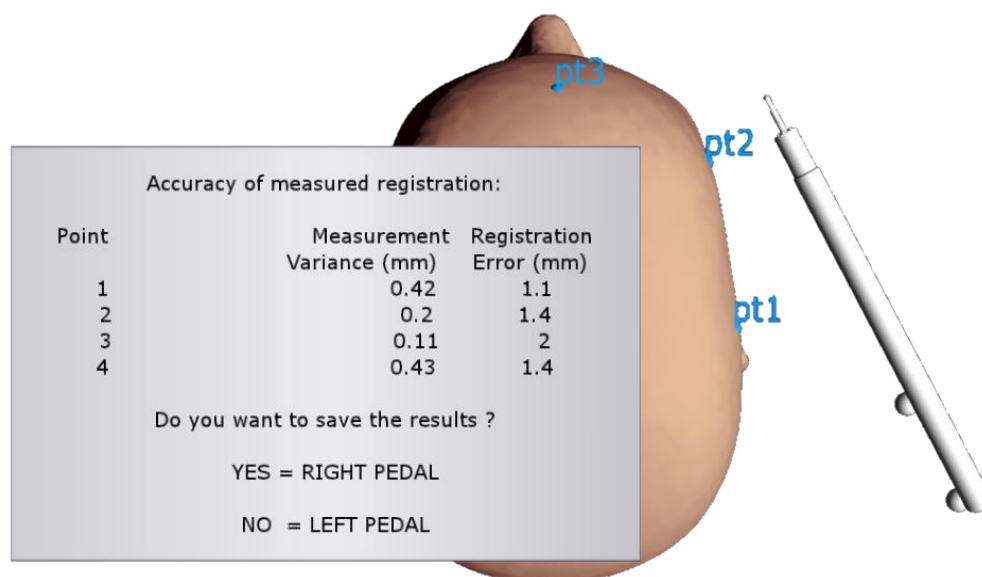


Figure 25: Results of NeuroTouch tool-mannequin registration process. Each point in the process corresponds to a physical recess on the mannequin head.

Each task involved planning, targeting and feedback. The planning stage utilized either the augmented reality guidance tool or a slice-based interface implemented in 3D Slicer. The augmented reality tool allowed the user to view the ellipsoids rendered directly on the mannequin head. The tablet could be held directly or mounted to a deformable spine to allow for hands-free operation. A stereoscopic display was not employed, so depth cues were largely gathered using cues from movement. Movement was constrained by the need to keep the image-based marker in the field of view of the device, but this still permits a rather large range of configurations in the approach. The alternative planning approach made use of a slice-based interface. Using this interface, users were able to view axial slices showing the intersection of the mannequin head and the current ellipsoid with each plane, which users could explore freely. Following the planning stage, users performed targeting. Targeting involved placing the tip of the mechanically tracked tool onto the mannequin head and indicating the chosen trajectory with its orientation. All targeting approaches originated from the right side as is standard for ventriculostomy procedures. Though the mechanically tracked arm limited tool range of motion, all

ellipsoids were oriented to provide an accessible targeting path. A foot pedal was used for final selection of the trajectory. Though a haptic device was used, no haptic feedback was given as our task is a visuospatial abstraction of surgical procedures simulated by the platform. Finally, feedback was given after each task. The open-source 3D modelling software Blender was used for rendering feedback. In the feedback stage, the user's trajectory was rendered alongside a mesh of the mannequin head and views from 2 camera sources were shown. All components of the module were implemented on the NeuroTouch computer and automation between stages was achieved using AutoHotKey, which is an operating system level scripting program to automate tasks. See Chapter 3 for a thorough discussion of simulator automation. Figure 26 displays screenshots of the planning and feedback stages.

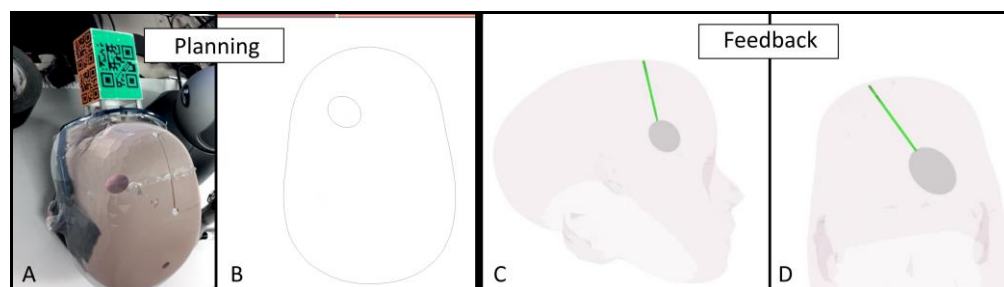


Figure 26: A) Screenshot taken from the augmented reality tool during the planning stage. B) Screenshot taken from Slicer 3D during the slice-based planning stage. C) Screenshot from feedback stage showing right sagittal view. The user's trajectory is rendered in green. D) Screenshot from feedback stage showing anterior coronal view.

4.3.4 Experimental Design

The user study involved the recruitment of 18 subjects with no experience in the medical domain beyond basic anatomical knowledge. There was an even gender split and a single subject was left hand dominant. Subjects were randomly assigned to 2 groups; the first group completed all 30 tasks with the slice-based interface and then completed 30 tasks with the AR system, and the second group completed the AR component first followed by the slice-based interface. All users were first allowed time to become accustomed to the haptic tool. For each task, the position on the mannequin, the direction of the

trajectory and the time of completion were all recorded. Video was recorded of the tool space in order to examine how users interact with the different approaches. Participation time never exceeded an hour – our subjects experienced minimal fatigue and maintained their enthusiastic motivation. Instructions were given in the form of a video tutorial and users were tasked to target the ellipsoids quickly and accurately. No additional instructions were given concerning specific techniques to approach each problem, permitting uninhibited learning and natural development of technique. Feedback was collected from each user following the tasks, but no formal questionnaire was administered. A user is shown using the AR interface with the system in Figure 27.

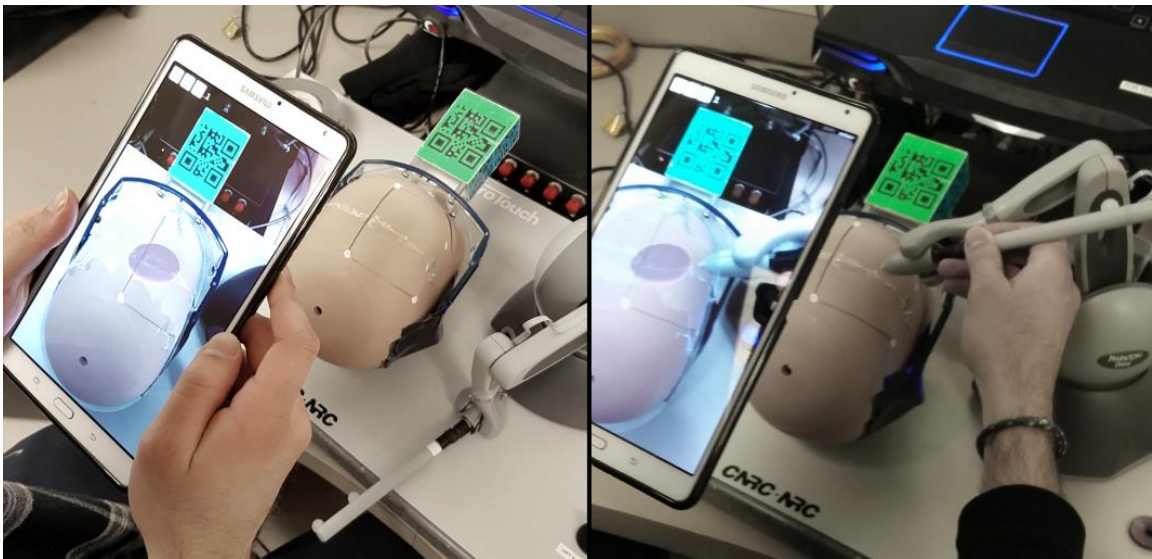


Figure 27: A user guides their trajectory using the AR interface.

4.3.5 Analysis

Analysis of user trajectories required development of objective metrics of performance. Bourdel et al. examined AR performance for a conceptually similar targeting task, but analysis did not cover broad metrics to fully characterize performance [23]. Similar studies have been performed by Abhari et al. [9,24]. In Abhari's examinations, rotational and translational errors were examined, but there was no attempt to qualify the type of error and examine specific biases. Rather than focus on overall compound performance, we decompose metrics to examine directional tendencies imparted by different approaches. We considered the time of the task, the angle of deviation from the longest

axis (rotational error) and the distance of the trajectory from the ellipsoid (translational error). The time and angle of deviation are relatively straightforward, but there are numerous candidate measures for the error between the correct trajectory to the ellipsoid and the user's trajectory. We explored various approaches. In the simplest case, we take the shortest distance between the trajectory line and the ellipsoid's centroid. This measure is valid for trajectories internal and external to the ellipsoid. In addition to the overall distance, we measured the distance in each Cartesian coordinate in order to examine targeting bias. The closest distance of the line to the surface was also considered, but this measure is only relevant when the trajectory does not enter the ellipsoid. The engagement is a measure that has been employed to evaluate ventriculostomy trajectories [25] and was adapted to ellipsoids in this study. The engagement is the length of the line segment that passes through the ellipsoid, which results in a value of 0 for trajectories that do not intersect the target. We adapted this measure as relative engagement, which takes the ratio of the engagement relative to the length of the target ellipsoid's longest axis. This allows for direct comparison of performance between ellipsoid cases. An additional relative measure is the relative distance from the centroid. To calculate this distance, we construct a line using the centroid and the closest point on the trajectory and we determine the distance from the centroid to the point on the ellipsoid that this new line intersects, which we term the inner distance. We calculate the relative distance by dividing the shortest distance from the centroid to the trajectory with this new value. A final measure was conceived using a piecewise function to penalize error in distance to a greater extent when trajectories are outside the target ellipsoid compared to those inside. The function is described by the following equation:

$$f(D) = \begin{cases} \frac{D_C}{D_I} & \text{Trajectory Intersects} \\ \ln(D_S + e) & \text{No Intersection} \end{cases} \quad (4)$$

In the equation, D_I is the inner distance, D_C is the distance from the center, and D_S is the distance from the surface. Inside the ellipsoid, a linear measure for error in position does not grow as largely as the logarithmic function outside. The function is continuous as a trajectory precisely on the surface of the ellipsoid will yield the same value on each side of the piecewise function. This function was derived to model the risk in clinical

environments with a particular focus on ventriculostomy procedures. A small misplacement of the catheter in a ventriculostomy will likely still result in a functioning drain as long as it enters a lateral ventricle [26]. A minor miss of the ventricle will result in a non-functional drain and necessitate a second pass, but may not result in additional injury [27]. However, a trajectory and placement of the catheter that deviates significantly from standard clinical approaches can lead to severe injury in the patient [28]. A number of measures for trajectories internal and external to the ellipsoids are depicted in Figure 28.

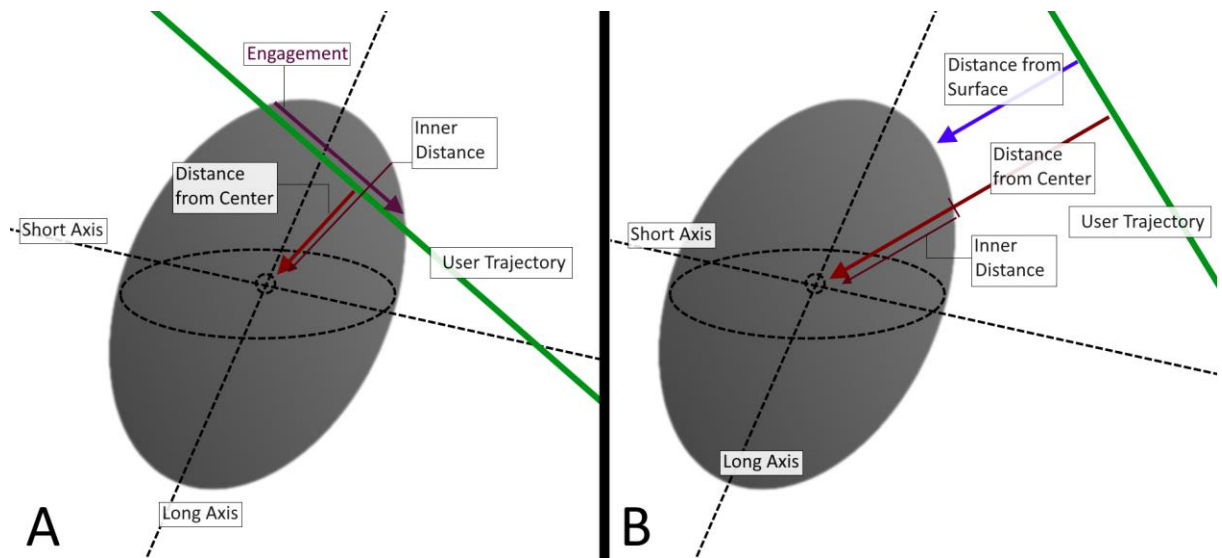


Figure 28: A) Measures for a trajectory that intersect the ellipsoid. B) Measures for targets that do not intersect the ellipsoid.

4.4 Results

4.4.1 AR System Accuracy

The mean targeting accuracy of users was $3.43 \pm 1.53\text{mm}$.

4.4.2 User Performance

Table 2 and

Table 3 compare various measures from all users using the AR interface compared to the slice-based interface. The X, Y, and Z, axes are based on medical planes; X is along sagittal slice planes, Y is along coronal slice planes, and Z is along axial slice planes.

For each comparison, sample data was examined for normality using a Kolmogorov-Smirnov (K-S) test to examine goodness of fit to a normal distribution. Where applicable, a two tailed t-test was used to compare means. When samples were considered non-normal ($p < 0.05$ for K-S test), a two tailed Mann-Whitney U test was applied to examine statistical significance.

Table 2: Mean values with standard deviations from all users for each modality for the percentage of intersecting trajectories, the deviation in the angle from the ellipsoid orientation, and the relative engagement. Two tailed t-tests are indicated by ‘*’ and Mann-Whitney U tests are indicated by ‘’.**

	Hit %	Angle Error (degrees)	Relative Engagement
Augmented Mean:	0.59 ± 0.072	20.84 ± 17.03	0.41 ± 0.38
Slice-based Mean:	0.35 ± 0.075	28.03 ± 20.56	0.21 ± 0.31
Significance:	Yes* ($p < 0.001$)	Yes** ($p < 0.01$ and U = 110845, N=540)	Yes** ($p < 0.01$, U=189.212, N=540)

Table 3: Mean values with standard deviations for measures examining translational error. Positive X is towards the right side of the mannequin head, positive Y is towards the anterior, and positive Z is towards the superior direction. Two tailed t-tests are indicated by ‘’ and Mann-Whitney U tests are indicated by ‘***’.**

	Translation Error (Eq. 4)	X Bias (mm)	Y Bias (mm)	Z Bias (mm)
Augmented Mean:	1.45 ± 1.18	-5.72 ± 11.54	-1.09 ± 14.14	11.69 ± 15.67
Slice-based Mean:	2.01 ± 1.18	-5.10 ± 12.57	-16.91 ± 16.53	9.40 ± 16.17
Significance:	Yes** (p < 0.001, U=105158, N=540)	No	Yes* (p < 0.05)	No

There was no significant difference in time taken between interfaces. Over the order of cases, a weak learning curve was observed for rotational error ($r = -0.36$ with 95% C.I. [0.25-0.46] for slice-based and $r = -0.23$ with 95% C.I. [0.11-0.34] for AR) and there was a weak-moderate reduction in the task completion time ($r = -0.46$ with 95% C.I. [0.36-0.55] for slice-based and $r = -0.59$ with 95% C.I. [0.51-0.66] for AR). As ellipsoid volume increased, there was a weak correlation with rotational error for AR targeting ($r = 0.37$ with 95% C.I. [0.26-0.47]), which was nonexistent for slice-based guidance. A moderate correlation was seen between ellipsoid volume and relative engagement for AR ($r = 0.60$ with 95% C.I. [0.52-0.67]) and slice-based ($r = 0.61$ with 95% C.I. [0.53-0.68]) targeting. A moderate correlation was also seen between volume and the distance measure presented in equation 1 for AR ($r = 0.57$ with 95% C.I. [0.48-0.65]) and slice-based ($r = 0.60$ with 95% C.I. [0.52-0.67]). The ellipsoid eccentricity (ratio of the longest

to the shortest axis) was not seen to affect any measure of performance. Finally, it was found that mean hit percentage for each user strongly correlates with their mean bias in the Z direction (superior to the mannequin) using the AR ($r = 0.96$, 95% C.I. [0.85 0.97]) and moderately with the slice-based interface ($r = 0.71$, 95% C.I. [0.20 0.79]), indicating that users with lower performance tend to underestimate depth, especially with AR.

4.5 Discussion and Future Work

While many clinical procedures demand sub-millimetre accuracy, there are some considerations to keep in mind. In this scenario, users had no reference information outside of the AR supplied rendering. In a clinical scenario, an AR guidance system such as this one could act as a rough estimate while allowing reliance on traditional anatomical cues to perform fine targeting. Additionally, precise localization of anatomical features is not the only perceptual advantage offered by AR guidance. The platform is also able to render anatomical features in the context of patient anatomy in an intuitive way; it renders 3D reconstructions of anatomy that convey spatial information. This spatial information may aid an operating surgeon in additional ways. Even without perfect accuracy and minute precision, the anatomy maintains its topological structure. However, the reliability of this information must take into account the form and quality of the preoperative imaging, the accuracy of the anatomical segmentations, and (in neurosurgical applications) the impact of brain shift and ventricle expansion.

By nearly all measures, AR guidance showed superior accuracy compared to the use of a traditional slice-based interface. No strong learning effects were seen, likely due to the nature of the experiment; it was relatively short and all tasks were performed in a single block session. Of particular interest was the examination of bias. Differences between approaches were seen in the Y and Z directions, which correspond to the posterior-anterior and inferior-superior planes in the mannequin, respectively. In the Z direction, large bias was seen with both approaches, but the bias is most prevalent among user with low hit ratios. Users who consistently miss the target using the AR guidance are nearly always underestimating the depth by choosing a trajectory that crosses the ellipsoid on the superior side. Difficulty estimating depth using augmented reality displays is a well-documented and investigated issue [28,29]. Similar to most

implementations, our device is hindered by the problem of occlusion and lighting discrepancies. Additionally, our approach uses monoscopic rendering, necessitating strategic use of the tool to obtain depth cues from motion - a use that may not be intuitive to all users. While this is a pervasive problem, numerous consumer products aim to remedy the shortcomings, such as the Microsoft HoloLens™ which will be released to market later this year and will be a target platform of our future work.

In the Y direction, little bias was seen using AR compared to a strong bias towards the anterior using the slice-based approach. This phenomenon warrants further investigation into the perceptual qualities of slice-based interfaces in the absence of anatomical structures to act as landmarks.

Although no formal survey was administered, user feedback indicated that the constraints imposed by the mechanical haptic device impeded tool movement.

Future work will involve improvements to the tracking and display of our AR system including targeting modern head-mounted platforms. Additionally, while our tracking system and registration is effective for our static mannequin, a more robust system is desired for translation to a clinical environment. An approach that tracks facial features is desirable, but these techniques have yet to prove robust. AR systems with as little as 2mm of registration error have been noted to cause drastic performance degradation [29]. A free-hand tracking system will be incorporated into future revisions.

Replication of this work can be performed using commodity technology. A mannequin skull could be printed from publicly available datasets and used as the target for any number of spatially tracked tools. Similar augmented reality applications can be developed using Vuforia and Android development tools.

4.6 Conclusion

In this chapter, we presented an approach to objectively evaluate targeting using various interface modalities. We found that our AR interface was more intuitive and resulted in more accurate targeting using numerous metrics, but caused users to underestimate target

depth. While the AR tracking used was suitable for the static mannequin, a more robust approach is desirable for clinical applications. We hope that our evaluation approach can be adapted to additional environments and AR platforms as a standard for objective analysis of user performance.

Bibliography

1. Armstrong, R., Wright, T., Eagleson, R., de Ribaupierre, S.: Augmented reality for Neurosurgical Guidance: An Objective Comparison of Planning Interface Modalities. Lecture Notes in Computer Science. 9805, (2016).
2. Kramers, M., Armstrong, R., Bakhshmand, S.M., Fenster, A., de Ribaupierre, S., Eagleson, R.: Evaluation of a mobile augmented reality application for image guidance of neurosurgical interventions. Stud Health Technol Inform. 196, 204–8 (2014).
3. Kramers, M., Armstrong, R., Bakhshmand, S.M., Fenster, A., de Ribaupierre, S., Eagleson, R.: A Mobile Augmented Reality Application for Image Guidance of Neurosurgical Interventions. American Journal of Biomedical Engineering. 3, 169–174 (2013).
4. Ungi, T., Yeo, C.T., Paweena, U., McGraw, R.C., Fichtinger, G., others: Augmented reality needle guidance improves facet joint injection training. SPIE Medical Imaging. p. 79642E–79642E (2011).
5. Sauer, F., Khamene, A., Bascle, B., Rubino, G.J.: A head-mounted display system for augmented reality image guidance: Towards clinical evaluation for imri-guided neurosurgery. Proceedings of the 4th International Conference on Medical Image

- Computing and Computer-Assisted Intervention. pp. 707–716 (2001).
6. Rosenthal, M., State, A., Lee, J., Hirota, G., Ackerman, J., Keller, K., Pisano, E.D., Jiroutek, M., Muller, K., Fuchs, H.: Augmented reality guidance for needle biopsies: an initial randomized, controlled trial in phantoms. *Medical image analysis*. 6, 313–320 (2002).
 7. Chu, M.W., Moore, J., Peters, T., Bainbridge, D., McCarty, D., Guiraudon, G.M., Wedlake, C., Lang, P., Rajchl, M., Currie, M.E., others: Augmented reality image guidance improves navigation for beating heart mitral valve repair. *Innovations: Technology and Techniques in Cardiothoracic and Vascular Surgery*. 7, 274–281 (2012).
 8. Krempien, R., Hoppe, H., Kahrs, L., Daeuber, S., Schorr, O., Eggers, G., Bischof, M., Munter, M.W., Debus, J., Harms, W.: Projector-based augmented reality for intuitive intraoperative guidance in image-guided 3d interstitial brachytherapy. *International Journal of Radiation Oncology* Biology* Physics*. 70, 944–952 (2008).
 9. Abhari, K., Baxter, J.S., Chen, E.S., Khan, A.R., Wedlake, C., Peters, T.M., de Ribaupierre, S., Eagleson, R.: Use of a Mixed-Reality System to Improve the Planning of Brain Tumour Resections: Preliminary Results. *Augmented Environments for Computer-Assisted Interventions*. pp. 55–66. Springer (2013).
 10. CHARBEL, F.: Second generation haptic ventriculostomy simulator using the ImmersiveTouch™ system. *Medicine Meets Virtual Reality 14: Accelerating Change in Healthcare: Next Medical Toolkit*. 119, 343 (2006).
 11. Leblanc, F., Champagne, B.J., Augestad, K.M., Neary, P.C., Senagore, A.J., Ellis, C.N., Delaney, C.P., Group, C.S.T.: A comparison of human cadaver and augmented reality simulator models for straight laparoscopic colorectal skills acquisition

- training. *Journal of the American College of Surgeons*. 211, 250–255 (2010).
12. Botden, S.M., Buzink, S.N., Schijven, M.P., Jakimowicz, J.J.: Augmented versus virtual reality laparoscopic simulation: what is the difference? *World journal of surgery*. 31, 764–772 (2007).
 13. Lahanas, V., Loukas, C., Smailis, N., Georgiou, E.: A novel augmented reality simulator for skills assessment in minimal invasive surgery. *Surgical endoscopy*. 29, 2224–2234 (2015).
 14. Sutherland, I.E.: A head-mounted three dimensional display. *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. pp. 757–764 (1968).
 15. Bajura, M., Fuchs, H., Ohbuchi, R.: Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *ACM SIGGRAPH Computer Graphics*. pp. 203–210 (1992).
 16. Shamir, R., Joskowicz, L., Shoshan, Y.: An augmented reality guidance probe and method for image-guided surgical navigation. *5th International Symposium on Robotics and Automation*. pp. 25–28 (2006).
 17. Kockro, R.A., Tsai, Y.T., Ng, I., Hwang, P., Zhu, C., Agusanto, K., Hong, L.X., Serra, L.: DEX-RAY: Augmented Reality Neurosurgical Navigation with a Handheld Video Probe. *Neurosurgery*. 65, 795–808 (2009).
 18. Mischkowski, R.A., Zinser, M.J., Kübler, A.C., Krug, B., Seifert, U., Zöller, J.E.: Application of an augmented reality tool for maxillary positioning in orthognathic surgery - a feasibility study. *J Craniomaxillofac Surg*. 34, 478–83 (2006).

19. Fischer, J., Eichler, M., Bartz, D., Straßer, W.: Model-based Hybrid Tracking for Medical Augmented Reality. EGVE. pp. 71–80 (2006).
20. Peters, T.M.: Image-guidance for surgical procedures. *Physics in medicine and biology*. 51, R505 (2006).
21. Delorme, S., Laroche, D., DiRaddo, R., Del Maestro, R.F.: NeuroTouch: A Physics-Based Virtual Simulator for Cranial Microneurosurgery Training. *Neurosurgery*. 71, ons32–ons42 (2012).
22. Luciano, C., Banerjee, P., Florea, L., Dawe, G.: Design of the ImmersiveTouch™: a high-performance haptic augmented virtual reality system. *Proceedings of the 11th international conference on human-computer interaction, Las Vegas, Nevada* (2005).
23. Bourdel, N., Collins, T., Pizarro, D., Bartoli, A., Da Ines, D., Perreira, B., Canis, M.: Augmented reality in gynecologic surgery: evaluation of potential benefits for myomectomy in an experimental uterine model. *Surgical endoscopy*. 1–6 (2016).
24. Abhari, K., Baxter, J.S., Chen, E.C., Khan, A.R., Peters, T.M., de Ribaupierre, S., Eagleson, R.: Training for planning tumour resection: augmented reality and human factors. *IEEE Transactions on Biomedical Engineering*. 62, 1466–1477 (2015).
25. Muirhead, W.R., Basu, S.: Trajectories for frontal external ventricular drain placement: virtual cannulation of adults with acute hydrocephalus. *British journal of neurosurgery*. 26, 710–716 (2012).
26. Kakarla, U.K., Chang, S.W., Theodore, N., Spetzler, R.F., Kim, L.J.: Safety and accuracy of bedside external ventricular drain placement. *Neurosurgery*. 63, ONS162–ONS167 (2008).

27. Hsieh, C.-T., Chen, G.-J., Ma, H.-I., Chang, C.-F., Cheng, C.-M., Su, Y.-H., Ju, D.-T., Hsia, C.-C., Chen, Y.-H., Wu, H.-Y., Liu, M.-Y.: The misplacement of external ventricular drain by freehand method in emergent neurosurgery. *Acta Neurol Belg.* 111, 22–8 (2011).
28. Berning, M., Kleinert, D., Riedel, T., Beigl, M.: A study of depth perception in hand-held augmented reality using autostereoscopic displays. *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on.* pp. 93–98 (2014).
29. Lee, C.-G., Oakley, I., Ryu, J.: Exploring the impact of visual-haptic registration accuracy in augmented reality. *Haptics: Perception, Devices, Mobility, and Communication.* pp. 85–90. Springer (2012).

Chapter 5

5 Augmented Reality for Ventriculostomy Guidance

This chapter is adapted from ‘An Examination of Factors that Impact Performance on a Patient-Specific VR-Based Ventriculostomy Simulator’ submitted to *Simulation* (Sage Journals).

My contribution to this chapter involved (i) study design, (ii) recruiting subjects and conducting the study, (iii) analyzing data, and (iv) writing the manuscript.

5.1 Introduction

While some professions have been using virtual reality for decades to train students [1], the majority of surgical skills continue to be trained through the Halstedian paradigm [2], where exposure to operative techniques is gained through an apprenticeship program where trainees learn by doing on live patients. However, providing a virtual environment for junior trainees to learn basic procedures is becoming recognized as essential to increase patient safety and allow for standardized and routine training [3]. The placement of an external ventricular drain (EVD) is a procedure commonly performed in an emergency setting by junior residents to relieve intracranial pressure from hydrocephalus [4]. The procedure is generally performed freehand and without image-guidance, resulting in a relatively high rate of malplacement [5].

We present our module for simulation of burr-hole and catheter trajectory selection tasks for EVD placement. A small number of Ventriculostomy simulators exist [6,7], providing a safe training environment where residents can be exposed to a variety of cases with unique anatomical features and receive immediate feedback. On our extended platform, we developed a testing platform for external ventricular drain (EVD) insertion. Additionally, we investigated the impact of patient-specific variation by identifying important factors that influence performance. Performance, in the strictest sense, was analyzed using an extension of Fitts’s methodology [8], taking into consideration both speed and accuracy of the task, as well as the natural trade-off

between the two measures. In this chapter, we derive extensions to these objective metrics which are gaining acceptance in the literature. Most notably, we consider the uncertainty of these metrics, and their variability metrics. We apply these metrics to characterize the user performance over a set of unique simulation scenarios derived from clinical imaging data. We hypothesize that these measures will be able to discern between expert and novice performance. In addition, we expect to see trends between the difficulty of cases and various anatomical measures. For example, ventricle volume is expected to impact performance; in models of human performance of simple targeting tasks, larger targets are invariably easier to target with all other factors held constant.

The study is divided into two phases. In the first phase, we use the simulation platform to evaluate an in-house developed mobile augmented-reality (AR) image-guidance system targeting deployment in an ICU setting (see Chapter 4 for an in-depth discussion). A Likert scale based measure for accuracy was developed relying on an expert surgeon's scoring of each user trajectory.

In the second phase, the analytical metrics are expanded upon. We adapt measures that allow for comprehensive and object analysis of trajectories for ventriculostomy catheterizations (as well as other general ventriculostomy procedures where the anterior horn must be targeted). AR is not examined, but subjects perform additional cases to further examine the role of anatomical variations in user performance.

5.2 Background

The insertion of an external ventricular drain involves choosing an appropriate burr-hole on the skull and guiding a catheter without neuro-navigation through the burr-hole and intermediate brain matter into a lateral ventricle in order to drain cerebrospinal fluid and relieve intracranial pressure. The placement of an EVD is often performed in the context of a traumatic brain injury; the technique is generally indicated for cases of acute hydrocephalus that require immediate care, whereas other ventriculostomy procedures, such as shunt placement, are intended for long-term management [4]. While some cases involve prototypical anatomy that facilitates optimal placement of the catheter, traumatic brain injuries may cause significantly deformed anatomical features [9]. Cases may

involve small or displaced ventricles, making accurate placement a difficult task. Frequently performed in emergency situations by the bed-side in the ICU [10], the use of state-of-the-art image-guidance systems is commonly precluded as a result of immobility of the patient and system and the preparation required. The resident (or practicing surgeon) must review orthogonal slices of the patient's brain acquired from pre-operative imaging (commonly axial CT-scans) to formulate a mental representation that would allow them to specify a trajectory to the ventricle relative to the skull of the patient and based on known anatomical landmarks (eyes, ears, etc.) and their estimation of the 3D localization of the ventricles. Ultimately, the procedure is complex and requires sufficient manual dexterity, spatial processing, and repeated practice to achieve proficiency. This complexity is compounded by the large variation of anatomical features seen among patients, motivating the need for the development of diverse training scenarios.

Although free-hand ventriculostomy is a common neurosurgical procedure and generally considered safe, numerous reports in the literature have identified that there is room for improvement. Catheter malplacement involves the localization of the distal tip within an extraventricular space or the progression of the tip through critical brain structures. Clinical studies have seen malplacement rates from as low as 12.3% to as high as 60.1%, although we speculate that discrepancies in malplacement criteria are one source of the large range [11-13]. Extraventricular placement necessitates replacement, which not only causes additional damage to brain tissue, but has been reported to result in increased complications and hemorrhaging [14]. It should be noted that teaching institutions have a high proportion of resident drain placements, likely inflating the rate of malplacement due to learning curve effects, which have not been thoroughly examined in the literature. There is also increasing clinical evidence indicating that anatomical variations affect outcome, but no work has examined creating models to predict risk based on anatomical measures [12,13,15].

Numerous ventriculostomy simulators have been proposed or implemented, but for the sake of brevity, we will only examine three modern implementations: the ImmersiveTouch [16], University of Florida's (UoF) ventriculostomy simulator [7], and the NeuroTouch [17]. The ImmersiveTouch is an augmented-reality simulator that uses a

haptic tool to interact with the overlaid virtual scene, allowing a user to guide a catheter into the ventricular system. The tool and internal anatomy can be rendered to provide feedback. The simulator has been shown to be accurate and to improve clinical performance through practice and feedback [18,19]. A number of anatomically varying cases have been incorporated (normal, slit, and shifted ventricles) with initial data illustrating that such topological changes affect performance [7]. Yudkowsky et. al, examined a larger set of cases also illustrating the effect of patient-specific variation on targeting accuracy [19]. The drawbacks of the platform are the cost, lack of physical presence, such as a mannequin, and lack of versatility in the targeting task. Though it varies clinically, the burr-hole location is fixed in the first generation ImmersiveTouch and is limited to three predetermined locations in the second generation implementation [20]. Additionally, there is no physical head for users to landmark – a feature requested by resident users of the system [7]. Uoff’s platform consists of a physical phantom to model tissue with a separate VR display to visualize internal anatomy. Due to the physical nature of the simulator, inclusion of patient-specific cases is more challenging and less cost-effective than for a strictly virtual system. Finally, the NeuroTouch is a haptic-enabled virtual-reality simulator that makes use of a full-sized mannequin head to simulate a number of neurosurgical procedures, including ventriculostomies [21]. The NeuroTouch is of comparable cost to the ImmersiveTouch. While the NeuroTouch supports burr-hole and trajectory selection tasks, it provides only a single case and does not yet allow for patient-specific cases. We have extended the functionality of the NeuroTouch to allow for inclusion of patient-specific cases with enhanced case preparation and feedback. Technical details are provided in section 3: Methodology.

In addition to providing a safe training environment, simulators allow us to test novel techniques and tools prior to clinical validation. AR guidance systems are gaining traction within research and clinical settings. As early as 2002, an AR guidance system for needle biopsies (a task similar to ventriculostomies) was evaluated using phantom models [22]. While image-guidance systems exist for ventriculostomy procedures, they have not gained widespread acceptance by surgeons, presumably due to added setup time and complexity [23]. Our group developed a mobile augmented-reality image-guidance tool for use in ventriculostomies that runs on consumer hardware, requires minimal setup

time, and does not impede the procedural workflow. The system also has the potential to simulate tasks currently constrained to the ImmersiveTouch or NeuroTouch systems at a comparably minor cost. The technical aspects and a more thorough review of related AR technology can be found in the previous chapter.

5.3 Methods

5.3.1 Simulator Design

The technical requirements involved incorporation of patient-specific anatomy into an environment that simulates the clinical process of preoperative planning and selection of a burr-hole location and catheter trajectory. As this was intended to be a training activity, the system was to provide immediate descriptive feedback to the user in the form of 3D renderings. A physical mannequin head was required for user landmarking as well as for use with image-guidance systems; it provides a common frame of reference for aligning scenes and allows clinicians to perform familiar landmarking approaches to guide their trajectory and burr-hole selection.

The simulator module focuses on a sub-task of the overall ventriculostomy, in which the participant must localize a desired burr-hole location and indicate the trajectory to the lateral ventricles. This phase of the task isolates the initial spatial reasoning component of the trajectory estimation crucial for targeting the ventricles. This clinical phase of the task requires no exploratory tactile navigation through the brain, and so the haptic tool was used as an input manipulandum (not for force feedback at this stage). The burr-hole location estimation, and angle-of-entry estimation, are perceptual-motor in nature. There are three components to each patient-derived case which proceed in the order listed.

1. The user examines either segmented axial slices of the case prior to the task or they use the AR guidance tool.
2. The user uses the mechanically-tracked pointing tool to indicate a burr-hole location as well as a trajectory into the ventricles on the mannequin head.
3. The user is shown feedback of their approach through a rendering of their trajectory within the head.

All scenarios used were derived from volumetric MR and CT imaging data. Primary features of interest were the lateral and third ventricles; these features were segmented using a combination of manual and semi-automatic techniques. Live-wire [24] and region growing [25] techniques were used to speed up the segmentation process, but the final contours were inspected manually. While automatic segmentation algorithms exist [26], they are not robust over strong anatomical variations, particularly within the third ventricle and with low-resolution and low-contrast clinical images. Rather than viewing the raw medical images, users are only shown the ventricular system within an outline of the mannequin head. Figure 29 presents a view of the slice-based interface. The interface is restricted to viewing axial slices as this is the common preoperative planning approach for catheter insertion.

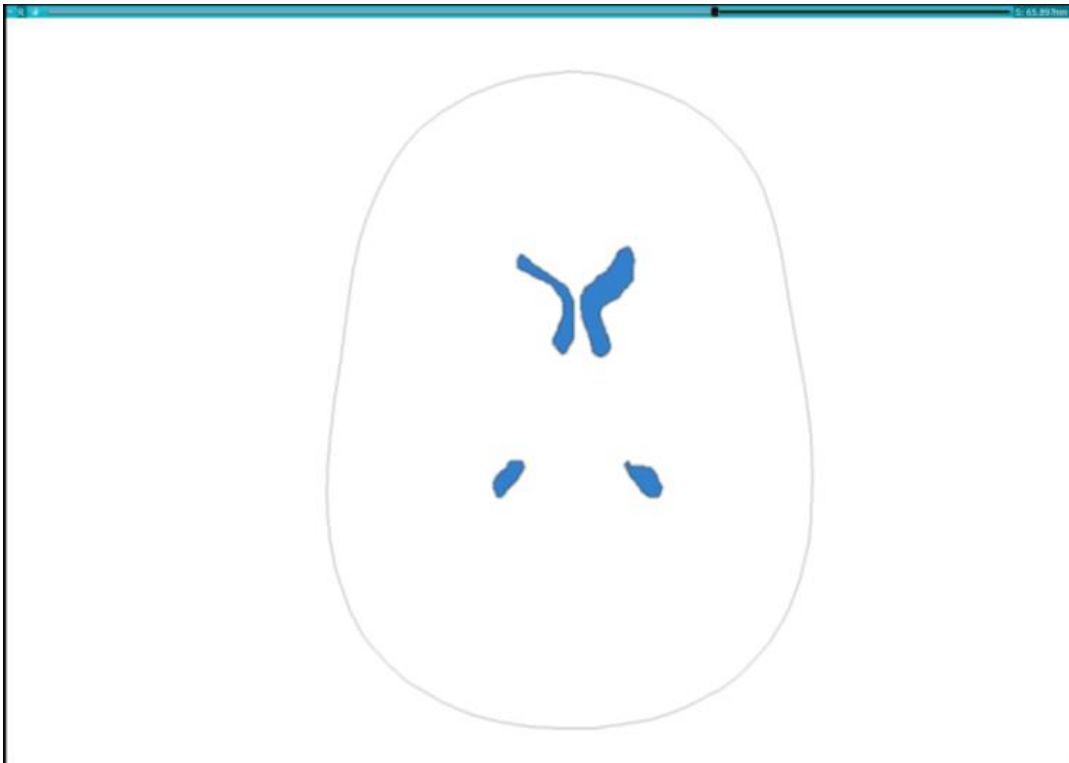


Figure 29: A screenshot of the slice-based interface. Unlike conventional radiographs, the right side of the image corresponds to the right side of the head.

Our module is integrated into the NeuroTouch – a virtual reality surgical simulator developed by the National Research Council of Canada, providing simulation modules for various neurosurgical procedures [17]. Although the NeuroTouch contains a module

for simulating burr-hole placement and trajectory selection, the simulator does not allow for incorporation of custom scenarios and feedback. To use the simulator's interface, we registered our custom scenarios to the coordinate system of the platform's mannequin. A neurosurgeon provided manual registration to align the virtual ventricles within the visual space of the mannequin head.

In addition to incorporating our custom scenarios, we integrated a mobile augmented reality image-guidance system developed by our group and discussed in the previous chapter [27]. The system tracks the position of an image-based marker (a coloured cube with quick response codes) mounted to the mannequin and overlays three-dimensional renderings of the ventricles corresponding to the current scenario. The scenario is chosen manually on the device based on the experimental ordering. The system is portable, lightweight, low cost, and smoothly operable on consumer hardware. Although there is potential for using the AR display as a low cost training tool, our aim is to use the simulator platform to evaluate the use of the AR tool as a clinical aid through image-guidance. While a manual registration is still required for optimal correspondence between the virtual and physical scenes, the system does provide a common landmark between the mannequin and patient imaging at the nasion (See Figure 30, which depicts the marker, the landmark and example ventricles). Figure 31 depicts the NeuroTouch with the guidance system.

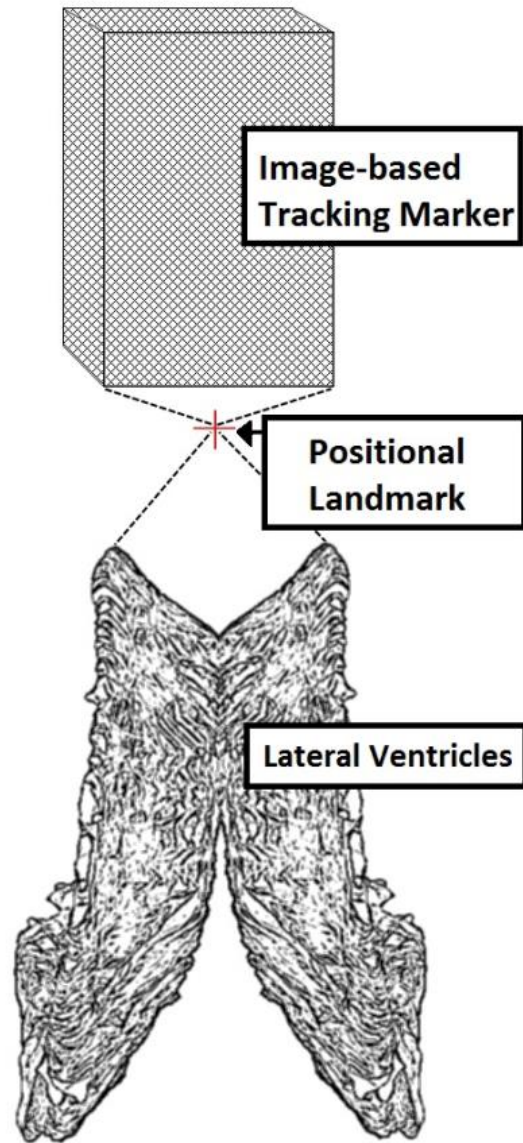


Figure 30: An illustration of the image-based marker cuboid, the positional landmark at the nasion, as well as patient-derived lateral ventricles.

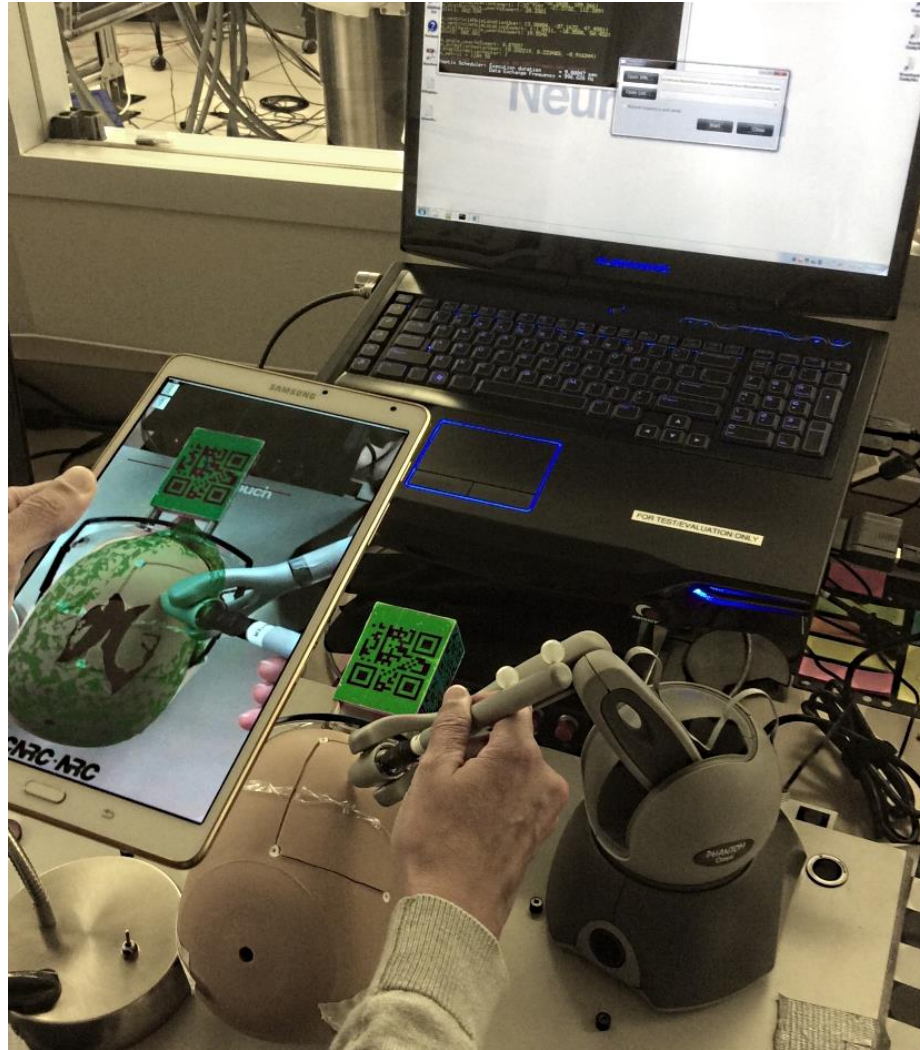


Figure 31: The NeuroTouch used in conjunction with the AR guidance tool. The ventricles can be seen overlaid on the mannequin head in addition to a green mesh showing alignment between the virtual head and the physical model.

In order to provide users with meaningful feedback of their performance (allowing them to be confident of a correct or incorrect trajectory and adjust as necessary), we developed a feedback module using the open source 3D modelling and animation suite Blender. The module takes the user output from the NeuroTouch and renders the user's trajectory through a transparent rendering of the mannequin with the internal ventricles corresponding to the scenario. To illustrate the ideal target, the right anterior horn is highlighted in red. Performance metrics are not provided through the feedback module (the users are never made aware of their score); only a visual

illustration of the user's trajectory in relation to the target is shown. An example can be seen in Figure 32.

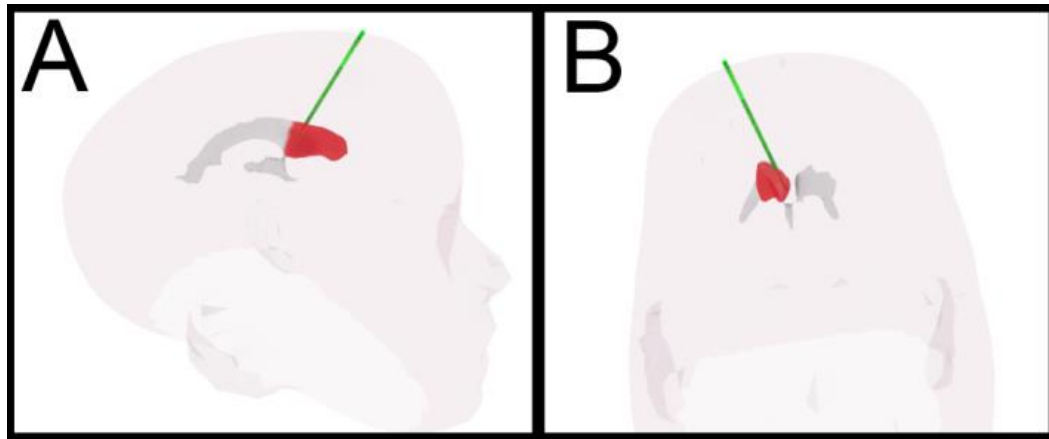


Figure 32: Two renderings providing feedback. A view from the right side (A) and from the front (B) are provided during experiments. The green line represents the user's trajectory. The right anterior horn of the lateral ventricles can be seen highlighted red – this roughly indicates the target area.

5.3.2 Experimental Design

This chapter colligates two separate but related experimental designs. The first study compared traditional ventriculostomy targeting to AR assisted targeting using expert scoring as the primary performance metric. To build on the repertoire of reliable and objective accuracy measures for part-task ventriculostomy procedures and to further examine the role that anatomical variations play in performance of the task, the second study examined additional targeting tasks with neurosurgical residents and medical novices (subjects with sufficient medical and anatomical knowledge, but minimal experience performing ventriculostomies) using further developed objective measures for targeting accuracy.

5.3.2.1 The Role of AR in Part-Task Ventriculostomies: Phase I

In total, 14 scenarios were created from 13 unique patients. The additional scenario was created from a case with significant midline shift that was flipped along the midline. To inform our initial arrangement of scenarios, an expert neurosurgeon

classified each scenario into 4 distinct levels of difficulty based on the size of the ventricles, as well as their shape and localization in space due to external mass effects as sometimes seen in traumatic brain injury. The classifications are: simple (large ventricles seen in hydrocephalus), mild-moderate (enlarged ventricles, but smaller than the simple cases), moderate (ventricle of normal size with no deformation), and complex (ventricles of normal size but with added deformation due to local mass effect). To convey the anatomical variation present in the simulated cases, Figure 33 depicts a cross section of all ventricles within the mannequin head.

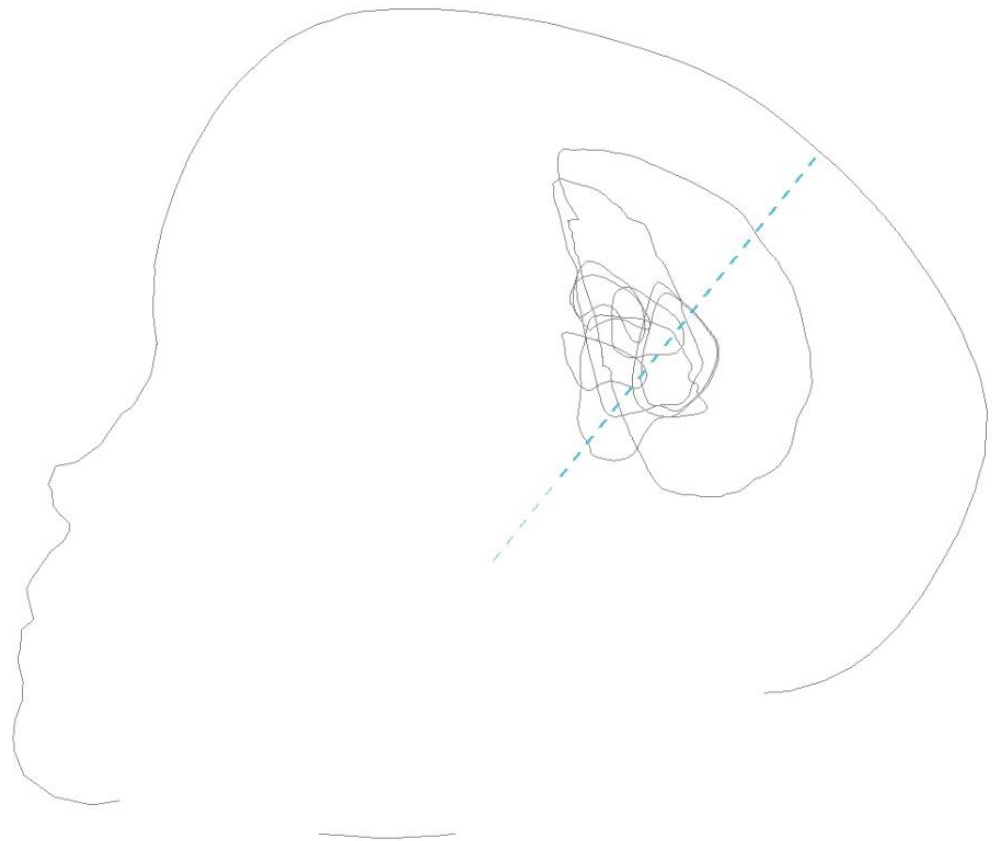


Figure 33: Sagittal cross section of mannequin head with all ventricle cases overlaid.

For each case, the total ventricular volume, Evan's ratio, and midline shift were measured to examine how these traits affect performance of the procedure. Evan's ratio is a ratio of the maximal width of the anterior horns on an axial plane to the maximal width of the

head that has been used as a measure of the severity of hydrocephalus, although it is not particularly reliable [28]. Figure 34 depicts the measuring of the widths required to determine the Evan's ratio.

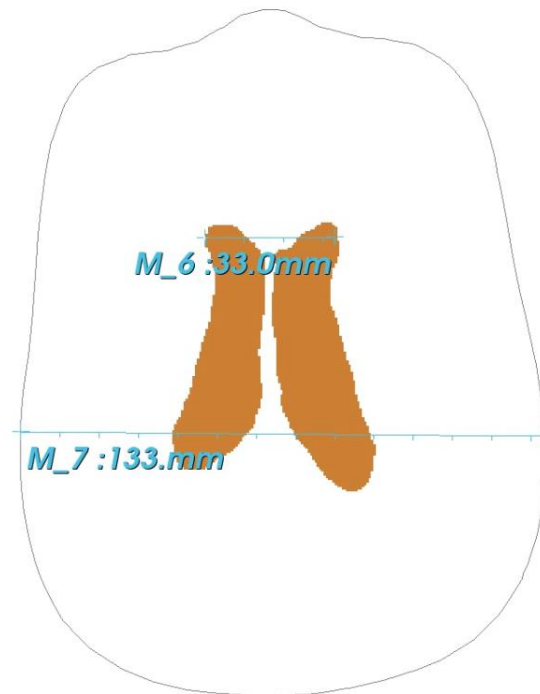


Figure 34: Measures taken along an axial slice to determine Evan's ratio.

The midline shift is an examination of lateral shift of the ventricles. As with Evan's ratio, midline shift is used as an indirect measure of intracranial pressure. In some cases, the magnitude of midline shift is used as an indicator for the need for surgical intervention as well as a marker for outcomes [29]. The magnitude of midline shift is measured in various ways, often using the septum pellucidum, pineal gland, or third ventricle as anatomical reference points. Deviation of these features is measured from the side of each skull side or from an ideal midline which divides the two brain hemispheres into (ideally) symmetric halves. The midline shift is a suitable candidate as a performance factor, as shift of the ventricle system drastically alters the trajectory from the prototypical approach. In this study, we examine midline shift using the third ventricle as a reference

point and an ideal midline as a reference. While this approach may not fully describe the shift of the anterior horns (the most common catheter target), it is a measure thoroughly used in the clinic. The ventricles used in this study are split into two groups: nominal midline shift (shift of less than a single standard deviation from the norm) and severe midline shift (shift greater than a single standard deviation). While there are existing automated methods to extract the midline shift from imaging, we measured shift manually as the number of cases used was not excessive [30].

Experiments consisted of a progression of scenarios utilizing all patient-derived cases. The participants consisted of seven residents and one expert neurosurgeon and were categorized based on years of residency (PGY1-PGY6) and the expert surgeon was placed in a distinct category. As EVD procedures are commonly performed in residency, this is a reasonable measure of exposure to the procedure, so we consider residency experience to coincide with procedure experience. Each user was tasked with performing all 14 scenarios without AR guidance in a set order, and each scenario with AR guidance after a brief practice period with the AR device. Before beginning the ventricle targeting, users were given the opportunity to gain familiarity with the system by targeting 5 ellipsoid cases. Afterwards, users were divided into two groups with either a simple to complex or a pseudo-random ordering of cases to examine any possible learning impact of the ordering. Users were further subdivided into performing the AR guidance component first or following the base tasks. In addition to recording the burr-hole location and trajectory of each task, the entire task is also timed from the beginning of the pre-task image analysis to the final selection of trajectory. Figure 35 presents a graphical representation of the user study.

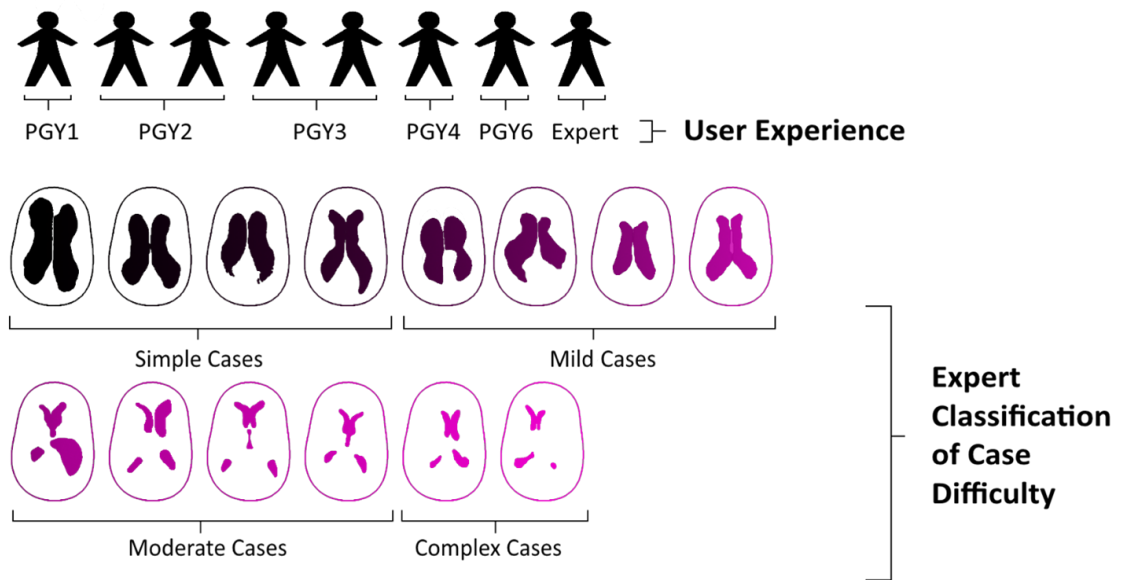


Figure 35: Summary of user study.

In order to score users, suitable metrics must be used that relate performance on the simulator to the user's ability to perform the procedure clinically under similar conditions. A naïve approach would involve selecting a single correct burr-hole and catheter tip location and scoring based on deviation from this trajectory (as commonly seen in similar simulators), as this does not account for the anatomical variation present.

From a geometrical perspective, there are a number of possible trajectories from the skull to the ventricles. Clinically – particularly when dealing with anatomical variations – there is generally no single path that can be identified as “correct” for a successful placement. Indeed, there are many possible variations that would result in a functional placement without risking damage to eloquent tissue. Traditionally, through the didactics of training, Kocher's point is sought as providing a desired burr-hole location, from which a trajectory can be estimated leading to the entrance of the ipsilateral foramen of Monro [9]. We are able to render this path and provide it as a reference when providing feedback to users, but deviation may still result in a perfectly scored placement. In their clinical evaluation, Kakarla et. al. [13] use a 3-grade scoring system which takes into account general tip location, functionality of the drain, and

damage to eloquent tissue. The difficulty with this scoring approach is that it does not discern between different manners of successful and unsuccessful placements. For example, there is no difference between a trajectory that misses the ventricles by a small margin compared to one that is not at all close. Yudkowsky et. al. [19] examined whether the tip was successfully placed in the ipsilateral ventricle and how far it is from the foramen of Monro. Abnormal anatomical variations, however, often necessitate catheter placements that do not target the foramen of Monro area. Hooten et. al. [7] developed a compound score factoring time, distance to foramen of Monro, proximity to Kocher point, and a number of point reductions for multiple attempts and passage through critical structures. It is the most expansive scoring metric we have encountered, but relies on a number of seemingly arbitrary values summed to a final score.

In this first investigation, we utilized expert scoring of accuracy by developing a 7-point Likert scale with objective guidelines for anchor points. The measure is termed the Index of Accuracy (IA) with a 1 indicating a perfect trajectory (one that will result in a successful catheter tip placement without damaging eloquent tissue) and a 7 indicating the worst scored trajectory (deviating from all standards and causing damage to eloquent tissue). Example trajectories and associated scores are illustrated in Figure 36. The Index of Performance (IP) takes into consideration both time and accuracy, following a Fitts's methodology. In this simple case, the IP is formulated as follows:

$$IP = \frac{1}{(IA)*time} \quad (5)$$

In addition to expert scoring of trajectories, the automated and objective metrics discussed as part of the second phase of this work was also applied to the data from the first phase to examine the relationship between these measures and expert scoring.

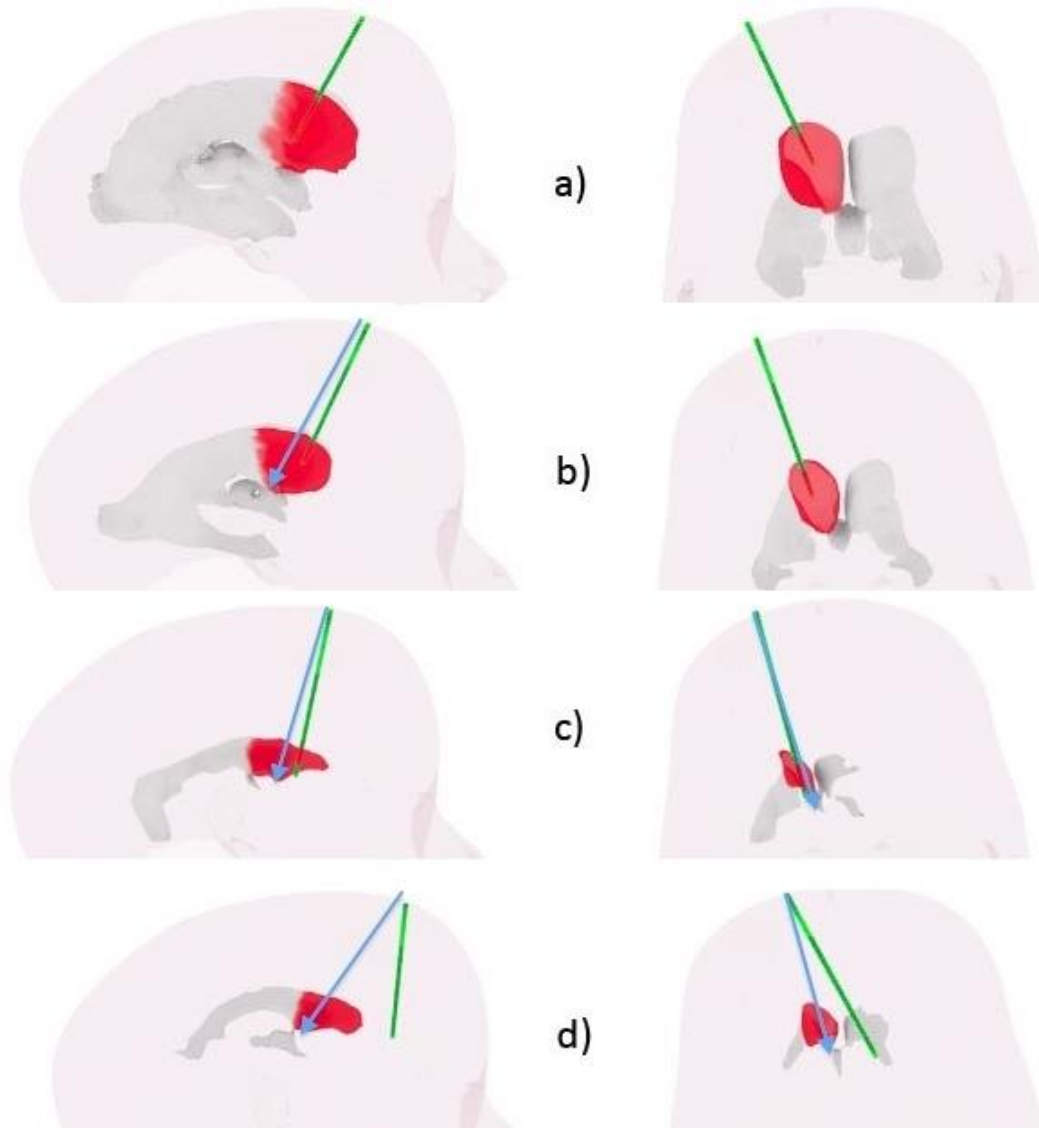


Figure 36: In green, the trajectory pointed by the user, in blue, the ideal trajectory.

A) Index of accuracy (IA) of 1 with a perfect trajectory. Note that the blue trajectory is occluded by the user's trajectory which is perfectly coincident. B) IA of 2 with a small angle change in one of the planes. C) IA of 2 with a small modification of angle in both planes, resulting in the trajectory coming out of the ventricle. D) IA of 6 with a poor trajectory.

5.3.2.2 Objective Examination of Ventriculostomies: Phase II

For this examination, there were 21 unique ventricle cases derived from patient imaging. The cases were presented in a pseudo-random order and then repeated in the same order, so a total number of 42 tasks were performed by users following an initial practice period. System automation and withdrawal of AR tasks allowed for the simulation of such a large number of trials per user within time constraints. In this study, 3 neurosurgical residents and 3 novices (no surgical training, but strong anatomical knowledge) were recruited.

This phase used lessons learned from the previous approach to further develop comprehensive and objective measures of ventriculostomy accuracy. Our accuracy measures are extended from previous work and derived from Muirhead et al. [9]. While Muirhead et al. used their measures to determine the most optimal computed trajectories based on various landmarking, we adapted their work to evaluate user performance of freehand trajectories chosen by users. The paper describes four measures of accuracy: engagement, relative sagittal angulation, relative coronal angulation, and error margin. Engagement is the length of the line segment that is created by the intersection of the surgical trajectory and the anterior ventricle horn, the partition of which is described by Lind et al. [31]. The relative coronal and sagittal angulations are the angles between the trajectory and the coronal and sagittal components of the Foramen of Monro. Finally, the error margin is the smallest angle the trajectory can be deviated for the midpoint of the intra-horn segment to no longer be contained in the anterior horn. This measure fails to account for the various ways the topology can vary; in some cases, the calculated trajectory can be placed outside of the anterior horn and yet still be a more viable surgical route. We have adapted this measure to what we call the projected area, which is the area of the anterior horn projected onto the observer's view plane from the chosen burr-hole location and view in line with the trajectory. This helps relate the task to a Fitts's methodology by producing a measure of target area. Figure 37 illustrates the measures considered. These measures are used to examine user performance in relation to expert classification of scenario difficulty; each case is categorized into simple, mild, moderate, or complex cases.

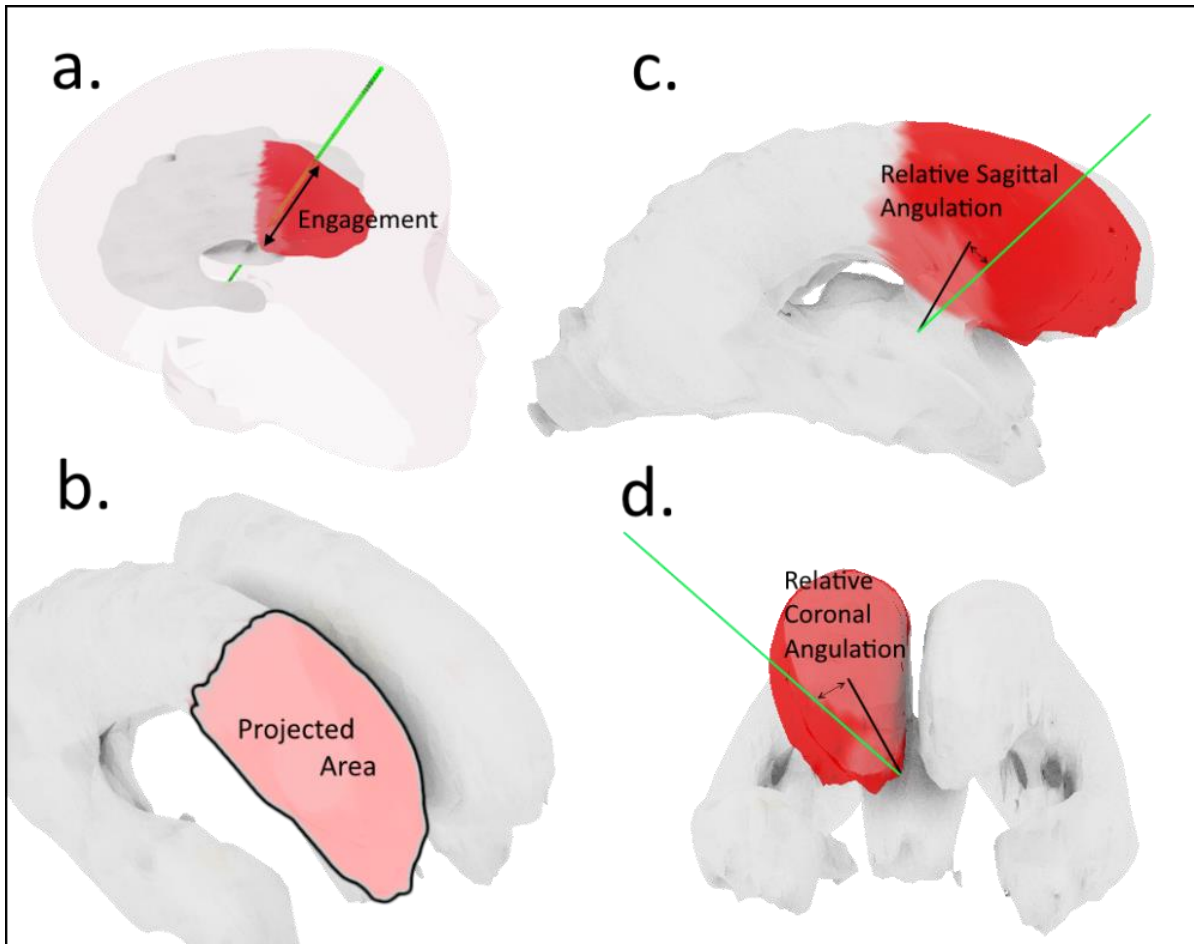


Figure 37: An illustration of the accuracy metrics. Engagement (a) is the length of the trajectory intersecting the anterior horn. The projected area (b) is the area from projecting the anterior horn onto the viewing plane from the burr-hole. The relative sagittal (c) and coronal (d) angulations are the angles the trajectory makes with the Foramen of Monro's respective path.

5.4 Results

5.4.1 The Role of AR in Part-Task Ventriculostomies (Phase I)

The accuracy of each trajectory was scored by a blinded expert for both AR and non-AR tasks. Distributional differences between user metric variance (time, IA, IP) was not seen applying a nonparametric Leven's test ($p > 0.15$). To examine for the presence of user

difference, a Kruskal-Wallis test was run across users for time, IA, and IP. Differences between users was seen across each of the aforementioned measures ($p < 0.001$, $p < 0.05$, and $p < 0.001$, respectively). Using the result Chi-Square values, effect size estimates were determined, suggesting that differences between users accounts for 57.35% of the variation in time, 6.46% of the variation in IA and 32.81% of the variation in IP.

5.4.1.1 Without AR Guidance

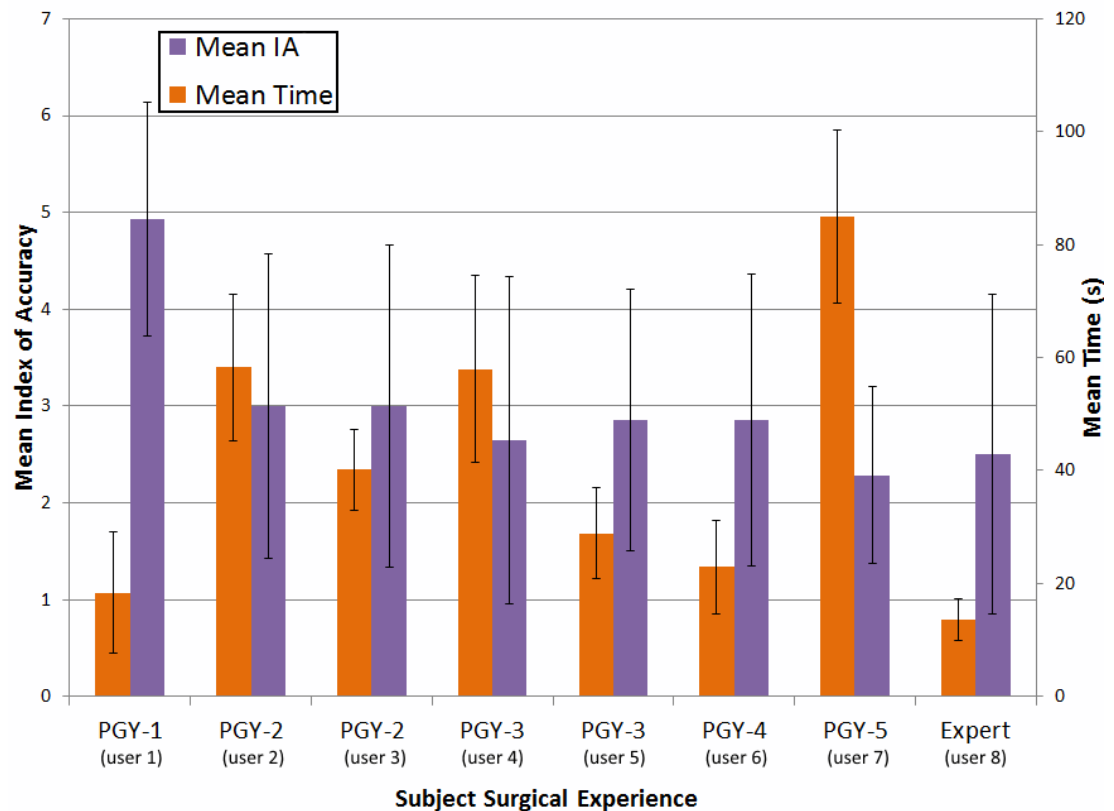


Figure 38: Mean time and IA for each user without AR guidance. The users are numbered and their class of experience is indicated. For both time and IA, a lower score is an indicator of higher performance. Standard deviations are visualized on the graph.

Figure 38 reports the mean IA and time by subject without AR guidance. An interesting result is the large task time variance between users, but low intra-user variance between tasks. Taking into account Fitts's law, which respects the trade-off between speed and accuracy, we calculated IP for each task (bits/seconds), but restrict this

analysis to the overall phase that involves visual inspection, tool positioning, and final angle selection. As a result, the IP values provided in Table 4 provide relative scores for the user and case metrics arranged by expert-assigned difficulty classifications for non-AR tasks.

There was a strong inverse correlation between years of experience among residents with IA ($r = 0.93$ with 95% C.I. [0.90-0.95]), but only a moderate correlation between experience and IP ($r = 0.71$ with 95% C.I. [0.62-0.78]). As can be seen in Table 4, the overall performance of all subjects was significantly better for the classified simple cases (IP = 0.023 bits/sec) than for the complex ones (IP = 0.006 bits/sec) based on expert classification ($p < 0.05$ using a two tailed t-test as neither group was found to deviate from normality using K-S test where each $p > 0.15$).

For performance of for all non-AR guided cases compared to anatomical measures, see Table 4.

Table 4: Performance of users for all cases (without AR guidance) divided by difficulty classification with anatomical measures. Standard deviations are supplied with means.

Difficulty Classification	Mean Volume (mL)	Mean Evan's Ratio	Mean Task Time (s)	Mean IA	Mean IP (bits/s)
Simple (n = 4)	212.94 ± 69.11	0.43 ± 0.054	38.22 ± 26.20	2.43 ± 1.27	0.023 ± 0.024
Mild-Moderate (n = 4)	240.77 ± 241.53	0.39 ± 0.20	37.62 ± 24.02	2.79 ± 1.52	0.018 ± 0.017
Moderate (n = 4)	33.76 ± 0.26	0.26 ± 0.03	39.26 ± 28.24	3.33 ± 1.53	0.014 ± 0.016
Complex (n = 2)	14.81 ± 0.00	0.21 ± 0.00	40.13 ± 23.36	4.5 ± 1.67	0.009 ± 0.007

5.4.1.2 With AR Guidance

When analyzing the individual AR results, the first 5 cases showed poor accuracy in comparison to later cases, although still not significant compared to the non-AR approach. Experience was not found to be predictive of performance using the AR system, as seen in Figure 39.

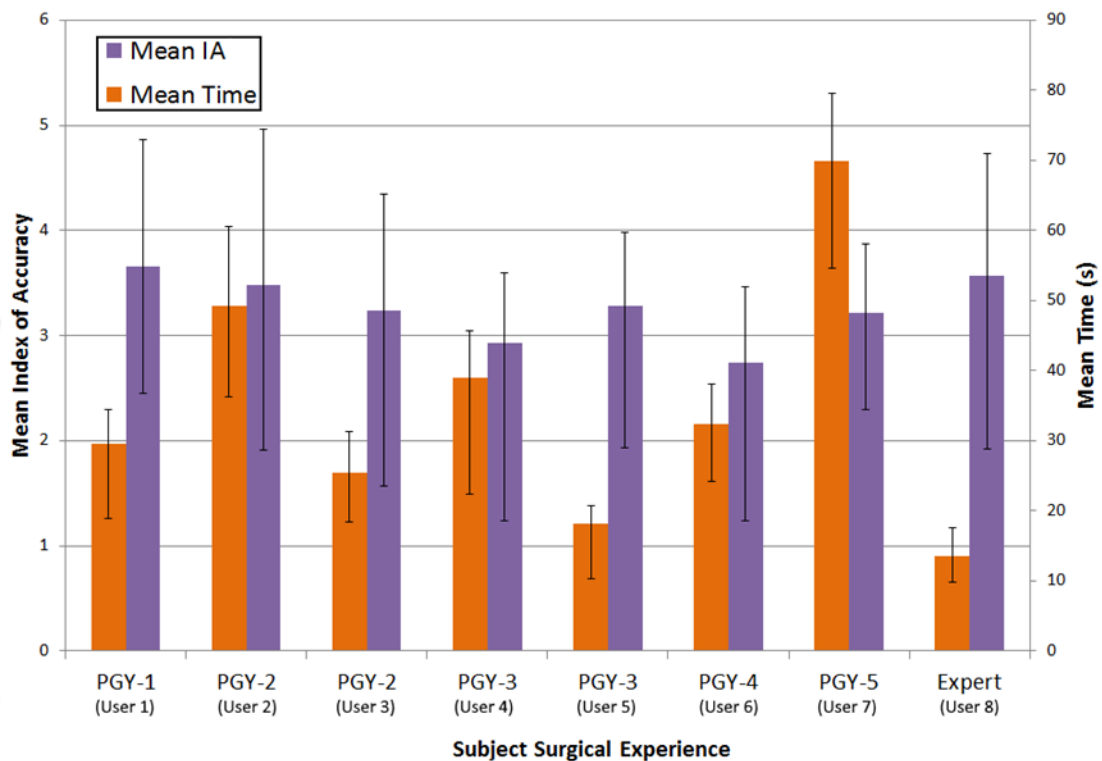


Figure 39: Mean time and IA for each user with AR guidance. The users are numbered and their class of experience is indicated. Standard deviations are visualized on the graph.

Table 5 holds performance results for AR guidance scenarios in relation to anatomical measures.

Table 5: Performance of users for all cases (with AR guidance) divided by difficulty classification with anatomical measures. Standard deviations are supplied with means.

Difficulty Classification	Mean Volume (mL)	Mean Evan's Ratio	Mean Task Time (s)	Mean IA	Mean IP (bits/s)
Simple (4 cases)	212.94 ± 69.11	0.43 ± 0.054	29.37 ± 26.20	3.39 ± 1.75	0.023 ± 0.028
Mild-Moderate (4 cases)	240.77 ± 241.53	0.39 ± 0.20	36.46 ± 24.02	2.75 ± 1.43	0.022 ± 0.020
Moderate (4 cases)	33.76 ± 0.26	0.26 ± 0.03	32.81 ± 28.24	3.00 ± 1.15	0.032 ± 0.057
Complex (2 cases)	14.81 ± 0.00	0.21 ± 0.00	34.65 ± 23.36	4.12 ± 2.10	0.010 ± 0.005

Examining the influence of patient-specific factors, we examined the correlations between IP and ventricle volume ($r^2 = 0.58$ with 95% C.I. [0.47-0.68]) and IP and Evan's ratio ($r = 0.62$ with C.I. [0.51-0.71]). The correlation between expert classification of difficulty and IP was seen to be the strongest ($r = 0.73$ with 95% C.I. [0.65-0.80]). Examining midline shift, the mean midline shift was 4.03 ± 5.73 mm (SD reported). Only two cases exhibited midline shift outside of the first standard deviation (the cases classified as complex). In comparing the accuracy of targeting ventricles with midline shift outside of 1SD to those within, accuracy was found to be non-normal for each group (K-S test $p < 0.01$), but A significant difference was seen for AR ($p < 0.05$ and $U=421$ using Mann-Whitney U test) and non-AR ($p < 0.05$ and $U=345$ using Mann-Whitney U test) targeting. Performance values were not seen to be non-normal (K-S test with $p >$

0.15 for each group) and were significantly different for large and small midline shifts (grouped by the 1 SD cutoff) using a two tailed t test for both AR and non-AR ($p < 0.05$). There was also a moderate overall correlation between IP and midline shift ($r = -0.75$ with 95% C.I. [0.67-0.81]). No learning effects were seen in response to the ordering of cases.

There was a significant difference for all users over engagement between simple and complex cases ($p < 0.05$ and $U=82$ using a Mann Whitney U test as samples were found to be non-normal using K-S test with $p>0.15$). No significant findings resulted from examining the additional objective metrics in this phase. No trend was observed relating any measure with experience.

5.4.2 Objective Examination of Ventriculostomies (Phase II)

The engagement measures of each user group was found to be non-normal using a K-S goodness of fit test ($p < 0.05$). As can be seen in Table 1, the engagement measure was best able to distinguish between novices and residents (residents performed significantly better for simple, moderate and complex cases, $p < 0.05$ and $U=2821$ using the Mann-Whitney U test) and whether it was a simple or complex case in question for experts ($p < 0.05$ and $U=0$ using the Mann-Whitney U test) and novices ($p < 0.05$ and $U=2.5$ using the Mann-Whitney U test). Weak learning curve trends were seen using relative angulations and projected areas, but otherwise these metrics were unremarkable.

Table 6: The average engagement for residents and novices over each difficulty class. Standard deviations are supplied with means.

Case Difficulty	Resident Mean Engagement(mm)	Novice Mean Engagement(mm)
Simple (6 cases)	24.01 ± 11.18	15.50 ± 14.01
Mild (6 cases)	22.95 ± 22.43	21.81 ± 28.27
Moderate (6 cases)	4.41 ± 7.07	2.88 ± 5.57
Complex (2 cases)	2.57 ± 6.01	0 ± 0

5.5 Discussion and Future Work

In this chapter, we presented a module to train residents to place EVDs in patient-specific scenarios implemented on a surgical simulation platform that is already deployed and in use at a number of teaching institutions. Beyond training, the platform can be utilized to examine various factors that influence user performance – all within a safe test bed. Further to this, we integrated a previously developed AR image-guidance system to evaluate its efficacy in guiding EVD placement simulated part-tasks. We observed a correlation between surgical experience and task performance, providing evidence that our initial measures of performance (namely the IP and IA) are able to discriminate between experience levels (criteria validity). A higher correlation was seen between experience and IA than experience and IP, indicating that IA may be a more appropriate measure to assess skill. Indeed, while the subjects were instructed to consider the speed of their targeting, the weighting of time in determining IP is largely arbitrary, leading us to propose decomposing the task into sub-tasks of visual inspection, tool movement, and angle specification. An interesting trend to note is the high inter-subject variability of task time compared to the relatively low intra-subject variability (Figure 38 and Figure 39). The likely explanation is that each subject has a preferred preparation and targeting approach that characterizes their response.

We also saw greater performance for simple cases compared to more challenging cases based on expert classification, which seems to outperform simple metrics such as ventricular volume and Evan's ratio in determining the difficulty of a given case. This indicates that the expert is relying on additional cues when making judgments relating to the difficulty. This is clear when examining our more difficult cases; these cases contain significant mass effects, leading to shifted ventricles and other deformations which are not captured by our current metrics. Further investigation is warranted, but – in line with Fitts's model – measures of ventricle size are at least somewhat predictive of the difficulty to target as they correspond to the size of the target. The shortcoming of a global volume measurement, however, is that it is not descriptive of ventricle topology. The midline shift was also seen to significantly impact performance, though the available

cases lacked diversity among shifted ventricles and a more clinically representative sample may produce different results.

A steep learning curve was seen with the use of AR guidance, which is expected when incorporating novel guidance techniques [32]; however, after the learning curve the AR system seemed to provide some benefits in terms of speed and accuracy together (IP), but we failed to see a significant difference to the non-AR case in terms of accuracy alone (IA). While no difference in performance was seen when isolating and comparing the simple cases, there was a significant advantage to using AR for moderately difficult and complex cases, indicating that AR guidance may be most appropriate for challenging cases that deviate from clinical norms. Unlike in a clinical setting, we did not augment the procedure with the AR, but we replaced the traditional preoperative planning of examining the patient's images. It may be that the combination of approaches allows for the greatest perspective and therefore superior performance. Such an approach may also better facilitate learning with the AR tool, as it would pair the novel technique within a familiar context.

A drawback of this study is the low sample size. While this is often expected of surgical simulation research, future work should continue data collection and analysis to improve the scope of the data to test additional hypotheses. Further examination in the relationship between experience and performance as well as the role of AR guidance in improving user performance is suggested. While AR guidance did not augment performance as significantly as predicted, the study was limited, the subjects did not receive extensive training on the platform (especially important considering the learning curve) and the implementation is under constant development. Indeed, we would expect to see a learning curve with any new technique and it may be that these experiments only targeted the initial segment. We also hope to further examine various objective metrics of ventricular geometry that impact performance. The measured developed for the second phase of this study did not provide noteworthy insight when retroactively applied to the original data from the first phase. The weak relationship between the engagement and the expert rating may be a good indication that the metrics are not sufficiently examining the factors that contribute to a successful trajectory, especially when considering the value of

the experts ratings in uncovering expected trends (as well as the precedent of using surgical expertise to evaluate simulator and OR performance).

In the second phase of the study, novel metrics for ventriculostomy simulation were examined. The data indicates the engagement measure appears to be valuable for examining ventriculostomy trajectory performance. While no strong trends were seen with the remaining metrics, the study suffers from a low sample size. The metrics may be inherently deficient, however. For example, from the data used there was very little variation in the sagittal and coronal angulations of the Foramen of Monro. As neurosurgeons generally localize the burr-hole with little variation between tasks (and the same is true in the clinic) the measures may hold no meaningful information. In future work, we will examine the use of additional metrics as well as investigate learning effects, experience differences, and how anatomical variations impact performance. By exploring these domains, we hope to further characterize and quantify the geometric aspects of ventriculostomy trajectories that affect proficiency and safety. The development of valid metrics will allow for automation, objectivity and consistency in scoring simulation scenarios. Additionally, they provide real-time numerical grading which may facilitate learning when supplied as feedback [33], although this would warrant an investigation.

As with all training approaches, the true test of a curriculum involves a rigorous application of the gold standards of evaluation. Additional research will examine the outcome of extended training in the environment, particularly concerning skill transfer into clinical settings. By providing diverse patient-derived scenarios, we are able to expose trainees to a wide-range of possible cases prior to their clinical experiences. A thorough evaluative methodology and arsenal of metrics will allow for the data mining of these training scenarios to investigate performance factors inherent to various neurosurgical procedures and part-tasks.

The safety of the simulation environment allows us to examine the intersecting effects of unique or rare anatomical variations and user surgical experience. By determining the role that these variables play in the difficulty of a case, we can make

clinically relevant predictions regarding accuracy that could inform the preoperative planning process, ultimately improving patient outcome. Future work will involve investigations into predictive models that can be validated in the safety of the simulation environment.

As with last chapter, a mannequin skull could be printed from publicly available datasets and used as the target for any number of spatially tracked tools. Similar augmented reality applications can be developed using Vuforia and Android development tools. Public imaging datasets are available for study. Imaging targets could be acquired from public imaging datasets and segmented as needed.

5.6 Conclusion

In this chapter, a novel patient-specific simulator platform was presented for performing part-task ventriculostomies. Two user studies were performed to examine performance metrics, human factors, the role of augmented reality guidance, and the impact of patient-specific anatomical variations. AR guidance shows promise, but research into further training is necessitated. Automated metrics are advantageous due to their objective and automatic nature, but further investigation is required to outperform expert ratings. Finally, we found that ventricle volume, Evan's ratio, and midline shift are all measures of anatomical variation that influence performance of the procedure.

Bibliography

1. Quest, D.O.: Naval aviation and neurosurgery: traditions, commonalities, and lessons learned. (2007).
2. Sealy, W.C.: Halsted is dead: time for change in graduate surgical education. *Current Surgery*. 56, 34–39 (1999).
3. Satava, R.M.: Surgical education and surgical simulation. *World journal of surgery*. 25, 1484–1489 (2001).
4. Z. Roitberg, N.K.M.A.T.H.F.C.J.A.B.: Bedside external ventricular drain placement for the treatment of acute hydrocephalus. *British journal of neurosurgery*. 15, 324–327 (2001).
5. Huyette, D.R., Turnbow, B.J., Kaufman, C., Vaslow, D.F., Whiting, B.B., Oh, M.Y.: Accuracy of the freehand pass technique for ventriculostomy catheter placement: retrospective assessment using computed tomography scans. (2008).
6. Lemole Jr, G.M., Banerjee, P.P., Luciano, C., Neckrysh, S., Charbel, F.T.: Virtual reality in neurosurgical education: part-task ventriculostomy simulation with dynamic visual and haptic feedback. *Neurosurgery*. 61, 142–149 (2007).
7. Hooten, K.G., Lister, J.R., Lombard, G., Lizdas, D.E., Lamprotang, S., Rajon, D.A., Bova, F., Murad, G.J.: Mixed Reality Ventriculostomy Simulation: Experience in Neurosurgical Residency. *Neurosurgery*. 10, 576–581 (2014).
8. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*. 47, 381 (1954).

9. Muirhead, W.R., Basu, S.: Trajectories for frontal external ventricular drain placement: virtual cannulation of adults with acute hydrocephalus. *British journal of neurosurgery*. 26, 710–716 (2012).
10. Schödel, P., Proescholdt, M., Ullrich, O.-W., Brawanski, A., Schebesch, K.-M.: An outcome analysis of two different procedures of burr-hole trephine and external ventricular drainage in acute hydrocephalus. *Journal of Clinical Neuroscience*. (2011).
11. Saladino, A., White, J.B., Wijdicks, E.F., Lanzino, G.: Malplacement of ventricular catheters by neurosurgeons: a single institution experience. *Neurocritical care*. 10, 248–252 (2009).
12. Toma, A.K., Camp, S., Watkins, L.D., Grieve, J., Kitchen, N.D.: External ventricular drain insertion accuracy: is there a need for change in practice? *Neurosurgery*. 65, 1197–1201 (2009).
13. Kakarla, U.K., Chang, S.W., Theodore, N., Spetzler, R.F., Kim, L.J.: Safety and accuracy of bedside external ventricular drain placement. *Neurosurgery*. 63, ONS162–ONS167 (2008).
14. Arabi, Y., Memish, Z.A., Balkhy, H.H., Francis, C., Ferayan, A., Al Shimemeri, A., Almuneef, M.A.: Ventriculostomy-associated infections: incidence and risk factors. *American journal of infection control*. 33, 137–143 (2005).
15. Taylor, S., Bayless, L., Klein, M.W., Janzen, J.M., others: Placement of external ventricular drains and intracranial pressure monitors by neurointensivists. *Neurocritical care*. 10, 241–247 (2009).
16. Luciano, C., Banerjee, P., Lemole, G., Charbel, F., Charbel, F.: Second generation haptic ventriculostomy simulator using the ImmersiveTouch™ system. *Studies in*

- health technology and informatics. 119, 343 (2005).
17. Delorme, S., Laroche, D., DiRaddo, R., Del Maestro, R.F.: NeuroTouch: A Physics-Based Virtual Simulator for Cranial Microneurosurgery Training. *Neurosurgery*. 71, 32–42 (2012).
 18. Banerjee, P.P., Luciano, C.J., Lemole Jr, G.M., Charbel, F.T., Oh, M.Y.: Accuracy of ventriculostomy catheter placement using a head-and hand-tracked high-resolution virtual reality simulator with haptic feedback. (2007).
 19. Yudkowsky, R., Luciano, C., Banerjee, P., Schwartz, A., Alaraj, A., Lemole Jr, G.M., Charbel, F., Smith, K., Rizzi, S., Byrne, R., others: Practice on an augmented reality/haptic simulator and library of virtual brains improves residents' ability to perform a ventriculostomy. *Simulation in Healthcare*. 8, 25–31 (2013).
 20. Alaraj, A., Charbel, F.T., Birk, D., Tobin, M., Luciano, C., Banerjee, P.P., Rizzi, S., Sorenson, J., Foley, K., Slavin, K., others: Role of cranial and spinal virtual and augmented reality simulation using immersive touch modules in neurosurgical training. *Neurosurgery*. 72, 115 (2013).
 21. Choudhury, N., Gélinas-Phaneuf, N., Delorme, S., Del Maestro, R.: Fundamentals of neurosurgery: virtual reality tasks for training and evaluation of technical skills. *World neurosurgery*. 80, e9–e19 (2013).
 22. Rosenthal, M., State, A., Lee, J., Hirota, G., Ackerman, J., Keller, K., Pisano, E.D., Jiroutek, M., Muller, K., Fuchs, H.: Augmented reality guidance for needle biopsies: an initial randomized, controlled trial in phantoms. *Medical image analysis*. 6, 313–320 (2002).
 23. O'Neill, B.R., Velez, D.A., Braxton, E.E., Whiting, D., Oh, M.Y.: A survey of ventriculostomy and intracranial pressure monitor placement practices. *Surgical*

- neurology. 70, 268–273 (2008).
24. Hamarneh, G., Yang, J., McIntosh, C., Langille, M.: 3 D live-wire-based semi-automatic segmentation of medical images. Proc. SPIE. pp. 1597–1603 (2005).
 25. Adams, R., Bischof, L.: Seeded region growing. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 16, 641–647 (1994).
 26. Schnack, H., Hulshoff Pol, H., Baaré, W.F.C., Viergever, M., Kahn, R.: Automatic segmentation of the ventricular system from MR images of the human brain. Neuroimage. 14, 95–104 (2001).
 27. Kramers, M., Armstrong, R., Bakhshmand, S.M., Fenster, A., de Ribaupierre, S., Eagleson, R.: A Mobile Augmented Reality Application for Image Guidance of Neurosurgical Interventions. American Journal of Biomedical Engineering. 3, 169–174 (2013).
 28. Weglinski, T., Fabijanska, A.: The concept of image processing algorithms for assessment and diagnosis of hydrocephalus in children. Prace Instytutu Elektrotechniki. 155–177 (2011).
 29. Tu, P.-H., Liu, Z.-H., Chuang, C.-C., Yang, T.-C., Wu, C.-T., Lee, S.-T.: Postoperative midline shift as secondary screening for the long-term outcomes of surgical decompression of malignant middle cerebral artery infarcts. Journal of Clinical Neuroscience. 19, 661–664 (2012).
 30. Wenan, C.: Automated measurement of midline shift in brain CT images and its application in computer-aided medical decision making. (2010).
 31. Lind, C.R., Tsai, A.M., Law, A.J., Lau, H., Muthiah, K.: Ventricular catheter trajectories from traditional shunt approaches: a morphometric study in adults with

hydrocephalus. (2008).

32. Lollis, S.S., Roberts, D.W.: Robotic catheter ventriculostomy: feasibility, efficacy, and implications. *Journal of Neurosurgery*. 108, 269–74 (2008).
33. Kluger, A.N., DeNisi, A.: The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin*. 119, 254 (1996).

Chapter 6

6 Closing Remarks

The underlying paradigms of medical training are undergoing a shift driven by technological advances in computer hardware and simulation techniques. The traditional models of training using static textbook-based references and the learning-by-doing approach of surgical apprenticeship are being superseded by tailored simulations that offer unique perspectives of medical processes combined with the capability to objectively examine human performance. This is not just driven by technological advances, but also clinical need; traditional training lacks versatility, is more costly than simulation training, and puts patients at risk by exposing them to the learning curve [1,2]. In addition to simulation technologies, image-guidance has shown great potential in augmenting surgical accuracy for tasks where critical anatomy is occluded to the surgeon's view. Emerging from the co-development and coupling of these technologies is a symbiotic relationship; high fidelity surgical simulators can provide both a test bed for new guidance technologies as well as a training platform.

An area of ongoing research that has seen little focus to date is the inclusion of patient-specific and diverse scenarios in anatomical training and surgical simulation technologies. One limiting factor is the process of creating such scenarios. The first step towards creating a scenario modeled from a patient is obtaining 3D medical imaging depicting structural anatomy. Typically, this comes in the form of MRI and CT images. The anatomy required for simulation must then be segmented from the volume(s). Depending on the requirements and imaging approach, desired anatomy may be difficult to distinguish and extract. Some anatomy may not be present in imaging; this is often the case with small vasculature and other miniscule features. Once extracted, anatomical features must be transformed into a format suitable for simulation. This often entails the creation of graphical meshes in the shape of the feature's surface. An important quality missing from the meshes at this point is fine detail on the surface, or texturing. This includes colouring of the surface, small perturbations of surface smoothness, and some vasculature. This information cannot be extracted from the patient without internal colour

images, generally necessitating the use of a priori information involving general assets and an assumption of commonality between patients [3]. This commonly takes the form of colour images that are mapped on to the anatomical meshes either automatically or by user intervention. Even when patient-specific internal colour imaging is available, the process of mapping the textures onto anatomical topology is a complicated process and ongoing area of inquiry [4]. At this point, a realistic anatomical model can be visualized, but its likeness to the patient is not guaranteed. This does not include a discussion of recreating the lighting and general visual environment within the patient during a procedure or the optimization of the mechanical properties of the tissue, which may also exhibit patient specificity. There are numerous other techniques and challenges that are beyond the scope of this section [5]. A pipeline for content generation is discussed in Appendix B.

The benefits of patient-specific simulation and anatomy training certainly warrant further investigation. After all, is a simulator truly representative of a procedure if it does not convey the variations and nuance that are seen clinically? Consider the related field of flight simulation. If flight simulators did not expose pilots to extreme conditions and abnormal operating conditions, they would not be adequately prepared for the reality of day-to-day operations. A challenge for every simulator project is the demand for validation. Indeed, validation is an important component for determining the worth and safety of a simulator system, but there will at some point have to be concession that certain approaches are acceptable standards that do not warrant the imposition of exhaustive validation criteria. This discussion is beyond the scope of this work, but the sentiment is echoed in Smith and Pell's critique [6].

Chapter 2 presented a novel tool for neuroanatomical training and evaluation. The tool was optimized to best convey the spatial relationships of neuroanatomical structures (making use of multiple depth cues, including stereopsis) and to provide an intuitive user interface to perform 3D targeting (prehension) tasks. The platform consisted of dual polarized projectors and a graphics architecture developed in OpenGL. While the development was a significant undertaking at the time, a similar system can now be constructed by much simpler means using Unity in conjunction with a modern VR

headset. However, the platform provides one advantage over modern headsets; the system supports multiple simultaneous viewers, allowing a student-instructor approach to anatomy training. Following development, user performance was examined in relation to visuospatial abilities and anatomical knowledge. Evaluating users, we saw a trend towards greater accuracy by individuals with higher spatial reasoning scores, but it was not entirely predictive of performance. The resident outperformed users on all tasks, indicating that anatomical knowledge may be the mediating factor. By showing a significant learning effect between trials, this subject demonstrated that performance may be influenced by 3D anatomical training or adaptation to the control mechanism of the system. Stereoscopy improved performance, but did not significantly impact accuracy as expected.

Chapter 3 presented the conceptualization and development of an automation and data management system for running simulator studies. The endeavor makes running complex part-task ventriculostomy studies feasible by eliminating nearly all between-task time resulting from the ecosystem required to support patient-specific scenarios on a closed-source neurosurgical simulator. The system is open-source and can be adapted to a multitude of simulation environments.

Chapter 4 outlined the development of a mobile AR image-guidance system, an abstracted simulator task implemented on the NeuroTouch and automated using the work from Chapter 3, and a novel methodology in order to examine all aspects of human performance around ellipsoid targeting. Ellipsoids are suitable targets for a number of reasons: they offer a unique targeting trajectory, are easy to render, and are conceptually simple to target; all that matters for choosing the perfect trajectory is aligning it with the longest axis. The user study analysis identified a number of factors that influenced subject performance. Subjects performed much better overall using the AR system, but there was a stronger trend to underestimate the depth of the ellipsoid. This is a trend common to AR interfaces as the system is only able to overlay virtual anatomy that should appear projected inside the patient (in this case the mannequin). The magnitude of a user's mean underestimation (Z -bias) was a very strong predictor of targeting

performance, indicating that individuals have different propensity to perceive and target the ellipsoids at their appropriate depth.

Chapter 5 involved the examination of factors that influence performance of simulated part-task ventriculostomies on patient-specific scenarios. The first phase examined performance of neurosurgical subjects of varying experience with and without the use of augmented reality guidance. Trajectories were initially evaluated using an expert graded system. Overall, AR improved targeting time and general performance, but there was no strong difference in targeting accuracy. AR offered the most improvement for more difficult cases (as classified by an expert surgeon). Experience strongly correlated with accuracy (an indicator of construct validity), but only moderately with performance, indicating that procedure time does not decrease with experience. Subject performance correlated with experience. A mild learning curve was observed with the AR approach, indicating that additional training may prove effective for improve subjects' ability to target using AR. The second phase followed up by examining additional cases with traditional preoperative planning only (no AR) using a comprehensive objective evaluative methodology. The engagement measure was able to significantly differentiate between residents and medical novices. In addition, the mean engagement across all users for cases classified as complex was significantly less than for cases classified as simple by an expert surgeon.

The work in this thesis can be summarized as follows:

1. Both anatomical knowledge and spatial reasoning play a role in the use of VR-based anatomical training and simulation scenarios. Immersive VR anatomy trainers may provide advantages in learning complex spatial relations of anatomical features.
2. There is a need for modularizing and streamlining surgical simulators for diverse use and user study management. Our work approached this problem by creating a data management and automation system to make user studies feasible and facilitate safe and efficient collection and analysis of data.

3. AR image-guidance interfaces show potential for improving human performance in targeting tasks, but many implementations suffer from perceptual artefacts that misrepresent depth. AR ellipsoid targeting performance can be predicted by a user's tendency to underestimate depth. Additional training may allow accommodation and improved performance.
4. Surgical simulators continue to show evidence of validity in differentiating surgical skill. Our implementation demonstrated that AR guidance for ventriculostomies is a promising technology that can improve targeting performance in the simulator environment. In addition, we expanded upon ventriculostomy simulator metrics to develop a comprehensive methodology to evaluate trajectories. Analysis of user trajectories indicated that ventricle anatomy influences performance. The use of image-guidance technologies may be better informed by the difficulty of each individual case, as users performed better on difficult cases using AR guidance.

In conclusion, various factors influence surgical performance in a simulation scenario, ranging from patient-specific anatomical variations, to surgical experience and cognitive abilities. Based on the evidence collected examining simulator validity in this thesis and elsewhere, these trends are likely to be generalizable to the clinical setting. As new image-guidance tools and techniques enter the clinic, we must determine the role that they fill; many of these devices are expensive and have steep learning curves, making them ill-suited for many cases. Simulator technology can be utilized to not only demonstrate the performance gains from utilizing such technology, but determine when inherent risks are present that would warrant their use (such as anatomical variations that impact performance).

6.1 Future Directions

The work presented in this thesis provides a foundation for future exploration more than it offers closure to existing research questions; the development of novel tools and methodology for examining surgical education and proficiency in patient-specific contexts creates new branches from existing research problems.

As commodity VR and AR hardware improves, the technical development path becomes clearer as it is driven by the industry. The Microsoft HoloLens is a consumer-level AR device that provides a true see-through display as well as stereoscopic VR rendering. In addition, the device comes with an onboard computer and a complex IR camera system for computer vision tasks. The device can infer its position and orientation in a room and provides built-in functionality for user gestures, providing an out-of-the-box intuitive interface. While still in active development, it is clear that devices such as the HoloLens will have much to offer AR implementations for medical training as well as image-guidance.

While the HoloLens offers significant improvements over existing hardware, the cost may be a barrier for widespread adoption and in some cases a VR interface may be more appropriate than full AR. Recent development efforts and consumer adoption of VR technologies (primarily VR HMDs) has driven the entry price of many devices to negligible amounts. A leading contender is Google Cardboard (as well as third party implementations) which is nothing more than a head mount with lenses that uses a mobile phone to provide the screen and positional/rotational tracking. These devices may be of particular interest to medical institutions with limited resources to allocate for training simulators. In Appendix A, an implementation of a simple ventriculostomy simulator is presented using Google Cardboard with a consumer level mobile phone. The platform makes use of an AR toolkit and image-based markers, using the camera to infer position of the HMD (by default, only orientation is determined using the mobile device's onboard gyroscope).

Machine learning is a technique of regression analysis to build predictive models that may have a role to play in surgical simulation and image-guided interventions. With patient-specific surgical simulations, we can examine human performance as a function of numerous factors. This thesis explored image-guidance and anatomical variations as factors that were demonstrated to influence targeting performance (See chapters 4 and 5). If sufficient data is gathered, predictive models can be trained to determine the difficulty of a given procedure based on preoperatively measured anatomical variations. By objectively classifying surgical cases by difficulty, recommendations for certain

techniques or use of image-guidance can be informed on objective measures. Assuming the surgical simulator for a given procedure reliably simulates the inherent anatomical variations, the predictive models can be trained by recruiting surgeons and residents to perform simulated tasks on a diverse assortment of scenarios.

Bibliography

1. Evgeniou, E., Loizou, P.: Simulation-based surgical education. ANZ journal of surgery. (2012).
2. Kunkler, K.: The role of medical simulation: an overview. The International Journal of Medical Robotics and Computer Assisted Surgery. 2, 203–210 (2006).
3. Kerr, J., Knapp, D., Frake, B., Sellberg, M.: “True” color surface anatomy: mapping the Visible Human to patient-specific CT data. Computerized medical imaging and graphics. 24, 153–164 (2000).
4. Totz, J., Chung, A.J., Yang, G.-Z.: Patient-specific Texture Blending on Surfaces of Arbitrary Topology.
5. Radetzky, A., Nürnberger, A.: Visualization and simulation techniques for surgical simulators using actual patient’s data. Artificial Intelligence in Medicine. 26, 255–279 (2002).
6. Smith, G.C., Pell, J.P.: Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. British Medical Journal. 327, 1459 (2003).

Appendix A: A Mobile VR Simulation Environment

Introduction

Immersive virtual reality (VR) environments are gaining traction in various medical disciplines with applications ranging from teaching anatomy [1-3] to performing complex surgical tasks [4-6]. Although the technology is bridging the gap between VR and physical reality with the vast majority of research focused on improved realism through advanced graphics [7], haptic interfaces [8], tissue deformation [9], and patient-specific simulation [10], many of these developments increase the cost of simulation platforms and make them inaccessible for individuals and many institutions.

It has been noted that one of the barriers to widespread use of current generation simulators is the high initial cost which is often a result of expensive proprietary hardware [6]. Until recently, these approaches were necessitated; consumer-level VR hardware is only now becoming accessible. A number of low cost head-mounted virtual reality devices are now on the market, with additional devices in development. The Google Cardboard is a low cost system that combines optical lenses with a cardboard shell in order to use mobile phones as the foundation for a head-mounted display (HMD). By itself, Google Cardboard lacks a comprehensive interface for human-computer interaction, especially when considering the requirements for medical and surgical simulations. In parallel with consumer-level virtual reality developments, however, there has been a recent drive in the development of commodity interface hardware, particularly implementations making use of gestures and spatial information.

In this appendix, we outline the development of an affordable virtual reality environment for medical simulation making use of a stereoscopic display with head tracking and free form hand interaction. We add positional tracking to Google Cardboard using Vuforia to track an image-based marker that acts as the simulation tabletop platform. Image-based markers are also affixed to a pointing tool in order to track a stylus with 6 degrees of freedom. To evaluate the system, we recruited subjects to perform a simple prehension task involving the localization of points inside ellipsoids within the context of a human head. We analyze the results using a Fitts's law framework.

Background

VR in medicine is not a new concept; systems in various forms have been developed over the years [6,11]. While the advantages of these systems over traditional approaches to learning continue to be examined, several studies have reported successful learning outcomes when compared to traditional approaches such as didactic instruction. Generally, it has been seen that immersive virtual environments can improve skill retention for spatial tasks [12]. In the medical domain, interactive 3D models have been demonstrated to aid in learning skills and knowledge [13]. Additionally, numerous surgical simulators have been demonstrated to reliably assess and train surgical skills [14]. However, the cost and accessibility of such systems remains a pervasive issue.

The costs are associated with technology that aids in the immersivity and interactivity of the virtual environment. Stereoscopic displays, for example, have been demonstrated to provide advantages in prehension tasks [15] as well as numerous additional benefits in the medical domain [16,17]. These displays have been prohibitively expensive until recently, traditionally requiring a costly headset or dual projector system. In addition to the costs, portability was an often overlooked feature, requiring platforms to be setup in fixed locations. Tracking head movements provides additional challenges that required electromagnetic or optical tracking systems [18]. In addition to the visualization system, user interaction provides a number of unique challenges. Some systems make use of haptic tools, sacrificing range of motion for realistic physical feedback [19], while other systems use free hand interactions to allow for greater flexibility without the sensation of touch [20].

Various systems in the past have attempted to reduce the economic entry barrier to immersive virtual reality. In 2008, an economical VR system for education was developed which included spatial input, stereoscopic vision and head tracking for the cost of \$4000 USD [21]. The system used numerous custom components and an expensive electromagnetic tracking system. While more accessible than most implementations, the system is still not within reach of most end consumers, especially considering the required domain knowledge for setting up the system in a classroom or home environment.

In recent years, there has been a growing effort for commercialization of commodity virtual reality hardware. Though not officially released, the Oculus Rift has been at the forefront of the movement, providing a high resolution stereoscopic head mounted display with orientation tracking using a high performance gyroscope. Google's contender, the Google Cardboard, makes use existing mobile phone hardware, enabling a properly equipped smart phone to become a virtual reality platform with a stereoscopic display and orientation tracking using the phone's gyroscope. While the Oculus Rift provides superior performance, it comes at a much steeper price and requires a separate computing platform in addition to the head mount. As far as we know, a full surgical simulation platform has yet to be implemented using Google Cardboard, but mobile phones are beginning to be targeted as economic platforms for medical simulation [22].

Methodology

The methodology is divided into two sections. The first section discusses general development of the system, while the second section discusses the setup of the user study to evaluate performance on the system.

System Description

The requirements of the system included a head mounted stereoscopic display, head tracking (both position and orientation) and a free form spatial input that allows the user to select positions in space using a virtual tool as a 3D cursor as well as trajectories. Google Cardboard in conjunction with a mobile phone was chosen as the ideal platform for our head mounted display due to the low price and accessibility of the product (also considering the mass availability of powerful mobile phones). By default, Google Cardboard provides stereoscopic rendering using the Cardboard SDK as well as orientation tracking using the gyroscope of the embedded mobile phone. While phone gyroscopes are low performance in relation to other head mounted displays, medical simulations don't require low latency tracking for rapid head movements. In order to add position tracking to the system, we make use of the phone's camera to track image-based markers that act as the tabletop of the virtual environment. This is enabled through the use of the augmented reality SDK Vuforia. All development is done on a Windows

platform using the Unity 5 game development engine. Unity was chosen as it integrates seamlessly with the Cardboard SDK as well as the Vuforia SDK and provides a graphical environment which simplifies prototyping and development. The Android platform was targeted in our initial developments (specifically Samsung Galaxy Alpha and Samsung Nexus 5 phones), but Unity's cross-platform capabilities make portability to iOS platforms possible with minor changes.

By default, all input to the app through the Cardboard SDK comes from the orientation of the headset or the pressing of a button on the device which taps the phone's touchscreen. These options are somewhat limiting for medical simulations that involve exploring complex anatomy and performing targeting tasks. To provide greater freedom in the human-computer interface, we used image-based marker tracking to track a stylus with 6 degrees of freedom for pointing and selection tasks. In our initial implementation, a pen is affixed to a flat, double sided marker, which is used to infer the orientation of the pen or the location of the tip. Our initial approach involved the integration of the Leap Motion hand tracking hardware, but mobile processing limitation prohibited effective use. See Figure 40 for an example of the system in use.



Figure 40: The mobile system depicted in its entirety (mobile device obstructed by headset).

User Study

The user study was designed as an abstract task to examine targeting performance as well as look at user perception of usability in order to determine the future direction of development.

Many surgical procedures can be decomposed into simplified tasks that require positional targeting or selection of a trajectory within anatomical contexts. In our examination of users' ability to target structures within the system, we abstracted the targeting of anatomy to generic shapes. Specifically, we examined the performance of users in localizing a point within ellipsoids of various shapes and positions within a transparent human head derived from an existing surgical simulator.

A total of 8 participants were recruited through online advertisement. All participants (by self-report) were in good physical and mental health and were free from

any impairments that would impact their effectiveness in using our immersive virtual environment. Each participant performed 15 ellipsoid targeting tasks on unique ellipsoids and then repeated the tasks in the same order for a total of 30 point localizations. Ellipsoids were created as graphical meshes using the open source modeling tool Blender 3D. Participants were instructed to maximize accuracy (correctly placing the tool tip within each ellipsoid) while minimizing the task time. For each task, the success of the placement was recorded along with the time.

Performance was examined using a Fitts's law framework [23]. We explored the relationship between target size (volume of the ellipsoid calculated using MeshLab), user accuracy and targeting time.

A Likert scale based questionnaire was used in order to obtain qualitative feedback from participants regarding use of the system. The scale is based on work by Witmer and Singer in order to determine user's sense of presence within the virtual environment [24]. The survey used is shown below:

1. How much were you able to control events?

NOT AT ALL		SOMEWHAT			COMPLETELY

2. How responsive was the environment to actions that you initiated (or performed)?

NOT RESPONSIVE		MODERATELY RESPONSIVE			COMPLETELY RESPONSIVE

3. How natural did your interactions with the environment seem?

EXTREMELY ARTIFICIAL		BORDERLINE NATURAL			COMPLETELY

4. How much did the visual aspects of the environment involve you?

NOT AT ALL		SOMEWHAT			COMPLETELY

5. How natural was the mechanism which controlled movement through the environment?

EXTREMELY ARTIFICIAL		BORDERLINE NATURAL			COMPLETELY

6. How compelling was your sense of objects moving through space?

NOT AT ALL COMPELLING		MODERATELY COMPELLING			VERY

7. How much did your experiences in the virtual environment seem consistent with your real world experiences?

_____	_____	_____	_____	_____	_____
NOT		MODERATELY		VERY	
CONSISTENT		CONSISTENT		CONSISTENT	

8. Were you able to anticipate what would happen next in response to the actions that you performed?

_____	_____	_____	_____	_____
NOT AT ALL		SOMEWHAT		COMPLETELY

9. How completely were you able to actively survey or search the environment using vision?

_____	_____	_____	_____	_____
NOT AT ALL		SOMEWHAT		COMPLETELY

10. How compelling was your sense of moving around inside the virtual environment?

_____	_____	_____	_____	_____
NOT		MODERATELY		VERY
COMPELLING		COMPELLING		COMPELLING

11. How closely were you able to examine objects?

_____	_____	_____	_____	_____
NOT AT ALL		PRETTY		VERY
		CLOSELY		CLOSELY

12. How well could you examine objects from multiple viewpoints?

_____	_____	_____	_____	_____
NOT AT ALL		SOMEWHAT		EXTENSIVELY

13. How involved were you in the virtual environment experience?

_____	_____	_____	_____	_____
NOT		MILDLY		COMPLETELY
INVOLVED		INVOLVED		ENGROSSED

14. How much delay did you experience between your actions and expected outcomes?

_____	_____	_____	_____	_____
NO DELAYS		MODERATE		LONG
		DELAYS		DELAYS

15. How quickly did you adjust to the virtual environment experience?

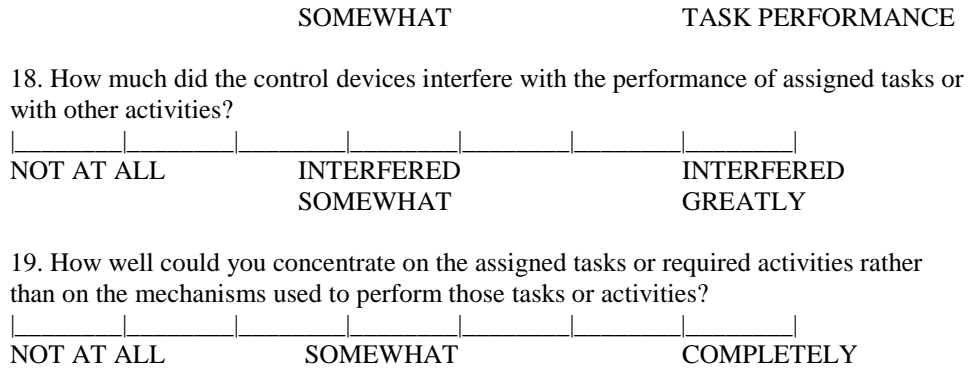
_____	_____	_____	_____	_____
NOT AT ALL		SLOWLY		LESS THAN
				ONE MINUTE

16. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?

_____	_____	_____	_____	_____
NOT		REASONABLY		VERY
PROFICIENT		PROFICIENT		PROFICIENT

17. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

_____	_____	_____	_____	_____
NOT AT ALL		INTERFERED		PREVENTED



Results

For a screenshot of the system in use, see Figure 41.

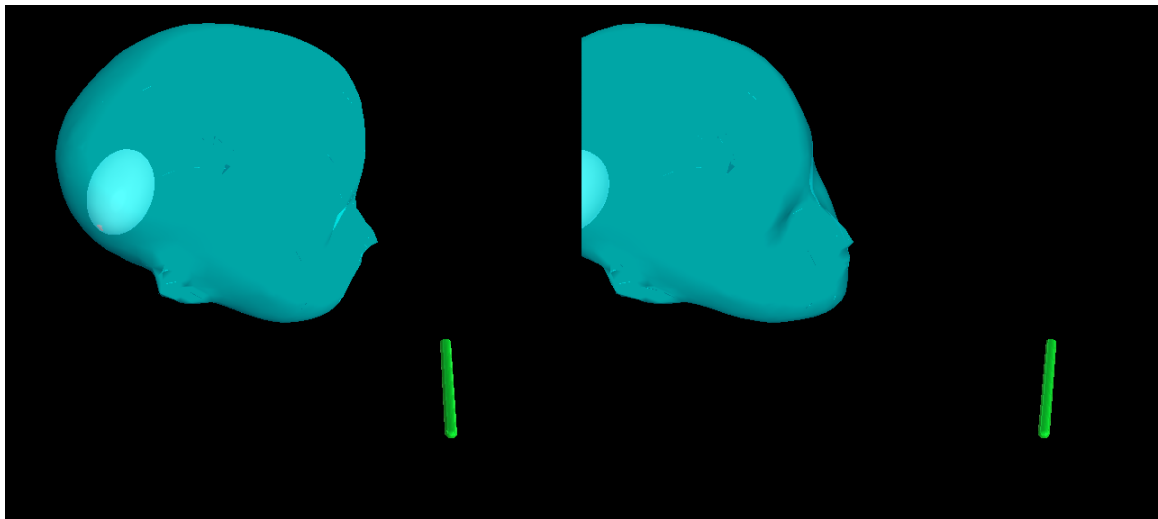


Figure 41: A screenshot of the mobile simulator in use.

For all users, the overall targeting accuracy was 63% and the average targeting time was 23.8 seconds. There were no clear relationships between targeting accuracy, task time and ellipsoid size. No clear learning trends were observed. There was high between-user variance, but low intra-user variance for task time, indicating that ellipsoid size and placement did not significantly affect task time, which was more reliant on each user’s individual approach. When examining the data for a speed-accuracy trade-off, no clear trend could be seen.

No strong opinions were observed by the questionnaire. On a scale ranging from 1 to 7, where 7 is the most positive outcome, users felt that the sense of objects moving through space was compelling (mean of 5.25), that the input mechanism was natural (mean of 5.13) and that the visual aspects of the environment were immersive (mean of 5.12). Users only somewhat felt that they were able to actively search the environment using vision (mean of 3.63). Additional comments were made regarding the robustness of the tracking system. Some users felt that the interruptions due to marker occlusion disrupted the immersive experience and negatively affected their performance.

Discussion

Results from this initial study indicate that the system is functional and immersive, but further developments in usability and robustness are required. Contrary to expectations, no significant relationship was seen between ellipsoid size and user performance. In a number of cases, marker tracking was disrupted and users had to realign their view in order for the rendering and tracking to resume. This may have been a source of noise in the data, resulting in unclear results. Additionally, there may have not been enough variation in position and size of the ellipsoids; as we intend to extend the system further into the medical domain, we restricted the task to an area confined within a to-scale head, which constrained the task. As we further develop the system, we hope to improve the precision of the tracking to allow targeting of smaller structures.

One of the issues noted by participants was the occasional loss of marker tracking, resulting in an untracked target or tool. Using multiple image-based markers on a tabletop platform makes occlusion difficult to avoid; movement of the tool often obstructed the camera's view. There were also issues with continuing to hold the tool in such a way that the camera could continuously view the image. The tool had to be held in such a way that the image was always facing the camera and did not allow for much change in the rotation of the users' hand. This meant that some of the time, the tool has to be held in an orientation that did not feel natural to the movement that was being performed. In addition, the user's hands and arms would often occlude the targets. Use of a device such as the Leap Motion would avoid this complication, but we found that the increased computational demands were too great for current generation mobile platforms. In future

developments, we will explore the use of haptic devices integrated with mobile platforms. While this will increase the cost of the simulator platform, it will greatly improve the usability and will still be relatively inexpensive and portable. An additional solution to the marker tracking and some perceptual and usability issues will be the inclusion of augmented reality contextualization information. McGill et al. found that provided an augmented reality reference helped overcome some of these drawbacks [25].

A common drawback attributed to various implementations of stereoscopy is the fatigue from eyestrain or headaches [26], but no user reported fatigue of any form. This is likely due to the absence of fatigue-inducing triggers such as bulky eyewear, quickly moving objects, and poor rendering [27]. However, participants did not spend an excessive amount of time using the system; the longest time spent using the system by any participant was around 20 minutes.

In addition to system improvements, future studies will aim to recruit additional subjects as the sample size in this study is a clear shortcoming. The task will also be extended to examine targeting accuracies of simulated ventriculostomy part-tasks to evaluate the potential of the tool for use in surgical training.

Conclusion

In this study, we were able to develop a cost-effective and portable simulation system with potential for application in surgical training and medical education. The results indicate that the system is usable but requires further developments to improve overall robustness.

Bibliography

1. Nicholson, D.T., Chalk, C., Funnell, W.R.J., Daniel, S.J.: Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. *Medical Education*. 40, 1081–1087 (2006).
2. Plumley, L., Armstrong, R., De Ribaupierre, S., Eagleson, R., others: Spatial ability and training in virtual neuroanatomy. *Studies in health technology and informatics*. 184, 324–329 (2012).
3. Armstrong, R., de Ribaupierre, S., Eagleson, R.: A software system for evaluation and training of spatial reasoning and neuroanatomical knowledge in a virtual environment. *Comput Methods Programs Biomed*. 114, 29–37 (2014).
4. Clarke, D.B., D'Arcy, R.C.N., Delorme, S., Laroche, D., Godin, G., Ghosh Hajra, S., Brooks, R., DiRaddo, R.: Virtual Reality Simulator: Demonstrated Use in Neurosurgical Oncology. *Surgical Innovation*. (2012).
5. Tavakol, M., Mohagheghi, M.A., Dennick, R.: Assessing the skills of surgical residents using simulation. *Journal of surgical education*. 65, 77–83 (2008).
6. Chan, S., Conti, F., Salisbury, K., Blevins, N.H.: Virtual reality simulation in neurosurgery: technologies and evolution. *Neurosurgery*. 72, A154–A164 (2013).
7. Oishi, M., Fukuda, M., Ishida, G., Saito, A., Hiraishi, T., Fujii, Y.: Presurgical simulation with advanced 3-dimensional multifusion volumetric imaging in patients with skull base tumors. *Neurosurgery*. 68, ons188 (2011).
8. Shamsunder, S., Manivannan, M., others: Haptic guided laparoscopy simulation improves learning curve. *Studies in health technology and informatics*. 132, 454 (2008).

9. Kim, J., Ahn, B., De, S., Srinivasan, M.A.: An efficient soft tissue characterization algorithm from in vivo indentation experiments for medical simulation. *The international journal of medical robotics and computer assisted surgery*. 4, 277–285 (2008).
10. Soler, L., Marescaux, J.: Patient-specific surgical simulation. *World journal of surgery*. 32, 208–212 (2008).
11. Satava, R.M., Jones, S.B.: Current and future applications of virtual reality for medicine. *Proceedings of the IEEE*. 86, 484–489 (1998).
12. Waller, D., Miller, J.: A desktop virtual environment trainer provides superior retention of a spatial assembly skill. *CHI 98 Conference Summary on Human Factors in Computing Systems*. pp. 339–340 (1998).
13. Kockro, R.A., Hwang, P.Y.: Virtual temporal bone: an interactive 3-dimensional learning aid for cranial base surgery. *Neurosurgery*. 64, ons216–ons230 (2009).
14. Dunkin, B., Adrales, G., Apelgren, K., Mellinger, J.: Surgical simulation: a current review. *Surgical endoscopy*. 21, 357–366 (2007).
15. Melmoth, D.R., Grant, S.: Advantages of binocular vision for the control of reaching and grasping. *Experimental brain research*. 171, 371–388 (2006).
16. Van Beurden, M.H.P.H., van Hoey, G., Hatzakis, H., Ijsselsteijn, W.A.: Stereoscopic displays in medical domains: A review of perception and performance effects. *SPIE Electronic Imaging*. p. 72400A–72400A (2009).
17. Held, R.T., Hui, T.T.: A guide to stereoscopic 3D displays in medicine. *Academic radiology*. 18, 1035–1048 (2011).

18. Abhari, K., Baxter, J.S., Chen, E.S., Khan, A.R., Wedlake, C., Peters, T.M., de Ribaupierre, S., Eagleson, R.: Use of a Mixed-Reality System to Improve the Planning of Brain Tumour Resections: Preliminary Results. *Augmented Environments for Computer-Assisted Interventions*. pp. 55–66. Springer (2013).
19. Gélinas-Phaneuf, N., Choudhury, N., Al-Habib, A.R., Cabral, A., Nadeau, E., Mora, V., Pazos, V., Debergue, P., DiRaddo, R., Del Maestro, R.F.: Assessing performance in brain tumor resection using a novel virtual reality simulator. *International journal of computer assisted radiology and surgery*. 9, 1–9 (2014).
20. Cooperstock, J.R., Wang, G.: Stereoscopic display technologies, interaction paradigms, and rendering approaches for neurosurgical visualization. *Proceedings of SPIE on Stereoscopic Displays and Applications XX*. 7237, (2009).
21. Adams, J.C., Hotrop, J.: Building an economical VR system for CS education. *ACM SIGCSE Bulletin*. pp. 148–152 (2008).
22. Kramers, M., Armstrong, R., Bakhshmand, S.M., Fenster, A., de Ribaupierre, S., Eagleson, R.: A Mobile Augmented Reality Application for Image Guidance of Neurosurgical Interventions. *American Journal of Biomedical Engineering*. 3, 169–174 (2013).
23. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*. 47, 381 (1954).
24. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*. 7, 225–240 (1998).
25. McGill, M., Boland, D., Murray-Smith, R., Brewster, S.: *A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays*. (2015).

26. Van Bergen, P., Kunert, W., Bessell, J., Buess, G.: Comparative study of two-dimensional and three-dimensional vision systems for minimally invasive surgery. *Surgical endoscopy*. 12, 948–954 (1998).
27. Lambooij, M.T., IJsselsteijn, W.A., Heynderickx, I.: Visual discomfort in stereoscopic displays: a review. *Electronic Imaging 2007*. p. 64900I–64900I (2007).

Appendix B: A Pipeline for Creating Patient-Specific ETV Scenarios

This section is adapted from “Patient-Specific Pipeline to Create Virtual Endoscopic Third Ventriculostomy Scenarios” [1].

Introduction

We propose to conceptualize and implement a pipeline for patient-specific simulation of a specific neurosurgical procedure: endoscopic third ventriculostomy (ETV) – one of two treatments offered for hydrocephalus. ETV is a prime candidate for simulation as it involves a number of surgical steps which have relatively straightforward geometries, and few degrees of freedom for the surgical tool. Despite this, due to lack of training, there are significant reports of complications and failures [2]. As it is an endoscopic procedure, the constrained space simplifies implementation of an interface compared to open surgeries. ETV is also a good candidate for patient-specific simulation due to high inter-patient geometrical variation [3].

Currently, no system exists that provides a pipeline for patient-specific content creation and simulation for the ETV procedure. The current state-of-the-art platform for the ETV procedure simulation is the commercial NeuroTouch software developed by the National Research Council of Canada [4] – research collaborators on this project. While the system provides a realistic rendering and deformation engine coupled with a natural haptic interface, they are restricted to a single generic ETV scenario partially constructed from imaging data and partially synthetic. Other patient-specific ETV systems have been proposed [5,6], but development has either been discontinued or the proposals have not come to fruition. A number of techniques have been developed for patient-specific virtual endoscopy [7-10], but the amount of interaction offered is sufficient only for exploration and is not suitable for practicing surgical skills. Additionally, patient-specific surgical scenarios are employed by simulators for other procedures [11-15], but the vast differences in implementations of the content creation pipelines and simulations themselves permits little overlap.

The purpose of this work is to develop the user-centric tools required to create patient-specific scenarios for use in ETV simulations. Although many integrated platforms exist for visualization and segmentation of medical data, there exists no solution that is optimized for the pipeline required by this procedure. Additionally, the majority of these tools have associated learning curves that are too large to allow for seamless incorporation into clinical and educational institutions. Indeed, many target end-users are clinicians and educators who are likely unfamiliar with image segmentation and processing algorithms.

Methodology

User-centricity is the primary concern in the development of the platform – facilitating accurate and timely content creation. For such a tool to be adopted, it must be accessible to users of all skill levels. Additionally, it must be efficient and require minimal user guidance. If it is too time-consuming of a process to create a patient-specific scenario, the platform will not fulfill its role of providing surgeons the means to pre-surgical rehearsal. It will also limit the possibility of creating a large wealth of scenarios for trainees. As such, the platform was developed as a software wizard, where users are presented with a sequence of steps in order to produce the required output through a number of well-defined tasks. The main task involved is segmentation of medical image data (often MR or CT images) in order to create the geometric meshes for rendering as well as the finite element meshes for simulation of soft-tissue deformation; however, the user must also guide the software in aligning the final meshes and texturing the meshes appropriately. Figure 42 illustrates this pipeline. The final output is constructed for incorporation into the NeuroTouch ETV simulation, which allows for user-defined scenarios.

After each stage in the wizard, the user must be able verify that the previous task was completed correctly. Due to the high inter-patient anatomical variability inherent to hydrocephalus, atlas-based comparison and validation is insufficient to determine that the appropriate segmentation was performed. This is where user knowledge is necessary. The user must have sufficient knowledge of the anatomy involved in order to determine if the structure is correctly segmented or if the task must be repeated. The visualization

platform is used to allow the user to easily verify that the anatomy is correctly segmented. The visualization is discussed in the following section.

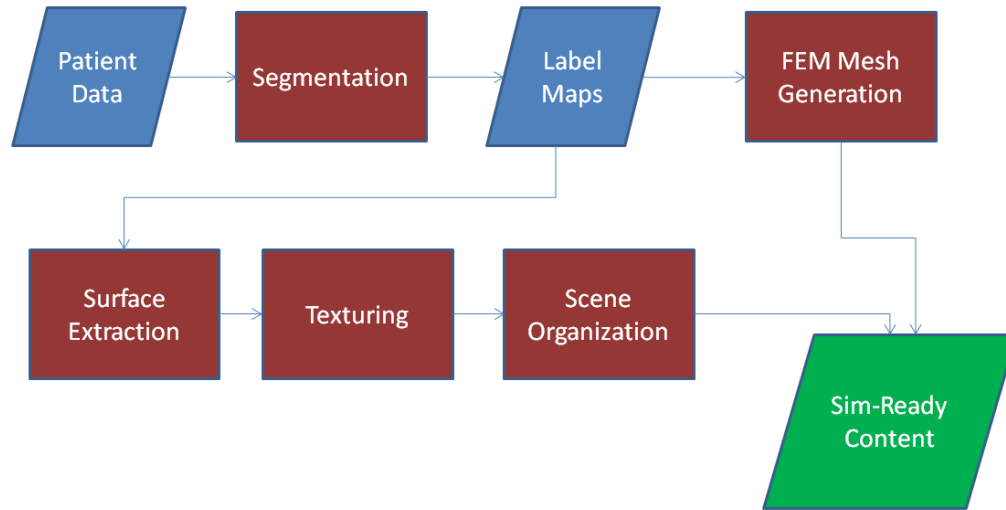


Figure 42: Pipeline for creation of patient-specific content from raw medical data.

Visualization

To facilitate user validation of segmented structures, a custom visualization module was developed for use in the pipeline. This is our primary contribution to user centricity for the required task. Rather than constructing isosurfaces from segmented volumes, we decided to use a direct volume rendering approach. Not only does this allow for much faster processing times (distinct from rendering time), but a novel implementation allows for a descriptive scene. While T1-weighted MR images are most often acquired preoperatively for the ETV procedure, other imaging modalities may be available and may be useful in creation of the required content. The system was developed with multimodality imaging in mind. Figure 43 illustrates a typical rendering.

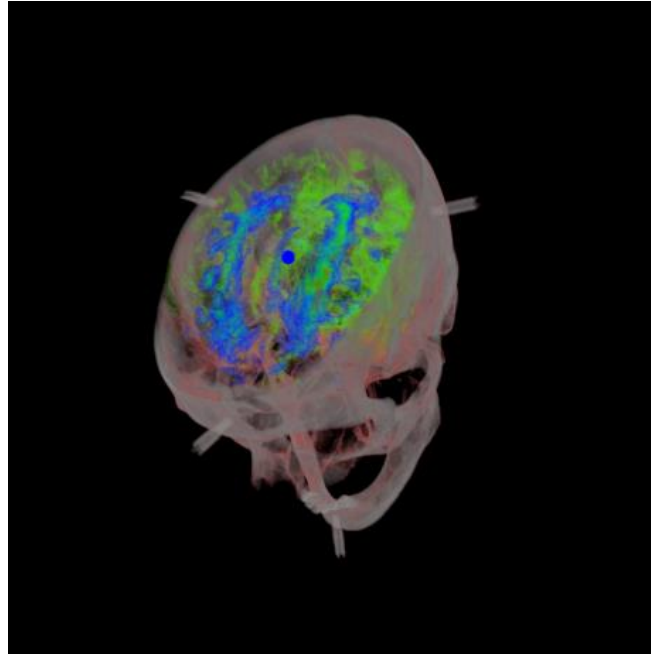
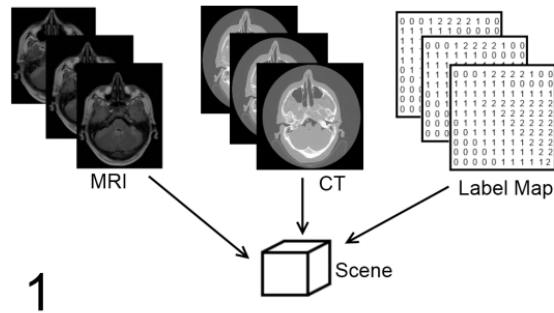


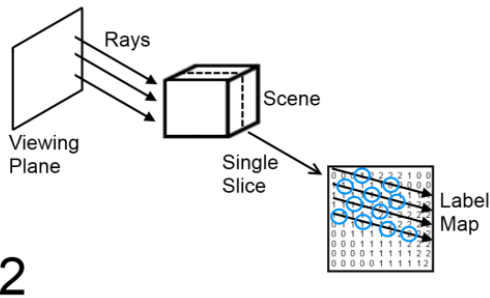
Figure 43: Rendering of skull from CT with encased lateral ventricles and grey matter. The central blue circle represents user input.

Our algorithm is implemented using GPU-based ray casting with the CUDA application programming interface. At initialization, the required volumetric images are loaded into video memory as textures, providing there is sufficient memory. Empty space skipping and early ray termination are both implemented as optimizations in addition to dynamic sampling dependent on user views. Sampling step size, maximum steps and opacity threshold for early ray termination are all user-adjustable (though automated for simplicity). Multiple modalities must be pre-registered prior to initialization and are aligned in normalized 3D texture space (0, 1) along each axis. In addition to the medical images represented using texture memory, a segmentation label map is also used with resolution of the highest resolution image in the scene. Each voxel in the map is an integer value that corresponds to a structure in the scene. During rendering, rays cast through the volume sample the value in the map in order to determine which image volume to sample from and which transfer function to use. Single dimensional transfer functions are generated automatically, with colours spread over the distribution of intensity values within the region. With transparency, this provides an advantage over opaque isosurfaces, as it allows the user to see strong variations in the intensity values

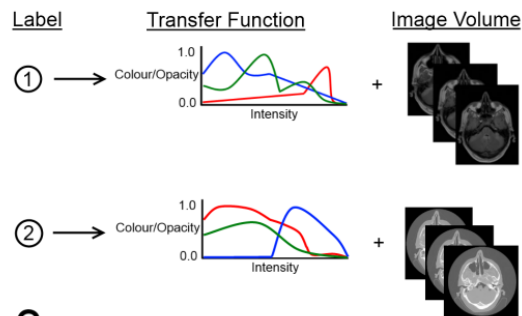
throughout the region – information lost in surface extractions. This information gives insight into areas where the segmentation algorithm used may have produced incorrect results. The transfer functions can also be set to opaque to examine the structure of the segmented regions. Figure 44 illustrates the entire rendering process.



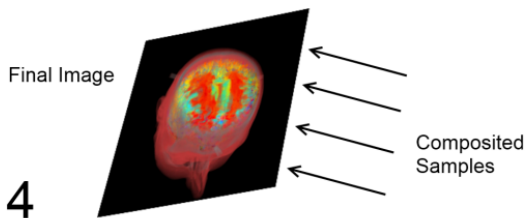
1



2



3



4

Figure 44: Rendering workflow. 1 – Images and label map are co-registered within the scene. 2 – Rays are cast into the scene for sampling. 3 – The label at each sample directs the algorithm to a transfer function and associated image. 4 – the result is composited to the viewing plane.

Segmentation

A separate segmentation stage was created for each anatomical feature required, as separate algorithms and parameters were needed for each. The following anatomical features were deemed necessary by experts to include in each simulation: The lateral and third ventricles, the fornix, the choroid plexus, and the basilar artery. A difficulty immediately evident is that the choroid plexus and (often) the basilar artery cannot be extracted from preoperative medical images. In this case, the user must place generic data into the scene in order to finalize it. While this effectively reduces the amount of patient-specific content in the scenario, the location of the choroid plexus and basilar artery can often be inferred from the geometry of the ventricular system. In any case, the majority of the procedure involves navigating the ventricles using spatial cues based on the geometry of the ventricular anatomy.

The segmentation workflow varies from structure to structure. For each, the user is instructed on where they must localize points within a slice-based view of the volume. Parameters are auto-tuned for the optimal result. After the required input is collected, the algorithm proceeds and immediately displays results to the user through the visualization. If the segmentation is satisfactory, the user can proceed to the next structure, otherwise the process must be repeated while fine-tuning the parameters of the specific algorithm. Guidelines are given on what is required to achieve a desired result. The Insight Segmentation and Registration Toolkit [16] is integrated into the platform to support a number of standard segmentation algorithms.

For the ventricular system, a seeded region growing algorithm is sufficient for the overall structure of the ventricles, with a live-wire algorithm being required for more fine detail in connecting the third ventricle to the lateral ventricles. A live-wire based algorithm was also chosen to segment the fornix. Though an appropriate segmentation can often easily be obtained, more work must be done examining the algorithms on images of varying quality as well as patients with abnormal geometry.

Once all of the required structures in the scene have been segmented and the user is satisfied with the results, they may proceed to extract the isosurfaces and finite element meshes required for the simulation.

Scene Construction

Before the content can be utilized in the simulation, it must be finalized. While future work will develop a custom platform for this task this is currently performed using a number of open source tools. Here, the user must import the final geometric models and ensure that all of the meshes make contact within the scene. If there are gaps between structures due to under-segmentation, it must be evaluated whether the structures must be re-segmented or if this can be corrected for simply by translating the anatomy in the scene as long as minimal structure displacement occurs. Additionally (if desired), the user must place the choroid plexus as well as the basilar artery into the scene at the correct locations. To finalize the scene, textures must be applied to the models. Generic vasculature and flesh textures can be used for simplicity, or intra-operative images can be mapped onto the geometry for increased realism. Once these steps are complete, the scene is ready to be imported into a custom scenario for simulation.

Conclusion and Future Work

While the design focus has emphasized user centricity, the ability of users on our platform to create content must be evaluated in terms of speed and accuracy compared to other generic platforms.

A large component of future work lies in validation of the segmentation and creation of content, especially if the tools are to be used to create pre-surgical rehearsal simulations. For example, while the process may lead to the creation of a realistic scenario, it may not necessarily accurately represent the patient. By rehearsing on such a simulation, the surgeon may be given an unrealistic representation of the patient and, as a result, make incorrect assumptions about the procedure prior to the operation. In addition to examining face validity, we plan on developing objective metrics to evaluate the ability of our scenarios to accurately represent the patients they are created from. This is

particularly important in evaluating the use of non-patient-specific information used within the scene, such as structures that cannot be imaged using current clinical techniques. An additional limitation exists in ensuring that segmented tissue boundaries lay in mutual contact. A fuzzy segmentation approach will be implemented in the future to overcome this issue.

Future work will also allow us to examine how patient-specific simulation impacts performance of trainees who will have access to a greater wealth of training scenarios, as well as surgeons who will have access to rehearsal. We can also examine how specific variations in patient anatomy affect the performance of the procedure to better determine where the focus should be placed on training for the procedure.

Bibliography

1. Armstrong, R., Eagleson, R., Ribaupierre, S.D.: Patient-Specific Pipeline to Create Virtual Endoscopic Third Ventriculostomy Scenarios. *Medicine Meets Virtual Reality 21: NextMed/MMVR21*. 196, 14 (2014).
2. Hellwig, D., Grotenhuis, J.A., Tirakotai, W., Riegel, T., Schulte, D.M., Bauer, B.L., Bertalanffy, H.: Endoscopic third ventriculostomy for obstructive hydrocephalus. *Neurosurgical Review*. 28, 1–34 (2005).
3. Yadav, Y.R., Parihar, V., Pande, S., Namdev, H., Agarwal, M.: Endoscopic third ventriculostomy. *Journal of Neurosciences in Rural Practice*. 3, 163 (2012).
4. Delorme, S., Laroche, D., DiRaddo, R., Del Maestro, R.F.: NeuroTouch: A Physics-Based Virtual Simulator for Cranial Microneurosurgery Training. *Neurosurgery*. 71, 32–42 (2012).
5. Tang, C.Y., Chin, W., Chui, Y.P., Poon, W.S., Heng, P.A.: A Virtual Reality-based Surgical Simulation System for Virtual Neuroendoscopy. *Integration Technology, 2007. ICIT'07. IEEE International Conference on*. pp. 253–258 (2007).
6. Brown, N., Natsupakpong, S., Johannsen, S., Manjila, S., Cai, Q., Liberatore, V., Cohen, A.R., Cavusoglu, M.C.: Virtual environment-based training simulator for endoscopic third ventriculostomy. *Studies in health technology and informatics*. 119, 73 (2005).
7. Nakajima, N., Wada, J., Miki, T., Haraoka, J., Hata, N.: Surface rendering-based virtual intraventricular endoscopy: Retrospective feasibility study and comparison to volume rendering-based approach. *Neuroimage*. 37, S89–S99 (2007).
8. Schreyer, A.G., Herfarth, H., Kikinis, R., Seitz, J., Schölmerich, J., Geissler, A., Feuerbach, S.: 3D modeling and virtual endoscopy of the small bowel based on

- magnetic resonance imaging in patients with inflammatory bowel disease. *Investigative radiology*. 37(9), 528-33 (2016).
9. Auer, D.P., Sendtner, P., Schneider, G., Auer, L.M.: Evaluation of virtual endoscopy for application in clinical neurosciences. *International Congress Series*. pp. 141–151 (2002).
 10. Burtscher, J., Dessl, A., Bale, R., Eisner, W., Auer, A., Twerdy, K., Felber, S.: Virtual endoscopy for planning endoscopic third ventriculostomy procedures. *Pediatric Neurosurgery*. 32, 77–82 (2000).
 11. Krombach, G., Rohde, V., Haage, P., Sruffert, T., Kilbinger, M., Thron, A.: Virtual endoscopy combined with intraoperative neuronavigation for planning of endoscopic surgery in patients with occlusive hydrocephalus and intracranial cysts. *Neuroradiology*. 44, 279–285 (2002).
 12. Willaert, W.I.M., Aggarwal, R., Van Herzeele, I., Cheshire, N.J., Vermassen, F.E.: Recent Advancements in Medical Simulation: Patient-Specific Virtual Reality Simulation. *World journal of surgery*. 1–10 (2012).
 13. Soler, L., Marescaux, J.: Patient-specific surgical simulation. *World journal of surgery*. 32, 208–212 (2008).
 14. Willaert, W., Aggarwal, R., Bicknell, C., Hamady, M., Darzi, A., Vermassen, F., Cheshire, N.: Patient-specific simulation in carotid artery stenting. *Journal of Vascular Surgery*. 52, 1700–1705 (2010).
 15. Sierra, R., Dimaio, S., Wada, J., Hata, N., Székely, G., Kikinis, R., Jolesz, F., others: Patient specific simulation and navigation of ventriculoscopic interventions. *Studies in health technology and informatics*. 125, 433 (2006).

16. Clarke, D.B., D'Arcy, R.C.N., Delorme, S., Laroche, D., Godin, G., Ghosh Hajra, S., Brooks, R., DiRaddo, R.: Virtual Reality Simulator: Demonstrated Use in Neurosurgical Oncology. *Surgical Innovation*. (2012).

17. Yoo, T.S., Ackerman, M.J., Lorensen, W.E., Schroeder, W., Chalana, V., Aylward, S., Metaxas, D., Whitaker, R.: Engineering and algorithm design for an image processing api: a technical report on itk-the insight toolkit. *Studies in health technology and informatics*. 586–592 (2002).

Appendix C: Published Works Reprint Permissions

Chapter 2:

For the paper ‘A Software System for Evaluation and Training of Spatial Reasoning and Neuroanatomical Knowledge in a Virtual Environment’ [116], Elsevier permits content reuse rights to authors for theses.

For the paper ‘Spatial Ability and Training in Virtual Neuroanatomy’ [117], IOS press does not retain not retain copyright; it belongs to the author for reuse under any circumstance.

Chapter 4:

For the paper ‘Augmented reality for Neurosurgical Guidance: An Objective Comparison of Planning Interface Modalities’ [147], permission from Springer to reprint is granted under License #4004890165551.

For the paper ‘Evaluation of a Mobile Augmented Reality Application for Image-guidance of Neurosurgical Interventions’ [148], IOS press does not retain not retain copyright; it belongs to the author for reuse under any circumstance.

For the paper ‘A Mobile Augmented Reality Application for Image-guidance of Neurosurgical Interventions’ [143], Sage grants an exclusive license, allowing authors to reuse works for any purpose.

Chapter 5:

For the paper ‘An Examination of Factors that Impact Performance on a Patient-Specific VR-Based Ventriculostomy Simulator’, Sage grants an exclusive license, allowing authors to reuse works for any purpose.

Appendix B:

For the paper ‘Patient-Specific Pipeline to Create Virtual Endoscopic Third Ventriculostomy Scenarios’ [214], IOS press does not retain not retain copyright; it belongs to the author for reuse under any circumstance.

Permission References

1. Armstrong, R., de Ribaupierre, S., Eagleson, R.: A software system for evaluation and training of spatial reasoning and neuroanatomical knowledge in a virtual environment. *Comput Methods Programs Biomed.* 114, 29–37 (2014).
2. Plumley, L., Armstrong, R., De Ribaupierre, S., Eagleson, R.: Spatial ability and training in virtual neuroanatomy. *Stud Health Technol Inform.* 184, 324–9 (2013).
3. Armstrong, R., Wright, T., Eagleson, R., de Ribaupierre, S.: Augmented reality for Neurosurgical Guidance: An Objective Comparison of Planning Interface Modalities. *Lecture Notes in Computer Science.* 9805, (2016).
4. Kramers, M., Armstrong, R., Bakhshmand, S.M., Fenster, A., de Ribaupierre, S., Eagleson, R.: Evaluation of a mobile augmented reality application for image guidance of neurosurgical interventions. *Stud Health Technol Inform.* 196, 204–8 (2014).
5. Kramers, M., Armstrong, R., Bakhshmand, S.M., Fenster, A., de Ribaupierre, S., Eagleson, R.: A Mobile Augmented Reality Application for Image Guidance of Neurosurgical Interventions. *American Journal of Biomedical Engineering.* 3, 169–174 (2013).
6. Armstrong, R., Eagleson, R., RIBAUPIERRE, S.D.: Patient-Specific Pipeline to Create Virtual Endoscopic Third Ventriculostomy Scenarios. *Medicine Meets Virtual Reality 21: NextMed/MMVR21.* 196, 14 (2014).

Appendix D: Research Ethics Approval

Research involving human subjects was guided and performed under the following approvals.



Use of Human Participants - Ethics Approval Notice

Principal Investigator: Sandrine de Ribaupierre
 Review Number: 17888E
 Review Level: Delegated
 Approved Local Adult Participants: 20
 Approved Local Minor Participants: 0
 Protocol Title: Realistic Computer Graphics Rendering and Task Analysis for Neurosurgery Simulation – Neuroendoscopy Project
 Department & Institution: Clinical Neurological Sciences, University of Western Ontario
 Sponsor:
 Ethics Approval Date: March 24, 2011 Expiry Date: June 30, 2014
 Documents Reviewed & Approved & Documents Received for Information:

Document Name	Comments	Version Date
UWO Protocol		
Letter of Information & Consent	Assent Letter	
Letter of Information & Consent		

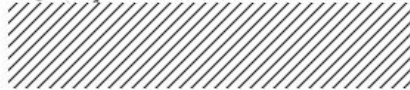
This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research

Involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/ICH Good Clinical Practice Practices: Consolidated Guidelines, and the applicable laws and regulations of Ontario has reviewed and granted approval to the above referenced revision(s) or amendment(s) on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.

The ethics approval for this study shall remain valid until the expiry date noted above assuming timely and acceptable responses to the HSREB's periodic requests for surveillance and monitoring information. If you require an updated approval notice prior to that time you must request it using the UWO Updated Approval Request Form.

Members of the HSREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the HSREB.

The Chair of the HSREB is Dr. Joseph Gilbert. The UWO HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.



Ethics Officer to Contact for Further Information

<input type="checkbox"/> Janice Sutherland (jsutherland@uwo.ca)	<input type="checkbox"/> Elizabeth Wambolt (ewambolt@uwo.ca)	<input checked="" type="checkbox"/> Grace Kelly (grace.kelly@uwo.ca)
--	---	---

This is an official document. Please retain the original in your files.

The University of Western Ontario
 Office of Research Ethics
 Room 5150, Support Services Building • London, Ontario • CANADA - N6A 3K7
 PH: 519-661-3036 • F: 519-850-2466 • ethics@uwo.ca • www.uwo.ca/research/ethics

Figure 45: Research ethics approval titled “Realistic Computer Graphics Rendering and Task Analysis for Neurosurgery Simulation – Neuroendoscopy Project”



LAWSON FINAL APPROVAL NOTICE

RESEARCH OFFICE REVIEW NO.: R-12-465

PROJECT TITLE: Learning with virtual environments

PRINCIPAL INVESTIGATOR: Dr. Sandrine de Ribaupierre

LAWSON APPROVAL DATE: December 20, 2012

Health Sciences REB#: 102917

Please be advised that the above project was reviewed by the Clinical Research Impact Committee and the project:

Was Approved

PLEASE INFORM THE APPROPRIATE NURSING UNITS, LABORATORIES, ETC. BEFORE STARTING THIS PROTOCOL. THE RESEARCH OFFICE NUMBER MUST BE USED WHEN COMMUNICATING WITH THESE AREAS.

Dr. David Hill
V.P. Research
Lawson Health Research Institute

All future correspondence concerning this study should include the Research Office Review Number and should be directed to Sherry Paiva, CRIC Liaison, Lawson Health Research Institute, 750 Baseline Road, East, Suite 300.

cc: Administration

Figure 46 Research ethics approval titled “Learning with virtual environments.”



**Western University Non-Medical Research Ethics Board
NMREB Amendment Approval Notice**

Principal Investigator: Dr. Sandrine de Ribaupierre
Department & Institution: Schulich School of Medicine and Dentistry/Clinical Neurological Sciences, Western University

NMREB File Number: 102917
Study Title: Learning with virtual environments

NMREB Revision Approval Date: November 21, 2016
NMREB Expiry Date: December 12, 2017

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Revised Western University Protocol	Received Nov 17, 2016	


The Western University Non-Medical Science Research Ethics Board (NMREB) has reviewed and approved the amendment to the above named study, as of the NMREB Amendment Approval Date noted above.

NMREB approval for this study remains valid until the NMREB Expiry Date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario.

Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.


Ethics Officer, on behalf of Dr. Riley Hinson, NMREB Chair

Ethics Officer: Erika Basile Katelyn Harris Nicole Kaniki Grace Kelly Vikki Tran Karen Gopaul

Figure 47: Research ethics amendment titled “Learning with virtual environments.”



**Western University Health Science Research Ethics Board
HSREB Amendment Approval Notice**

Principal Investigator: Dr. Sandrine de Ribaupierre
Department & Institution: Schulich School of Medicine and Dentistry\Clinical Neurological Sciences, Western University

HSREB File Number: 102917
Study Title: Learning with virtual environments
Sponsor:

HSREB Amendment Approval Date: June 23, 2014
HSREB Expiry Date: August 31, 2017

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Change in Study Personnel	Adding co-investigators as more students are working on parts of the study - Lauren Allen - PhD candidate - lallen22@uwo.ca - Ryan Armstrong - PhD candidate - rarmst2@uwo.ca - Dayna Noltie - medical student -	

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the amendment to the above named study, as of the HSREB Initial Approval Date noted above.

HSREB approval for this study remains valid until the HSREB Expiry Date noted above, conditional to timely submission and acceptance of HSREB Continuing Ethics Review. If an Updated Approval Notice is required prior to the HSREB Expiry Date, the Principal Investigator is responsible for completing and submitting an HSREB Updated Approval Form in a timely fashion.

The Western University HSREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Guideline for Good Clinical Practice Practices (ICH E6 R1), the Ontario Personal Health Information Protection Act (PHIPA, 2004), Part 4 of the Natural Health Product Regulations, Health Canada Medical Device Regulations and Part C, Division 5, of the Food and Drug Regulations of Health Canada.

Members of the HSREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.



Ethics Officer to Contact for Further Information

<input type="checkbox"/> Erika Basile ebasile@uwo.ca	<input checked="" type="checkbox"/> Grace Kelly grace.kelly@uwo.ca	<input type="checkbox"/> Mina Mekhail mmekhail@uwo.ca	<input type="checkbox"/> Vikki Tran vikki.tran@uwo.ca
---	---	--	--

This is an official document. Please retain the original in your files.

Figure 48: Research ethics amendment titled “ Learning with virtual environments.”



Curriculum Vitae

Name:	Ryan Armstrong
Post-secondary Education and Degrees:	<p>The University of Western Ontario London, Ontario, Canada 2007-2011 B.M.Sc.</p> <p>The University of Western Ontario London, Ontario, Canada 2011-2016 Ph.D (expected)</p>
Honours and Awards (Western):	<p>3MT Engineering Winner 2016</p> <p>3MT Second Place at Institution 2015</p> <p>Ontario Graduate Scholarship 2013-2016</p> <p>NSERC CAMI 2011-2013</p> <p>WGRS 2011-2016</p> <p>Forester's Competitive Scholarship 2007-2011</p> <p>The Laurene Paterson Estate Scholarship 2010-2011</p> <p>Queen Elizabeth II Aiming for the Top Scholarship 2007-2011</p> <p>Faculty of Science Dean's Honour List 2007-2011</p> <p>Faculty of Science in-course Scholarship 2007-2008</p>

Western Scholarship of Excellence
2007-2008

Related Work Experience	Teaching Assistant	
	The University of Western Ontario	
	2011-2016	
	<i>Teaching Assistantships:</i>	
	Theoretical Foundations of Software Engineering	(2012-2016)
	Programming Fundamentals for Engineers	(2015)
	Computer-Human Interface Design	(2012-2013)
Software Construction	(2012)	
Software Design	(2011)	

Publications:

R. Armstrong, R. Eagleson, S. de Ribaupierre, Automation of a Closed Source Surgical Simulator: Making User Studies Feasible, submitted to The American Journal of Biomedical Engineering.

R. Armstrong, D. Noltie, R. Eagleson, S. de Ribaupierre, An Examination of Factors that Impact Performance on a Patient-Specific VR-Based Ventriculostomy Simulator, submitted to Simulation.

R. Armstrong, R. Eagleson, S. de Ribaupierre, Augmented reality for Neurosurgical Guidance: An Objective Comparison of Planning Interface Modalities", due to appear in Volume 9805 of the Lecture Notes in Computer Science series.

E. Omid, L. Fuetterer, S. Mousavi, R. Armstrong, L. Flynn, A. Samani, Characterization and assessment of hyperelastic and elastic properties of decellularized Human adipose tissues. Journal of Biomechanics. (2014). 28;47(15):3657-63.

M. Kramers, R. Armstrong, S.M. Bakhshmand, A. Fenster, S. de Ribaupierre, R. Eagleson, Evaluation of a Mobile Augmented Reality Application for Image-guidance of Neurosurgical Interventions. Studies in Health and Technology and Informatics. (2014). 196:204-86.

R. Armstrong, R. Eagleson, S. de Ribaupierre, Patient-Specific Pipeline to Create Virtual Endoscopic Third Ventriculostomy. *Studies in Health and Technology and Informatics*. (2014). 196:14-6.

R. Armstrong, S. de Ribaupierre, R. Eagleson, A Software System for Evaluation and Training of Spatial Reasoning and Neuroanatomical Knowledge in a Virtual Environment. *Computer Methods and Programs in Biomedicine*. (2014). 114(1):29-37.

M. Kramers, R. Armstrong, S. Bakhshman, A. Fenster, S. de Ribaupierre, R. Eagleson, A Mobile Augmented Reality Application for Image-guidance of Neurosurgical Interventions. *American Journal of Biomedical Engineering*. (2013). 3(6):169-174.

L. Plumley, R. Armstrong, S. de Ribaupierre, R. Eagleson, Spatial Ability Training in Virtual Neuroanatomy. *Studies in Health and Technology and Informatics*. (2013). 184:324-329.

Conference Podium Presentations:

R. Armstrong, R. Eagleson, S. de Ribaupierre, Augmented reality for Neurosurgical Guidance: An Objective Comparison of Planning Interface Modalities", MIAR, Bern, Switzerland, 2016.

R. Armstrong, T. Wright, R. Eagleson, S. de Ribaupierre, A VR system for medical simulation using Google Cardboard and Vuforia, CARS, Heidelberg, Germany, 2016.

R. Armstrong, M. Kramers, D. Noltie, R. Eagleson, S. de Ribaupierre, An Examination of Factors that Impact Performance on a Patient-Specific VR-Based Ventriculostomy Simulator, CARS, Barcelona, Spain, 2015.

R. Armstrong, R. Eagleson, S. de Ribaupierre, A Software Pipeline for Creation of Patient-Specific Scenarios for Neurosurgical Simulation, GRAND NCE, Toronto, Ontario, Canada, May 2013.

R. Armstrong, R. Eagleson, S. de Ribaupierre, Novel Interaction Techniques for Immersive Exploration of Volumetric Medical Data Sets, Western Research Forum, London, Ontario, Canada, March 2012.

Technical Experience

Languages: C++, GLSL, C#, Java, HTML/CSS, JavaScript, Lua, Python, MATLAB, PHP, RegEx, AutoHotKey, SQL

Platforms: MS Windows, Linux, UNIX, Android, Raspberry Pi, Arduino

APIs, Toolkits: OpenGL, WebGL, DirectX, CUDA, Vuforia, Google Cardboard Win32, Android SDK, VTK and ITK, Qt, Git, MySQL

Software: MS Visual Studio, Android Studio, Aptana, Eclipse, Dreamweaver CS5.5, MS Office, Blender, GIMP, Paint.net

Music: Ableton Live, Max for Live, SoftStep Advanced Editor, Massive

Professional Affiliations, Committees, Volunteering

2015-2016	Let's Talk Science Volunteer Educator
2015-2016	UWO Biomedical Engineering Students Council Sports Coordinator
2012-2015	Institute of Electrical and Electronics Engineers (IEEE) Student Member
2011-2014	Graphics, Animation and New Media (GRAND) NCE Member of Graduate Student Committee

Extracurricular Activities

2011-2013

**Captain of Men's Cross Country and Track and Field Teams
(CIS All-star)**

Western University