2016

Virtual Views: Exploring the Utility and Impact of Terrestrial Laser Scanners In Forensics and Law

Ryan Andrew Mullins
University of Windsor

Recommended Citation

Mullins, Ryan Andrew, "Virtual Views: Exploring the Utility and Impact of Terrestrial Laser Scanners In Forensics and Law" (2016). *Electronic Theses and Dissertations*. 5855.
https://scholar.uwindsor.ca/etd/5855

VIRTUAL VIEWS: EXPLORING THE UTILITY AND IMPACT OF TERRESTRIAL LASER SCANNERS IN FORENSICS AND LAW

by

Ryan A. Mullins

A Thesis
Submitted to the Faculty of Graduate Studies
through Criminology
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts at the
University of Windsor

Windsor, Ontario, Canada

Virtual Views: Exploring the Utility and Impact of Terrestrial Laser Scanners in Forensics and Law

by

Ryan A. Mullins

APPROVED BY:

D. Tanovich
Faculty of Law

G. Cradock
Department of Sociology, Anthropology, and Criminology

J. Albanese, Advisor
Department of Sociology, Anthropology, and Criminology

7 September 2016

AUTHOR'S DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Terrestrial laser scanners are sophisticated measurement devices that have relatively recently become forensic tools. Despite a respectable breadth on the topic in existing literature, issues surrounding the use of this technology on-scene, the validity of its accuracy, and the legal evidentiary implications of its use remain largely unexplored. This research explored these gaps via experimentation with the Faro Focus^{3D} 330X's accuracy and forensic anthropological utility (the latter tested through biological characteristic estimation regression models). This experiment was situated in a broader legal evidentiary context that involved review of relevant case law and a case study of the first use of TLS-produced evidence in Canada. Results show that the device has an error rate of approximately 3 mm. Measurements obtained and inputted into the forensic anthropological regression models yielded results quite similar to those estimated by control (i.e., by hand) measurements, despite statistically significantly different means.

DEDICATION

I dedicate this work, my first major academic endeavour, to my parents, Michael and Mary Lou. Without your support and encouragement, none of this would have been possible.

ACKNOWLEDGMENTS

I would like to sincerely thank the three members of my thesis committee, namely: my advisor, Doctor John Albanese, for his boundless commitment to this project and his invaluable insight in all aspects of my research, Doctor Gerald Cradock, for his assistance in better understanding the sociology of law and science and his impeccable literature recommendations, and Professor David Tanovich, for his expertise on the law of evidence and criminal law in general. Your time and commitment to this project ensured its success.

I would also like to sincerely thank John Johnston and Melinda Hanneson from Faro Technologies for providing access to equipment, assisting with data collection and providing information on terrestrial laser scanners, and Abi Yousufi from the Toronto Police Services' Forensic Identification Service for explaining how their forensic investigators utilize terrestrial laser scanning equipment.

TABLE OF CONTENTS

AUTHOR'S DECLARATION OF ORIGINALITY iii

ABSTRACT iv

DEDICATION v

ACKNOWLEDGMENTS vi

LIST OF TABLES ix

LIST OF ILLUSTRATIONS x

CHAPTER

I. INTRODUCTION 1

II. LITERATURE REVIEW AND BACKGROUND
Introduction 3
Terrestrial Laser Scanning for Forensics 4
Nature and Admissibility of TLS-Produced Evidence 7
The Suitability of Forensic Anthropology to Explore TLS Measurement Accuracy 13
Conclusion 14

III. METHODS AND MATERIALS
Introduction 15
The Scene 16
Virtual Forensic Data Collection Procedures 17
Measurements 22
Statistical Models 24
Conclusion 27

IV. EXPERIMENT RESULTS
Introduction 27
Results 27
Conclusion 30

V. DISCUSSION
Introduction 30
Experiment Results 31
Terrestrial Laser Scanning in Forensics 34
Terrestrial Laser Scanning in Law 36
Conclusion 41

VI. CONCLUSION 41
Research Limitations 43
Future Research 43

APPENDIX A: Data Tables 45

APPENDIX B: Illustrations 50

REFERENCES 53

VITA AUCTORIS 58

LIST OF TABLES

TABLE 1 - Stature Estimate Trials Descriptive Statistics 45

TABLE 2 - MD and MAD Amongst Measurement Trials 46

TABLE 3 - Bone Measurement Trials Descriptive Statistics 47

TABLE 4 - One-Way Repeated Measures ANOVA - Bone Measurement Means 48

TABLE 5 - One-Way Repeated Measures ANOVA - Stature Estimate Means 49

TABLE 6 - Intra-observer TEM 49

TABLE 7 - Sex Estimation Case Study 49

LIST OF ILLUSTRATIONS

FIGURE 1 – The Scene 50

FIGURE 2 – An Example of the Virtual Iliac Breadth Measurement 51

FIGURE 3 – An Illustration of Incorrect Point Selection 52

Chapter I: Introduction

Forensic analysis – typically performed at a crime scene – comprises a large and important part of the overall inquiry into the occurrence of certain criminal offences. Crime scene investigation refers to the coordinated and systematic examination of a crime scene wherein official personnel – which, depending on the offence, may include crime scene investigators (CSI), forensic identification officers, Office of the Fire Marshal (OFM) investigators, and the coroner – respond to a select set of occurrences including homicide, arson, and accidents resulting in death or grave injury (which may involve criminally negligent conduct), among others. Crime scenes are dynamic environments: once a crime has occurred, the scene that contained it immediately becomes vulnerable to contamination and the deterioration of evidence quality (Barazzetti et al., 2012). Of course, police work necessarily entails the destruction of the crime scene as evidence is gathered and the scene is returned to normal. It is usually the case that you do not get a second chance at properly documenting and collecting evidence. In fact, investigators are often pressured to return the scene to its normal use as quickly as possible (Topol et al., 2008). As such, issues of scene security and temporality are of utmost importance to a policing unit whose sole function is to process a crime scene in attempts to reconstruct the events that transpired and gather evidence to that effect. The investigation process frequently includes referring evidentiary displays or objects to various third parties, including: scientific experts for analysis; Crown attorneys and defence counsel alike, who rely on the accuracy of evidence gathered by the police in arguing their respective case, and; most importantly, judges and jurors, who hold a power over life, sometimes fatally.

Those who must know the minute details of the case (especially the third parties mentioned above) often have to rely on oral and visual representations of a crime scene once the scene has been fully processed and released. In recent years, technological advances have prompted a number of police departments to adopt various mechanical instruments as tools of crime scene investigation. Amongst the pool of most recent additions, we see a drastic change in metrological technologies.

Terrestrial laser scanners are sophisticated measurement devices and have become the instrument of choice for scene documentation in a minority, but increasing number, of police departments (Bucheli et al., 2014). These organizations have only relatively recently (i.e., within approximately the last 15 years) begun to use terrestrial laser scanners (TLS)[1] as tools of forensic investigation. Even more recent is the admission of this form of evidence into courts of law in a limited number of jurisdictions. TLS technology has the ability to extract reality from an environment and reproduce it as an almost perfect reflection. These devices utilize controlled laser pulses to capture the dimensions and physical properties of the space in which they are deployed. The data captured by TLS can be rendered into scale models of the scene *in situ*, within which objects and space can be measured accurately and noninvasively via computer software. As such, police hypotheses can be tested empirically within a virtual reconstruction of the crime scene. As this thesis will discuss, this form of evidence can be highly salient to the fact-finding function of a criminal court.

The primary focus of this research involves an evaluation of terrestrial laser scanners in the context of forensic investigation and the subsequent legal scenarios that may utilize their evidence. Due to the number of TLS currently available on the forensic market, I have chosen to limit investigation to a specific TLS – the Faro Focus^{3D} 330X – currently employed by the Toronto Police Services (TPS) in their forensic investigations. Aside from conclusions made regarding the accuracy and forensic utility of this specific device, the conclusions produced through this evaluation of TLS can be extrapolated to the general use of this technology in forensics and, concomitantly, law.

This research involves an extensive review and synergy of the body of literature on TLS as is used in forensics. In doing so, the author highlights limitations in the sparse body of literature, which includes a scarcity – or, in some cases, a complete absence – within the literature to validate TLS measurement accuracy, inquire as to how police specifically utilize these devices in forensic contexts, and consider the broader legal evidentiary implications. Furthermore, the author explores the first Canadian criminal trial wherein TLS-produced evidence was adduced.

This research also involved a substantial experimental section wherein the author, with the assistance of Faro Technologies' employees and his thesis supervisor, tested the accuracy capabilities of the Faro Focus^{3D} 330X (hereafter referred to as the 'Focus^{3D}')

---

[1] TLS may refer to "terrestrial laser scan(s)", "terrestrial laser scanner(s)", and "terrestrial laser scanning" based on what is grammatically appropriate in any given context.

through an innovative forensic anthropological application. The experiment entailed laser scanning a succession of bones to obtain landmark-based osteometric data from hipbones and femora, which were inputted into regression models to estimate sex[2] and stature, respectively, and juxtaposed against control (i.e., taken by hand) measurements. Measuring the skeletal remains of an unknown individual is a plausible forensic scenario. It uniquely satisfies the objective of exploring and validating the accuracy of the Focus^{3D} while also specifically testing the ability of laser scanning equipment to obtain subtle measurements.

Overall, the author addresses important questions such as: in which situations are TLS deployed by police, are error rates known and reliably accurate, what forms of evidence do these devices produce, and how have common law jurisdictions evaluated this novel forensic scientific evidence thus far? As the use of TLS in forensics becomes routinized, and its evidence increasingly proffered before the court, it is essential that measurement accuracy is tested and shown to be valid and reliable. Moreover, the dynamic and technical nature of this evidence necessitates that it be well understood by evidentiary scholars.

Chapter II: Literature Review and Background

Introduction

Forensic investigators are continually looking for new, efficient ways to process crime scenes and solve crimes. This endeavour is increasingly involving the use of scientific equipment and techniques (birthing the discipline of forensic science). Furthermore, forensic investigators are interested in scene documentation methods that are efficient, noninvasive, and easily displayed to fact-finders. Terrestrial laser scanners seem to satisfy the above criteria – they are highly efficient pieces of surveying technology that can scan the physical space in which they are deployed in a matter of minutes. They are completely noninvasive; besides being situated within, or proximately close to, the crime scene, they do not need to physically touch any object to measure it,

---

[2] It is useful to make a distinction between sex – a biological trait based on anatomical or chromosomal categories distinguishing males and females – and gender – the socially constructed roles related to sexual distinctions. In this thesis, I refer only to the former.

eliminating any chance of evidence contamination. The virtual environment data collected by TLS can be measured via computer software, rendered into scale models for display, and can provide the basis for virtual crime scene walkthroughs, animations, and simulations. These qualities make TLS very enticing to police forces that routinely deal with crime scenes such as homicide, which can take enormous amounts of time and resources to process with traditional methods. In addition, the ability to virtually preserve the crime scene provides an opportunity to reevaluate the evidence *in situ* in the future. As such, police agencies rely on TLS primarily for reasons of efficiency, long-term cost reduction, and comprehensiveness.

Terrestrial Laser Scanning for Forensics

As the name indicates, terrestrial laser scanners are LiDAR (light detection and ranging) devices that use laser light to measure distances. A laser beam is emitted from the device, reflected off the surface in its path, and transmitted back to the device. While all laser scanners operate on this basic principle, the technology can take two different forms: time-of-flight or phase-shift scanners, with the Focus^{3D} falling into the latter category (Mihandoost, 2015). Phase-shift scanners operate by modulating the wavelength of a single laser beam, which is projected repeatedly into the environment at a variety of angles off of a rotating mirror. According to Mihandoost (2015), "[t]he distance between the object or structure is then calculated based on the analysis of the phase shifts in the wavelength of the returning beam compared to the emitted beam" (p. 11). This process generates x, y, and z (latitude, longitude, elevation) coordinates for the millions of points in the environment touched by laser light. This mesh of individual points is referred to as a 'point cloud'. In addition, the Focus^{3D} is equipped with a camera that takes pictures (85 total) of the immediate environment, capturing a pseudo-panorama of the device's line of sight. This provides users with the option of mapping the photographs onto the existing point cloud, rendering them in colour.

TLS allow police agencies to complete their work more efficiently: the base 'resolution' (i.e., point volume and spacing) setting of the Focus^{3D} captures millions of measurements *in potentia* in as little as seven seconds. The increase in efficiency that this technology affords, without sacrificing accuracy, can save police valuable time and expense. Moreover, the use of TLS allows police a larger margin of error concerning what should be recorded on scene. Baber and Butler's (2012) study found that the investigator's level of expertise plays a large role in deciding what has evidentiary value at the scene. However, even experts make mistakes: such is human nature. Since TLS can capture virtually everything and anything, investigators will not have to worry if they realize that they failed to make a crucial measurement after the scene has been released. In other words, the subjectivity involved with determining what measurements should be taken on scene can be minimized (Eyre et al., 2014) and TLS use can be considered a safety net as far as scene documentation and reconstruction is concerned (Siuri, 2004). Provided that scans were properly taken from the outset, and contemporary data formats do not quickly become obsolete, police can easily revisit the scan data months or years later and the robustness of such data lends itself well to *post hoc* scene comprehension.

Terrestrial laser scanning technology is not a new invention. Typically it has been applied to other fields such as land-surveying, manufacturing, and construction. However, its application to forensics within approximately the last 15 years is novel. In Canada, only a handful of policing agencies use TLS equipment. The Toronto Police Services (TPS) began using TLS for forensics in 2010 (Stancu, 2010). The Focus^{3D} has become an integral part of their homicide investigation process (it is also used occasionally for other criminal investigations) (A. Yousufi, personal communication, November 17, 2015). Besides preserving a virtual record of the crime scene, the data gathered by the Focus^{3D} is used to create hypotheses, test said hypotheses and, occasionally, corroborate the narrative of Crown prosecutors.

Academics, too, have realized the forensic potential of TLS and have generated a modest body of literature on the subject, though Elmes et al. (2014) argue not enough. Due to the wide range of situations in which TLS can be used, the literature has been largely disparate and disjointed. Authors have focused on testing or discussing TLS in specific forensic contexts, such as blood spatter analysis (Buck et al., 2011; Hakim & Liscio, 2015), virtual shoeprint-impression casting (Komar et al., 2012), perpetrator height estimation (Johnson & Liscio, 2015), clandestine grave investigation (Jorge et al., 2012), crime or accident scene reconstruction (Agosto et al., 2008; Barazzetti et al., 2012; Buck et al., 2013; Roedl et al., 2014; Eyre et al., 2014; Wozniak, 2015), crime scene preservation (Siuri, 2004), systematic TLS measurement error detection (Cuartero et al., 2010; Wang et al., 2012), and post-mortem examination (Bucheli et al., 2012; Komar et al., 2012; Schweitzer et al., 2013). Some have contrasted the efficacy of TLS against photogrammetry – accurate methods for collecting measurements from photographs, often for land-survey applications – or have called for the coordinated use of both metrological methods (Barazzetti et al., 2012; Buck et al., 2011; Buck et al., 2013).

Despite the breadth of topics involving TLS, there has been a large-scale absence of scientific inquiry into the specific ways police utilize TLS technology during crime scene investigations. This is surprising: after all, the above-mentioned experiments are undertaken in the name of forensic literature, which necessarily exists because of policing. If researchers are interested in applying TLS as police do (i.e., in forensic contexts), they should be cognizant of any standards or best practices. For example, TPS investigators typically place reference targets – normally in the form of white spheres or small, paper checker patterns – within the crime scene environment when using TLS. The purpose of this is to ensure accurate registration, the process wherein individual scans are stitched together into one cohesive body. Notwithstanding, some scanners, including the Focus^{3D}, can perform accurate, automatic registration without targets. This is made possible through computer software with the ability to automatically recognize geometric planes in the environment. Ultimately, as long as there is sufficient overlap within separate point clouds, whether targets or natural geometric features, the software can easily register disparate scans into one cohesive point cloud. The targets prove most useful when there are no easily recognized geometric planes, which may be the case in outdoor scenes. The use of targets should be considered a fail-safe and a police best practice, especially in environments lacking distinctive geometric features.

In addition, authors have specifically noted a lack of focus on validating the accuracy and reliability of TLS data (Sholts et al., 2010; Mihandoost, 2015) and the computer-generated reconstructions they produce – this is not something that should be taken for granted. Indeed, based on an extensive and comprehensive search for relevant literature, there are only a handful of studies that have endeavoured to validate TLS measurements directly (Boehler and Marbs, 2003; Mihandoost, 2015) or tangentially (Hakim & Liscio, 2015; Johnson & Liscio, 2015). Mihandoost's (2015) study specifically explored the Focus^{3D} and two related processing softwares and found measurement error averages over double the manufacturer's specifications, at approximately 5 mm (Mihandoost notes that this amount of error may be partially attributable to human error arising from obtaining the virtual measurements via software) (R. Mihandoost, personal communication, October 20, 2015). Despite being double the specification value, it is worth noting that 5 mm is still incredibly accurate. In the U.S., the National Institute for Standards and Technology (NIST) recommends crime scene measurements be accurate to within one-quarter inch, or approximately 6.4 mm (NIST, 2013). This margin of error was selected to account for slight variations in tape measurers that may not be NIST certified (K. Lothridge, personal communication, October 23, 2015). The average measurement error attributed to the Focus^{3D} in Mihandoost's (2015) study still falls comfortably within this boundary.

In Canada, forensic investigators are not bound by any official measurement standard, nor has any recognized body recommended one; however, my correspondent in the TPS's Forensic Identification Service (FIS) noted that they generally ensure traditionally obtained measurements (i.e., by hand) are accurate to within one-eighth of an inch, or approximately 3.2 mm (A. Yousufi, personal communication, November 17, 2015). It is likely that this also applies to TLS generated measurements.

Nature and Admissibility of TLS-Produced Evidence

Some authors have insightfully asked whether TLS-produced evidence would be admissible in court (Komar et al., 2012; Barazzetti et al., 2012) – this is an extremely important question. Even though some jurisdictions have been admitting TLS-produced evidence in court for years, there has been virtually no scholarship on the legal implications of the use of this technology. This means a dearth of information regarding the forms this evidence can take (i.e., real or demonstrative), the legal tests it may be required to satisfy for admissibility (i.e., concerning novel science and expert evidence), and the general implications of relying on this technology in courts of law (i.e., prejudicial versus probative effect). Forensic authors cannot necessarily be faulted for a failure to address these important issues; however, it must then fall onto the shoulders of critical criminological and legal scholars. Indeed, a number of critical legal scholars have already devoted great amounts of energy towards questioning the use of currently – and often uncritically – accepted forensic science technologies and techniques.

Forensic science refers to the general application of scientific method to legal processes, particularly criminal investigation. Readers may be familiar with DNA profiling, blood spatter point-of-origin analysis, and fingerprint comparison: all are typical examples of forensic science. The purpose of forensic scientific experimentation and analysis is to attempt, with some degree of scientific certainty, to discern the truth of legal questions that come before the court. Does the blood sample taken at the scene match the primary suspect's blood? Is the defendant's story of shooting in self-defense corroborated by bullet trajectory analysis? These are the types of questions that forensic science is often called upon to answer. Cloaked in the prestige and certainty associated with science, forensic science evidence has historically been regarded by the courts as a highly probative, superior form of evidence (Williams and Saks, 2015).

In recent years, however, the discipline of forensic science has undergone a withering criticism from practitioners and academics alike. Inquiries into miscarriages of justice, including a string of notable wrongful convictions, have cast a great deal of doubt upon well-recognized forensic scientific methods that have been accepted into courts for decades with little-to-no impediment (see Shelton, 2010). A number of authors (see Shelton, 2010; Edmond and Roach, 2011; Edmond, 2015; Williams and Saks, 2015) have parroted and elaborated upon the scathing conclusion given in the American National Research Council's (2009) inquiry into the state of forensic science:

> [t]he bottom line is simple: In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem. (p. 53)

Gary Edmond has been a particularly outspoken critic about the failure of common law trial judges to act as gatekeepers against unproven forensic scientific evidence and consistently apply scientific evidentiary thresholds such as those found in *Daubert*, the leading American case on the admissibility of novel scientific evidence (Edmond and Roach, 2011; Edmond, 2011; Edmond, 2015).

The point of mentioning recent criticisms of forensic science is not to suggest that terrestrial laser scanners are 'junk science', because that is certainly not the case. Rather, the point is to develop a context that highlights the importance of scrutinizing scientific and highly technical procedures and equipment that permeate the legal sphere. As Hill et al. (2013) note, as parties increasingly proffer and rely upon the substantive truths of technological media, the intersection between technology and the rules of evidence becomes critical. Moreover, the principle of *stare decisis* has historically precluded all motivation for critical retrospection on forensic science that is admitted consistently and/or prior to the imposition of stricter evidentiary standards. With that said, where *do* terrestrial laser scanners – or, more accurately, the evidence they produce – fit into the evidentiary landscape? There is no simple answer to this question.

The data provided by TLS, and the varied uses for it, ensures that this form of evidence is not easily classified. For example, is it demonstrative or real evidence? The short answer is that it can be either or both. Demonstrative evidence is: "[t]hat evidence addressed directly to the senses without intervention of testimony… Such evidence… [that] illustrate[s] some verbal testimony and *has no probative value in itself*" (Mnoonkin, 1998, p. 67, emphasis added). Mnoonkin's (1998) exploration of the origins of photography in law gives a particularly interesting account of the evolution of demonstrative evidence, the conception of which has historically oscillated between 'mere representation' and 'intrinsically, scientifically substantive'. Modern courts have largely settled on the former conception of demonstrative evidence. Examples of demonstrative evidence include, but are not limited to, charts, graphs, and re-enactments.

In contrast, real evidence is adduced to help establish facts in issue and it "provides the trier of fact with an opportunity to draw a relevant first hand sense impression" (Hill et al., 2013, p. 23-1). The modern conception of real evidence is somewhat difficult to fully operationalize. In some cases, it seems to refer only to physical, tangible evidence (e.g., a weapon, blood sample, bullet casings). In other cases, it refers broadly to evidence with inherently substantive properties. Bex (2011) cites *State v. Famber* when referring to real, or direct, evidence as that "evidence which, if believed, proves the existence of the fact in issue without inference or presumption" (p. 13). The inclusion of both real and direct evidence in this definition is somewhat confusing for two reasons: 1) it appears to conflate two different ideas, as direct evidence (e.g., witness testimony that she physically saw the accused shoot the victim) may be considered real evidence by some authorities (Law and Martin, 2014), but not all real evidence (e.g., a bloodied shirt) is direct evidence; and 2) if believed, direct evidence does not require inference, but real evidence will likely require it. For example, the introduction of a bloodied shirt – considered real, not direct, evidence – is only relevant to the proceedings if we infer that its condition is the direct result of the crime in question and not the product of some unrelated event, such as an accident. Furthermore, Delisle et al. (2015) posit "[r]eal evidence, the gun, the narcotics, the bloodstained shirt, tendered as an object within the courtroom *is not a helpful aid but rather is evidence itself*" (p. 418, emphasis added). Whatever the definition, all authors seem to be in agreement that real evidence is that which *substantively* represents the truth of a given proposition. Its probative value resides within its own characteristics.

Whether demonstrative or real, evidence must satisfy certain criteria before being accepted by the court. Hill et al. (2013) offer the following qualifications: "[t]he primary issue for [the admittance of] real evidence is relevance and authentication while for demonstrative evidence the issues are largely fairness and prejudice" (p. 23-2). At its base, all evidence is subject to relevancy – if it is not relevant to the proceedings at hand, it is of no value to the trier of fact. This is especially true with real evidence, which is likely to be afforded greater weight due to its intrinsically substantive value. According to Hill et al. (2013), evidence is relevant "if it tends to make the existence of the fact more probable or likely" (p. 4-2). For the authentication of real evidence, the onus falls on the party seeking to adduce it. Authentication will normally be determined *prima facie*, rendering this a fairly light threshold to satisfy (Hill et al., 2013).

The fairness of demonstrative evidence is determined through reference to the accuracy of the depictions in question. Because demonstrative evidence has no intrinsic probative value, its worth lies in how accurately (or fairly) it represents the testimony or evidence upon which it is based. Moreover, the evidence must not bring the integrity of the court into disrepute (i.e., by unfairly favouring one side over the other). When considering prejudice, the court must weigh the potential prejudicial effect against the probative value of the evidence. Delisle et al. (2015) cite McCormick's (1992) text on evidence when defining prejudicial as: "decision on an improper basis, commonly, *though not necessarily*, an emotional one" (p. 420, emphasis added). The dangers of prejudice increase when demonstrative evidence represents only one party's version of the facts in dispute (Delisle et al., 2015). Ultimately, these issues may go to the weight, as opposed to the admissibility, of the evidence in question.

The final issue concerning the use of TLS in law involves its novel scientific application, how the court handles such matters, and whether expert opinion evidence is necessary to explain the mechanics of the device to the trier of fact. When scientific evidence is adduced, particularly novel science, it must satisfy certain evidentiary thresholds. Additionally, the court almost always requires scientific (or highly technical) evidence be accompanied by expert opinion evidence on the underlying science or technical aspects in question. The use of expert witnesses helps the trier of fact to understand complicated matters, lessening the possibility that certain facts or evidence will be improperly weighted and/or misconstrued. Delisle et al. (2015), in accordance with Canadian case law, cite the steps for the admission of expert opinion evidence as outlined in the leading judgment on the issue, *White Burgess Langille Inman*:

> At the first step, the proponent of the evidence must establish the threshold requirements of admissibility. These are the four *Mohan* factors (relevance, necessity, absence of an exclusionary rule and a properly qualified expert) and in addition, in the case of an opinion based on novel or contested science or science used for a novel purpose, the reliability of the underlying science for that purpose (p. 888).

Subsequently, for all expert evidence, the court must also consider specifically whether the prejudicial effect outweighs the probative value by again applying factors such as relevance, necessity, reliability, and the absence of bias (Delisle et al., 2015). Determining when the *Mohan* factors apply is somewhat complicated by the fact that "expert evidence, simply unopinionated special knowledge testimony, is not subject to the [*Mohan*] opinion evidence rule" (Hill et al., 2005, p. 12-23).

In assessing the reliability of novel scientific evidence, Canadian common law has devised a non-exclusive and somewhat flexible test that stems from the *Daubert* "reliable foundation" standard (Hill et al., 2013). Delisle et al. (2015) note that the courts were explicit about their conception of reliability: "when they [the court] spoke of the need for reliability they were speaking of *evidentiary reliability, i.e., trustworthiness, and that reliability would be based upon scientific validity*" (p. 884, emphasis added).

The Canadian equivalent of *Daubert* was outlined in *J.-L.J.,* which is considered the leading case on the matter of novel scientific evidence. Delisle et al. (2015) summarize the test as follows: 1) can the theory or technique be tested and has it been tested? 2) has the theory or technique been subjected to peer review and publication? 3) what is the potential rate of error? and 4) has the technique been generally accepted? These factors underlie a rational approach to considering the admissibility of novel scientific evidence. Again, the potential prejudice of the evidence must also be considered when determining admissibility, even after weighing the evidence on the above-four criteria.

It seems apparent that a device that measures objects remotely via laser light is scientific in nature; however, courts have handled TLS-produced evidence differently. In America, a pamphlet by Leica Geosystems (a producer of TLS devices) displays a judge's affirmative ruling on a *Daubert* motion seeking to introduce TLS-produced evidence. The judge ruled that the defendants "have laid a proper scientific and technical foundation for the admissibility of this evidence" and that, barring a subsequent finding of irrelevance or uselessness by the trier of fact, it would be admitted (Leica Geosystems, n.d., p. 8). This trial judge clearly considered TLS-produced evidence to fall within the scope of novel scientific evidence.

Conversely, a Canadian court did not share the same opinion; however, this was based on the capacity in which the TLS-produced evidence was introduced in this particular case. As of 2013, *R. v. Doodnaught* is the first and only case in Canada that has had to address the question of TLS-produced evidence. The *voir dire* hearing concerning this evidence was extremely rigorous, and the trial transcripts are particularly revealing of the grey area within which TLS-produced evidence falls. There are two interesting findings to draw from this trial: 1) the *voir dire* involved both the Crown and the defence calling "experts" to speak about TLS data, while simultaneously clarifying that they did not consider this to be expert opinion evidence. Through no ill intent, both parties effectively dodged the expert opinion evidence thresholds; 2) TLS use in the forensic context was described as "novel" on at least one occasion during trial. The judge inquired whether it had been barred from any other courts, how accurate the measurements were (i.e., error rates), and how exactly the technology worked. While these inquiries quite clearly parallel the novel science thresholds, the TLS-produced data was not regarded as novel scientific evidence and was not subjected to the evidentiary thresholds associated with it. The contrast between the two examples supports the notion that TLS-produced evidence is difficult to categorize and must undergo further scrutiny.

The Suitability of Forensic Anthropology to Explore TLS Measurement Accuracy

Forensic anthropology, the discipline underlying the experimental portion of this research, is essentially a subset of physical anthropology applied specifically to legal cases (Klepinger, 2006). Physical anthropologists study skeletal variation to estimate biological characteristics such as age, sex, and stature. In essence, the known boundaries of skeletal variation allow forensic anthropologists to utilize various methods, in this case regression models, to estimate biological characteristics, often with a great amount of accuracy. For example, Albanese's (2003) metric logistic models allowed the author to estimate biological sex with an allocation accuracy that exceeds 95%.

Since TLS are purposed primarily for gathering the metric characteristics of large environments, research involving laser scanning in forensic anthropology has typically relied on devices purposed towards scanning *objects* (i.e., not TLS) as opposed to whole *environments*. Some examples include virtual craniometry (Park et al., 2006; Sholts et al., 2010) and child-death investigation (Davy-Jow et al., 2013). In a forensic anthropological context, the use of a TLS would be more appropriate in situations involving mass data collection, such as clandestine graves, mass graves, and ossuaries.

I specifically chose to test measurement accuracy within the discipline of forensic anthropology for three important reasons: 1) the availability of an expert (Dr. Albanese) in the department with the resources and expertise to advise this research, 2) the sparsity of literature concerning TLS and forensic anthropology, and 3) most importantly, the unique suitability of forensic anthropology to test measurement accuracy through the obtainment of baseline, standardized measurements, which can easily be compared against control (i.e., by hand) measurements and inputted into sex and stature estimation regression models as a secondary determination of accuracy and forensic utility. The effect of small measurement errors can be particularly revealing when osteological measurements are involved, as even small errors may result in inaccurate biological characteristic estimation. Thus, this discipline provides an excellent context in which to explore the validity and reliability of virtual tools such as TLS, the results of which, however, are not constrained solely to forensic anthropology.

When analyzing the results of this research (outlined in Chapter IV) it is important to be cognizant of the sources of TLS error. In the case of TLS, errors can result from both systematic error (of the Focus^{3D} and SCENE software) and human error. Systematic error concerns the generation of *coordinate values* that are inaccurate. Conversely, human error concerns the generation of *measurements* that are inaccurate. The former may exacerbate the extent of the latter. Systematic error can occur in the scanning stage and the scan registration stage: in the former, error is simply the result of using imperfect technology (the Focus^{3D} manufacturer's specifications list an error rate of +/- 2 mm up to a range of 25 m when calibrated). In the latter, inaccuracies can occur during scan registration if separate scans do not correctly merge into a cohesive point cloud and/or if there is insufficient data overlap (SCENE provides a registration report which displays mean point error).

In this particular measurement context, human error can arise from technical error of measurement (TEM), data insufficiency,[3] accidental incorrect point selection, and an inability to rotate the virtual viewpoint while obtaining measurements. If Mihandoost's intuition is correct, human error may significantly contribute to overall measurement error.

There is no way to avoid measurement error (this applies to any form of measurement, virtual or otherwise) so it is important to observe the degree of inaccuracy. This is particularly salient both in a forensic investigation and for an evaluation of novel scientific evidence, wherein knowledge of error rates is one aspect of the legal threshold for admissibility.

Conclusion

Despite a respectable breadth, existing literature has done very little to assess the validity of TLS-produced evidence in a forensic context. Furthermore, there is almost a complete absence of inquiry into the ways in which police use this technology. The

---

[3] Data was deemed insufficient if the lack or sparsity of data points made it impossible to visualize an osteological landmark in question.

absence of any discussion concerning the fluid nature of TLS-produced evidence, and the potential legal implications surrounding this fluidity, is equally stark. Now that TLS are being purposed specifically for forensics, scholarship must follow the line through to its completion. Similar to certain legal analyses of photography (see Mnoonkin, 1998), once a technology becomes sufficiently relied upon in legal settings, any and all limitations should be explored. I argue that this is doubly important when the technology purports to provide substantive information (i.e., measurements), such as in the case of TLS. The point is not to be unduly skeptical but to remain diligent in guarding the legal process from novel and marginally explored technologies.

Chapter III: Materials and Methods

Introduction

There were essentially two major parts to this research: firstly, an experimental aspect which quantitatively tested the accuracy and utility of TLS-produced measurements in a unique forensic anthropological context. This involved obtaining repeat osteological measurements in two different manners – traditionally and virtually – so as to compare measurement accuracy and observe any differences in the results of the forensic anthropological regression models based on said measurements.

Secondly, a case study and exegesis of the *R. v. Doodnaught* judgment and relevant sections of the trial transcripts – in tandem with relevant legal evidentiary literature and case law – to unpack and theorize on the implications of using TLS-produced evidence in court. *R. v. Doodnaught* provides an example of one of many ways in which TLS-produced evidence may be relied upon in court. Due to the versatility of this technology, and the fluidity[4] of the evidence it produces, it is important to consider the various forms its evidentiary products can take, as this will trigger different admissibility thresholds and considerations.

---

[4] I use fluidity to describe the potentiality, changeability, and versatility of TLS-produced evidence. Terrestrial laser scanner data (i.e., the point cloud) can stand alone as a source to draw measurements from or, depending on the volume and completeness of the data, can serve as an illustration of the overall environment. Conversely, it can also serve as the basis for computer-generated 3D models (from which measurements can also be obtained), animations, and simulations, each of which serves a unique purpose in court.

This chapter deals almost exclusively with the forensic anthropological experimental aspect. The second (i.e., legal) aspect, which is more of an exercise in literature review and analysis, is dealt with primarily in Chapter V.

The Scene

The experimental portion of this research involved simulating a forensic scene (i.e., a scene reasonably suspected to have involved a crime and/or to contain evidence relating to a suspected crime) involving skeletal remains, examples of which may include homicide scenes, clandestine graves, mass graves or any uncovering of skeletal remains, malevolent or otherwise. The purpose of this simulation is not to replicate the aftermath of any particular criminal offence, including the ones listed above. Indeed, the care in object placement involved in this experiment distinguishes it from the haphazardness one may expect to find in an actual forensic scene. Rather, the objective is to validate the accuracy – defined by Wang et al. (2012) as "the closeness of agreement between a measured quantity value and a true quantity value of a measurand" (p. 187) – of the Focus^{3D} through an innovative forensic anthropological application. The accuracy of an evidence-producing technology is relevant to the success of a forensic investigation and to both demonstrative and real evidence admissibility considerations. The most salient feature – the ability to virtually measure forensic objects (i.e., bones) – is present in both this simulated scene and in legitimate scenes, allowing for generalization. However, as a terrestrial laser scanner can only measure objects in its line of sight, the disarray of an actual crime scene *in situ* may provide additional challenges to forensic investigators measuring noninvasively with TLS.

The author created a "hot spot" – within Dr. Albanese's laboratory at the University of Windsor – by placing a number of bones into a space of approximately eight square meters demarcated by brown packaging paper resting on top of foam pads (see Figure 1 for an illustration of the scene). This "hot spot" is analogous to the most salient area(s) of a crime scene. It is not always initially clear what will be of most importance to an investigation, so saliency in this context refers to the area(s) with obvious evidentiary value (e.g., the immediate area surrounding a cadaver, or a mass grave pit).

Dr. Albanese and I placed 10 pairs of femora and hipbones (40 bones total) in succession in the “hot spot”. These bones are part of the teaching collection used in anthropology and forensics courses at the University of Windsor. It is worth noting that there are no premortem records of the sex and stature of these individuals. Typical experiments involving sex and stature estimation compare estimates against “forensic” sex and stature information (i.e., that reported in official records) to assess the forensic utility of the models (Cardoso et al., 2016); however, since I am assessing the impact of measurement error (i.e., the differences that arise between control and virtual trials), premortem records are irrelevant. In other words, the amount of change (if any) between the two measurement methods is of principal importance, not the closeness of agreement between virtual estimates and “forensic” information.

The hipbones were laid down on the iliac crest and approximately the distal end of the pubis symphysis, with the acetabulum facing upwards. The femora were placed anterior side upwards. Bone placement was determined based on a subjective assessment of which positions would render the relevant landmarks most visible to the TLS, ultimately yielding the best data and the best virtual measurements. The specimens were laid out in pairs by skeleton number from 01 to 12 (04 and 10 were absent) as if the individuals were lying supine.

Dr. Albanese and I also laid out two skeletons (one complete, one postcranial) and a variety of bones with distinctive features (e.g., fused vertebrae, bones torqued by muscle mass) with the intent to undertake a qualitative assessment of the laser scan data to determine if these features would be virtually recognizable. This aspect was ultimately abandoned due to data insufficiency and irrelevancy in terms of the main objective of this research. Bone measurements were only taken from the femora and hipbones (N=40) series.

Virtual Forensic Data Collection Procedures

When evaluating the accuracy of the Focus^{3D} (and planning data collection), I wished to pay particular attention to the procedures that shape the use of TLS in the forensic context. This endeavour is somewhat difficult due to the void in literature on how specifically police use TLS on scene, and the nonexistence of any recognized standard or guide for forensic TLS use. The latter may be attributable, in part, to the novelty of this application and the variety of TLS devices available on the market. Ultimately, I based the experimental procedures around how the TPS utilize TLS in forensics (though I was unable to perfectly adhere to all of their espoused protocols). In essence, procedural considerations include: number of scans and their positions, device settings levels, and the use of reference targets as a fail-safe for scan registration. These considerations are guided by two important forensic investigation objectives: accuracy and efficiency.

My correspondent in the TPS FIS indicated that terrestrial laser scanners, while a drastic change in technology, have not drastically altered forensic investigative procedures. This is largely because scene mapping with the Focus^{3D} is completed prior to the commencement of traditional crime scene investigation and documentation tasks (e.g., forensic photography, evidence collection, etc.) (A. Yousufi, personal communication, November 17, 2015). The objective in doing so is to capture the scene undisturbed, *in situ*. To capture as much information – in the form of data points – as possible, forensic investigators typically take multiple scans of a crime scene from a variety of positions, all the while being careful not to disturb the as-found environment. Exact positioning of the device in the scene is situation-specific. With the goal of capturing as much comprehensive data as possible in mind, the Focus^{3D} is placed around the room or, at the very least, around the object(s) or areas of importance (i.e., “hot spot(s)”). For the TPS, the average number of scans taken for an indoor scene is 20. For an outdoor scene, this may increase to 40 or even 60 depending on the size of the crime scene (A. Yousufi, personal communication, November 17, 2015). Similar to the positioning, ultimately, scene characteristics dictate how many scans are necessary.

The Focus^{3D} settings are also important to note, as the ‘resolution’ and ‘quality’ of the scans can be adjusted. The TPS, for example, typically use a resolution setting of 1/5th and a quality setting of four (of eight) (A. Yousufi, personal communication, November 17, 2015). As resolution and quality are increased, the time it takes to perform a single scan increases exponentially. The resolution setting modifies both the point volume and the distance between points: as resolution increases, the Focus^{3D} obtains a greater volume of points, which are spaced closer and closer together (Faro, 2011; J. Johnston, personal communication, April 14, 2015). The quality setting represents the number of times the laser strikes a single spot, the average values (x,y,z) of which become the reported coordinate values (J. Johnston, personal communication, April 14, 2015). The TPS have determined that the above settings levels result in high quality, comprehensive (in terms of data collection) scans while also keeping scan time to a minimum (approximately 9 minutes) (A. Yousufi, personal communication, November 17, 2015). While these settings do affect point cloud aesthetics, their true importance lies in their effect on data collection.

Lastly, as a fail-safe, the TPS almost always include reference targets in the crime scene, as they can be useful for automatic registration (A. Yousufi, personal communication, November 17, 2015). As I noted, they are not necessary for accurate registration *if* there is sufficient data overlap amongst scans. The quantitative amount of overlap necessary for accurate registration is not fixed – as long as there is enough data for the software to recognize identical geometric planes or shapes shared amongst different scans, it will be able to register without targets. Moreover, the Focus^{3D} has a built in inclinometer, compass, and altimeter, the data from which is collected with each scan and can be used by the software as an aid in the registration process. Nonetheless, the Focus^{3D} operating manual recommends the use of reference targets within the scene, as "you will usually achieve more precise registration results when you have manually placed such targets" (Faro, 2011, p. 53). The positioning of reference targets in the environment is determined by the line of sight of the Focus^{3D} (i.e., do not place them behind objects) and by a need to avoid disturbing or contaminating physical evidence. Ultimately, it may be prudent to use reference targets whenever possible so as to completely eliminate the possibility that registration cannot be performed.

In terms of the procedures for my own research, we collected three separate sets of scans, each at different settings levels. Each *set* of scans was comprised of three *individual* scans, all of which were centered around the "hot spot" described above.

The first two *sets* of scans *each* involved three scans total: two scans with the Focus^{3D} mounted on the tripod at identical heights, taken from the periphery of the "hot spot" area and one scan, with the Focus^{3D} placed on a wooden board on the floor, taken from the middle of the "hot spot" area. The final scan *set* was comprised of three separate scans, each taken with the Focus^{3D} mounted on the tripod at identical heights, around the periphery of the “hot spot” area. The three above-mentioned sets of scans ultimately resulted in three cohesive point clouds. Unfortunately, the first two sets of scans were ultimately discarded, as they both showed evidence of motion contamination (i.e., indications that at least one of the scan positions had shifted slightly during scanning, as evidenced by duplicated objects slightly off-position of the originals). This may have been the result of placing the scanner on the wooden board (resting on packaging paper and foam pads) in the middle of the “hot spot” area, which was not as stable as the tripod mount.

The resolution and quality settings for the third and only useable set of scans were set at 1/5th and three, respectively. This yielded scan times of approximately six minutes each and approximately 28.5 million data points per scan, or 85.5 million data points total in the registered point cloud.

The first two sets of scans included the placement of three white reference spheres in and around the “hot spot” area at varying heights and distances from the three scan positions. Reference targets were not available for the third set of scans, though registration was not hindered due to the natural geometry of the laboratory space.

The discrepancies between the TPS procedures I outline at the beginning of this subsection and the procedures I adhere to in my own experiment are the result of a combination of subjective assessment and resource constraints. These discrepancies include: performing less than the average (20) amount of scans, using resolution and quality settings below TPS’s typical settings levels, and not using reference targets in my final set of scans. While I take care to note these discrepancies, and elaborate briefly on their respective causes below, there is nothing to indicate that adhering more closely to the TPS-espoused procedures would have significantly increased the accuracy of my results. As I have mentioned, there are no recognized standards for the use of TLS in forensics. As such, reference to TPS procedures simply serves the purpose of having at least some guideline to compare to and, more importantly, reminds us that the way we operate this technology has implications on data collection and accuracy. Ultimately, adhering to the TPS’s terrestrial laser scanner forensic data collection procedures (to give it a name) is not a necessary condition for comprehensive data collection, as is indicated by my results.

Since we were originally obtaining sets of scans at different settings levels, it was decided that three scans per set would be both sufficient and efficient – after all, my major focus was on the "hot spot", not the entirety of the room. For the most part, three scans per set proved adequate; however, more scans would have undoubtedly increased the virtual aesthetics of the environment by providing more data points, theoretically eliminating all data insufficiency.[5] More specifically, more scans *at varying elevations* (in this context, lower to the ground would have been beneficial) would likely have improved the virtual aesthetics of the skeletal elements. Quantitatively, collecting high volumes of data points means increasing the potential for more precise measurements, which may affect accuracy; for example, you cannot precisely measure from the most superior point on the iliac crest *unless* a laser strike(s) has captured that specific coordinate value. As such, increasing the point volume also increases the probability that the exact point you wish to measure from is captured. To capture more points, the user must increase the number of scans (or the resolution setting, or both); however, this also increases time spent on scene, decreasing overall efficiency and generating large point cloud files. Qualitatively, as point volume increases, the image quality and clarity of the virtual environment also increases. As I noted earlier, ultimately, the situation determines the number of scans – as a rule of thumb, two scans is the minimum for 3D visualization.

The settings levels for the final set of scans (and, ultimately, the only useable set of scans), while not exact to what the TPS uses, are still very comprehensive. There are no "correct" settings levels – even the base settings will provide valuable information. Similarly to increasing scan volume, resolution and quality do not *directly* affect accuracy; rather, they *indirectly* provide a greater opportunity to obtain the most accurate data possible by increasing the volume of points obtained while decreasing the spaces between them, and providing a greater number of laser strikes to average coordinates from, respectively. Ultimately, time constraints and operating on the premise that there

---

[5] To be clear, eliminating all data insufficiency does not equate to increasing measurement accuracy. Increasing scan volume may have allowed me to obtain measurements that were otherwise unobtainable due to data insufficiency but this is not to say overall accuracy would automatically increase.

were other sets of scans to rely upon (which either matched or exceeded TPS's settings levels) contributed to the decision to operate the scanner below TPS's settings levels.

Finally, reference targets were not used in the final set of scans simply because they were not in the possession of the Faro Technologies employee at the time of scanning. Ultimately, this had no effect on the process.

Measurements

Each femur yielded three well-recognized, landmark-based measurements: maximum femur length, maximum femur head diameter, and epicondylar breadth. The hipbones each yielded three well-recognized, landmark-based measurements: hipbone height, iliac breadth, and superior pubis ramus length (SPRL) (see Figure 2 for an illustration of the virtual iliac breadth measurement). Albanese (2003) devised the final measurement, SPRL, as a more reliable alternative to the traditional pubic bone length measurement (Bass, 2005). The descriptions of all of these measurements can be found in Albanese's (2003) paper.

Dr. Albanese, a forensic anthropologist with a major research interest in the use of skeletal measurement to assess skeletal variation, and who possesses the relevant experience necessary to accurately collect the measurements, obtained all of the traditional measurements prior to placement in the "hot spot". These traditional measurements (i.e., by hand) are the control measurements against which the virtual measurements are compared. These measurements were collected in accordance with standard measurement practices using calibrated calipers and an osteometric board.

The virtual data was collected via the Focus^{3D}. Two Faro Technologies employees assisted in data collection by operating the scanner, describing relevant settings levels, and providing instruction on the use of the Faro SCENE software version 6.0 (hereafter referred to as "SCENE") (Faro, Lake Mary, FL). After scanning was complete, the Faro Technologies employees used SCENE to register the scans into cohesive point clouds.

Prior to data collection, I first spent a day familiarizing myself with SCENE and taking practice measurements, which I compared informally against the control measurements to acquire an indication of preliminary validity (i.e., that I was measuring what I thought I was measuring). This entailed experimenting generally with the measurement feature of SCENE, the point size setting levels (e.g., small, medium, or large), and determining the virtual viewpoints most conducive to obtaining the landmark-based measurements. Based on my preliminary work with SCENE, I enabled the “Gap Filler” setting, which fills the gaps between points (i.e., where no data was collected) that are physically close together (Faro, 2016) and set the point size to ‘small, adaptive’. These setting choices were determined by a subjective assessment of what was most conducive to the aesthetics of the virtual objects. In other words, rendering the best depictions of the morphology of the bones.

In addition, I had an in-depth conversation about the relevant skeletal landmarks on the femora and hipbones with Dr. Albanese, studied photographs of the skeletal landmarks taken by Dr. Albanese, and relied on Bass’ (2005) human osteology field manual to ensure that I was extremely familiar with the relevant sections of the femora and hipbones being measured. Three days later, I began formal data collection.

Using a trial version of SCENE provided by Faro Technologies, I formally collected the virtual skeletal measurements over a period of approximately two weeks. The only change I made to the original data was the removal of the data points comprising the walls and roof of the laboratory, which were not required for measurement and often obstructed the view of the “hot spot” area. This had no effect on measurement.

The SCENE software measurement function allows users to measure data point to data point within the virtual environment. Once the first measurement point is selected, the usually rotatable viewpoint becomes locked in place. This necessitates choosing a viewpoint wherein both points of measurement are visible *prior* to beginning the measurement process. It also necessitates that, after the measurement function is cancelled, the user rotate their viewpoint around the measurement they took, which is illustrated by a virtual line, to ensure that they have not unwittingly selected the wrong data point (i.e., on a different surface) (see Figure 3 for an example of incorrect point selection). This is an essential step because it can sometimes appear as if the user has selected the points they wanted, only to rotate the view and find that they have actually selected a stray point off of the intended object, or one on an object in the background.

The measurement process yields horizontal, vertical, and total measurements. Vertical distance was never needed. Between horizontal and total distance, I always reported the latter. The total distance best captured the true distances between the skeletal landmarks, which, due to the irregularity of their morphology (especially the hipbones), were often at varying heights and distances.

Statistical Models

Due to the nature of this experiment, I used a number of different statistical tests and models to assess the accuracy of the measurements collected, the differences between the means of the measurement trials (control, Trial 1, and Trial 2) and their errors, and their usefulness in estimating biological characteristics through forensic anthropological regression models.

I relied on Dr. Albanese's generic models (i.e., non-group-specific) to estimate sex and stature, respectively (Albanese, 2003; Albanese et al., 2016). The stature estimation model utilizes univariate linear regression and only requires the maximum femur length. Estimation was possible for all 20 femora (descriptive statistics for the stature estimates can be found in Table 1). The model provides estimated stature in centimeters and also determines an upper and lower bracket based on the standard error of the estimate (SEE). Albanese et al.'s (2016) results showed the femur to be the best (long bone) univariate predictor of stature, but the authors recommended a multivariate approach to further increase precision whenever possible.

The biological sex estimation model utilizes logistic regression and requires hipbone height, iliac breadth, SPRL, maximum femur head diameter, and femur epicondylar breadth. Albanese's (2003) study found this multivariate model to be the best-fit for sex estimation, while noting that it is not always the case that all of these skeletal elements will be available. Due to missing data, I was only able to use this model on 16 of 20 cases. The model estimates biological sex by expressing a p-value indicating the probability that sex was correctly allocated. P-values greater than 0.5 indicate male while values less than 0.5 indicate female. The p-values also signify the degree of statistical probability regarding accurate group classification: values closest to the extremes of the continuum (i.e., zero for females and one for males) suggest that a given case has been correctly estimated with a very high degree of statistical probability. As an example, a p-value of 0.98 indicates that there is a 98% probability the case is male and, subsequently, only a 2% chance it is actually female.

To assess the accuracy of the virtual measurements, I calculated the mean difference (MD) and mean absolute deviation (MAD) between the control group and Trials 1 and 2, respectively (inter-observer), and between Trials 1 and 2 (intra-observer). A summary of these statistics can be found in Table 2. MD is the average of the difference between two sets of measurements (e.g., control and Trial 1, or Trial 1 and Trial 2). MD can be useful for highlighting directionality in aggregate error; however, it is also limited by the fact that positive and negative errors may cancel each other out, obscuring the true magnitude of the difference between measurements. MAD, on the other hand, is the mean of the *absolute value* of the differences between two sets of measurements. In other words, positive and negative is ignored, yielding the true magnitude of the difference between measurements. MD and MAD were calculated in millimeters and also converted to percentages. The rationale behind using percentages is that any given size of metric error will have a different meaning depending on the context. For example, 2 mm error is much more significant for the SPRL measurement (a smaller distance) than the maximum femur length measurement (a larger distance). As such, percentages are comparable due to their standardization.

To determine whether there existed any statistically significant differences between the control measurements and the TLS-produced measurements, I utilized the SPSS statistical package version 23 (SPSS Inc., Chicago, IL) to run one-way Repeated Measures ANOVA, which is built specifically for datasets that contain measurements on the same subjects but over time or under variable conditions. One-way Repeated Measures ANOVA allows users to statistically assess whether there exist any significant differences between group means, and a *post hoc* Bonferroni correction shows exactly which means are significantly different from one another. For all three sets of trials (control, Trial 1, and Trial 2), I used this statistical test to check for differences between the measurement means and the stature means produced through the linear regression model.

In addition, I conducted a technical error of measurement (TEM) assessment to determine the level of variability attributable to TEM between my two measurement trials (i.e., intra-observer). TEM accounts for variability resulting from technical shortcomings, such as variations in locating landmarks or inconsistently executing a measurement technique in repeated measurement sets (Perini et al., 2005). The goal is to minimize TEM as much as possible and, at the very least, to consider its effect on measurement variability when assessing repeated measurement accuracy. Inter-observer TEM is also possible, but it is only applicable in situations where the measurement tool is the same across observers (Perini et al., 2005). Since Dr. Albanese used traditional methods and I used virtual methods, the inter-observer TEM does not apply. For beginners in anthropological measurement, Perini et al. (2005) recommend a 1.5% intra-observer threshold, above which values are deemed unacceptable.

It is worth noting some data limitations when considering the results of the one-way Repeated Measures ANOVA tests and the forensic anthropological regression models discussed in Chapter IV. These limitations include: 1) the bones included in this experiment were not randomly selected so as to be representative of a wide range of human variation. Resource constraints required that I use the bones available to me, providing a convenience sample. Since my primary focus was on measurement accuracy – and the forensic anthropological regression models can be generically applied (with great accuracy) regardless of the demographic characteristics of the sample – this is not a significant problem. 2) Even though I collected 230 measurements in total, their division into measurement groups meant that I was comparing means whose Ns ranged from 17 to 20. These sample sizes are relatively small, potentially decreasing the power of my results. However, these regression models are purposed to apply well to specific cases (i.e., when a single skeleton is found). Moreover, other research in this discipline has determined the success of these models with much larger N sizes (see Albanese, 2003 & Albanese et al., 2016). 3) In some cases, the assumptions – including approximately normally distributed dependent variables, sphericity, and the absence of any outliers – underlying one-way Repeated Measures ANOVA were violated. The Shapiro-Wilk statistic tests whether a sample is normally distributed. It may be more difficult to achieve normality with small sample sizes. Fortunately, one-way Repeated Measures ANOVA is particularly robust to violations of normality (i.e., slightly non-normal samples should have no negative effect). Non-normal samples are specified in Chapter IV. None of the samples violated Mauchly’s Test of Sphericity and there were no outlier values. Working with real-world data often includes violating assumptions and it is important to note them always and correct for them whenever possible.

Conclusion

In this particular experiment, the accuracy of the measurements is wholly dependent on TLS technology. Subsequently, the accuracy of skeletal measurements is critical to the success of the above-mentioned regression models. Depending on the degree of error, and the model in question, the use of incorrect measurements can result in inaccurate stature estimation, misclassification of biological sex, and an element of uncertainty expressed by weak p-values. Albanese (2003) found that, in sex estimation, pelvic measurement errors of approximately 2% can make the difference between a correct or incorrect allocation of sex. In the forensic anthropological context, biological characteristic estimation errors can confound or even ruin an investigation. Inaccurate measurements can have the same effect in the general forensic context.

Chapter IV: Experiment Results

Introduction

This chapter provides an overview of my findings concerning the accuracy of the Focus^{3D} and its ability to capture virtual measurements conducive to biological characteristic estimation, which is utilized in forensic anthropology. I outline relevant MD and MAD values, the results of the one-way Repeated Measures ANOVA tests, the results of the forensic anthropological regression models, and TEM considerations. I then discuss the implications of my results in Chapter V.

Results

Formal data collection entailed systematically obtaining landmark-based measurements from the TLS-produced data over two separate trials, two days each. On the first day, I collected femora measurements. On the second day, I collected pelvic measurements. To assess intra-observer error, this identical process was repeated a week later. I entered measurements into an Excel datasheet as I collected them. Ultimately, this yielded 230 measurements in total, 115 measurements each trial. The total should have been 240 measurements, but areas of insufficient data on one femur and three hipbones made certain measurements impossible. Descriptive statistics for the landmark-based measurements can be found in Table 3.

The inter-observer error between the control measurements and Trials 1 and 2, respectively, was quite small on average (see Table 2). Across all bone metrics, average MD and MAD between control and Trial 1 was 3.19 mm and 3.28 mm, respectively: hipbone height had the highest error (MD=4.81 mm, MAD=5.03 mm) while maximum femur head diameter had the lowest error (both MD & MAD=2.03 mm). Across all bone metrics, average MD and MAD between control and Trial 2 was 2.87 mm and 3.12 mm, respectively: SPRL had the highest error (MD & MAD=5.18 mm) while epicondylar breadth had the lowest error (MD=1.01 mm, MAD=1.74 mm). Considering measurement error in terms of percentages, the SPRL metric had the highest inter-observer measurement error in both trials (control-Trial 1: MAD=6.75%, and control-Trial 2: MAD=7.5%). Overall, intra-observer measurement error was also extremely small across all bone metrics (MAD=1.42 mm). SPRL had the highest intra-observer MAD at 2.52 mm (3.76%).

I performed separate one-way Repeated Measures ANOVA tests for each sample group (i.e., the bone measurements and the stature estimates). For the first group, this was broken down further into bone metric categories, wherein each category was subjected to separate one-way Repeated Measures ANOVA tests. For example, the three hipbone height measurement trials comprised a single one-way Repeated Measures ANOVA test, the three SPRL measurement trials comprised a single one-way Repeated Measures ANOVA test, and so on. It is possible to include all of the relevant factors into the model simultaneously, but this results in a listwise exclusion of cases, unnecessarily decreasing N sizes to reflect the smallest N size amongst all samples.

Shapiro-Wilk tests indicate that five of 21 samples in the ANOVA tests (separated by bone metric and/or trial) were reported as non-normal (most were borderline). These included: Trial 1 maximum femur length (p=0.044) and epicondylar breadth (p=0.045), Trial 2 maximum femur length (p=0.037), and the stature estimates for Trial 1 (p=0.044) and 2 (p=0.037).

Ultimately, every bone measurement group metric (Table 4) and the stature estimation group (Table 5) showed statistically significant differences between trials as reported by the one-way Repeated Measures ANOVA. The *post hoc* Bonferroni correction revealed the means for the vast majority of pairs were statistically significantly different. Tables 4 and 5 summarize the relevant statistics for each one-way Repeated Measures ANOVA test.

In the overall *bone measurements* group, all of the inter-observer means (between control and Trials 1 and 2, respectively, for each bone metric) were statistically significantly different except for the epicondylar breadth between control and Trial 2 (p=0.172). Hipbone height (p=0.034), epicondylar breadth (p=0.023), and maximum femur head diameter (p=0.001) intra-observer means were also statistically significantly different. Conversely, iliac breadth (p=1.000), SPRL (p=0.749), and maximum femur length (p=1.000) intra-observer means were not statistically significantly different.

In the *stature estimation* group, both inter-observer means were statistically significantly different (Trial 1: p<0.000, Trial 2: p<0.000) while the intra-observer mean was not statistically significantly different (p=1.000).

Table 1 shows that the means and standard deviation values of all three stature estimation trials are remarkably close together. This indicates that there are negligible differences between the control and virtual stature estimation trials. As the maximum femur length MAD between control and Trial 1, and control and Trial 2, is 2.83 mm (0.68%) and 3.08 mm (0.74%), respectively, this is not surprising. When expressed as percentage errors of the control-estimated stature, 90% of the cases have stature error values that are less than 1%.

The biological sex estimation logistic regression results indicate close agreement between the three trials. Sex allocation is consistent across all trials (Males=10, Females=6) but with some variation in the p-values. Using the control measurement estimations as a baseline, of the 16 useable cases, only three (03L, 03R, and 09R) cases show p-value changes greater than 3% across any trial, inter- and intra-observer alike. For case 03L, the p-value dropped from approximately 98% to 96% between control and Trial

1, and then increased to 99% in Trial 2. For case 03R, the p-value dropped from approximately 95% to 89% between control and Trial 1, and then increased to 99% in Trial 2. For case 09R, the p-value dropped from approximately 99% to 97% between control and Trial 1, and decreased further to 95% in Trial 2. Of the 48 p-value errors (i.e., the p-value differences between trials), 35 values (73%) express less than a 1% change.

Finally, intra-observer TEM calculations were performed on each bone metric in the overall *bone measurements* group. These values range from high (SPRL relative TEM=3.56%) to low (maximum femur length relative TEM=0.34%) with an average of 1.47% across all bone metrics. The relevant summary statistics for the TEM calculations can be found in Table 6.

Conclusion

My research found the Focus^{3D} to be highly accurate with an ability to generate measurements useful to forensic anthropological biological characteristic estimation. In the majority of the cases, virtual trials performed just as well as the control trial in estimating sex and stature. In this context, an expert forensic anthropologist would be able to come to largely the same conclusions regardless of whether they were using traditional or virtual data. MD values also illuminated a chronic tendency to under measure. These results cannot be generalized to other TLS devices.

Chapter V: Discussion

Introduction

In this chapter, I undertake a discussion of the results found in Chapter IV. This includes interpreting the closeness of agreement between the three (control, Trial 1, and Trial 2) sex and stature estimation trials despite statistically significantly different sample means, discussing the implications of chronically under measuring in this particular forensic anthropological context, and considering the applicability of TEM thresholds proposed by Perini et al. (2005) to my own research.

I also expand upon the discussions I began in Chapter II concerning the practical and technical considerations that affect the accuracy of TLS-produced measurements, and the evidentiary classification and implications of using TLS-produced evidence in court. Ultimately, I argue that these implications do not operate in a vacuum: how TLS-produced measurements or displays are generated is relevant to legal admissibility considerations, and whether TLS technology is widely adopted in forensics may be determined, at least in part, by admissibility decisions.

Experiment Results

As the results indicate, the virtual measurements obtained by the Focus^{3D} are extremely accurate. The average inter-observer error across 230 measurements is approximately 3 mm (MD and MAD, respectively), which is exceptional considering the low quantity of scans and the subtle, landmark-based measurements sought. The average intra-observer error across 230 measurements is approximately 1.5 mm (MAD). In terms of pure accuracy, this technology and software performed quite well in aggregate.

Notwithstanding the high degree of accuracy, the MD values illuminate a tendency to under measure when compared against the control measurements. The tendency to under measure when using the Focus^{3D} and SCENE is consistent with Mihandoost's (2015) findings (Mihandoost also found this to be the case when using a different measurement software).

When considering the dual purpose of this experiment, the directionality of the errors poses implications for the success of the forensic anthropological regression models. The nature of this experiment allowed me to assess not only accuracy, but also forensic utility, the latter which was solely dependent upon the former. Ultimately, despite a chronic tendency to under measure, in the majority of cases, the virtual measurements from both trials performed just as well as the control measurements when inputted into the forensic anthropological regression models. This is, in part, a testament to both the accuracy of the Focus^{3D} and the robusticity of the generic forensic anthropological models developed by Albanese (2003) and Albanese et al. (2016).

In the case of biological sex estimation, it is also the by-product of using a multivariate model. In this particular context, a tendency to under measure will skew the p-values – depending on the measurement in question – towards either male or female allocation. For example, male bones are, on average, *absolutely* larger than female bones.

However, *relatively speaking*, female hipbones are greater in size compared to their overall body proportions (due to sexual dimorphism related to childbirth). As such, certain pelvic measurements are *relatively* larger in females than in males. For example, Albanese's (2003) research reported negative coefficients for iliac breadth and SPRL (i.e., when holding the other predictors constant, unit increase in iliac breadth or SPRL equates to decreasing the likelihood of being classified as male). All other coefficients showed positive directionality (i.e., when holding the other predictors constant, unit increase in hipbone height, maximum femur length, epicondylar breadth, and maximum femur head diameter equates to increasing the likelihood of being classified as male). Therefore, a tendency to under measure the above-mentioned bones, whose coefficient directions are opposite, should provide a counter-balancing effect. This may help to explain why, in aggregate, the sex estimation p-values changed very little despite the tendency to under measure and the presence of statistically significantly different means.

As an example in my own research, observe case 03R in the sex estimation model, the measurements of which are displayed in Table 7. Except for maximum femur head diameter in Trial 2, both virtual trials under measured all bone metrics. As we will recall, the p-value dropped from approximately 95% to 89% between control and Trial 1, and then increased by 10% to 99% in Trial 2. In all cases, the model is clearly estimating the individual is male. Nonetheless, the drastic difference between p-values bears investigating. The difference between the p-values of Trial 1 and 2 best exemplify the anthropological principles in the above discussion.

In Trial 2 (compared to Trial 1), hipbone height increases by 2.3 mm (1.16%), iliac breadth decreases by 0.4 mm (0.3%), SPRL decreases by 6.4 mm (9.4%), maximum femur head diameter increases by 1.6 mm (3.4%), and epicondylar breadth increases by 0.6 mm (0.7%).[6] This sequence of measurements perfectly exemplifies the potential for a compounding effect to occur (i.e., decreases in iliac breadth and SPRL, skewing towards male, and increases in hipbone height, maximum femur head diameter, and epicondylar breadth, also skewing towards male) despite chronic under measuring, which, in absolute terms, should skew towards female. It is prudent to be cognizant of this phenomenon when interpreting p-value changes across measurement trials.

---

[6] Dividing the error value by the respective control measurement and multiplying by 100 produced these percentages. They provide a better indication of the magnitude of measurement error.

The virtual stature estimation data, too, faired extremely well when compared against the control data. The maximum inter-observer stature error value was only 1.63 cm (approximately 1% of the control-estimated stature for that case). The closeness of agreement amongst all three trials is not surprising considering maximum femur length measurement errors were routinely below 1% for both virtual trials when compared to the control. It is, however, particularly interesting that the inter-observer means were shown to be statistically significantly different, as this difference did not reflect in the stature estimation values, nor in the means and standard deviations when compared across the three trials. The miniscule quantity of the maximum femur length errors explains the closeness of agreement between the stature estimation values; however, it does not explain why the inter-observer means were statistically significantly different. This seemingly paradoxical divide may be explained by considering the difference between the *accuracy* characteristics of the virtual measurements (i.e., chronically under measured), but which, at least in this case, had negligible negative impact on the *forensic utility* of the measurements for stature estimation. In other words, the inter-observer means were often found to be statistically significantly different but this hardly changed the results produced by the regression models.

Finally, the TEM thresholds recommended by Perini et al. (2005) further complicate the interpretation of the divide between accuracy and forensic utility. The relative TEM for three bone metrics (SPRL, epicondylar breadth, and maximum femur head diameter) fall above the relative TEM threshold of 1.5% proposed by Perini et al. (2005) (Table 6). Considering the results of my research, the TEM thresholds proposed by Perini et al. (2005) may not be useful for judging measurement utility in this particular context. In addition, these thresholds appear to be arbitrarily set and are relative to the anthropometric element in question. Nonetheless, it is likely the case that, in this context, the relative TEM values for the bone metrics that did exceed the 1.5% threshold could be reduced through continued measurement training, decreases in data sparsity, and the ability to rotate viewpoints during virtual measurement obtainment. Notwithstanding these observations, the forensic utility of the majority of the measurements remained virtually uncompromised when compared against the control measurements.

Terrestrial Laser Scanning in Forensics

It is undeniable that terrestrial laser scanners possess great potential in real-world, forensic scenarios. This potential is already being advantaged by a number of policing agencies. TLS have the ability to noninvasively and comprehensively capture and preserve environmental data, decrease police costs and increase efficiency, and provide compelling evidence in legal settings. As I have shown in this research, the Faro Focus^{3D} is a highly accurate metrological device. Due to the number of TLS available on the market, it should not be taken for granted that this applies equally to all of them.

When TLS are utilized in a forensic context, policing agencies should strive towards achieving uniformity in their procedures and practices. At the very least, policing agencies should consider the effect of scan quantity and positioning, resolution and quality settings, and the use of reference targets on point cloud quality. Academics that undertake forensic research with these devices should also consider the same. These considerations may vary from device to device but it is important to be cognizant of device-specific setting permutations. A terrestrial laser scanner only serves as a useful tool if it is used in a way that maximizes its accuracy; if accuracy is sacrificed, all other benefits, including efficiency, become irrelevant.

This experiment purposefully endeavoured to test the accuracy of the Focus^{3D} in a forensic anthropological context that was challenging, innovative and plausible. Accuracy results, however, can be generalized beyond the discipline of forensic anthropology. Obtaining baseline, standardized bone measurements served the dual purpose of testing accuracy and, subsequently, testing forensic utility through forensic anthropological regression models.

While the Focus^{3D} performed exceptionally well, it may be the case that traditional measurement techniques are most conducive to accurately capturing certain bone data. On a number of occasions I have mentioned some of the difficulties inherent in the virtual measurement process, such as data sparsity and an inability to rotate the virtual viewpoint while obtaining measurements. The latter limitation was only relevant to the obtainment of certain measurements, such as the SPRL and the hipbone height, but likely inflated the mean inter- and intra-observer measurement error for these metrics. In the case of the former, Albanese's (2003) intra-observer measurement error for SPRL was

0.57% as opposed to 3.76% in my study. Establishing a viewpoint that rendered completely visible both the landmark in the acetabulum and the landmark on the superior margin of the pubic symphysis simultaneously was often extremely difficult, contributing to measurement error for this metric. In the case of hipbone height, an overhead viewpoint was most conducive to selecting the measurement point at the most superior point on the iliac crest, whereas a side profile viewpoint was most conducive to selecting the measurement point at the most inferior point on the ischial tuberosity. It is impossible to simultaneously adopt both viewpoints and I ultimately utilized an overhead viewpoint, afterwards confirming point selection on the ischial tuberosity from the side profile viewpoint.

Data sparsity and viewpoint problems do not exist in a traditional measurement context. Furthermore, when the objective is to obtain *maximum* distance measurements (e.g., of the femur length and epicondylar breadth), the simple nature of the process of manipulating a bone with an osteometric board provides a greater opportunity to realize the true maximum dimensions of an object. In that context, traditional measurement is preferable to virtual measurement; however, it may not always be the case that the investigator will want to, or be able to, physically handle the object in question, which is the reason for (and advantage of) TLS in the first place.

Capturing subtle measurements, such as those common in forensic anthropology, is on the increased difficulty end of the spectrum in regards to what scenarios TLS may be deployed in. It is more likely the case that routine measurements will be captured in centimeters and meters – situations where up to 5 mm error values are negligible. Nonetheless, the types of experiments that truly test the bounds of our forensic technologies are the ones that assure us of their accuracy in less demanding situations. Policing and legal professionals should not take for granted the accuracy of these devices considering the critical circumstances in which the evidence may be used. For example, it is entirely plausible that situations may arise where the data captured by a TLS provides the only opportunity to obtain measurements after a scene has been released. In certain scenarios, the spatial interactions between various artifacts in a crime scene may be of great importance to the investigation (Maksymowicz et al., 2014). Since the substantive truths of the data produced by TLS may be relied upon in crucially important situations, its accuracy and limitations must be known. In turn, the extent to which considerations of accuracy and device limitations constrain the admissibility of this type of evidence is entirely dependent on how – or in what form – it is introduced in court.

While a full determination of the exact proportions attributable to each form of error – systematic and human – is beyond the scope of this research, I surmise that human error was likely the greatest overall contributor to total measurement error. Firstly, the Focus^{3D} requires a yearly calibration and was properly calibrated at the time of the experiment.[7] Secondly, SCENE reported a registration mean error value under 2 mm. Both of these observations indicate that the hardware and software were operating correctly. On the other hand, data point sparsity in some areas made it qualitatively difficult to precisely locate certain landmarks. This is reflected most prominently in the inflated inter-observer percentage MD and MAD of the SPRL, which was the most difficult virtual measurement to obtain.

Terrestrial Laser Scanning in Law

It is not always the case that TLS-produced data will be introduced in court (quite the opposite, as we see the first instance in Canada arriving three years after TPS began using the technology); however, when it is introduced, whether as demonstrative or real evidence, it is likely to have a great impact on the proceedings, especially in the case of the latter classification.

As a case study, I turned to *R. v. Doodnaught*, the only Canadian case to deal with TLS-produced evidence thus far. While the judge's ruling on the matter provides some insight into how TLS-produced evidence was used in this case, the trial transcripts offer a more direct look at the judge's minute assessment of this evidence, which allowed me to follow along as he wrestled with admissibility considerations during the *voir dire*.

In *R. v. Doodnaught*, the issue at hand was the alleged sexual assault of 21 female patients by a resident anesthesiologist, Dr. Doodnaught, over a span of years. All of these alleged assaults took place during the course of surgeries, in various hospital operating rooms (OR). The trial judge identified one of the two core points at issue as "the

---

[7] The proposition that the Focus^{3D} was operating in a state of proper calibration is one made by the Faro Technologies employees (and the accuracy results seem to support this proposition). I do not know what is involved with calibration, was not present when the device was calibrated, and do not know the circumstances that cause a device to become uncalibrated.

opportunity to commit the offences", which was broken down into two separate issues: 1) likelihood of detection, i.e., whether or not the anesthesiologist's colleagues, who were present during the surgeries, would have been able to witness the alleged acts taking place. The Crown posited that medical draping set up around the OR and operating table were sufficiently large enough to have obstructed Doodnaught's colleagues' lines of sight, allowing the acts to take place undetected - the defence argued the opposite. And 2) Physical impossibility, i.e., whether the height of the operating table, the physical dimensions of Dr. Doodnaught, the height and positioning of the complainants, and the presence of certain, obstructing medical equipment, or a combination of all or some of these factors, made it impossible for some of the acts to have taken place in the manner in which they were alleged. Again, the Crown argued the acts were possible while the defence argued they were impossible.

The very crux of the issue of opportunity is one of metrology. The positioning of OR personnel and the dimensions of various elements in the OR were needed to begin to establish the truth of the matter. This situation seems to be the perfect opportunity for the introduction of TLS-produced measurements; however, the Crown and the defence proffered only *demonstrative* evidence *based* on TLS-produced data, which the judge ultimately admitted as such (using them, however, to varying degrees). While all parties involved recognized the substantive potential of TLS, their evidentiary objectives (i.e., to use TLS-produced evidence to bolster witness testimony about positioning and draping) occasioned only the need for demonstrative evidence. This decision was also shaped, in part, by logistical concerns. Due to the variety of different surgeries the complainants underwent, the different draping setups for each, the changing personnel for each, and the great span of time over which these events took place, the trial judge determined it to be nearly impossible to substantively ascertain the exact extent (i.e., dimensions) of the draping with precision. In the words of the trial judge: "The best that could be expected was to obtain a reasonably accurate understanding of the draping in the ORs at the relevant times." (Doodnaught judgment, 2013, p. 13). To do so, the judge physically *took* *a view*[8] of the OR rooms but also relied upon the *virtual views* produced by the TLS. Interestingly, only the latter was used as evidence upon which to decide the fact in issue.

The Crown evidence was ultimately determined to be the "best evidence of the draping setups and general [OR] layouts … I [the judge] conclude that they [laser scans and pictures] provide a fairly accurate depiction of the typical setup for each of the procedures." (Doodnaught judgment, 2013, p. 13). Conversely, the defence "tendered 3-dimensional model video reconstructions of the OR setups, with medical personnel inserted in the videos in the locations asserted by the defence" where the trial judge "found them helpful … in understanding the defence arguments, but … [did] not find them helpful as stand-alone evidence concerning the height and width of the draping, or the positioning of the medical personnel." (Doodnaught judgment, 2013, p. 13). Ultimately, Dr. Doodnaught was convicted of all 21 counts, meaning that the judge determined the opportunity to commit these offences to be possible beyond a reasonable doubt. The judge did not arrive at this conclusion through the substantive, real evidence of TLS measurements, though the demonstrative displays produced by TLS did inform his decision.

R. v. Doodnaught is particularly interesting because it highlights how the judge's determination of the form of the evidence – demonstrative or real – is influenced by the particulars of the case. This, in turn, shapes the stringency of admissibility thresholds, even if the evidence appears *prima facie* to be novel scientific and, thus, subject to more rigorous scrutiny. My exegesis of the judge's decision and the trial transcripts is not to insinuate that the judge acted incorrectly – I do not believe that he did – but rather to better understand his thought process in dealing with TLS-produced evidence and to highlight the legal quagmire surrounding the use of TLS in forensics and law.

If the evidence had been introduced as real evidence on the basis of its measurement capabilities, it is likely the case that the novel scientific application (which the judge recognized) would be questioned and subjected to "gatekeeper scrutiny" and the admissibility tests in *J.-L.J* (Hill et al., 2013). It is probable that the TLS-produced evidence, at least if supplied by the Focus^{3D}, would satisfy all four criteria in *J.-L.J.* The

---

[8] 'Views' are occasions where the trier(s) of fact must go to see the real evidence when it is impossible to bring it into the courtroom. The discretion to order a view resides with the judge. TLS-produced virtual views may, in the future, replace the need for physical views.

judge would also have to be satisfied that the probative value of the evidence outweighed any prejudicial effect. It may be the case that TLS-produced evidence is actually less prejudicial in a measurement capacity as opposed to an illustrative (i.e., demonstrative) capacity. In the latter case, TLS's impressive ability to reproduce an almost perfect reflection of reality and display it to triers of fact may leave an unduly lasting impression. Agosto et al. (2008) go as far as to say "evidence gathered with laser scanning can be more compelling for juries" (p. 6282). Moreover, scholars have recognized the influential nature of visual aids as opposed to purely verbal aids (Fiedler, 2003; Tung et al., 2014). Whether this is helpful or prejudicial will be dependent on the particulars of the case, the form of the TLS evidence, and the narrative surrounding it. It will also depend on whether it is a judge or jury trial, as the former is more suited to disregarding undue weight.

When novel scientific evidence is introduced, expert opinion evidence may also be required. If the expert discussion on TLS-produced evidence exceeds the simple explanation of technical matters, the expert opinion evidence, too, would trigger its own thresholds. If it does not, the criteria espoused in *Mohan* and *White Burgess Langille Inman* need not apply. The crucial element is ultimately whether an *opinion* is given or just expertise in the form of scientific and/or technical knowledge. In the case of TLS, judges must be sure to avoid letting the expert usurp the fact-finding process through an opinion that infringes too much on the ultimate issue (Hill et al., 2013). Taking *R. v. Doodnaught* as an example, if an expert, armed with TLS data, gave the opinion that it would have been possible for Dr. Doodnaught to commit the assaults, this may come too close to deciding the ultimate issue of guilt or innocence, which must be left to the trier of fact.

While both parties did call "experts" to explain the minutiae of TLS-produced evidence, and the judge (rightly so) inquired about TLS accuracy, its novel application in forensics, and whether it had been barred from any prior judicial proceedings, it was admitted only as demonstrative evidence, under a lower threshold than expert opinion evidence and/or novel scientific evidence. Even though the judge's demonstrative evidence considerations overlapped typical novel scientific evidence considerations, the admittance of the TLS-based renditions (i.e., video animations, models) as demonstrative, and not real, evidence did not occasion the need to systematically weigh novel science threshold considerations. After all, in this case, the data was not relied on for its substantive truths but rather for its demonstrative capabilities.

Finally, I shift my critical gaze from TLS to my own research and pose an important question: what is the value of my scrutiny of TLS, a technology that has been used in some jurisdictions for over a decade and has already been accepted by some courts? As context, the dissenting justices in *Trochym* worried that the majority was establishing a regressive precedent – by rejecting hypnotically refreshed memories as evidence despite previous use in court – that would allow parties to call into question techniques and scientific evidence that had achieved general acceptance for decades (Delisle et al., 2015). With respect, I believe this analysis misses the mark. The decision in *Trochym* was not a *carte blanche* invitation to question even the most rudimentary techniques and scientific procedures (which the majority recognized). Rather, it was an important effort to truly act as a gatekeeper against unvalidated forensic science, *despite* prior (and less-critical) acceptance. If the underlying science is equivocal or unproven, it should not be trusted in such an important setting. Mistaking our acceptance of a science or technique as proof of its underlying validity is exactly what allows it to incorrectly proliferate in the first place (Latour, 1987). As Latour (1987) notes, *tacit knowledge* is rarely scrutinized. In a legal context, if it underwent appropriate scrutiny prior to becoming tacit, then critique may be unnecessary; however, if it did not, if it was uncritically accepted and repetitious acceptance cemented its authority, then it is the duty of the legal system to critically assess such knowledge.

As such, in the context of the relatively recent introduction of TLS-produced evidence in Canadian court, and the devastating critiques of widely accepted forensic science, I view this critical exercise as extremely important and necessary. It was never the case that I expected to find extraordinary issues with this technology. Nonetheless, its fluid nature, novel forensic scientific application, and potentially great impact upon the trial process necessitates definitively ruling out (or at least acknowledging) the presence of salient limitations. Terrestrial laser scanning is *not new science*; however, its routine use in forensics, and its use in law, *is new*. With any novel science, strict analysis should be conducted prior to its acceptance in court. If not conducted, prohibiting the ability to retrospectively analyze its validity (the principle espoused in the *Trochym* dissent) means that *stare decisis* reigns and that potentially unvalidated or erroneous science is allowed to continue to have a negative impact on the fact-finding process. This unequivocally negative impact may manifest itself in the form of miscarriages of justice, which we have already seen result from the use of unproven forensic science. Furthermore, as the *Trochym* majority notes: “the admissibility of scientific evidence is not frozen in time” (cited in Delisle et al., 2015, p. 911). It is prudent to continually re-test our forensic technologies and methods as they change and evolve.

Conclusion

While I have purposely discussed TLS in the separate contexts of forensics and law, ultimately, the use of this technology becomes most critical at the intersection between the two. Although the results concerning the accuracy and utility of the measurements generated by the Focus^{3D} cannot be generalized to other laser scanning devices (which should not be accepted without scrutiny), my critique of TLS as used in forensics and law can be generalized to this technological field as a whole.

Chapter VI: Conclusion

This research focused on the accuracy and forensic utility of one terrestrial laser scanner and software package – the Faro Focus^{3D} 330X and SCENE – and the forensic and legal implications of using terrestrial laser scanning technology in general. The accuracy and forensic utility results of this thesis apply only to the above-mentioned hardware and software. The discussions on the fluidity of TLS-produced evidence and the legal admissibility considerations, however, can be extrapolated to other TLS devices.

A literature review on the relevant material showed that TLS has received some attention amongst forensic scholars; however, almost no attention has been paid to validating the accuracy of these devices, considering their limitations, and considering the impact in law when TLS-produced evidence is introduced. As such, this research validated the accuracy of the Focus^{3D} by obtaining bone measurements, which were then inputted into forensic anthropological regression models. Despite being obtained in a specific forensic anthropological scenario, the results are applicable to any forensic context wherein precise measurements are relevant to the investigation.

When compared to the control measurements, which were taken by hand by Dr. Albanese, the average inter-observer error (both MD and MAD) produced by the Focus^{3D} was approximately 3 mm. While this exceeds the manufacturer's specifications of +/- 2 mm, it is still extremely accurate. It is likely the case that the obtainment of minute measurements such as these will be needed less frequently in other forensic scenarios (e.g., homicide or arson scenes), further mitigating the impact of errors on the investigatory and legal process. Moreover, by distinguishing between systematic and human error, I was able to deduce that it is likely the case that human error is a greater contributor than systematic error to overall measurement error. However, an empirical study to this effect was beyond the scope of this research.

Although one-way Repeated Measures ANOVA tests showed most of the inter-observer bone metric means to be statistically significantly different, this had negligible impact on the results of the sex and stature estimations. Despite a tendency to under measure in the virtual trials, in the majority of cases, the virtual measurements performed just as well as the control measurements. While this speaks to the accuracy of the measurements taken by the Focus^{3D} and the robusticity of the regression models used, in the sex estimation context, it is also attributable to a counter-balancing phenomenon that reflects the relative and absolute bone size differences amongst the biological sexes. In a univariate model, or depending on which measurements are used in a multivariate model, this may skew p-values towards one of the sexes incorrectly. In other words, the directionality of error values may complicate analysis depending on the context in question.

The fluid nature of TLS-produced evidence means it can take many forms. It can produce substantive (or real) evidence in the form of highly accurate measurements, which are situated in a visual display of the actual environment from which they were obtained. The point cloud can also serve as the foundation for demonstrative displays such as 3D models, animations, and simulations. Since TLS in the forensic context can be considered novel science (due to its relatively new application and lack of legal scrutiny), when called upon as real evidence, it should trigger the legal thresholds related to novel science. It may also be accompanied by expert opinion evidence, which triggers its own thresholds. If the technology is simply explained and no opinion is given, the latter threshold need not apply. When TLS-produced evidence is used demonstratively, it is important to remember that the accuracy of these displays is contingent on the accuracy of the underlying data. In the case of demonstrative evidence, inaccuracies may go to weight and not admissibility when considering the probative versus prejudicial effect. In the case of real evidence, a mistaken measurement can mean the difference between corroborating whether a witness truly could see a crime take place, or whether there was an opportunity for a crime to happen undetected. I utilized the Canadian case of *R. v. Doodnaught* as an example of one such situation where TLS-produced evidence was used to determine a fact in issue (though the evidence was only used as demonstrative). Ultimately, the form this evidence takes is dependent on the particulars of the case in question.

Research Limitations

Firstly, the data is a convenience sample and is not representative of a wide range of human variation. While I collected over 200 measurements, the grouping of the measurements reduced N sizes to a maximum of 20. It may be more difficult to detect statistical significance with smaller sample sizes.

Secondly, measurement obtainment may have also been constrained by the fact that I am a novice at collecting osteological measurements. In the case of certain measurements, such as the hipbone height, this could be compounded further by the imprecise nature of certain landmarks (i.e., the inferior point on the ischial tuberosity can be difficult to locate, as the pubis fuses onto the ischium during development, blurring the distinction between the two). Training and practice with the software may have offset the potential negative effect of this limitation.

Future Research

Future research should continue to validate this technology in various situations (all the while striving to adhere to police procedures). For example, I had included knives in the scene because highly reflective surfaces can pose problems for laser measurement devices. However, I ultimately decided that this was a tangential issue to this research and excluded this aspect.

Experimenting with various scan elevations would also be a worthwhile pursuit in determining how to maximize data collection. In this vein, directly comparing measurement results across different resolution and quality settings levels, and scan volumes, may illuminate the point at which data collection becomes redundant and counterproductive. Increasing the number of scans (and varying their positioning and elevation) and increasing the resolution setting may maximize the comprehensiveness of data in both a quantitative and, concomitantly, qualitative sense; however, it will also lead to decreases in efficiency.

An overview of the policing agencies that use TLS for forensics would also be a worthwhile pursuit. This research could endeavour to develop a set of standards and best practices for the use of TLS in forensics. Moreover, a systematic study of the legal cases that have relied upon TLS-produced evidence may yield more insight on the approach towards this evidence in different jurisdictions.

Finally, an empirical exploration of the magnitude of systematic versus human error would be worthwhile for pinpointing the sources of the errors in this equipment. For example, it is possible that measurement accuracy could increase if performed by a veteran forensic anthropologist. Considering the tendency to under measure, the development of a correction factor could also be tested to bring traditional and virtual measurements into greater agreement.

APPENDIX A: DATA TABLES

TABLE 1 - Stature Estimate Trials Descriptive Statistics

Trial	N	Mean (cm)	SD (cm)
C	20	156.73	7.55
T1	20	156.03	7.44
T2	20	155.97	7.61

C = Control, T1 = Trial 1, T2 = Trial 2

TABLE 2 - MD and MAD Amongst Measurement Trials

Trial	**Measurement**	**N**	**MD (%)**	**MAD (%)**	**MAD SD (mm)**
E1	Hipbone Height	20	4.81 (2.43)	5.03 (2.53)	2.36
E2		20	3.58 (1.82)	3.84 (1.94)	2.00
E3		20	-1.23 (-0.61)	1.86 (0.93)	1.34
Inter-observer AVERAGE			4.20	4.44	
E1	Iliac Breadth	19	3.42 (2.30)	3.42 (2.30)	1.42
E2		19	3.49 (2.36)	3.49 (2.36)	1.15
E3		19	0.07 (0.06)	0.65 (0.44)	0.62
Inter-observer AVERAGE			3.46	3.25	
E1	SPRL	19	4.33 (6.35)	4.56 (6.75)	2.65
E2		19	5.18 (7.50)	5.18 (7.50)	2.98
E3		19	0.85 (1.15)	2.52 (3.76)	1.92
Inter-observer AVERAGE			4.76	4.87	
E1	Max. Femur Length	20	2.52 (0.60)	2.83 (0.68)	1.56
E2		20	2.76 (0.67)	3.08 (0.74)	1.72
E3		20	0.25 (0.07)	1.37 (0.32)	1.45
Inter-observer AVERAGE			2.64	2.96	
E1	Epicondylar Breadth	17	2.05 (2.76)	2.22 (2.99)	1.42
E2		17	1.01 (1.37)	1.74 (2.39)	1.41
E3		17	-1.05 (-1.39)	1.24 (1.66)	1.25
Inter-observer AVERAGE			1.53	1.98	
E1	Max. Femur Head Diameter	20	2.03 (4.72)	2.03 (4.72)	0.96
E2		20	1.21 (2.78)	1.36 (3.13)	0.94
E3		20	-0.83 (-1.94)	0.9 (2.14)	0.79
Inter-observer AVERAGE			1.62	1.70	
E1 AVERAGE			3.19	3.28	
E2 AVERAGE			2.87	3.12	
TOTAL Inter-observer AVERAGE			**3.03**	**3.20**	
TOTAL Intra-observer AVERAGE			-0.32	1.42	

E1 = Control - Trial 1, E2 = Control - Trial 2, E3 (intra) = Trial 1 - Trial 2
Note: MD and MAD values shown in mm and (%).

TABLE 3 - Bone Measurement Trials Descriptive Statistics

Trial	Measurement	N	Mean (mm)	SD (mm)
C	Hipbone Height	20	199.80	17.33
T1		20	195.00	17.71
T2		20	196.23	17.96
C	Iliac Breadth	20	148.68	9.35
T1		19	145.26	9.21
T2		19	145.19	9.53
C	SPRL	20	67.16	5.65
T1		19	62.83	5.41
T2		19	61.98	4.07
C	Max. Femur Length	20	414.40	27.27
T1		20	411.89	26.84
T2		20	411.64	27.47
C	Epicondylar Breadth	20	74.88	5.95
T1		17	72.83	6.22
T2		17	73.88	6.45
C	Max. Femur Head Diameter	20	43.25	3.46
T1		20	41.22	3.57
T2		20	42.05	3.52

C = Control, T1 = Trial 1, T2 = Trial 2

TABLE 4 - One-Way Repeated Measures ANOVA - Bone Measurement Means

Measurement	Pair	N	SS	df	MS	F	F Sig.	Pairwise Sig.
Hipbone Height	C - T1	20	249.21	2	124.61	42.14	<0.001	<0.001
	C - T2	20						<0.001
	T1 - T2	20						0.034
Error			112.36	38	2.96			
Iliac Breadth	C - T1	19	151.51	2	75.75	108.97	<0.001	<0.001
	C - T2	19						<0.001
	T1 - T2	19						1.000
Error			25.03	36	0.70			
SPRL	C - T1	19	293.25	2	146.62	31.76	<0.001	<0.001
	C - T2	19						<0.001
	T1 - T2	19						0.749
Error			166.19	36	4.62			
Max. Femur Length	C - T1	20	93.35	2	46.68	21.41	<0.001	<0.001
	C - T2	20						<0.001
	T1 - T2	20						1.000
Error			82.86	38	2.18			
Epicondylar Breadth	C - T1	17	35.83	2	17.91	12.11	<0.001	<0.001
	C - T2	17						0.172
	T1 - T2	17						0.023
Error			47.35	32	1.48			
Max. Femur Head Diameter	C - T1	20	41.69	2	20.85	42.02	<0.001	<0.001
	C - T2	20						<0.001
	T1 - T2	20						0.001
Error			18.85	38	0.50			

C = Control, T1 = Trial 1, T2 = Trial 2

Note: SS=Type III Sum of Squares, MS=Mean Square

TABLE 5 - One-Way Repeated Measures ANOVA - Stature Estimate Means

Pair	N	SS	df	MS	F	F Sig.	Pairwise Sig.
C - T1	20	7.16	2	3.58	21.41	<0.001	<0.001
C - T2	20						<0.001
T1 - T2	20						1.000
Error		6.36	38	0.17			

C = Control, T1 = Trial 1, T2 = Trial 2

Note: SS=Type III Sum of Squares, MS=Mean Square

TABLE 6 - Intra-observer TEM

	Hipbone Height	Iliac Breadth	SPRL	Max. Femur Length	Epi. Breadth	Femur Head
N	20	19	19	20	17	20
Sum of Squares	103.36	14.98	187.13	76.99	50.78	28.23
Absolute TEM	1.61	0.63	2.22	1.39	1.22	0.84
VAV	195.61	145.23	62.40	411.76	73.35	41.63
Relative TEM	0.82	0.43	3.56	0.34	1.67	2.02

VAV = Variable average value

Note: See Perini et al. (2005) for TEM and VAV formulas.

TABLE 7 - Sex Estimation Case Study

Trial	Case	Hipbone Height	Iliac Breadth	SPRL	Femur Head	Epi. Breadth	P-Value	Sex
C	03R	198	155	68	47	81	0.952429352	M
T1	03R	192	152	65.5	46.00	80.20	0.892770677	M
T2	03R	194.3	151.6	59.1	47.60	80.80	0.9999978	M

C = Control, T1 = Trial 1, T2 = Trial 2

Note: Measurements in mm. Observe how p-values change amongst trials. See Discussion for details.

APPENDIX B: ILLUSTRATIONS

FIGURE 1 – The "hot spot" scene and the Faro Focus^{3D} 330X. Note the femora and hipbone series on the left side.

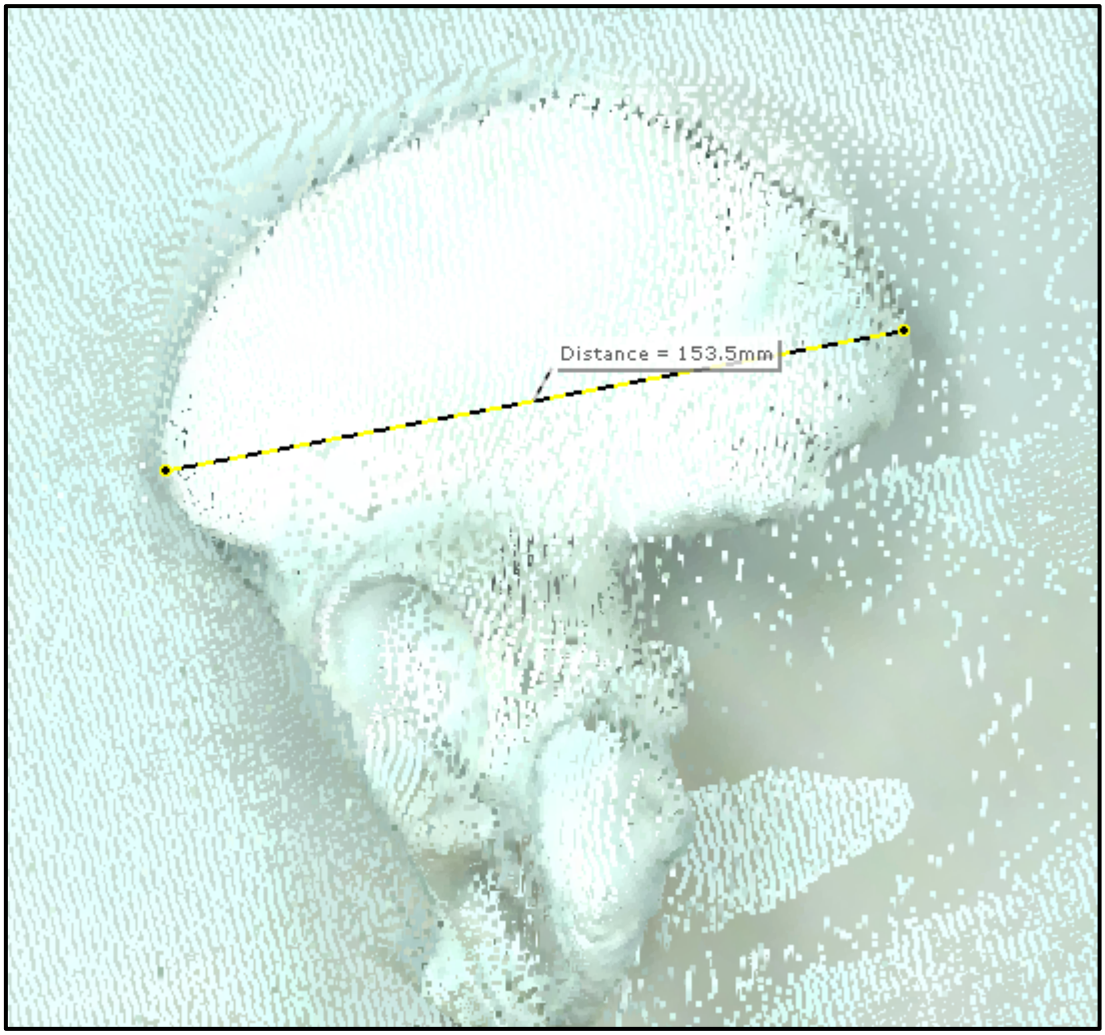

FIGURE 2 – The virtual iliac breadth measurement obtained via Faro SCENE software.

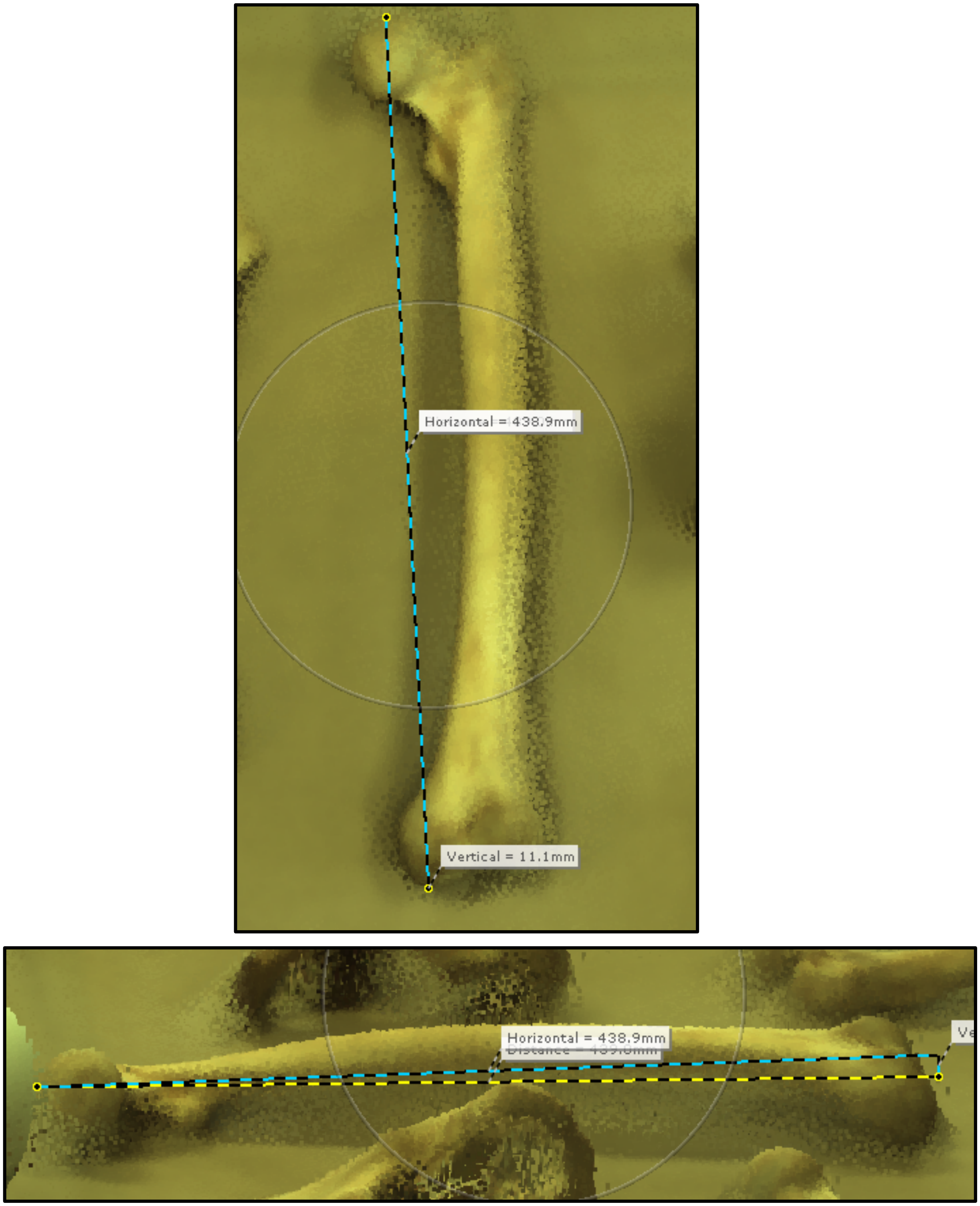

FIGURE 3 – An example of incorrect point selection. The overhead view (top picture) seems to show correct point selection for the maximum femur length measurement on the medial condyle. The side profile view (bottom picture), however, reveals that I have actually selected a data point that is off of the bone itself (pictured on the right side of the image).

REFERENCES

Agosto, E., Ajmar, A., Boccardo, P., Tonolo, F., & Lingua, A. (2008). Crime scene reconstruction using a fully geomatic approach. *Sensors,* 8, 6280-6302. doi: 10.3390/s8106280

Albanese, J. (2003). A metric method for sex determination using the hipbone and the femur. *Journal of Forensic Sciences,* 48(2), 1-11. Retrieved from www.astm.org

Albanese, J., Tuck, A., Gomes, J., & Cardoso, H. (2016). An alternative approach for estimating stature from long bones that is not population- or group-specific. *Forensic Science International,* 259, 59-68. http://dx.doi.org/10.1016/j.forsciint.2015.12.011

Baber, C., & Butler, M. (2012). Expertise in crime scene examination: Comparing search strategies of expert and novice crime scene examiners in simulated crime scenes. *Human Factors,* 54(3), 413-424. doi:10.1177/0018720812440577

Barazzetti, L., Sala, R., Scaioni, M., Cattaneo, C., Gibelli, D., Guissani, A., … Vandone, A. (2012). 3D scanning and imaging for quick documentation of crime and accident scenes. Proceedings from *SPIE: The International Society for Optical Engineering*. doi:10.1117/12.920728

Bass, W. M. (2005). *Human osteology: A laboratory and field manual* (5th ed.). MO: Missouri Archaeological Society.

Bex, F. J. (2011). *Arguments, stories and criminal evidence: A formal hybrid theory.* Netherlands: Springer.

Boehler, W., Vicent, M., & Marbs, A. (2003). Investigating laser scanner accuracy. *The International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, 34 (Part 5/C15), 696-701.

Bucheli, S., Pan, Z., Glennie, C., Lynne, A., Haarman, D., & Hill, J. (2014). Terrestrial laser scanning to model sunlight irradiance on cadavers under conditions of natural decomposition. *International Journal of Legal Medicine,* 128, 725-732. doi:10.1007/s00414-014-1013-1

Buck, U., Kneubuehl, B., Nather, S., Albertini, N., Schmidt, L., & Thali, M. (2011). 3D bloodstain pattern analysis: Ballistic reconstruction of the trajectories of blood drops and determination of the centres of origin of the bloodstains. *Forensic Science International,* 206, 22-28. doi:10.1016/j.forsciint.2010.06.010

Buck, U., Naether, S., Rass, B., Jackowski, C., & Thali, M. (2013). Accident or homicide – Virtual crime scene reconstruction using 3D methods. *Forensic Science International,* 225, 75-84. http://dx.doi.org/10.1016/j.forsciint.2012.05.015

Cardoso, H. F. V., Marinho, L., & Albanese, J. (2016). The relationship between cadaver, living and forensic stature: A review of current knowledge and a test using a sample of adult Portuguese males. *Forensic Science International,* 258, 55-63. http://dx.doi.org/10.1016/j.forsciint.2015.10.012

Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council. (2009). *Strengthening forensic science in the United States: A path forward* (Document No. 228091). Washington, DC: The National Academies Press.

Cuartero, A., Armesto, J., Rodriguez, P. J., & Arias, P. (2010). Error analysis of terrestrial laser scanning data by means of spherical statistics and 3D graphs. *Sensors,* 10, 10128-10145. doi:10.3390/s101110128

Davy-Jow, S. L., Lees, D. M. B., & Russell, S. (2013). Virtual forensic anthropology: Novel applications of anthropometry and technology in a child death case. *Forensic Science International,* 224, e7-e10. http://dx.doi.org/10.1016/j.forsciint.2012.11.002

Delisle, R., Stuart, D., Tanovich, D., & Dufraimont, L. (2015) *Evidence principles and problems* (11th ed.). Toronto, Ont.: Carswell.

Edmond, G. (2011). The admissibility of incriminating expert opinion evidence in the US, England and Canada. *Judicial Officers' Bulletin,* 23(8), 67-70.

Edmond, G. (2015) What lawyers should know about the forensic 'sciences'. *Adelaide Law Review,* 36, 33-100. Retrieved from http://heinonline.org

Edmond, G., & Roach, K. (2011). A contextual approach to the admissibility of the state's forensic science and medical evidence. *University of Toronto Law Journal,* 61, 343-410. doi:10.3138/utlj.61.3.343

Elmes, G., Roedl, G., & Conley, J. (2014). Concepts, principles, and definitions. In G. Elmes, G. Roedl, & J. Conley (Eds.), *Forensic GIS: The role of geospatial technologies for investigating crime and providing evidence* (pp. 3-18). Dordrecht: Springer.

Elmes, G., Roedl, G., & Conley, J. (2014). Geospatial technologies in the courtroom. In G. Elmes, G. Roedl, & J. Conley (Eds.), *Forensic GIS: The role of geospatial technologies for investigating crime and providing evidence* (pp. 19-38). Dordrecht: Springer.

Eyre, M., Foster, P., & Speake, G. (2014). Surveying the scene. *The Safety and Health Practitioner,* 32(9), 41-44.

Faro Technologies. (2011) Faro laser scanner Focus^{3D} manual. Retrieved from https://doarch332.files.wordpress.com/2013/11/e866_faro_laser_scanner_focus3d_manual_en.pdf

Faro Technologies. (2016) SCENE 6.0 user manual. Available from http://www.faro.com/products/faro-software/scene/downloads#Download

Faro Technologies. (2016). SCENE Software (Version 6.0) [Software]. Available from http://www.faro.com/products/faro-software/scene/downloads#Download

Fiedler, B. S. (2003). Are your eyes deceiving you?: The evidentiary crisis regarding the admissibility of computer generated evidence. *New York Law School Review,* 48, 295-321. Retrieved from http://heinonline.org

Hakim, N., & Liscio, E. (2015). Calculating point of origin of blood spatter using laser scanning technology. *Journal of Forensic Sciences,* 60(2), 409-417. doi:10.1111/1556-4029.12639

Hill, C., Tanovich, D. M., & Strezos, L. P. (2005). *McWilliams' Canadian criminal evidence* (4th ed.). Aurora, Ont.: Canada Law Book.

Hill, C., Tanovich, D. M., & Strezos, L. P. (2013). *McWilliams' Canadian criminal evidence* (5th ed.). Aurora, Ont.: Canada Law Book.

Johnson, M., & Liscio, E. (2015). Suspect height estimation using the Faro Focus 3D laser scanner. *Journal of Forensic Sciences*, 1-7. doi:10.1111/1556-4029.12829

Klepinger, L. L. (2006). *Fundamentals of forensic anthropology.* Hoboken, NJ: Wiley-Liss.

Komar, D., Davy-Jow, S. L., & Decker, S. (2012). The use of a 3-D laser scanner to document ephemeral evidence at crime scenes and postmortem examinations. *Journal of Forensic Sciences,* 57(1), 188-191. doi:10.1111/j.1556-4029.2011.01915.x

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society.* Cambridge, MA: Harvard University Press.

Law, J., & Martin, E.A. (2009). *A dictionary of law* (7th ed.). Retrieved from http://www.oxfordreference.com.ezproxy.uwindsor.ca/view/10.1093/acref/9780199551248.001.0001/acref-9780199551248-e-3234

Leica Geosystems. (n.d.). *Leica public safety solutions.* Retrieved from psg.leica-geosystems.us

Maksymowicz, K., Tunikowski, W., & Kosciuk, J. (2014). Crime event 3D reconstruction based on incomplete or fragmentary evidence material. *Forensic Science International,* 242, e6-e11. http://dx.doi.org/10.1016/j.forsciint.2014.07.004

Mihandoost, R. K. (2015). *A validation study of the measurement accuracy of SCENE and SceneVision 3D software programs*. Retrieved from ProQuest (UMI# 1589162).

Mnoonkin, J. L. (1998). The image of truth: Photographic evidence and the power of analogy. *Yale Journal of Law & the Humanities,* 10(1), 1-74. Retrieved from http://digitalcommons.law.yale.edu/yjlh/vol10/iss1/1

National Institute of Standards and Technology. (2013). *Crime scene investigation: A guide for law enforcement.* Largo, FL: National Forensic Science Technology Center.

Park, H-K., Chung J-W., & Kho, H-S. (2006). Use of hand-held laser scanning in the assessment of craniometry. *Forensic Science International*, 160(2), 200-206. doi:10.1016/j.forsciint.2005.10.007

Perini, T. A., Lameira de Oliveira, G., dos Santos Ornellas, J., & Palha de Oliveira, F. (2005). Technical error of measurement in anthropometry. *Revista Brasileira de Medicina do Esporte,* 11(1), 86-90. http://dx.doi.org/10.1590/S1517-86922005000100009

R. v. Doodnaught, 8022 Ontario Superior Court of Justice. (2013). Retrieved from https://www.canlii.org/en/on/onsc/doc/2013/2013canlii89551/2013canlii89551.html?autocompleteStr=doodnaught&autocompletePos=2

Roedl, G., Elmes, G. & Conley, J. (2014). Spatial technology applications. In G. Elmes, G. Roedl, & J. Conley (Eds.), *Forensic GIS: The role of geospatial technologies for investigating crime and providing evidence* (pp. 53-70). Dordrecht: Springer.

Schweitzer, W., Rohrich, E., Schaepman, M., Thali, M. J., & Ebert, L. (2013). Aspects of 3D surface scanner performance for post-mortem skin documentation in forensic medicine using rigid benchmark objects. *Journal of Forensic Radiology and Imaging,* 1, 167-175. http://dx.doi.org/10.1016/j.jofri.2013.06.001

Shelton, D. E. (2010). *Criminal adjudication: The challenges of forensic science evidence in the early 21st century.* Retrieved from ProQuest (UMI# 3415644).

Sholts, S., Warmlander, S., Flores, L., Miller, K., & Walker, P. (2010). Variation in the measurement of cranial volume and surface area using 3D laser scanning technology. *Journal of Forensic Sciences,* 55(4), 871-876. doi:10.1111/j.1556-4029.2010.01380.x

Siuri, B. (2004). Laser technology helps preserve crime scene. *Law & Order,* 52(5), 52-55.

Stancu, H. (2010, November 2). Coming soon to a crime scene near you: 3D cameras. *Toronto Star.* Retrieved from https://www.thestar.com/news/crime/2010/11/02/coming_soon_to_a_crime_scene_near_you_3d_cameras.html

Topol, A., Jenkin, M., Gryz, J., Wilson, S., Kwietniewski, M., Jasiobedzki, P., … Bondy, M. (2008). Generating semantic information from 3D scans of crime scenes. *Canadian Conference on Computer and Robot Vision,* 333-340. doi:10.1109/CRV.2008.27

Tung, N., Barr, J., Sheppard, D., Elliot, D., Tottey, L., & Walsh, K. (2015). Spherical photography and virtual tours for presenting crime scenes and forensic evidence in New Zealand courtrooms. *Journal of Forensic Sciences,* 60(3), 753-758. doi:10.1111/1556-4029.12736

Wang, J., Kutterer, H., & Fang, X. (2012). On the detection of systematic errors in terrestrial laser scanning data. *Journal of Applied Geodesy,* 6, 187-192. doi:10.1515/jag-2012-0025

Williams, K. E. G., & Saks, M. J. (2015). Why don't the gatekeepers guard the gates? comments prompted by Edmond. *Adelaide Law Review*, 36, 109-124. Retrieved from http://heinonline.org

Wozniak, K., Moskala, A., Rzepecka-Wozniak, E. (2015). Imaging for homicide investigations. *Italian Society of Medical Radiology,* 120, 846-855. doi:10.1007/s11547-015-0529-x

VITA AUCTORIS

Ryan A. Mullins was born in 1993 in London, Ontario. He graduated from Chatham-Kent Secondary School in 2011. From there he went on to the University of Western Ontario where he obtained a B.A.[H] in Criminology in 2015. From there he went on to the University of Windsor where he obtained a M.A. in Criminology in 2016. He is currently a candidate for the Juris Doctor degree in Law at Queen’s University and hopes to graduate in Spring 2019.