ABSTRACT

Skill Acquisition in Saw Kerf Class Identification

Kelsea Francis

Director: Dr. Joseph Ferraro

The qualifications of expert witnesses are under scrutiny following the *Daubert* ruling. However, there is little consensus as to what defines an expert in a forensic field. There is no designated definition of 'expert' that courts use to determine whether someone is qualified enough to be an expert witness, but level of education and years of experience are important factors. This work is a preliminary analysis of the rate of skill acquisition in regards to saw mark, or kerf mark, analysis. Some kerf marks could be learned and expertly identified by analysts in as little as 5 minutes, whereas other kerfs could not, even after 30 minutes of preparation. Our findings suggest that there are some concepts within kerf mark analysis that can be mastered with high accuracy and others that cannot.

APPROVED BY DIRECTOR OF HONORS THESIS:

____

Dr. Joseph Ferraro, Department of Anthropology

APPROVED BY THE HONORS PROGRAM:

____

Dr. Elizabeth Corey, Director

DATE: ______________________

SKILL ACQUISITION IN SAW KERF CLASS IDENTIFICATION

A Thesis Submitted to the Faculty of

Baylor University

In Partial Fulfillment of the Requirements for the

Honors Program

By

Kelsea Francis

Waco, Texas

May 2018

TABLE OF CONTENTS

ACKNOWLEDGMENTS

A huge thank you to Dr. Joseph Ferraro for making this study possible. Thank you for all of the time, energy, and care that you spent on this project and for teaching me diligence. In addition, I'd like to thank Sergeant James Huggins for providing me with volunteers. I'm also grateful for the assistance of Dr. Jack Tubbs in helping me with the statistics and for creating the graphs.

CHAPTER ONE

Introduction

The *Daubert* standard is used by trial judges to assess whether an expert witness's testimony is scientifically valid. The factors of determination are: whether the methodology used is testable and has been tested, whether it has been peer reviewed and published, the error rate of the methodology, whether there are established standards for the technique, and if it has widespread acceptance in the scientific community.[1] The *Daubert* ruling has motivated forensic scientists to test the accuracy of the methods and principles used for procuring evidence so that it may be considered admissible in court. It is of the utmost importance that investigators and scientists handle and test evidence in ways that yield dependable results with low error rates. However, few publications exist focused on what it takes to master skills in forensic fields. In forensics, taphonomy refers to the study of alterations made on the surfaces of bone. Scavenging damage and weather erosion are both examples of taphonomic alterations. The focus of this study will be on saw kerf mark analysis on bone. What does it take to become an expert kerf mark analyst? How much practice would one need in order to attain mastery-level skill? For the purpose of this study, expert-level accuracy is considered 95% or higher accuracy rate. The general trends of the learning curves will be reported and resulting implications will be discussed.

Literature Review on Expert-Level Skill Acquisition

Ericsson et al.'s (1993) analysis regarding the acquisition of expert performance is the most extensive and often cited review on the subject. In their meta-analysis based on a number of field-specific studies, the authors found that "the highest level of expert performance is displayed by individuals with more than 10 years of experience."[2] They report that while skill is acquired from practice and experience, it can only increase as a result of deliberate efforts to improve. Ten years of experience and exposure alone does not make one an expert in a field. One must receive extensive testing and training in the form of higher education. For many fields, this may come in the form of a doctorate. In addition to a PhD, forensic anthropologists have the American Board of Forensic Anthropology (ABFA) board certification tests, which they denote as being "the highest recognized level of professional qualification in the field of forensic anthropology."[3] Smaller, less-developed forensic fields such as kerf mark analysis do not have a set of agreed upon requirements for expert-level mastery so it isn't likely that a kerf mark analyst could be admitted as an expert witness in court. The National Academy of Science's publication on strengthening forensic science sums up the situation well: "In short, the quality of forensic practice in most disciplines varies greatly because of the absence of adequate training and continuing education, rigorous mandatory certification and accreditation programs, adherence to robust performance standards, and effective oversight" which threaten the "quality and credibility of forensic science practice".[4]

Even in some of the more developed forensic fields, the question of what constitutes expert status is still a problem. This is exemplified in Thompson et al.'s experiment in 2013,[5] which analyzed the potential error rate of expert fingerprint examiners. Five examiners with over 10 years of experience were covertly evaluated and given a print that, unbeknownst to them, they had already identified as a match in the past. They were told that the print was from a famous case of FBI misidentification. Only one examiner maintained his original conclusion. The results lead one to question whether current qualifications to be an expert in a forensic field are rigorous enough. The discussion of expert-level skill acquisition is one that the forensic community should revisit and address.

Background

Hand-powered saws, such as hacksaws and carpenter saws, are often used to disarticulate human remains, leaving characteristic saw mark damage known as kerf marks on bone. The study of these kerf marks is the subject of forensic anthropologist Dr. Steven Symes dissertation.[6] A fundamental concept of kerf mark analysis is that different types of saws vary in morphology producing predictable kerf morphologies which can be grouped into classes. Dr. Symes identifies four classes of kerf marks (A, B, C, and D) which can help analysts identify the type of saw used to produce the kerf.[6] The three main macroscopic features an analyst can observe to determine kerf class are: the width of the kerf; whether the kerf floor is flat, rounded, or contains a median “island”; and whether the floor corners are rounded or squared.

CHAPTER TWO

Materials and Methods

TABLE 1. A table of the corresponding saw, saw teeth, and kerf characteristics by kerf class.

Class	**Saws**	**Saw Teeth Characteristics**	**Kerf Characteristics**
A	Fine-toothed bow saws (i.e. hacksaws and coping saws)	Small teeth	Narrow kerf width, kerf floor is flat, 1-2 rounded floor corners
B	Rip/chisel saws (i.e. mitres, some keyholes)	Medium-sized teeth with a chiseling form	Narrow-medium kerf width; kerf floor is flat, stepped, or exhibits a small island; squared floor corners
C	Crosscut saws (i.e. tree pruning and log saws)	Large teeth, raker-set present, angled tooth filing	Wide kerf width, large bone islands, squared floor corners
E	Carbide-grit abrasion saws	No teeth	Very wide & round kerf width

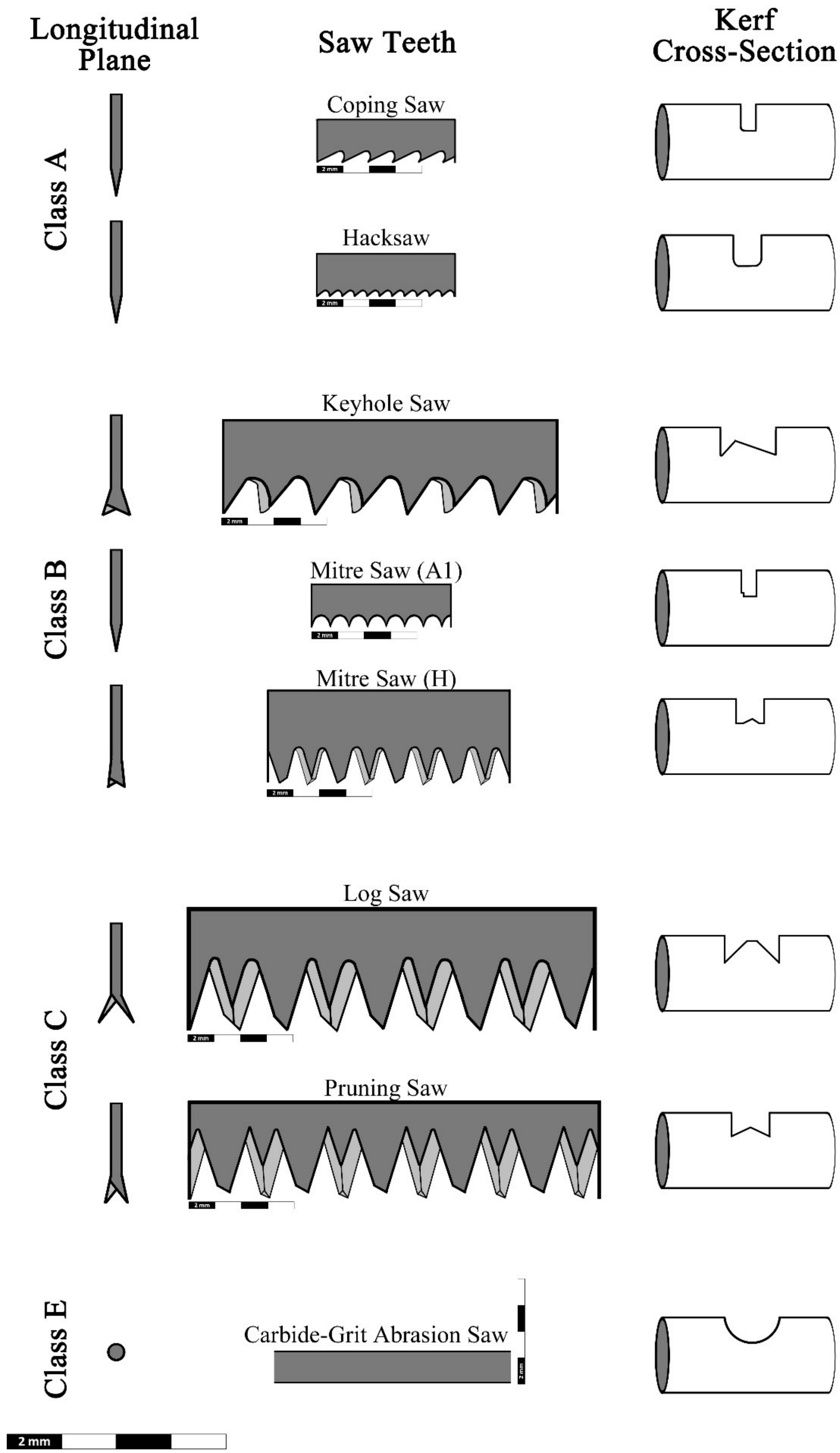
Saws and Kerf Classes

Eight saws were chosen and categorized into kerf classes A, B, C, and E based on the saw typology created by Symes.[6] Kerf class E is a newly defined class not present in Symes' dissertation constituting carbide-grit abrasion saws. Class D, the power saws, was excluded from the experiment. Table 1 above summarizes saw types by kerf class along with corresponding saw teeth and kerf characteristics. The saws used in this experiment were: a coping saw, hacksaw, keyhole saw, two differently sized mitre saws, a log saw, pruning saw, and a carbide-grit abrasion saw. Figure 1 on page 6 contains diagrams of the teeth shape, size, and kerf morphology for the saws used in the experiment.

The Bones

The bones used were Redbarn brand dog bones. They were purchased at a local pet store. They are pre-cleaned, bleached cow tibiae shafts with the epiphysis and metaphysis of the bones removed.

FIGURE 1. Diagrams of the saw longitudinal planes, teeth, and subsequent kerf marks.



Experimental Protocol

Five kerf marks were made on 10 bones, producing a total of 50 kerf marks. Twelve kerfs each were made by saws from classes A, B, and C. Fourteen class B kerfs were made. The kerfs were cut in a randomized order. To control the variable of kerf depth, all the kerfs were made by the same person using the same number of saw strokes for each kerf. The kerfs were cut just deep enough for all the relevant features to be visible.

Thirty undergraduate analysts from an introductory forensics course volunteered for the first study. Two individuals were excluded from the study for scoring 2 standard deviations below the mean; in other words, their scores were significantly below the normal distribution. In the second study, 15 different undergraduate students from an impression evidence course volunteered. Neither groups had any prior knowledge of kerf marks analysis.

For both studies, I first gave the analysts a PowerPoint presentation. The first slides introduced relevant terminology. I explained what a kerf mark is and where certain features, such as the kerf floor and corners, are located. I taught them that the kerf width, floor, and floor corners are the three macroscopic features they will be using to determine class. The next slides discussed each of the four classes and described the kerf width, floor, and floor corners for each. These included a rough sketch of each of the kerfs. In the last slide, I passed out the tibiae and told the analysts what they were supposed to do: match the kerfs on the bones to their appropriate classes using exclusion. I advised them that kerf depth should not be used to determine class. After distributing the bones, I taught the analysts how to angle them in order to view the kerf features. I also distributed an outline of the PowerPoint for them to use during the test (see the Appendix). The analysts were given 2 minutes per bone to identify the corresponding kerf class. For the first study, the analysts returned a week later to repeat the test with the outline but without viewing the presentation. In the second iteration, immediately after the analysts finished the test they were given as much time as they wanted to examine the saws, the kerfs, and make kerf marks themselves. The purpose of this was to measure if extended practice time would improve the analyst's mastery of kerf mark analysis. This lasted for about 30 minutes. During this time, I warned the analysts that based on the previous study, B and C kerfs were frequently confused as well as A and B kerfs. The analysts then repeated the test.

FIGURE 2. The exact agreement represents the kerfs that analysts assigned to class A in both weeks. Kerfs that were identified differently between weeks 1 and 2 are represented by the partial agreement. (Courtesy of Dr. Jack Tubbs)

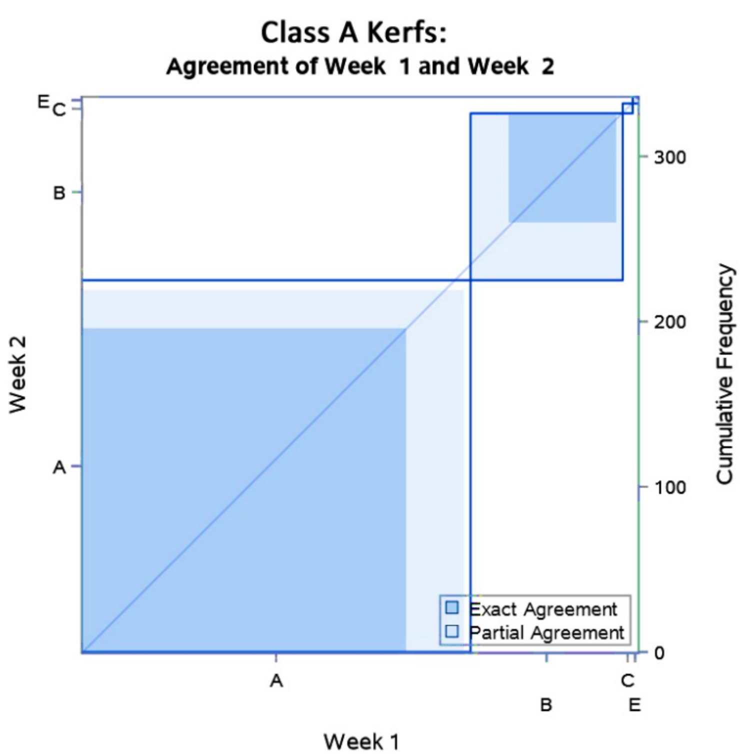


FIGURE 3. A graph representing the agreement within analysts of class B kerfs. (Courtesy of Dr. Jack Tubbs)

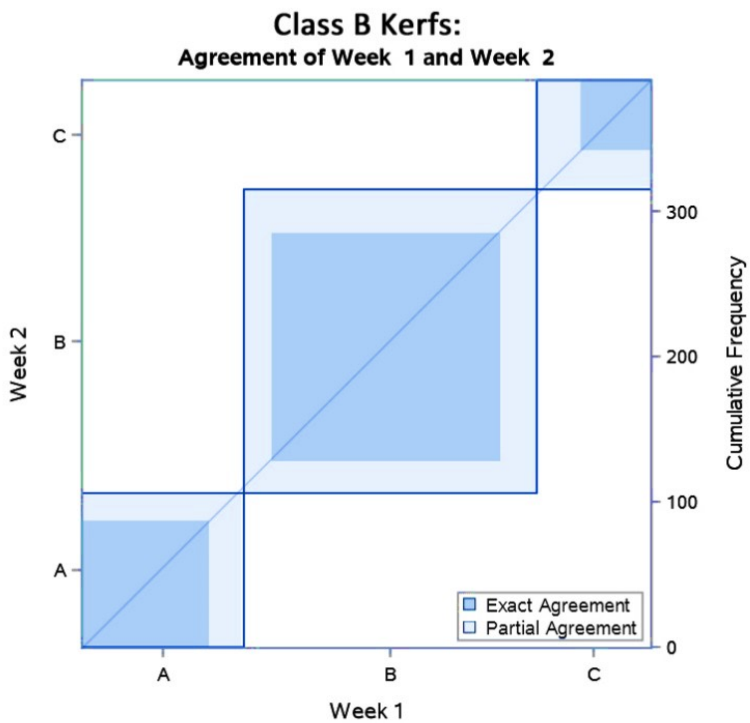
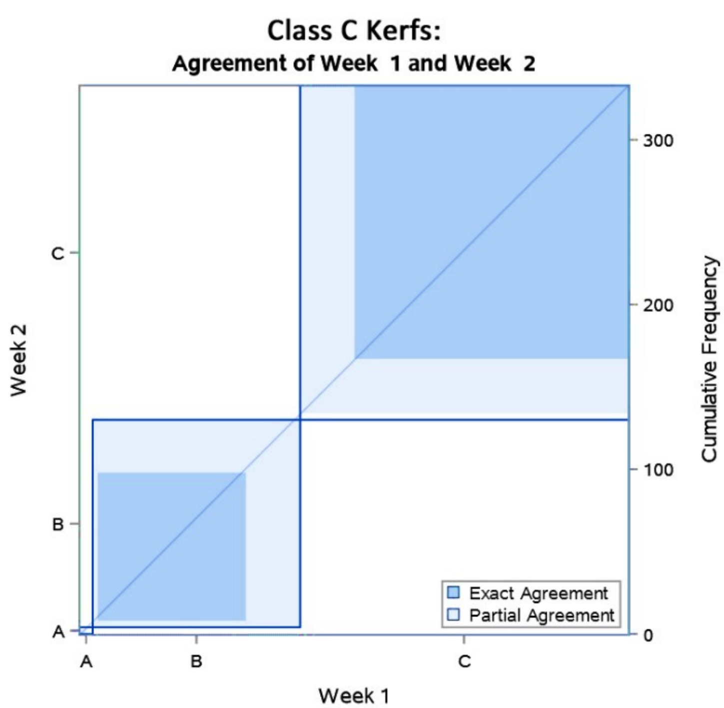


FIGURE 4. A graph representing the agreement within analysts of class C kerfs. (Courtesy of Dr. Jack Tubbs)



Interpretation

The kappa statistic was implemented to assess the agreement within analysts in choosing kerf classes. The kappa statistic measures inter-rater agreement for raters choosing between categories while taking chance into account.[7] For every class, the rater agreement was measured between week 1 and week 2. The graphs in Figures 2-4 represent the degrees of agreement by class between analysts in their class selections between week 1 and week 2. In Figure 2, the bigger, darkly-shaded square shows that a majority of the analysts consistently selected "Class A" when presented with a class A kerf both in week 1 and in week 2. The lighter-shaded band around that dark square represents those who selected class A in one of the weeks, but not both. Similarly, the smaller darkly-shaded square in that graph represents a smaller group of analysts who chose class B for class A kerfs consistently for both weeks. The squares that represent analyst selection of class C and class E for class A kerfs are extremely small because those classes were rarely chosen by analysts when presented with a class A kerf. The other graphs should be interpreted similarly. It is important to emphasize that the kappa measures the analysts' agreement—not accuracy—in class selections. The kappa statistic was not used for class E because of its consistently high rater agreement and low error rates.

CHAPTER THREE

Results

Week 1

After only a brief PowerPoint presentation, analysts masterfully identified class E kerfs about 95% of the time. The second most-successfully identified class was A at 70%. Classes C and B were selected much less successfully by analysts, at only 60% and 52% respectively. When confronted with a class A kerf, analysts chose class B almost 30% of the time. Similarly, analysts chose class B for class C kerfs with a frequency of 40%. However, classes A and C were only confused with each other 4% of the time and the frequency of incorrectly choosing class E was less than 2%.

Week 2

The results from week 2 did not show any significant differences from week 1. Class A and class E decreased in accuracy by 3% and 1% respectively, while classes B and C exhibited a 2% and 1% increase in accuracy. If nothing else, the analysts did appear to exhibit retention of the information based on the kappa scores. The simple kappa scores for classes A, B, and C measured between both weeks are 0.52, 0.59, and 0.53 respectively. According to the ranges dictated by Landis and Koch[7], the classes showed moderate to substantial agreement which proves that the analysts showed retention of the information between weeks 1 and 2.

Extended Preparation Time

The results of the first test that the 15 analysts took were similar to those of the previous study. The class with the highest identification success rate was still class E at 92%. The only class which showed a significant difference in accuracy rate compared to the first study was class A, which was only chosen correctly 53% of the time here as opposed to 70%. The second test the analysts took after their extended training exhibited only an overall 2% increase in average accuracy compared to the first test. Classes A and C were rarely confused with each other (only 3%), but continued to be confused with class B at a rate of 34% and 41% respectively. The tests confirm that these analysts struggled in choosing between the same classes as the analysts in the first study, even after their increased exposure and experience with analyzing the kerfs.

\

CHAPTER FOUR

Discussion

This study began with the main question of skill acquisition: would novice analysts be able to master kerf mark analysis? Would they be able to correctly identify kerf class with a high rate of accuracy after just having learned how? Although analysts showed mastery-level skill when identifying class E, other classes were only successfully identified 70% of the time. The majority of the kerf classes could not be expertly identified by analysts. The extension of our first question was whether or not analyst accuracy would improve with an extended 30 minutes of preparation prior to the second test. The results were the same: only class E was mastered. The 30 extra minutes of preparation time did not have a significant effect in improving the analysts' accuracy.

It became apparent over the course of this study that classes A and B and classes B and C were frequently confused with each other. Classes A and C were rarely ever confused with each other. The explanation for this lies within Symes' typology. According to his class typology, from A to B and from B to C the kerf width goes from generally narrow to generally wide, the kerf floor progresses from flat to the presence of a small or large island, and the corners progress from rounded to squared. The classes are arranged along a continuum of characteristics. Classes A and C are the farthest apart on the continuum so they have few shared traits and are unlikely to be confused. Class E exhibited a consistently high accuracy rate because it is not on Symes' kerf class continuum. Analysts studying kerf marks should be aware of these commonly confused classes.

At this point, it is unlikely that a trial judge would admit kerf mark analysis as evidence in court because it does not meet *Daubert* standards.[1] Not enough research has been done to establish protocol of methodology, error rates, or widespread consensus of the research within the forensic community. Bone damage as a result of gunshot or stabbing wounds will likely always be more commonly found in homicide investigations than kerf mark evidence. Nevertheless, we should still try to research and understand all that we can so that when investigators are confronted with this evidence, they will be prepared.

CHAPTER FIVE

Conclusion

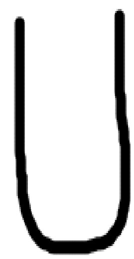
It is not clear at this point how long it would take to master kerf mark analysis. Further work is required to determine which type and amount of training is needed to become an expert forensic kerf mark analyst. This study demonstrates the importance of testing and defining accuracy rates for forensic practices to ensure that expert witnesses in court are supported by valid, reliable science as *Daubert* requires. It is our recommendation that the microscopic analysis of kerf marks should be subsequently studied for accuracy and reliability. All forensic fields should continue to research methodology error rates and revisit what it means to be an expert in a forensic field.

APPENDIX

Kerf Class Morphology Handout

Kerf Class A:

- Kerf Width: generally narrow
- Kerf Floor: flat
- Floor Corners: 1 or both rounded

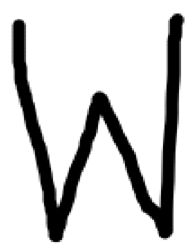


Kerf Class B:

- Kerf Width: generally wide
- Kerf Floor: flat or tiny island
- Floor Corners: non-rounded

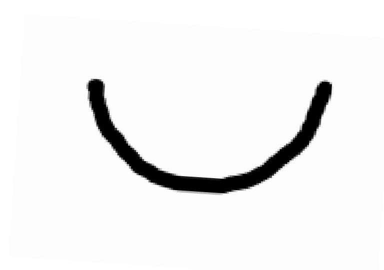
Kerf Class C:

- Kerf Width: generally wide
- Kerf Floor: large island
- Floor Corners: non-rounded



Kerf Class E:

- Kerf Width: wide
- Kerf Floor: rounded
- Floor Corners: both rounded



REFERENCES

[1] Legal Information Institute, Daubert Standard. https://www.law.cornell.edu/wex/daubert_standard, (accessed 1 April 2018).

[2] K.A. Ericsson, R.T. Krampe, C. Tesch- Römer, The role of deliberate practice in the acquisition of expert performance, *Psychological Review*. 100(1993) 363-406.

[3] The American Board of Forensic Anthropology. http://theabfa.org/, 2016 (accessed 15 March 2018).

[4] Committee on Identifying the Needs of the Forensic Sciences Community, Strengthening forensic science in the United States: a path forward. 2009, pp. 6.

[5] M.B. Thompson, J.M. Tangen, D.J. McCarthy, Expertise in Fingerprint Identification, *Journal of Forensic Science*. 58(2012) 1519-1530.

[6] S.A. Symes, Morphology of saw marks in human bone: identification of class characteristics. ProQuest Dissertations Publishing. 1992.

[7] J.R. Landis, G.G. Koch, A one-way components of variance model for categorical data, *Biometrics*. 33(1997) 671-679.