Document Title:	**The Statistical Evaluation of Torn and Cut Duct Tape Physical End Matching**
Author:	**Frederic A. Tulleners, M.A., Jerome V. Braun, Ph.D.**
Document No.:	**235287**
Date Received:	**July 2011**
Award Number:	**2009-DN-BX-K235**

The Statistical Evaluation of Torn and Cut Duct Tape Physical End Matching

Award Number **2009-DN-BX-K235**

Frederic A. Tulleners[1], MA, P.I.
Jerome V. Braun[2], PhD, Co P.I.

Graduate Student Researchers
Kaitlin McCabe, BS[1], Gillian Currie, BS[1], Esmeraldo Gorecho BS[1], Ka Lok Chan, MS[1]

This project was supported by Award No. 2009-DN-BX-K235 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the author(s) and do not necessarily reflect those of the Department of Justice.

1. University of California at Davis Forensic Science Graduate Program, 1333 Research Park Drive, Davis, CA 95618
2. University of California at Davis Department of Statistics, One Shields Avenue, Davis, CA 95616

Abstract

Duct tape is often found in association with criminal activity, most commonly in abductions, homicides, and the construction of explosive devices – all considered felony crimes. As such, forensic scientists are frequently asked to analyze and compare duct tape samples in order to establish possible evidentiary links between a suspect and victim, a suspect and a particular crime, or multiple crimes. Through duct tape end matching, analysts can reunite separated fragments. Based on the principle that each tear is unique, this type of matching has a significantly high evidentiary value and is considered to be one of the strongest associations in forensic science comparative examinations. Physical end matching of duct tape is common in crime laboratories, yet the process of physical end matching remains undefined and the reliability and error rates associated with these generally accepted procedures are unknown.

This study was designed to research duct tape physical end matching, including criteria to describe the matching process, a protocol for training analysts in physical end matching, and statistically evaluating the associated error rates and overall accuracy. Each trial evaluated the variation in inconclusive rates, error rates, and accuracy rates due to: differing brands, grades, and colors of duct tape; differing analysts; and differing separation methods. In addition, the design also helped to assess the reliability of the protocol as represented by the reproducibility of the end result provided by blind peer review, as well as allow assessment of conditions that could restrict its validity.

The experimental design consisted of a blind study looking at four different methods of separating duct tape: hand torn, Elmendorf machine torn, scissor cut, and cut with a box cutter knife. Three Graduate Student Researchers (GSRs) were selected to work as duct tape analysts for the duration of the study. The GSRs produced individual results from the same sets, and the

possible outcomes of a hypothetical peer review between analysts was assessed. Each GSR worked part-time to eliminate fatigue and maintain consistency among the analysts.

This study confirms that it is possible to use physical end matching to identify duct tape samples as matching or non-matching and that differences between analysts, brands, tape grades, tape color, and method of separation have varying contributions to misidentifications or inconclusive results. This research looked at 1800 torn tape specimens and 400 cut tape specimens. The mean accuracy observed ranged from 98.58 to 100.00 percent for torn tape and from 98.15 to 99.83 percent for cut tape. The mean false-positive rate ranged from 0.00 to 0.67 percent for torn tape and from 0.00 to 3.33 percent for cut tape. The mean false-negative rate ranged from 0.00 to 2.67 percent for torn tape and was 0.33 percent for both types of cut tape. In general, it seems that the brand and tape grade are more important than color in their effect on an analyst to correctly identify duct tape end matches. Scissor cut tapes appear more difficult to analyze than hand torn tapes, but there is no significant difference in difficulty between hand torn tape and tape cut with a box cutter knife. Finally, consistent tearing conditions do not seem to affect an analyst's ability to correctly identify duct tape end matches. This study also demonstrated the importance of peer review in duct tape analysis and its ability to greatly reduce the number of misidentifications made by analysts.

Table of Contents

Executive Summary

Duct tape is often found in association with criminal activity, most commonly in abductions, homicides, and the construction of explosive devices. As such, forensic scientists are frequently asked to analyze and compare duct tape samples, as well as establish possible evidentiary links between a suspect and victim, a suspect and a particular crime, or multiple crimes. Through duct tape end matching, analysts can reunite separated fragments. Based on the principle that each tear is unique, this type of matching has a significantly high evidentiary value and is considered to be one of the strongest associations in forensic science comparative examinations. Physical end matching of duct tape is common in crime laboratories, yet the process of physical end matching remains undefined and the reliability and error rates associated with this generally accepted procedure are unknown.

This study was designed to develop a protocol for duct tape physical end matching as well as a program for training analysts in physical end matching, and to statistically evaluate the associated error rates and overall accuracy. Each trial evaluated the variation in inconclusive rates, error rates, and accuracy rates due to differing brands, grades, and colors of duct tape, in addition to differing analysts and separation methods. This study also assessed the reliability of the protocol as represented by reproducibility of the end result and peer review.

Methods

The experimental design consisted of a blind study that observed four different methods of separating duct tape: hand torn, Elmendorf machine torn, scissor cut, and cut with a box cutter knife. Three paid Graduate Student Researchers (GSRs) from the Forensic Science Graduate program at the University of California, Davis, were selected to work as duct tape analysts to

examine the different methods of separating duct tape. The GSRs produced individual results as well as performed peer review on the other two analysts' work. Each GSR worked part-time to eliminate fatigue and maintain consistency among the analysts.

Training

Before experimental trials were conducted, the GSRs underwent extensive training and competency testing to ensure they were qualified in the area of duct tape end matching. The training program adopted is described in appendix A. Part of this training program was taken from a local forensic science laboratory training protocol for duct tape analysts. The GSRs had approximately 140 hours of practical training over a period of three to four months. During this time, they compared 12820 duct tape specimens.

Hand Torn Tape

The experimental design used two brands of duct tape (Nashua and 3M), with two duct tape grades from each brand (general and professional), and two colors from each grade (black and silver), totaling eight different types of duct tape. Two hundred exemplar and unknown sample pairs from each of the eight tape types were created by hand tearing, resulting in a total of 1600 torn pairs. The exemplar and unknown samples were randomly assigned numbers, and the exemplar numbers were preceded by the letter "E" to differentiate from the unknown tape samples. New tape pairs were created through randomization with half of the pairs matching and half of the pairs not matching and randomized into 1600 envelopes. Each GSR analyzed all 1600 envelopes in order from envelope 0001 through envelope 1600 to determine if each pair was a match, non-match, or inconclusive without consultation, and they were kept blind to each other's results.

Elmendorf Torn Tape

The second experimental trial examined the effect of a constant force and tear angle on analysis through the use of an Elmendorf Tear Tester. 3M general grade gray duct tape was used for this experiment. In total, 200 tape pairs were torn using the Elmendorf Tear Tester. Following separation the same process as in the hand torn trial was used for preparing the tape sets. Each GSR determined whether each envelope contained a matching, non-matching, or inconclusive tape pair, and once again the GSRs were without consultation and were blind to each other's results.

Scissor Cut Tape

The third experimental design of the study involved cutting the duct tape samples with scissors. 3M professional grade gray duct tape was used for this trial. Two hundred exemplar and unknown samples were prepared by cutting duct tape strips in half with scissors. The same process as in the hand torn trial was used for preparing the tape sets. Each GSR determined whether the content of each envelope was a match, a non-match, or inconclusive without consultation or knowledge of the other GSRs' results.

Box Cutter Knife Cut Tape

The final experimental design analyzed duct tape cut with a box cutter knife. 3M professional grade gray duct tape was used for this experiment. Two hundred exemplar and unknown samples were prepared by cutting duct tape strips in one fluid motion with a box cutter knife. After they were cut in half, the duct tape samples were prepared following the same process as in the previous three trials. Each GSR determined whether the tape pairs in each envelope were a match, non-match, or inconclusive without consultation or knowledge of the other GSRs' results.

Research Results

Hand Torn Tape Study

Overall Performance

Exhibit 1 displays performance over all hand torn samples for all GSRs. GSR #1 successfully matched 1542 of 1600 tape sets, correctly identifying 794 of 800 non-matching pairs (99.25%) and 748 of 800 matching pairs (93.50%). GSR #1 did not successfully match 13 of 1600 tape sets resulting in one false-positive identification and 12 false-negative identifications. Forty-five of the 1600 pairs were reported as inconclusive. Of the 45 inconclusive sets, five were true non-matching pairs and 40 were true matching pairs.

GSR #2 successfully matched 1555 of 1600 tape sets, correctly identifying 797 of 800 non-matching pairs (99.25%) and 758 of 800 matching pairs (93.50%). GSR #2 did not successfully match seven of 1600 tape sets resulting in no false-positive identifications and seven false-negative identifications. Thirty-eight of the 1600 pairs were reported as inconclusive. Of the 38 inconclusive sets, three were true non-matching pairs and 35 were true matching pairs.

GSR #3 successfully matched 1584 of 1600 tape sets, correctly identifying 790 of 800 non-matching pairs (98.75%) and 794 of 800 matching pairs (99.25%). GSR #3 did not successfully match six of 1600 tape sets resulting in five false-positive identifications and one false-negative identification. Ten of the 1600 pairs were reported as inconclusive. Of the 10 inconclusive sets, five were true non-matching pairs and five were true matching pairs.

Exhibit 1: GSR Performance over All Hand Torn Samples

	Reported Non-match	Reported Match	Reported Inconclusive	Total
GSR #1:				
True Non-match	794 (99.25%)	1 (0.13%)	5 (0.63%)	800
True Match	12 (1.50%)	748 (93.50%)	40 (5.00%)	800
GSR #2:				
True Non-match	797 (99.63%)	0 (0.00%)	3 (0.38%)	800
True Match	7 (0.88%)	758 (94.75%)	35 (4.38%)	800
GSR #3:				
True Non-match	790 (98.75%)	5 (0.63%)	5 (0.63%)	800
True Match	1 (0.13%)	794 (99.25%)	5 (0.63%)	800

Peer Review Analysis

Exhibit 2 displays tape pairs where two or more of the GSRs did not report a correct identification. Overall 93 pairs of duct tape had an incorrect or inconclusive report by at least one GSR. Of those 93 pairs, 71 pairs had an incorrect or inconclusive report by one GSR, 18 pairs had incorrect or inconclusive reports by two GSRs, and four pairs had incorrect or inconclusive reports by all three GSRs. The 22 tape pairs with multiple incorrect or inconclusive reports were made with tapes NGB, NPB, NPG, 3MGB, and 3MGG. Tapes NPB, 3MGB, and 3MGG had the largest number of multiple incorrect or inconclusive reports. No tape pair was misidentified by all three GSRs.

Exhibit 2: False-Positive (FP), False-Negative (FN), or Inconclusive (I) by Two or More GSRs

Tape	Exemplar	Unknown	GSR #1	GSR #2	GSR #3
NGB	E1018	1011	I	I	
NGB	E1122	1024		I	I
NPB	E3015	3141	I	I	
NPB	E3117	3194	I	I	
NPB	E3171	3095	I	I	
NPB	E3181	3022	I	I	
NPG	E4046	4161	FP		FP
3MGB	E5003	5073	I	FN	
3MGB	E5007	5013	I	I	I
3MGB	E5015	5104	I	I	
3MGB	E5022	5161	I	FN	I
3MGB	E5054	5041	FN		I
3MGB	E5106	5148	I	I	
3MGG	E6003	6052	I	I	
3MGG	E6015	6036	FN	I	
3MGG	E6032	6143	FN	I	FN
3MGG	E6038	6133	I	I	
3MGG	E6045	6048	I	I	
3MGG	E6046	6088	I	I	
3MGG	E6079	6045	FN	I	
3MGG	E6093	6074	I		I
3MGG	E6199	6087	FN	FN	I

Elmendorf Torn Tape Study

Overall Performance

Exhibit 3 displays performance over all Elmendorf torn samples for all GSRs. GSR #1 successfully matched 176 of 200 tape sets, correctly identifying 98 of 100 non-matching pairs (98.00%) and 78 of 100 matching pairs (78.00%). GSR #1 did not successfully match one of 200 tape sets resulting in one false-negative identification. Twenty-three of the 200 pairs were reported as inconclusive. Of the 23 inconclusive sets, two were true non-matching pairs and 21 were true matching pairs.

GSR #2 successfully matched 194 of 200 tape sets, correctly identifying 100 of 100 non-matching pairs (100.00%) and 94 of 100 matching pairs (94.00%). GSR #2 did not successfully match one of 200 tape sets resulting in one false-negative identification. Five of the 200 pairs were reported as inconclusive. Of the five reported as inconclusive sets, all five were true matching pairs.

GSR #3 successfully matched 195 of 199 tape sets, correctly identifying 99 of 100 non-matching pairs (99.00%) and 96 of 99 matching pairs (96.97%). GSR #3 did not successfully match one of 199 tape sets resulting in one false-negative identification. Three of the 200 pairs were reported as inconclusive. Of the three reported as inconclusive sets, one was a true non-matching pair and two were true matching pairs. One tape set was not analyzed giving GSR #3 a total sample size of 199.

Exhibit 3: GSR Performance over All Elmendorf Torn Samples

	Reported Non-match	Reported Match	Reported Inconclusive	Total
GSR #1:				
True Non-match	98 (98.00%)	0 (0.00%)	2 (2.00%)	100
True Match	1 (1.00%)	78 (78.00%)	21 (21.00%)	100
GSR #2:				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	1 (1.00%)	94 (94.00%)	5 (5.00%)	100
GSR #3:				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	1 (1.01%)	96 (96.97%)	2 (2.02%)	99

Peer Review Analysis

Exhibit 4 displays tape pairs where two or more of the GSRs did not report a correct identification. Overall 26 pairs of duct tape had an incorrect or inconclusive report by at least one GSR. Of those 26 pairs, 20 pairs had an incorrect or inconclusive report by one GSR, four pairs had incorrect or inconclusive reports by two GSRs, and two pairs had incorrect or inconclusive reports by all three GSRs.

Exhibit 4: False-Positive (FP), False-Negative (FN), or Inconclusive (I) by Two or More GSRs

Envelope	Exemplar	Unknown	GSR #1	GSR #2	GSR #3
17	E2127	2085	FN	I	I
44	E2016	2017	I	I	
80	E2158	2003	I	I	
108	E2139	2022	I	I	FN
136	E2055	2161	I	FN	
167	E2163	2159		I	I

Scissor Cut Tape Study

Overall Performance

Exhibit 5 displays performance over all scissor cut samples for all GSRs. GSR #1 successfully matched 197 of 200 tape sets, correctly identifying 99 of 100 non-matching pairs (99.00%) and 98 of 100 matching pairs (98.00%). Three of the 200 pairs were reported as inconclusive. Of the three inconclusive sets, one was a true non-matching pair and two were true matching pairs.

GSR #2 successfully matched 189 of 200 tape sets, correctly identifying 92 of 100 non-matching pairs (92.00%) and 97 of 100 matching pairs (97.00%). GSR #2 did not successfully match eight of 200 tape sets resulting in eight false-positive identifications. Three of the 200 pairs were reported as inconclusive. Of the three inconclusive sets, all three were true matching pairs.

GSR #3 successfully matched 197 of 200 tape sets, correctly identifying 98 of 100 non-matching pairs (98.00%) and 99 of 100 matching pairs (99.00%). GSR #3 did not successfully match three of 200 tape sets resulting in one false-negative identification and two false-positive identifications. None of the 200 pairs were reported as inconclusive.

Exhibit 5: GSR Performance over All Scissor Cut Samples

	Reported Non-match	Reported Match	Reported Inconclusive	Total
GSR #1:				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	0 (0.00%)	98 (98.00%)	2 (2.00%)	100
GSR #2:				
True Non-match	92 (92.00%)	8 (8.00%)	0 (0.00%)	100
True Match	0 (0.00%)	97 (97.00%)	3 (3.00%)	100
GSR #3:				
True Non-match	98 (98.00%)	2 (2.00%)	0 (0.00%)	100
True Match	1 (1.00%)	99 (99.00%)	0 (0.00%)	100

Peer Review Analysis

Exhibit 6 displays tape pairs where two or more of the GSRs did not report a correct identification. Overall 14 pairs of duct tape had an incorrect or inconclusive report by at least one GSR. Of those 14 pairs, 12 pairs had an incorrect or inconclusive report by one GSR, one pair had an incorrect or inconclusive report by two GSRs, and one pair had incorrect or inconclusive report by all three GSRs.

Exhibit 6: False-Positive (FP), False-Negative (FN), or Inconclusive (I) by Two or More GSRs

Envelope	Exemplar	Unknown	GSR #1	GSR #2	GSR #3
103	E3103	3035	I	FP	FP
140	E3167	3023		FP	FP

Box Cutter Knife Cut Tape Study

Overall Performance

Exhibit 7 displays performance over all box cutter knife cut samples for all GSRs. GSR #1 successfully matched 200 of 200 tape sets, correctly identifying 100 of 100 non-matching pairs (100.00%) and 100 of 100 matching pairs (100.00%). None of the 200 pairs were reported as inconclusive.

GSR #2 successfully matched 199 of 200 tape sets, correctly identifying 100 of 100 non-matching pairs (100.00%) and 99 of 100 matching pairs (99.00%). GSR #2 did not successfully match one of 200 tape sets resulting in one false-negative identification. None of the 200 pairs were reported as inconclusive.

GSR #3 successfully matched 200 of 200 tape sets, correctly identifying 100 of 100 non-matching pairs (100.00%) and 100 of 100 matching pairs (100.00%). None of the 200 pairs were reported as inconclusive.

Exhibit 7: GSR Performance over All Box Cutter Knife Cut Samples

	Reported Non-match	Reported Match	Reported Inconclusive	Total
GSR #1:				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100
GSR #2:				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	1 (1.00%)	99 (99.00%)	0 (0.00%)	100
GSR #3:				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100

Peer Review Analysis

Only one duct tape sample had a misidentification by one GSR.

Combined Performance Statistics

Exhibit 8 summarizes inconclusive rates for all separation methods and tape types across all GSRs. Exhibit 9 displays the mean GSR inconclusive rates. The mean GSR inconclusive rates ranged from 0.00 to 5.17 with a standard deviation ranging from 0.00 to 5.51. Across the hand torn tape, NGB, NPB, 3MGB, and 3MGG had the highest mean inconclusive rates and standard deviation. The Elmendorf torn 3MGG tape had the same mean inconclusive rate than the hand torn 3MGG tape, but a larger standard deviation. The box cutter knife cut 3MPG tape had the same mean inconclusive rate and standard deviation as the hand torn 3MPG tape, which was lower than the scissor cut 3MPG tape. In general, across all tape samples, GSR #1 had the most inconclusive reports and GSR #3 had the fewest inconclusive reports.

Exhibit 8: Combined Inconclusive Rates (%)

Sample	GSR #1	GSR #2	GSR #3	Mean	SD
Hand Torn NGB	2.50	3.50	1.50	2.50	1.00
Hand Torn NGG	1.50	0.00	0.00	0.50	0.87
Hand Torn NPB	6.00	6.00	0.00	4.00	3.46
Hand Torn NPG	0.00	0.50	0.00	0.17	0.29
Hand Torn 3MGB	4.00	3.50	1.50	3.00	1.32
Hand Torn 3MGG	8.00	5.50	2.00	5.17	3.01
Hand Torn 3MPB	0.50	0.00	0.00	0.17	0.29
Hand Torn 3MPG	0.00	0.00	0.00	0.00	0.00
Elmendorf Torn 3MGG	11.50	2.50	1.51	5.17	5.51
Scissors Cut 3MPG	1.50	1.50	0.00	1.00	0.87
Box Cutter Cut 3MPG	0.00	0.00	0.00	0.00	0.00

Exhibit 9: Mean GSR Inconclusive Rate

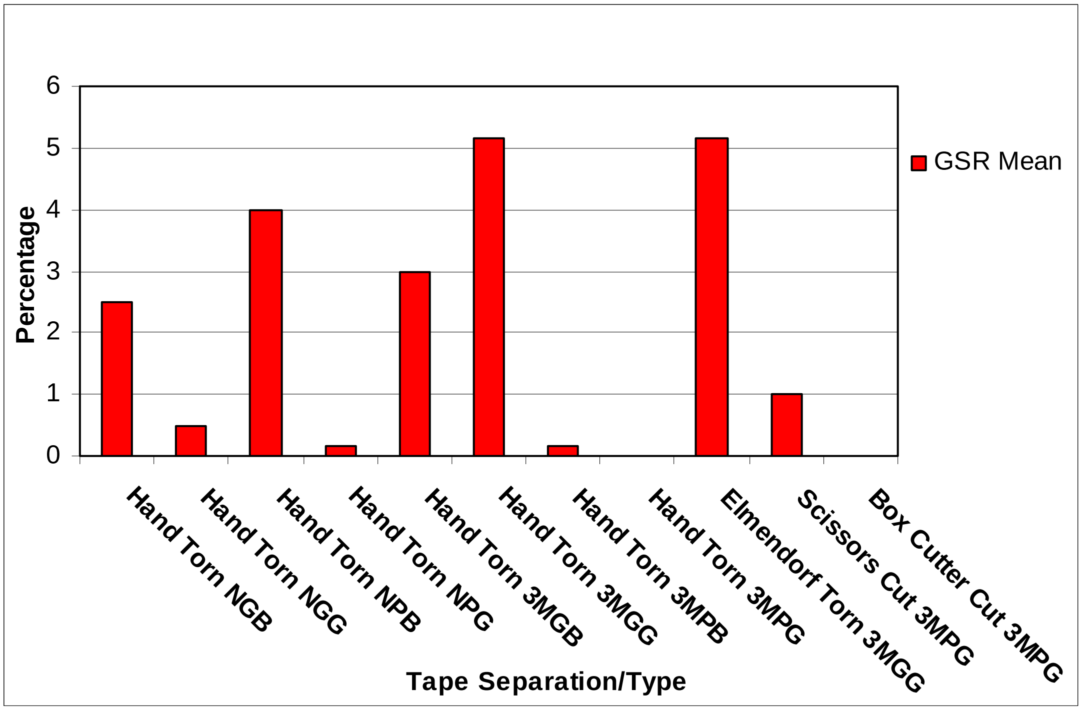


Exhibit 10 summarizes false-positive rates for all separation methods and tape types across all GSRs. Exhibit 11 displays the mean GSR false-positive rates. The mean GSR false-positive rates range from 0.00 to 3.33 with a standard deviation ranging from 0.00 to 4.16. In the hand torn tape set, all false-positives were reported with tapes NPB, NPG, 3MGG, and 3MPG. NPG and 3MPG had the same mean false-positive rate, but 3MPG had a slightly higher standard deviation. The Elmendorf torn 3MGG tape had a higher mean false-positive rate and standard deviation than the hand torn 3MGG tape. The box cutter knife cut 3MPG tape had a lower mean false-positive rate and standard deviation than the hand torn 3MPG tape, and the hand torn 3MPG tape had a lower mean false-positive rate and standard deviation than the scissor cut 3MPG tape. In general, across all tape samples, GSR #3 had the most false-positive identifications and GSRs #1 and #2 had few or no false-positive identifications.

Exhibit 10: Combined False-Positive Rates (%)

Sample	GSR #1	GSR #2	GSR #3	Mean	SD
Hand Torn NGB	0.00	0.00	0.00	0.00	0.00
Hand Torn NGG	0.00	0.00	0.00	0.00	0.00
Hand Torn NPB	0.00	0.00	1.00	0.33	0.58
Hand Torn NPG	1.00	0.00	1.00	0.67	0.58
Hand Torn 3MGB	0.00	0.00	0.00	0.00	0.00
Hand Torn 3MGG	0.00	0.00	1.03	0.34	0.60
Hand Torn 3MPB	0.00	0.00	0.00	0.00	0.00
Hand Torn 3MPG	0.00	0.00	2.00	0.67	1.15
Elmendorf Torn 3MGG	0.00	0.00	0.00	0.00	0.00
Scissors Cut 3MPG	0.00	8.00	2.00	3.33	4.16
Box Cutter Cut 3MPG	0.00	0.00	0.00	0.00	0.00

Exhibit 11: Mean GSR False-Positive Rates

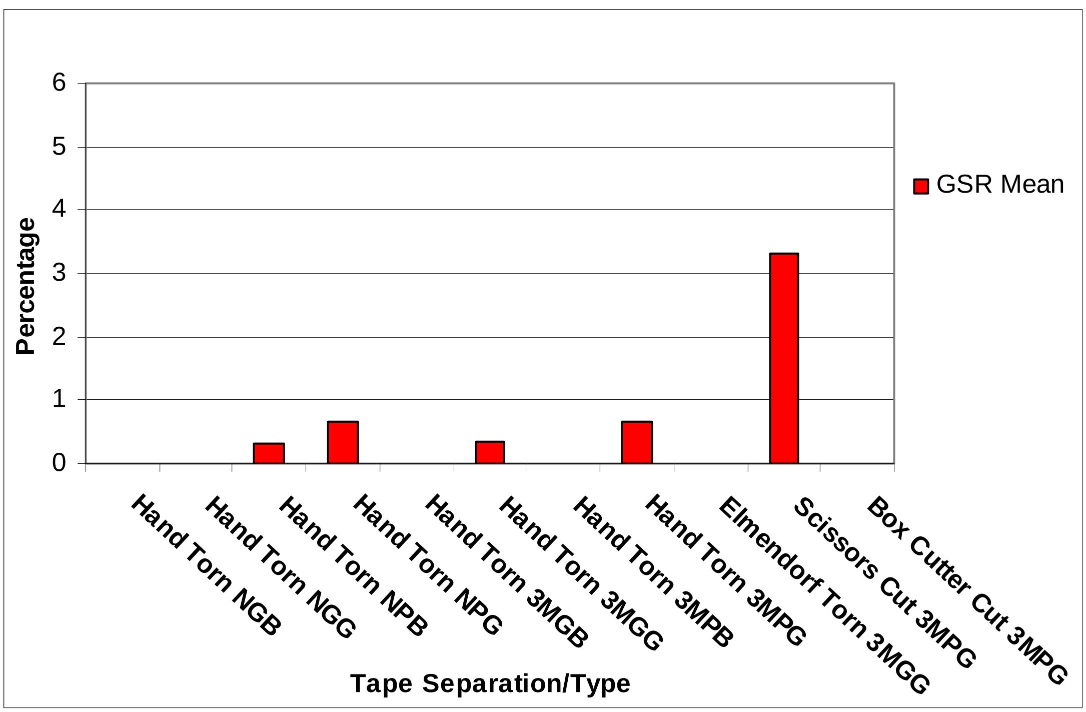


Exhibit 12 summarizes false-negative rates for all separation methods and tape types across all GSRs. Exhibit 13 displays the mean GSR false-negative rates. The mean GSR false-negative rates range from 0.00 to 2.67 with a standard deviation ranging from 0.00 to 2.78. In the hand torn tape set, NPB, 3MGB, and 3MGG had the highest mean false-negative rates and standard deviations. The Elmendorf torn 3MGG tape had a lower mean false-negative rate and standard deviation than the hand torn 3MGG tape. The box cutter knife cut 3MPG tape had the same mean inconclusive rate and standard deviation as the scissor cut 3MPG tape, which was higher than the hand torn 3MPG tape. In general, across all tape samples, GSR #3 had few or no false-positive identifications, while GSR #1 and GSR #2 had higher rates

Exhibit 12: Combined False-Negative Rates (%)

Sample	GSR #1	GSR #2	GSR #3	Mean	SD
Hand Torn NGB	0.00	0.00	0.00	0.00	0.00
Hand Torn NGG	0.00	1.00	0.00	0.33	0.58
Hand Torn NPB	3.37	1.12	0.00	1.50	1.72
Hand Torn NPG	0.00	0.00	0.00	0.00	0.00
Hand Torn 3MGB	4.26	3.19	0.00	2.48	2.21
Hand Torn 3MGG	5.88	1.12	1.01	2.67	2.78
Hand Torn 3MPB	0.00	1.00	0.00	0.33	0.58
Hand Torn 3MPG	0.00	0.00	0.00	0.00	0.00
Elmendorf Torn 3MGG	1.27	1.05	1.03	1.12	0.13
Scissors Cut 3MPG	0.00	0.00	1.00	0.33	0.58
Box Cutter Cut 3MPG	0.00	1.00	0.00	0.33	0.58

Exhibit 13: Mean GSR False-Negative Rates

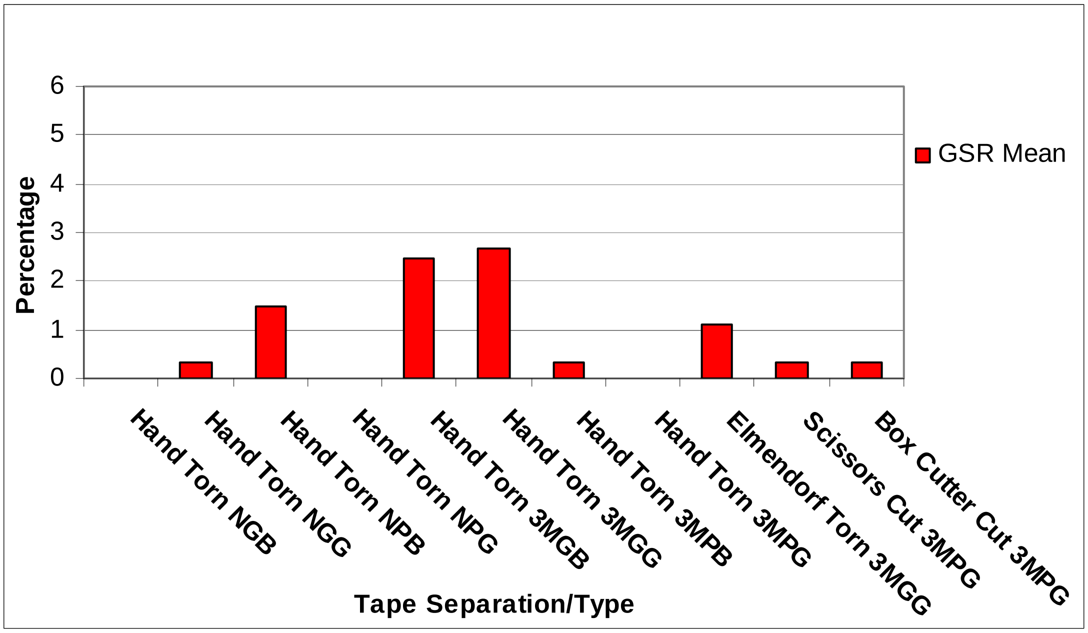
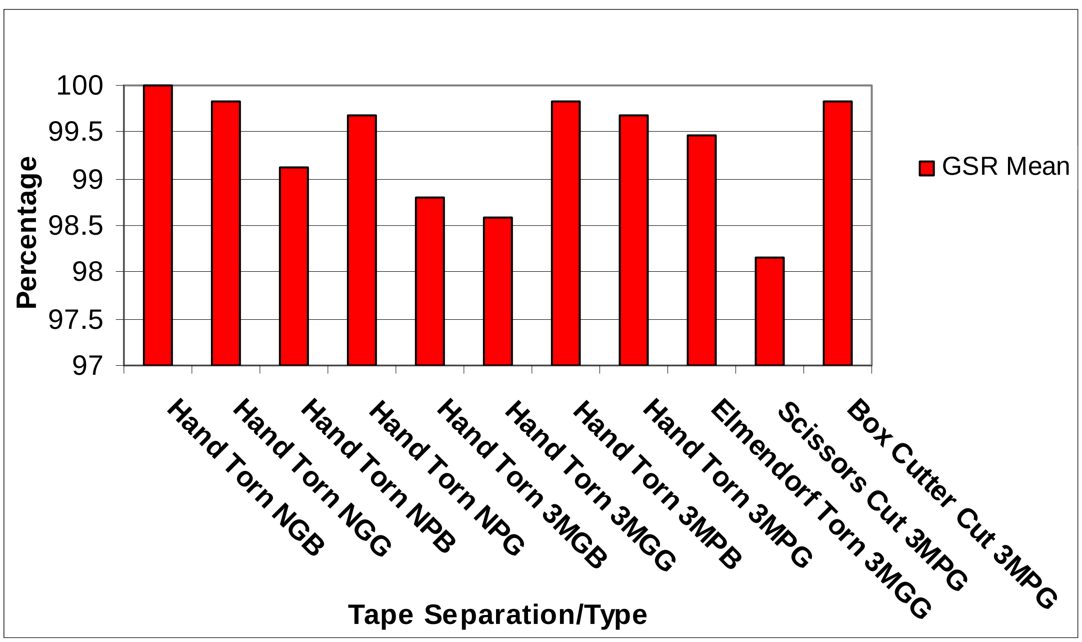


Exhibit 14 summarizes accuracy rates for all separation methods and tape types across all GSRs. Exhibit 15 displays the mean GSR accuracy rates. The mean GSR accuracy rates range from 98.15 to 100.00 with a standard deviation ranging from 0.00 to 2.05. In the hand torn tape set, 3MGB and 3MGG were the only tape types to have a mean accuracy below 99.00 percent. 3MGB and 3MGG also had the largest standard deviation. The Elmendorf torn 3MGG tape had a higher mean accuracy rate and standard deviation than the hand torn 3MGG tape. The box cutter knife cut 3MPG tape had a higher mean accuracy rate and standard deviation than the hand torn 3MPG tape, and the hand torn 3MPG tape had a higher mean accuracy rate and standard deviation than the scissor cut 3MPG tape. In general, all three GSRs had high accuracy rates across all tape samples.

Exhibit 14: Combined Accuracy Rates (%)

Sample	GSR #1	GSR #2	GSR #3	Mean	SD
Hand Torn NGB	100.00	100.00	100.00	100.00	0.00
Hand Torn NGG	100.00	99.50	100.00	99.83	0.29
Hand Torn NPB	98.40	99.47	99.50	99.12	0.63
Hand Torn NPG	99.50	100.00	99.50	99.67	0.29
Hand Torn 3MGB	97.92	98.45	100.00	98.79	1.08
Hand Torn 3MGG	97.28	99.47	98.98	98.58	1.15
Hand Torn 3MPB	100.00	99.50	100.00	99.83	0.29
Hand Torn 3MPG	100.00	100.00	99.00	99.67	0.58
Elmendorf Torn 3MGG	99.44	99.49	99.49	99.47	0.03
Scissors Cut 3MPG	100.00	95.94	98.50	98.15	2.05
Box Cutter Cut 3MPG	100.00	99.50	100.00	99.83	0.29

Exhibit 15: Mean GSR Accuracy Rates



Conclusion

This study confirms that it is possible to use physical end matching to identify duct tape samples as matching or non-matching and that differences between analysts, brands, tape grades, tape color, and method of separation have varying contributions to misidentifications or inconclusive results. The mean accuracy observed ranged from 98.58 to 100 percent for torn tape and from 98.15 to 99.83 percent for cut tape. The mean false-positive rate ranged from 0.00 to 0.67 percent for torn tape and from 0.00 to 3.33 percent for cut tape. The mean false-negative rate ranged from 0.00 to 2.67 percent for torn tape and was 0.33 percent for both types of cut tape.

Overall the three GSRs were consistent in their analysis. Given that an identification was made, all three GSRs showed an accuracy of 95.94 percent or higher for all tape types. A low percentage of errors were made and many of the errors that did occur were seen with similar tape types or on similar tape samples.

This study focused on two brands of duct tape, Nashua and 3M. In this study, 3M tapes had a higher percentage of misidentifications and inconclusive results than Nashua tapes. However, the 3M general grade tapes contributed to the majority of the inconclusive and false-negative identifications. Therefore, the combination of brand and grade likely interact to affect an analyst's ability to correctly identify duct tape end matches.

General grade tapes had a higher percentage of overall misidentifications and inconclusive results than professional grade tapes. This is likely due to the structural differences between the two grades, as professional grade tape is generally stronger and less prone to distortion or stretch upon separation. It is probable that tape grade affects an analyst's ability to conduct these examinations.

The total number of misidentifications in this study was similar for each tape color. The black tapes gave a slightly higher number of false-negative identifications, while the gray tapes gave a higher rate of false-positive identifications. The black tapes gave a higher percentage of inconclusive results. In general, brand and grade appeared to have a more significant effect on an analyst's ability to correctly identify duct tape end matches.

In this study, the Elmendorf torn tape results were compared with the 3M general grade gray hand torn tape results. Overall, the Elmendorf torn tape pairs and the hand torn tape pairs gave similar results. The hand torn tapes had a slightly higher percentage of false-positive and false-negative identifications, and the total number of reported inconclusive results was equal for both separation methods. As such, consistent tearing conditions do not appear to have a significant affect on an analyst's ability to correctly analyze duct tape physical end matches.

The cut tape results were compared with the 3M professional grade gray hand torn tape results. The scissor cut tape had a higher percentage of inconclusive and false-positive identifications, but the hand torn tapes had a slightly higher percentage of false-negative identifications. However, it should be noted that the majority of the false-positive identifications on the scissor cut tapes were made by a single analyst. The tapes cut with a box cutter knife had an equal number of inconclusive results compared with the hand torn tapes, but the hand torn tapes had a slightly higher percentage of both false-positive and false-negative identifications. In general, tape cut with scissors had a greater affect on an analyst's ability to correctly analyze duct tape physical end matches than hand torn tape. In contrast, tape cut with a box cutter knife did not appear to have a significant affect on an analyst's ability to correctly analyze duct tape physical end matches as compared with hand torn tape.

Implications for Policy and Practice

The following recommendations have been formed for analysts conducting duct tape end matching:

- Report inconclusive results for difficult to match samples whenever there is any doubt in order to avoid false-positive results.
- Develop a comprehensive training procedure for both the analyst and peer reviewer modeled after the one listed in the appendix unless new procedures can be assessed and shown to provide higher accuracy with fewer errors.
- Use a light box and magnification light during analysis to enhance analysis of duct tape physical end matching.
- All duct tape end matches should be subjected to a process of blind peer review.

Future Recommendations

It is recommended that studies be conducted using a larger number of analysts, consisting of both paid GSRs and analysts from the forensic science community. Reviewing a larger population of duct tape brands in combination with varying separation methods will help to fully evaluate whether the brand, tape grade, or separation method has a greater effect on analyst accuracy and error rates.

Finally, it is recommended that future studies be conducted on duct tape end matching using mathematical and computerized methods of analysis, and that those results be compared with this study to determine the abilities of computerized analysis versus human analysis. One such proposed method would mathematically profile the torn end of a duct tape and subsequently compare the resulting profile to other duct tape specimens that have been mathematically profiled.

Introduction

Duct tape was first invented by Harold Potter, a distributor for Arno Adhesive Tapes, in the mid-1940s (Johnston 2003). Originally used during World War II as a waterproof surgical tape to seal ammunition boxes, the industry waned at the end of the war when there was no longer a great need by the military (Smith 2007). Using the same technology, Potter replaced the olive, military-colored pigment with aluminum powder and carbon black and manufactured a waterproof tape for sealing duct systems. Over time, duct tape adapted a broad range of uses, and with its popularity came a variety of brands, grades, colors, widths, thicknesses, and materials that make up each individual tape (Johnston 2003).

Duct tape comprises three main components: the backing, fabric or yarn reinforcement also known as the scrim, and the adhesive. The backing is generally made from polyethylene and comes in a variety of thicknesses and colors. The scrim (exhibit 16) is a woven or gauze-like fabric, generally cotton or a polyester blend, which is used to strengthen the tape. The yarn running the length of the tape, or machine direction, is known as the warp, while the yarn running across the tape is referred to as the fill (Smith 2007).

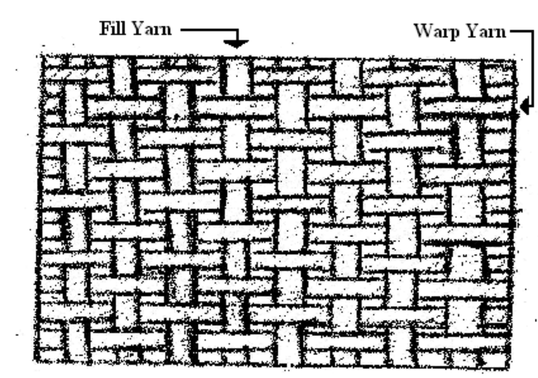


Exhibit 16: Warp and fill yarns in plain weaving (©UCD–MSFS)

The number, size, and type of yarn (exhibit 17) used to make up the reinforcement fabrics vary between brands and tape grades. Industrial grade tapes generally have a higher yarn count and the yarn may be thicker, increasing its overall strength, while tapes manufactured for general retail have a lower yarn count. Finally, the adhesive is made up of a combination of various elastomers, tackifying resins, and fillers that vary depending on the manufacturer and purpose for which the tape is being made (Smith 2007).

a b c

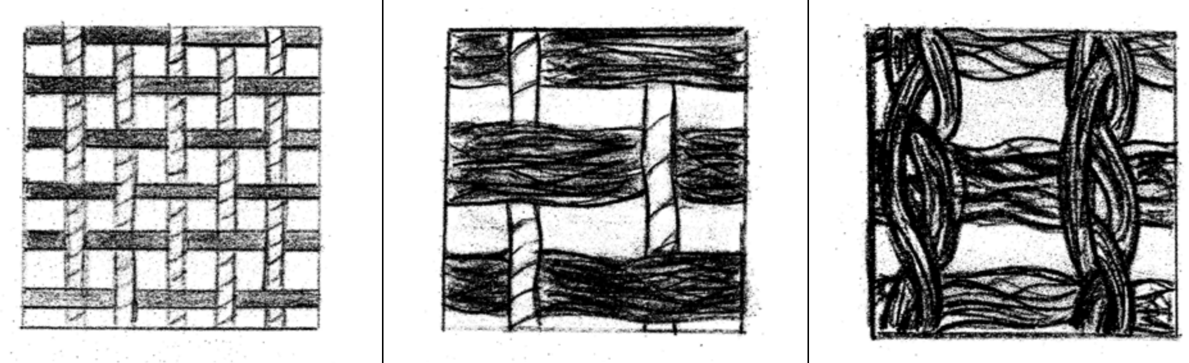


Exhibit 17a depicts a typical over-under woven scrim. Exhibit 17b demonstrates a woven scrim known as Polyfil. Exhibit 17c illustrates a typical weft insertion scrim pattern. (©UCD–MSFS)

Presently, duct tape is sold at most convenience stores in America. From a forensic science point of interest, duct tape is often found in association with criminal activity, most commonly in abductions and homicides as ligatures, restraints, and blindfolds and in the construction of explosive devices – all considered felony crimes. As such, forensic scientists are frequently asked to analyze and compare duct tape samples in order to establish possible evidentiary links between a suspect and victim, a suspect and a particular crime, or multiple crimes (Smith 2007).

Forensic evaluation of pressure-sensitive tapes can comprise a range of physical and chemical analyses. Extensive research in tape analysis has been conducted using methods such as microscopy, color analysis, fabric patterns, and chemical analysis, employing infrared spectroscopy, x-ray fluorescence, scanning electron microscopy energy-dispersive spectroscopy, inductively coupled plasma mass spectrometry, and pyrolysis gas chromatography (Smith 2007).

Although all of these methods can provide useful information as to the possible brand and grade of a tape sample or whether or not two samples share a common origin, one of the greatest evidentiary values comes from physical end matching.

Physical end matching, known as fracture matching when referring to more brittle substances, is performed by an analyst in an attempt to reunite separated fragments. Based on the principle that each fracture is unique, this type of matching has a significantly high evidentiary value and is considered to be one of the strongest associations in forensic science comparative examinations (Claytor 2010). This principle can be attributed to the nature of fracturing and the behavior of fractures, which are "an accumulation of random processes (including the stress itself, grain size of the material, and crystal structure), which when combined, produce a unique surface" (Claytor 2010: 324). Numerous studies have been done to show the value of fracture matching in comparative analysis and to prove the uniqueness of physical matching using a range of materials commonly encountered in criminal cases, such as torn metal-coated paper, silicon sheets, various metals, glass, and multiple types of tapes (Claytor 2010, Smith 2007, Tsach 2007). Argon and Schecter (1986) used photographic means to examine two methods of tearing electrical tape (tension and shearing) and determined that each tear was unique and non-reproducible. Although some of these studies have provided error rates for duct tape physical end matching, few samples have been examined.

Claytor and Davis (2010) illustrated the uniqueness of fractures in an experiment using hacksaw blades, and though the data pool was considered small, the study achieved a specificity rate of 98 percent and an error rate of two percent. The study looked at the fracture of metal to depict two things. First, they wanted to show that each fracture of consecutively manufactured hacksaw blades is unique and individualizing by examining fracture edges to find a "pattern fit"

(Claytor 2010). Second, they wanted to show that a fractured blade edge can be correctly associated to a cast of that edge through examination of topographical characteristics. Both of these hypotheses were found to be true with a small number of samples; however, further experimentation was required to determine if these ideals still held true with larger data pools (Claytor 2010).

Similar hypotheses were tested by Tsach *et al* (2007), looking at torn edges of various materials with different mechanical properties. This mechanical process is termed shearing and occurs upon application of a stress "greater than the internal strength of the material" (Tsach 2007: 77). The applied force and physical structure of the material influence the direction and shape of the shear. In this study, metal-coated paper and sheets of white and red silicon casts were torn with a tensile machine to create reproducible conditions for each tear. The results allowed the hypothesis to be accepted that all tears were unique, even under repetitive conditions, which was shown by the fact that experts were able to correctly associate all torn materials. No statistical error rate was given due to the limited number of samples in this experiment.

In the case of duct tape, shearing results from the failure of the elastomeric and fibrous materials of the polyethylene backing and the scrim fibers, resulting in two two-dimensional edges that can be matched based on individual characteristics. To perform the end matching of duct tape, the analyst assesses a number of factors. The first step is the evaluation of the overall correspondence of the duct tape characteristics, termed class characteristics. Class characteristics of a specimen are measurable features that are determined by the manufacturer, including but not limited to width, color, tape backing features, layer composition, adhesive type, and chemical composition (NIJ). After these features are found to match between two specimens, then individual and identifying characteristics are assessed. This involves examining the physical correspondence between the torn or cut edges and scrim alignment.

Bradley *et al* (2006) looked specifically at this process for duct tape end matching and made an assessment of error, though, as in the Claytor and Davis study, the data pool was small. This experiment set out to determine whether two or more samples could be linked to a common source by physical end matching and the error rates associated with that analysis. Four analysts were each asked to examine three hand torn tape sets and two scissor cut tape sets to determine the number of end matches in each set. No misidentifications were made in the initial study, but some matches were not identified and considered inconclusive. Blind peer review by the other three analysts was performed when an end match was not identified. Although there were no misidentifications in the hand torn samples during this process, there were two misidentifications with the scissor cut samples, suggesting that scissor cut ends present more of a challenge to the analysts as there are fewer points of comparison. This proved especially true the closer the angle of the cut edge is to 90 degrees (Bradley 2006).

Physical end matching of duct tape is common in crime laboratories, yet the process of physical end matching remains undefined and the reliability and error rates associated with this generally accepted procedure are unknown. As end matching is practiced today, each tear is assumed to be unique. Although studies such as Argon and Schecter (1986) and Tsach *et al* (2007) have shown that fractures and shearing seem unique, to date, only a single, unpublished study has looked at a moderately large number of torn duct tape specimens. Tulleners (2010) described the inter-comparison of 100 exemplar tapes with 100 unknown tapes. In this set, 94 tape pairs matched while 6 exemplar tapes did not match any of the unknowns. This exercise resulted in the comparison of 10000 tape pairs. Each analyst correctly matched each of the unknown tape samples to the corresponding exemplar sample, while correctly excluding the six blind specimens. Also, neither analyst was able to find multiple matches for a single exemplar specimen, demonstrating that each tear is unique.

The National Research Council has called for research to address the issues of accuracy, reliability, and the validity in the forensic sciences through studies establishing objective scientific criteria for the various forensic science methods (NRC 2009). This study was designed to expand upon Bradley's and Tulleners' research by evaluating the variation in inconclusive rates, error rates, and accuracy rates due to: differing brands, grades, and colors of duct tape; differing analysts; and differing separation methods. Other goals for this study included developing a protocol for duct tape physical end matching as well as a protocol for training analysts in physical end matching. This study was also designed to help assess the reliability of the protocol as represented by the reproducibility of the end result provided by blind peer review, as well as allow the assessment of conditions that could restrict its validity.

Methods

Research Design

The primary design examined the effects of brand, grade, and color on false positive and false negative error rates in hand torn duct tape. Two more limited designs were also studied. One design examined Elmendorf machine torn samples in comparison with one of the hand torn samples. The other design examined scissor cut and box cutter knife cut samples in comparison with another of the hand torn samples.

The main focus of the study was to obtain good quality estimates of the error rates and the potential impact of brand, grade, and color, as well as an assessment of the impact of analyst to analyst variation and that of different separation methods. A factorial design across brand, grade, and color was implemented to allow the assessment of possible interactions among these factors. To allow an assessment of peer review, each analyst evaluated the same envelopes in the same sequence.

The hand torn study in either matching or non-matching envelopes was designed as a single blinded completely randomized 2^3 factorial design with a binary response variable. The three factors are duct tape brand (Nashua and 3M), grade (general and professional), and color (black and gray). Analysts may be considered a repeated measures factor at each combination of brand, grade, and color.

For the 3M general grade gray tape only, the Elmendorf torn and the corresponding hand torn samples provided a single blinded blocked randomized one-way design. The single factor was separation method (Elmendorf torn versus hand torn). The blocking factor of time was completely confounded with separation method. Analysts may be considered a repeated measures factor for each separation method.

For the 3M professional grade gray tape only, the scissors cut, box cutter knife cut, and hand torn samples provided a single blinded blocked randomized design. The single factor was separation method (scissors cut versus box cutter knife cut versus hand torn). The blocking factor of time was completely confounded with separation method. Analysts may be considered a repeated measures factor for each separation method.

The sample size of 100 pairs for each combination of factor levels was chosen for each error rate (either false positive or false negative) in order to provide a 95 percent confidence of detecting an error rate of three percent or more (Louis 1981). This resulted in a total of 200 pairs for each combination of factor levels.

The factorial design coupled with the required sample size to estimate low error rates restricted the number of brands, grades, and colors that could be studied. The design was chosen to allow assessment of two and three factor interaction effects on error rates, as well as an initial assessment of the possible effects of brand. Although more brands could have been studied, as well as industrial grade and the many other possible colors, this would have produced a design that could not provide good information on either interaction effects or on the actual error rates.

The repeated measures design with respect to analysts would allow the assessment of the value of peer review on results. Although analysts could naturally be expected to vary in their error rates, a consistent training methodology was used to try to eliminate obvious confounders such as use of different criteria, different evaluation conditions, and different definitions of matching and non-matching. The three selected Graduate Student Researchers (GSRs) from the Forensic Science Graduate program at the University of California, Davis produced individual results. Each GSR worked part-time to eliminate fatigue and to maintain consistency among the analysts.

Training

Before experimental trials were conducted, the GSRs underwent extensive training and competency testing to ensure they were qualified in the area of duct tape end matching. The training program was adopted, in part, from a local forensic science laboratory training regiment for duct tape analysts. The training consisted of the following five phases:

Phase 1: Literature Review and Group Discussion

Each GSR was required to read journals and articles on a variety of topics related to duct tape, pressure sensitive tapes, fracture matching, and methods of physical and chemical analysis of tapes. Each piece of literature was read individually and later reviewed by the GSRs in a group discussion. The list of the training material reviewed by the GSRs in this study is available as appendix B.

Phase 2: Protocol for End Matching Exercises

Throughout their analysis, the GSRs were expected to record the following information:

- Exemplar #: the number assigned to the exemplar duct tape sample.
- Unknown #: the number assigned to the unknown duct tape sample.
- Distance of Matching Area: a measurement for the longest amount of matching area along the width of the tapes (the torn edge) in centimeters.
- % Match: distance of matching area measured divided by the total width of the tape (in centimeters).
- Match Category: a number 1-6 assigned to the tape pair where 1 = match, 2 = possible match, 3 = inconclusive, 4 = likely non-match, 5 = non-match, and 6 = exclusion due to different tape types.
- Description of Reasoning: notes or information explaining the match category selected.

An example of the training form used by the GSRs in this study is available as appendix C.

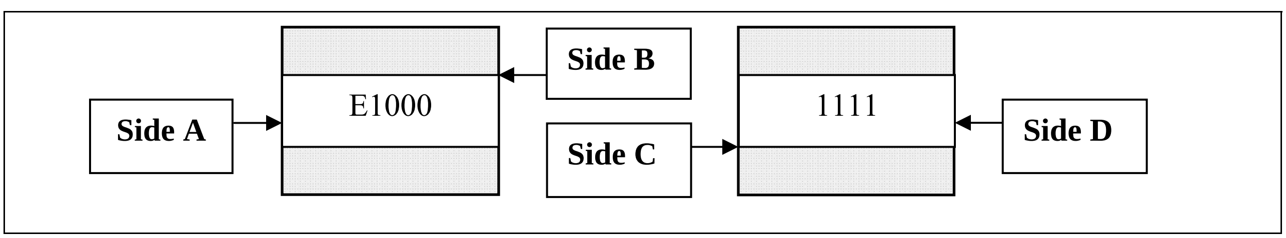
Phase 3: Practice Set One for Hand Torn Tape

Set one consisted of 20 envelopes, each containing four hand torn exemplar tape samples and four hand torn unknown tape samples. Each tape sample consisted of one torn edge and one cut edge, and only the torn edges were used for comparison. Each GSR compared all four exemplars to all four unknown samples in each envelope, giving a total of 16 comparisons per envelope and 320 total duct tape comparisons. After each GSR completed practice set one, the results were compared with one another and peer reviewed. The peer review consisted of reaching a consensus between the GSRs on distance of matching area, percent match, and match category for all 320 comparisons.

Phase 4: Practice Set Two for Hand Torn Tape

Set two included 25 envelopes, each containing five hand torn exemplar tape samples and five hand torn unknown tape samples. In this set, both ends of each piece of tape were torn. All five exemplars were compared with all five unknowns in each envelope; the order and orientation for each tape was unknown to the GSR, giving four possibilities for matching each set of tape. To differentiate tape order and orientation, it was agreed that when looking at an exemplar sample with the numbered label reading left to right, the left side of the tape would be side A and the right side would be side B. Similarly, looking at each unknown sample with numbered label reading left to right, the left side of the tape would be side C and the right side would be side D (exhibit 18).

Exhibit 18: Side Identification of Practice Set 2 Tapes



The possible combinations resulted in 100 comparisons per envelope and 2500 total duct tape comparisons. Throughout the analysis the GSRs were expected to compare all four sides for each set of tape, choosing the side with the largest distance of matching area to analyze. The GSRs followed the same protocol from practice set one and recorded the same information;

however, the GSRs also recorded whether sides A, B, C, or D were used in the final evaluation for each tape set. After each GSR completed practice set two, the results were compared with one another and peer reviewed. The peer review consisted of reaching a consensus between the GSRs on distance of matching area, percent match, and match category for all 2500 comparisons.

Phase 5: 100 Set Competency Test

The 100 set competency test consisted of 100 exemplar pieces of tape and 100 unknown pieces of tape. Out of the 100 pairs, 95 matched while five pairs did not have any matches. Each tape sample consisted of one torn edge and one cut edge, and only the torn edges were used for comparison. Each GSR compared all 100 exemplar pieces to all 100 unknown pieces, making a total of 10000 comparisons. In Phase 5, the GSRs recorded the paired matches found and submitted their final answers for external review. Each GSR was required to successfully complete this competency test with 100 percent accuracy in order to continue with the study. Overall, the GSRs completed approximately 140 hours of training over a 3 month period and conducted 12820 duct tape comparisons.

Hand Torn Tape

Upon the completion of training and achieving 100 percent accuracy on the competency test, the GSRs investigated hand torn duct tape. The experimental design used duct tape from two brands common to the area where the study was conducted: Nashua and 3M. Two duct tape grades were used from each brand (general and professional) and two colors from each grade (black and silver) to give a total of eight different types of duct tape. Exhibit 19 illustrates the process of preparing the tape sets. First, the duct tape samples were created by hand tearing tape from each of the eight different types of duct tape and placing the matching ends adjacent to one another on transparency films. The specimen on the left side of the acetate sheet was considered the exemplar sample, while the specimen on the right side was the unknown sample. Overall, 200 exemplar and unknown samples were hand torn from each of the eight different duct tapes for a total of 1600 torn pairs of tape to compare. The GSRs shared the task of tearing the tape samples. To maintain consistency, a single analyst was tasked with tearing all 200 samples for a single tape type.

Exhibit 19: Process of Preparing Tape Sets

Step 1: Create Samples/ Place on Transparency	**Step 2: Assign Numbers to Matching Pairs**	**Step 3: Create Pairs (1/2 match, 1/2 non-match)**	**Step 4: Assign Envelopes**
	E1001 1027	E1001 1027	Envelope 20
	E1002 1136	E1002 1116	Envelope 136
	E1003 1188	E1003 1094	Envelope 42
	E1004 1001	E1004 1001	Envelope 187

Next, while on the transparency the exemplars were labeled sequentially and the corresponding unknown samples were assigned random numbers. The numbers for the exemplars were preceded by the letter "E" to be differentiated from the unknown tape samples. An individual outside the study then paired the exemplars and unknowns with half of the sets matching and the other half randomly paired to be non-matching. The exemplars and unknowns were paired within their respective tape types; tape samples from differing tape types were never paired together. Finally, the pairs were randomized into 1600 envelopes. The 1600 envelopes were divided into sequential sets of 50 envelopes, and each GSR analyzed the envelopes of duct tape pairs set by set. In this way, each GSR analyzed all 1600 envelopes in order from envelope 0001 through envelope 1600 and determined if each pair was matching, non-matching, or inconclusive. An inconclusive result meant that the GSR was unable to conclude a match or a non-match, most likely due to distortion of the torn edges. In such a case, the GSR would generally observe enough corresponding features to suspect a match, but there were not enough matching features to report the pair as a definitive match. Each GSR analyzed the samples without consultation, and they were kept blind to each other's results.

Elmendorf Torn Tape

The second experimental trial examined machine torn duct tape using an Elmendorf Tear Tester. An Elmendorf Tear Tester is an instrument generally used to determine the internal tearing resistance of paper and board (exhibit 20). In this study, the Elmendorf instrument was used to tear duct tape under constant conditions – force and tear angle. The purpose of using the Elmendorf Tear Tester was to determine if constant force and tear angle would affect end matching analysis by the GSRs. The 3M general grade gray duct tape from the hand torn trial was selected to be used for this experiment. Each duct tape sample was measured, cut to four inches, and secured to the holders with the dual leveling screws on the Elmendorf instrument. The pendulum was raised and then released, causing the tape to be torn in half. Previous unpublished research using the Elmendorf method to tear duct tape mentioned making a small cut on the bottom of the tape before starting;

Exhibit 20: Elmendorf Tear Tester

however, this was deemed an unnecessary step and eliminating the cut was thought to create a more realistic scenario (Tulleners 2010). Two hundred torn pairs were created by a single GSR using the Elmendorf procedure. Following separation, the same process as in the hand torn trial was used for preparing the tape sets (exhibit 19). Each GSR determined whether each envelope contained a matching, non-matching, or inconclusive tape pair. Each GSR analyzed the 200 Elmendorf torn samples without consultation, and they were again blind to each other's results.

Scissor Cut Tape

The third experimental design of the study involved cutting the duct tape samples with scissors. The 3M professional grade gray duct tape from the hand torn set was selected for this trial. Once more a single GSR created the 200 exemplar and unknown samples by cutting approximately six-inch samples of duct tape in half with a pair of scissors for each duct tape sample. The same process as in the hand torn trial was used for preparing the tape sets (exhibit 19). Before beginning the experiment, each GSR was required to inter-compare 16 practice scissor cut specimens for a total of 256 comparisons. All three GSRs were able to match these specimens with 100 percent accuracy and continue with the experimental samples. The GSRs individually analyzed the 200 scissor cut samples and determined if each pair was a match, a non-match, or inconclusive The GSRs were again without consultation or knowledge of the other GSRs' results.

Box Cutter Knife Cut Tape

The final experiment in this design analyzed duct tape cut with a box cutter knife. The 3M professional grade gray duct tape that was used in the hand torn and scissor cut trials was also used for the box cutter experiment. A single GSR prepared 200 exemplar and unknown samples. The GSR first used scissors to cut approximately six-inches of duct tape. Next, half of the duct tape was stuck onto a surface while the other half was pulled tight and then cut in a fluid slicing motion with a box cutter knife. After being cut in half the duct tape samples were prepared following the same process as in the previous three trials (exhibit 19). Before they began the experiment, each GSR was required to inter-compare 16 box cutter cut knife cut tape pairs for a total of 256 comparisons. All three GSRs were able to match these specimens with 100 percent accuracy and allowed to continue with the experiment. Each GSR individually analyzed the 200 experimental tape pairs and determined whether each tape pair was a match, non-match, or inconclusive. These comparisons were completed without consultation, and each GSR was blind to the other GSRs' results.

Materials

The materials used in this study were purchased from local hardware stores. Exhibit 21 shows the tape types and their number assignments for the hand torn trial.

Exhibit 21: Tape Types and Number Assignments for Hand Torn Trial

Tape Description	Tape Vendor Number	Range of Assigned Number (Hand Torn Trial)
Nashua General Black (NGB)	394	1001-1200
Nashua General Gray (NGG)	394	2001-2200
Nashua Professional Black (NPB)	398	3001-3200
Nashua Professional Gray (NPG)	398	4001-4200
3M General Black (3MGB)	L255	5001-5200
3M General Gray (3MGG)	L255	6001-6200
3M Professional Black (3MPB)	6969	7001-7200
3M Professional Gray (3MPG)	6969	8001-8200

Eight different tape types were used throughout the experimental trials: Nashua 394 General Purpose Duct Tape in Black and Gray, Nashua 398 Max Duty Duct Tape in Black and Gray, 3M L255 All-Purpose Duct Tape in Black and Gray, and 3M 6969 Industrial Grade Duct Tape in Black and Gray. Other materials included 3M AF4300 Write-On Transparency Films, Office Depot All-Purpose Envelopes (3 5/8" x 61/2"), and Office Depot White Inkjet Labels (1/2" x 1 3/4"). Various instruments were used in the preparation of the tape specimens and during analysis, including an Elmendorf Tear Tester instrument, 10" x 7.5" light box, 5X portable magnification lamp, scissors, a box cutter knife, and a transparent ruler with centimeter and millimeter units. An 8-35x stereo microscope was also available to the analysts.

The light box and magnification lamp were experimented with during training and determined by the GSRs to be the most useful tools during analysis. The specific light box used in this study was an Impact Photographic Display System, PACS Design, Los Altos, CA 94022 (exhibit 22). The actual backlit display area measured 8.5" x 6", and the cost was approximately $30.00 each. The magnification lamp used in this study was a generic jeweler/ watchmaker's portable 5X magnification lamp (exhibit 23). The dimensions are approximately 5" x 6" with a 3.5" diameter lens. These can be found on the Internet in a great variety for approximately $30.00 a piece.

The listing of vendor names is not a product endorsement, but rather a mechanism to illustrate what is available at lower cost. The illumination trend in these units seems to be moving away from fluorescent illumination and changing to LED optics. The use of these LED

	
Exhibit 22: Light Box	Exhibit 23: 5X Magnification Lamp

optics and its effect on physical matching is unknown at this time.

Procedures

The case report used during the experimental trials was created by the GSRs following training. See appendix D for a copy of the case report used. An overall procedure for conducting duct tape end matching analysis is as follows:

- Record date and envelope number.
- Open envelope and record exemplar and unknown tape numbers.
- Observe the class characteristics of the duct tape: tape type, color, backing material, etc. In this study all exemplar samples were paired with unknown samples of the same tape type and the analysts were aware of this prior to the experiment; therefore, there was no need for comparison of class characteristics.
- Place the exemplar sample on the left and the unknown sample on the right of the light box.
- Observe overlaps and/or gaps and record on the protocol form. Observe overall edge alignment and record on the protocol form.
- Flip both tape samples and observe the scrim pattern. Record scrim alignment on the protocol form. Use magnification lamp if needed. If the tape samples have matching angle and scrim alignments and are believed to be a match, then measure tape width and the length of matching distance along edge. Calculate the percent match from these measurements.
- Make final determination if the tape pair is a match, non-match, or inconclusive.
- Record any notes or observations needed to explain analysis on tape set.

Following analysis by the GSRs, all results were compared by an individual outside the study and characterized as correct, false-positive, false-negative, or inconclusive. The implications for each of these results are as follows:

- Correct responses indicate the analyst either reported a match on a true matching pair or the analyst reported a non-match on a true non-matching pair. In real case work, a correct response may lead to clearing an innocent individual or the conviction of a criminal.
- False-positive responses indicate the analyst reported a match on a true non-matching pair. This is considered a Type I error and is the more serious type of error in forensic science; in real case work, a false-positive response could lead to the conviction of an innocent person.
- False-negative responses indicate the analyst reported a non-match on a true matching pair. This is considered a Type II error. In real case work, a false-negative response may lead to the misdirection of an investigation. Although Type II errors are unfortunate, they are less serious than a Type I error.
- Inconclusive responses indicate the analyst determined there were not enough characteristics present along the torn edge to conclusively report a match or non-match. In real case work, an inconclusive response is not considered an error but rather is a perfectly acceptable response which speaks to the quality of the sample examined.

Data Quality

Data quality was evaluated using frequency tabulations to check for incorrectly coded responses or extraneous white space, range checks for numeric values, and complete data listings as well as listings of inconclusive results to check for other anomalies. Exemplar and evidential labels entered by GSRs were compared with assigned labels to evaluate clerical error rates.

Statistical Analysis

For each GSR, and for each combination of factor levels, data was presented using contingency tables displaying frequency and percentages of true match or true non-match versus results reported as inconclusive, non-match, or match. Also for each GSR, and for each combination of factor levels, the following performance statistics were computed:

- The inconclusive rate was computed as the percentage of inconclusive results out of all true matching and non-matching pairs.
- The accuracy rate was computed as the percentage of correctly reported matches or non-matches out of all reported matching and non-matching pairs, excluding reported inconclusive pairs.
- The false-positive rate was computed as the percentage of incorrectly reported matches out of true non-matching pairs only.
- The false-negative rate was computed as the percentage of incorrectly reported non-matches out of true matching pairs only.

Tabulations of samples that were misidentified or reported as inconclusive by two or more GSRs were used to evaluate the possible results of a peer review process applied to the end matching protocol. To provide an assessment of performance across GSRs for each combination of factor levels, each performance statistic was summarized over GSRs using mean and standard deviation. Formal statistical methods were not applied due to technical issues arising from the very low error rates observed in the study.

Results

Study Conduct

Protocol Deviations

The GSRs participating in this study established individual interpretations of several steps of the protocol. These differences among GSR procedure may account for differences in their results and are noted here:

- GSR #1 and GSR #3 used a ruler along the top of the tape to level the two samples when needed (especially during the scissor cut trial).
- GSR #1 did not use the magnification lamp for the hand torn or Elmendorf torn sets, but did use it for the scissor cut and box cutter knife cut sets.
- GSR #2 marked tape sets inconclusive if unsure, waited until the next day, and re-examined all of the recorded inconclusive tape sets.
- GSR #2 reported a non-match if discrepancies in scrim or angle alignment could not be explained by distortion.
- GSR #3 measured the match length and calculated percent match for reported matches only.
- GSR #3 examined each tape pair in a single set of 50 envelopes and if unsure about any would set those pairs aside and re-examine them at the end.

Clerical Errors

A variety of clerical errors were noted upon comparison of the case report data with the envelope assignment records. These errors included entering incorrect envelope numbers, incorrect exemplar numbers, or incorrect evidential numbers, as well as one case of a missing identification.

Specifically:

- GSR #1 had 14 clerical errors during the hand torn trial, one clerical error during the Elmendorf torn trial, and two clerical errors during the scissor cut trial for a total of 17 clerical errors.
- GSR #2 had four clerical errors during the hand torn trial and one clerical error during each of the other three trials for a total of seven clerical errors.
- GSR #3 had 15 clerical errors during the hand torn trial and three clerical errors during the Elmendorf torn trial for a total of 18 clerical errors.
- GSR #2 noted measurements in centimeters instead of millimeters during all four trials, and GSR #3 noted measurements in centimeters instead of millimeters during the hand torn trial and the Elmendorf torn trial.
- GSR #3 gave no result (a blank) for one sample pair in the Elmendorf torn trial.

All clerical errors were corrected before statistical analysis was performed.

Hand Torn Tape Study

Overall Performance

Exhibit 24 displays performance over all hand torn samples for all GSRs. GSR #1 successfully matched 1542 of 1600 tape sets, correctly identifying 794 of 800 non-matching pairs (99.25%) and 748 of 800 matching pairs (93.50%). GSR #1 did not successfully match 13 of 1600 tape sets resulting in one false-positive identification and 12 false-negative identifications. Forty-five of the 1600 pairs were reported as inconclusive. Of the 45 inconclusive sets, five were true non-matching pairs and 40 were true matching pairs.

GSR #2 successfully matched 1555 of 1600 tape sets, correctly identifying 797 of 800 non-matching pairs (99.25%) and 758 of 800 matching pairs (93.50%). GSR #2 did not successfully match seven of 1600 tape sets resulting in no false-positive identifications and seven false-negative identifications. Thirty-eight of the 1600 pairs were reported as inconclusive. Of the 38 inconclusive sets, three were true non-matching pairs and 35 were true matching pairs.

GSR #3 successfully matched 1584 of 1600 tape sets, correctly identifying 790 of 800 non-matching pairs (98.75%) and 794 of 800 matching pairs (99.25%). GSR #3 did not successfully match six of 1600 tape sets resulting in five false-positive identifications and one false-negative identification. Ten of the 1600 pairs were reported as inconclusive. Of the 10 inconclusive sets, five were true non-matching pairs and five were true matching pairs.

Exhibit 24: GSR Performance over All Hand Torn Samples

	Reported Non-match	Reported Match	Reported Inconclusive	Total
GSR #1:				
True Non-match	794 (99.25%)	1 (0.13%)	5 (0.63%)	800
True Match	12 (1.50%)	748 (93.50%)	40 (5.00%)	800
GSR #2:				
True Non-match	797 (99.63%)	0 (0.00%)	3 (0.38%)	800
True Match	7 (0.88%)	758 (94.75%)	35 (4.38%)	800
GSR #3:				
True Non-match	790 (98.75%)	5 (0.63%)	5 (0.63%)	800
True Match	1 (0.13%)	794 (99.25%)	5 (0.63%)	800

Performance by Tape Type

Exhibit 25: GSR #1 Performance on Hand Torn Samples by Tape Type

	Reported Non-match	Reported Match	Reported Inconclusive	Total
NGB				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	95 (95.00%)	5 (5.00%)	100
NGG				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	97 (97.00%)	3 (3.00%)	100
NPB				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	3 (3.00%)	86 (86.00%)	11 (11.00%)	100
NPG				
True Non-match	99 (99.00%)	1 (1.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100
3MGB				
True Non-match	98 (98.00%)	0 (0.00%)	2 (2.00%)	100
True Match	4 (4.00%)	90 (90.00%)	6 (6.00%)	100
3MGG				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	5 (5.00%)	80 (80.00%)	15 (15.00%)	100
3MPB				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100
3MPG				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100

Exhibit 25 displays GSR #1's performance by tape type. GSR #1 correctly identified at least 98.00 percent of non-matching pairs and a minimum of 80.00 percent of all matching pairs for each tape type. No misidentifications were reported for tapes NGB, NGG, 3MPB, or 3MPG. All false-negative identifications occurred with tapes NPB, 3MGB, and 3MGG, and one false-positive identification occurred with tape NPG. Tapes NPB, 3MGB, and 3MGG had the largest number of inconclusive pairs. No inconclusive pairs were reported for tapes NPG or 3MPG.

Exhibit 26: GSR #2 Performance on Hand Torn Samples by Tape Type

	Reported Non-match	**Reported Match**	**Reported Inconclusive**	**Total**
NGB				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	0 (0.00%)	94 (94.00%)	6 (6.00%)	100
NGG				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	1 (1.00%)	99 (99.00%)	0 (0.00%)	100
NPB				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	1 (1.00%)	88 (88.00%)	11 (11.00%)	100
NPG				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	99 (99.00%)	1 (1.00%)	100
3MGB				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	3 (3.00%)	91 (91.00%)	6 (6.00%)	100
3MGG				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	1 (1.00%)	88 (88.00%)	11 (11.00%)	100
3MPB				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	1 (1.00%)	99 (99.00%)	0 (0.00%)	100
3MPG				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100

Exhibit 26 displays GSR #2's performance by tape type. GSR #2 correctly identified at least 99.00 percent of non-matching pairs and a minimum of 88.00 percent of all matching pairs for each tape type. No misidentifications were reported for tapes NGB, NPG, or 3MPG. Tapes NGG, NPB, 3MGG, and 3MPB each had one false-negative identification, and tape 3MGB had three false-negative identifications. No false-positive identifications were reported. Tapes NGB, NPB, 3MGB, and 3MGG had the largest number inconclusive pairs. No inconclusive pairs were reported for tapes NGG, 3MPB, or 3MPG.

Exhibit 27: GSR #3 Performance on Hand Torn Samples by Tape Type

	Reported Non-match	Reported Match	Reported Inconclusive	Total
NGB				
True Non-match	98 (98.00%)	0 (0.00%)	2 (2.00%)	100
True Match	0 (0.00%)	99 (99.00%)	1 (1.00%)	100
NGG				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100
NPB				
True Non-match	99 (99.00%)	1 (1.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100
NPG				
True Non-match	99 (99.00%)	1 (1.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100
3MGB				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	97 (97.00%)	3 (3.00%)	100
3MGG				
True Non-match	96 (96.00%)	1 (1.00%)	3 (3.00%)	100
True Match	1 (1.00%)	98 (98.00%)	1 (1.00%)	100
3MPB				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100
3MPG				
True Non-match	98 (98.00%)	2 (2.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100

Exhibit 27 displays GSR #3's performance by tape type. GSR #3 correctly identified at least 96.00 percent of non-matching pairs and a minimum of 97.00 percent of all matching pairs for each tape type. No misidentifications were reported for tapes NGB, NGG, 3MGB, or 3MPB. The one false-negative identification occurred with tape 3MGG. Tapes NPB, NPG, and 3MGG each had one false-positive identification, and 2 false-positive identifications were reported with tape 3MPG. Tapes NGB, 3MGB, and 3MGG contained all of the reported inconclusive pairs.

Performance Statistics

Exhibit 28 displays the statistical performance for GSR #1. Over all samples, GSR #1 had an inconclusive rate of 22.50 percent, a false-positive rate of 0.13 percent, a false-negative rate of 1.58 percent, and an accuracy rate of 99.16 percent. GSR #1's accuracy ranged from 97.28 to 100.00 percent over the eight tape types and was 100.00 percent for four of the eight tape types.

Exhibit 28: GSR #1 Performance Statistics on Hand Torn Samples

	Inconclusive Rate (%)	False Positive Rate (%)	False Negative Rate (%)	Accuracy (%)
All Samples	22.50	0.13	1.58	99.16
NGB	2.50	0.00	0.00	100.00
NGG	1.50	0.00	0.00	100.00
NPB	6.00	0.00	3.37	98.40
NPG	0.00	1.00	0.00	99.50
3MGB	4.00	0.00	4.26	97.92
3MGG	8.00	0.00	5.88	97.28
3MPB	0.50	0.00	0.00	100.00
3MPG	0.00	0.00	0.00	100.00

Exhibit 29 displays the statistical performance for GSR #2. Over all samples, GSR #2 had an inconclusive rate of 19.00 percent, a false-positive rate of 0.00 percent, a false-negative rate of 0.92 percent, and an accuracy rate of 99.5 percent. GSR #2's accuracy ranged from 98.45 to 100.00 percent over the eight tape types and was 100.00 percent for three of the eight tape types.

Exhibit 29: GSR #2 Performance Statistics on Hand Torn Samples

	Inconclusive Rate (%)	False Positive Rate (%)	False Negative Rate (%)	Accuracy (%)
All Samples	19.00	0.00	0.92	99.55
NGB	3.50	0.00	0.00	100.00
NGG	0.00	0.00	1.00	99.50
NPB	6.00	0.00	1.12	99.47
NPG	0.50	0.00	0.00	100.00
3MGB	3.50	0.00	3.19	98.45
3MGG	5.50	0.00	1.12	99.47
3MPB	0.00	0.00	1.00	99.50
3MPG	0.00	0.00	0.00	100.00

Exhibit 30 displays the statistical performance for GSR #3. Over all samples, GSR #3 had an inconclusive rate of 5.00 percent, a false-positive rate of 0.63 percent, a false-negative rate of 0.13 percent, and an accuracy of 99.62 percent. GSR #3's accuracy ranged from 98.98 to 100.00 percent over the eight tape types and was 100.00 percent for four of the eight tape types.

Exhibit 30: GSR #3 Performance Statistics on Hand Torn Samples

	Inconclusive Rate (%)	False Positive Rate (%)	False Negative Rate (%)	Accuracy (%)
All Samples	5.00	0.63	0.13	99.62
NGB	1.50	0.00	0.00	100.00
NGG	0.00	0.00	0.00	100.00
NPB	0.00	1.00	0.00	99.50
NPG	0.00	1.00	0.00	99.50
3MGB	1.50	0.00	0.00	100.00
3MGG	1.50	1.03	1.01	98.98
3MPB	0.00	0.00	0.00	100.00
3MPG	0.00	2.00	0.00	99.00

Peer Review Analysis

Exhibit 31 displays tape pairs where two or more of the GSRs did not report a correct identification. Overall 93 pairs of duct tape had an incorrect or inconclusive report by at least one GSR. Of those 93 pairs, 71 pairs had an incorrect or inconclusive report by one GSR, 18 pairs had incorrect or inconclusive reports by two GSRs, and four pairs had incorrect or inconclusive reports by all three GSRs. The 22 tape pairs with multiple incorrect or inconclusive reports were made with tapes NGB, NPB, NPG, 3MGB, and 3MGG. Tapes NPB, 3MGB, and 3MGG had the largest number of multiple incorrect or inconclusive reports. No tape pairs had misidentifications by all three GSRs.

Exhibit 31: False-Positive (FP), False-Negative (FN), or Inconclusive (I) by Two or More GSRs

Tape	Exemplar	Unknown	GSR #1	GSR #2	GSR #3
NGB	E1018	1011	I	I	
NGB	E1122	1024		I	I
NPB	E3015	3141	I	I	
NPB	E3117	3194	I	I	
NPB	E3171	3095	I	I	
NPB	E3181	3022	I	I	
NPG	E4046	4161	FP		FP
3MGB	E5003	5073	I	FN	
3MGB	E5007	5013	I	I	I
3MGB	E5015	5104	I	I	
3MGB	E5022	5161	I	FN	I
3MGB	E5054	5041	FN		I
3MGB	E5106	5148	I	I	
3MGG	E6003	6052	I	I	
3MGG	E6015	6036	FN	I	
3MGG	E6032	6143	FN	I	FN
3MGG	E6038	6133	I	I	
3MGG	E6045	6048	I	I	
3MGG	E6046	6088	I	I	
3MGG	E6079	6045	FN	I	
3MGG	E6093	6074	I		I
3MGG	E6199	6087	FN	FN	I

Difficult Samples

Exhibits 32 and 35 display hand torn tape pairs that were misidentified or reported as inconclusive across the GSRs. Photographs of these tape pairs and their original matching pairs can be viewed in exhibits 33, 34, and 36. Of the 1600 pairs of hand torn tape, four pairs were particularly difficult to evaluate and are examined more closely here. Three of these four pairs were misidentified or reported as inconclusive by all three GSRs, and the other tape pair was reported as a false-positive by two of the three GSRs.

Exhibit 32: Misidentified Hand Torn Non-Matching Tape Pair

Envelope Pair	Envelope Number	Expected Results	GSR Results	Original Specimen
E4046 - 4161	462	Non – match	2 False-positive, 1 Correct	E4046 - 4151

Exhibit 33: Misidentified Hand Torn Non-matching Tape Pair

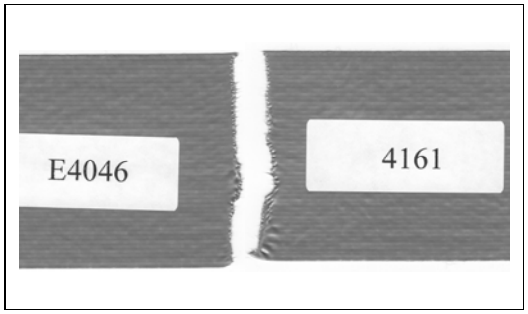


Exhibit 34: Original Matching Pairs – E4046 with 4151 and E4053 with 4161

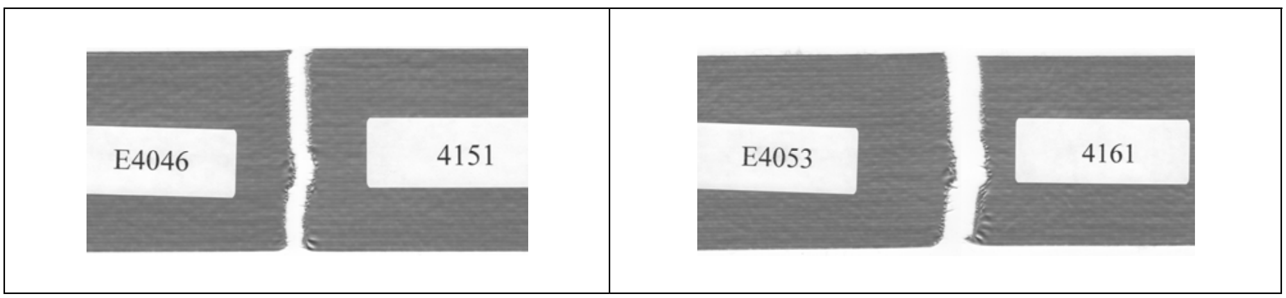
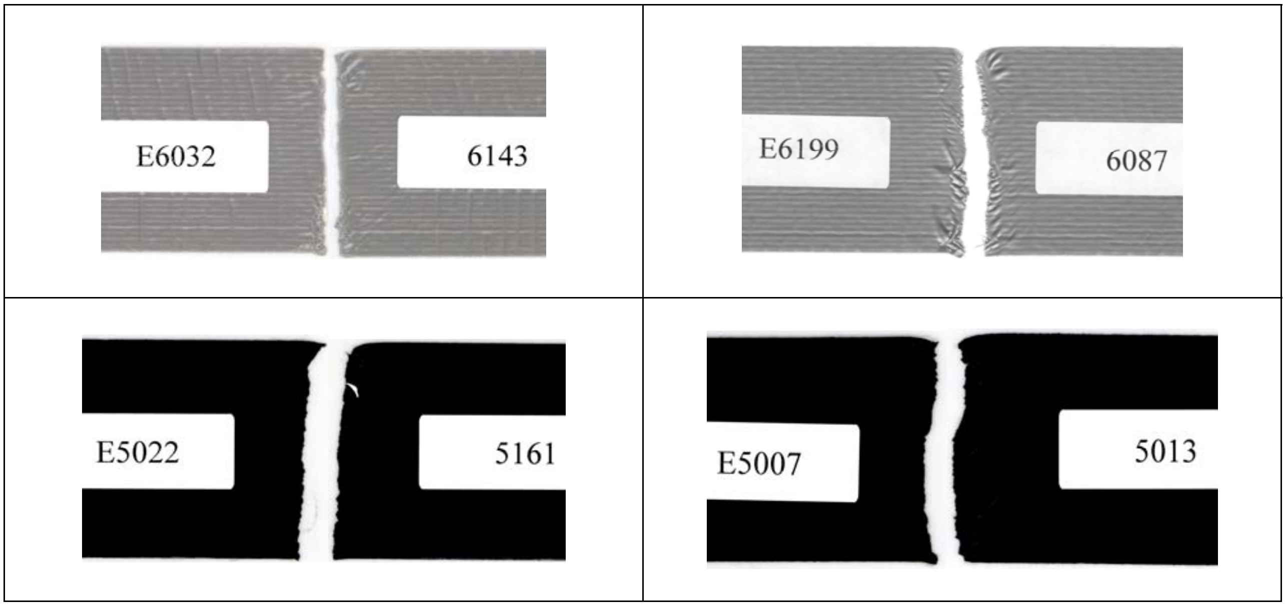


Exhibit 35: Misidentified or Inconclusive by All GSRs: Hand Torn Matching Tape Pairs

Envelope Pair	Envelope Number	Expected Results	GSR Results	Original Specimen
E6032 - 6143	112	Match	2 False-negative, 1 Inconclusive	E6032 - 6143
E6199 - 6087	404	Match	2 False-negative, 1 Inconclusive	E6199 - 6087
E5022 - 5161	954	Match	1 False-negative, 2 Inconclusive	E5022 - 5161
E5007 - 5013	219	Match	3 Inconclusive	E5007 – 5013

Exhibit 36: Misidentified or Inconclusive by All GSRs: Hand Torn Matching Tape Pairs



Elmendorf Torn Tape Study

Overall Performance

Exhibit 37 displays performance over all Elmendorf torn samples for all GSRs. GSR #1 successfully matched 176 of 200 tape sets, correctly identifying 98 of 100 non-matching pairs (98.00%) and 78 of 100 matching pairs (78.00%). GSR #1 did not successfully match one of 200 tape sets resulting in one false-negative identification. Twenty-three of the 200 pairs were reported as inconclusive. Of the 23 reported inconclusive sets, two were true non-matching pairs and 21 were true matching pairs.

GSR #2 successfully matched 194 of 200 tape sets, correctly identifying 100 of 100 non-matching pairs (100.00%) and 94 of 100 matching pairs (94.00%). GSR #2 did not identify one of 200 tape sets resulting in one false-negative identification. Five of the 200 pairs were reported as inconclusive. Of the five inconclusive sets, all five were true matching pairs.

GSR #3 successfully matched 195 of 199 tape sets, correctly identifying 99 of 100 non-matching pairs (99.00%) and 96 of 99 matching pairs (96.97%). GSR #3 did not successfully match one of 199 tape sets resulting in one false-negative identification. Three of the 200 pairs were reported as inconclusive. Of the three inconclusive sets, one was a true non-matching pair and two were true matching pairs. One tape pair was not analyzed giving GSR #3 a total sample size of 199.

Exhibit 37: GSR Performance over All Elmendorf Torn Samples

	Reported Non-match	Reported Match	Reported Inconclusive	Total
GSR #1:				
True Non-match	98 (98.00%)	0 (0.00%)	2 (2.00%)	100
True Match	1 (1.00%)	78 (78.00%)	21 (21.00%)	100
GSR #2:				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	1 (1.00%)	94 (94.00%)	5 (5.00%)	100
GSR #3:				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	1 (1.01%)	96 (96.97%)	2 (2.02%)	99

Performance Statistics

Exhibit 38 displays the Elmendorf tear statistical performance for all GSRs. Across GSRs, the inconclusive rate ranged from 2.00 to 12.00 percent, the false-positive rate was 0.00 percent for all GSRs, the false-negative rate ranged from 1.03 to 1.27 percent, and the accuracy ranged from 99.44 to 99.49 percent.

Exhibit 38: GSR Performance Statistics on Elmendorf Torn Samples

	Inconclusive Rate (%)	False Positive Rate (%)	False Negative Rate (%)	Accuracy (%)
GSR #1	12.00	0.00	1.27	99.44
GSR #2	3.00	0.00	1.05	99.49
GSR #3	2.00	0.00	1.03	99.49

Peer Review Analysis

Exhibit 39 displays tape pairs where two or more of the GSRs did not report a correct identification. Overall 26 pairs of duct tape were reported as incorrect or inconclusive by at least one GSR. Of those 26 pairs, 20 pairs had an incorrect or inconclusive report by one GSR, four pairs had incorrect or inconclusive reports by two GSRs, and two pairs had incorrect or inconclusive reports by all three GSRs.

Exhibit 39: False-Positive (FP), False-Negative (FN), or Inconclusive (I) by Two or More GSRs

Envelope	Exemplar	Unknown	GSR #1	GSR #2	GSR #3
17	E2127	2085	FN	I	I
44	E2016	2017	I	I	
80	E2158	2003	I	I	
108	E2139	2022	I	I	FN
136	E2055	2161	I	FN	
167	E2163	2159		I	I

Difficult Samples

Exhibit 40 displays Elmendorf torn tape pairs that were misidentified or reported as inconclusive across the GSRs. Photographs of these tape pairs can be seen in exhibit 41. Of the 200 pairs of Elmendorf torn tape, two pairs were particularly difficult to evaluate and are examined more closely here. Both of these tape sets were identified incorrectly or reported as inconclusive by all three GSRs.

Exhibit 40: Misidentified or Inconclusive by All GSRs: Elmendorf Torn Matching Tape Pairs

Envelope Pair	Envelope Number	Expected Results	GSR Results	Original Specimen
E2127 - 2085	17	Match	1 False-negative, 2 Inconclusive	E2127 - 2085
E2139 - 2022	404	Match	1 False-negative, 2 Inconclusive	E2139 - 2022

Exhibit 41: Misidentified or Inconclusive by All GSRs: Elmendorf Torn Matching Tape Pairs

Scissor Cut Tape Study

Overall Performance

Exhibit 42 displays performance over all scissor cut samples for all GSRs. GSR #1 successfully matched 197 of 200 tape sets, correctly identifying 99 of 100 non-matching pairs (99.00%) and 98 of 100 matching pairs (98.00%). Three of the 200 pairs were reported as inconclusive. Of the three reported inconclusive sets, one was a true non-matching pair and two were true matching pairs.

GSR #2 successfully matched 189 of 200 tape sets, correctly identifying 92 of 100 non-matching pairs (92.00%) and 97 of 100 matching pairs (97.00%). GSR #2 did not successfully match eight of 200 tape sets resulting in eight false-positive identifications. Three of the 200 pairs were reported as inconclusive. Of the three reported inconclusive sets, all three were true matching pairs.

GSR #3 successfully matched 197 of 200 tape sets, correctly identifying 98 of 100 non-matching pairs (98.00%) and 99 of 100 matching pairs (99.00%). GSR #3 did not successfully match three of 200 tape sets resulting in one false-negative identification and two false-positive identifications. No inconclusive pairs were reported.

Exhibit 42: GSR Performance over All Scissor Cut Samples

	Reported Non-match	Reported Match	Reported Inconclusive	Total
GSR #1:				
True Non-match	99 (99.00%)	0 (0.00%)	1 (1.00%)	100
True Match	0 (0.00%)	98 (98.00%)	2 (2.00%)	100
GSR #2:				
True Non-match	92 (92.00%)	8 (8.00%)	0 (0.00%)	100
True Match	0 (0.00%)	97 (97.00%)	3 (3.00%)	100
GSR #3:				
True Non-match	98 (98.00%)	2 (2.00%)	0 (0.00%)	100
True Match	1 (1.00%)	99 (99.00%)	0 (0.00%)	100

Performance Statistics

Exhibit 43 displays the scissor cut statistical performance all GSRs. Across the GSRs, the inconclusive rate ranging from 0.00 to 1.50 percent, the false-positive rate ranged from 0.00 to 8.00 percent, the false-negative rate ranged from 0.00 to 1.00 percent, and the accuracy ranged from 95.94 to 100.00 percent.

Exhibit 43: GSR Performance Statistics on Scissor Cut Samples

	Inconclusive Rate (%)	False Positive Rate (%)	False Negative Rate (%)	Accuracy (%)
GSR #1	1.50	0.00	0.00	100.00
GSR #2	1.50	8.00	0.00	95.94
GSR #3	0.00	2.00	1.00	98.50

Peer Review Analysis

Exhibit 44 displays tape pairs where two or more of the GSRs did not report a correct identification. Overall 14 pairs of duct tape had an incorrect or inconclusive report by at least one GSR. Of those 14 pairs, 12 pairs had an incorrect or inconclusive report by one GSR, one pair had incorrect or inconclusive report by two GSRs, and one pair had an incorrect or inconclusive report by all three GSRs.

Exhibit 44: False-Positive (FP), False-Negative (FN), or Inconclusive (I) by Two or More GSRs

Envelope	Exemplar	Unknown	GSR #1	GSR #2	GSR #3
103	E3103	3035	I	FP	FP
140	E3167	3023		FP	FP

Difficult Samples

Exhibit 45 displays scissor cut tape pairs that were commonly misidentified across the GSRs. Photographs of these tape pairs and their original matching pairs can be viewed in exhibits 46 through 49. Of the 200 pairs of scissor cut tape, two pairs were particularly difficult to evaluate and are examined more closely here. One tape pair was misidentified or reported as inconclusive by all three GSRs, and the other tape pair was misidentified by two of the three GSRs.

Exhibit 45: Misidentified or Inconclusive Scissor Cut Matching Tape Pairs

Envelope Pair	Envelope Number	Expected Results	GSR Results	Original Specimen
E3103 - 3035	103	Non-match	2 False-positive, 1 Inconclusive	E3103 - 3050
E3167 - 3023	140	Non-match	2 False-positive, 1 correct	E3167 - 3157

Exhibit 46: Misidentified or Inconclusive by All GSRs: Scissor Cut Non-matching Tape Pair

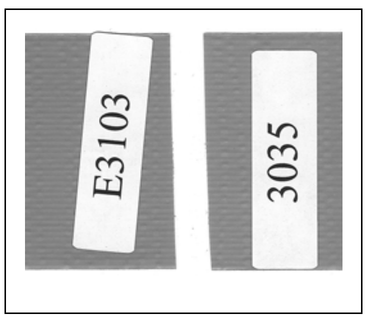


Exhibit 47: Original Matching Pairs – E3103 with 3050 and E3197 with 3035

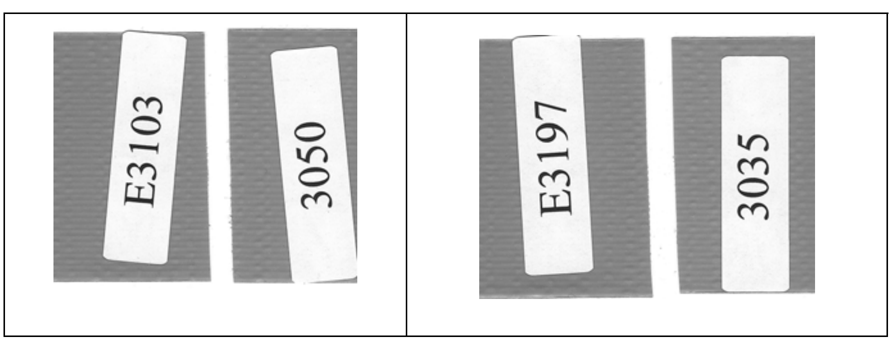


Exhibit 48: Misidentified Scissor Cut Non-matching Tape Pair

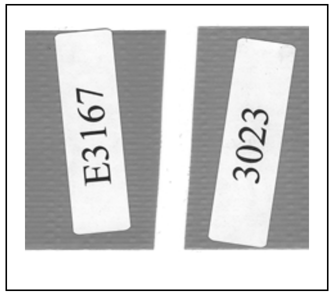
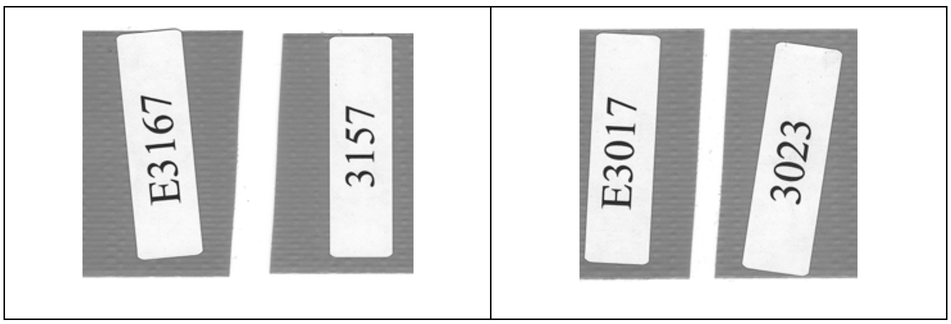


Exhibit 49: Original Matching Pairs – E3167 with 3157 and E3017 with 3023



Box Cutter Knife Cut Tape Study

Overall Performance

Exhibit 50 displays performance over all box cutter knife cut samples for all GSRs. GSR #1 successfully matched 200 of 200 tape sets, correctly identifying 100 of 100 non-matching pairs (100.00%) and 100 of 100 matching pairs (100.00%). There were no reported inconclusive pairs.

GSR #2 successfully matched 199 of 200 tape sets, correctly identifying 100 of 100 non-matching pairs (100.00%) and 99 of 100 matching pairs (99.00%). GSR #2 did not successfully match one of 200 tape sets resulting in one false-negative identification. There were no reported inconclusive pairs.

GSR #3 successfully matched 200 of 200 tape sets, correctly identifying 100 of 100 non-matching pairs (100.00%) and 100 of 100 matching pairs (100.00%). There were no reported inconclusive pairs.

Exhibit 50: GSR Performance over All Box Cutter Knife Cut Samples

	Reported Non-match	Reported Match	Reported Inconclusive	Total
GSR #1:				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100
GSR #2:				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	1 (1.00%)	99 (99.00%)	0 (0.00%)	100
GSR #3:				
True Non-match	100 (100.00%)	0 (0.00%)	0 (0.00%)	100
True Match	0 (0.00%)	100 (100.00%)	0 (0.00%)	100

Performance Statistics:

Exhibit 51 displays the box cutter knife cut statistical performance for all GSRs. Across the GSRs, the inconclusive rate was 0.00 percent for all GSRs, the false-negative rate ranged from 0.00 to 1.00 percent, the false-positive rate was 0.00 percent for all GSRs, and the accuracy ranged from 99.50 to 100.00 percent.

Exhibit 51: GSR Performance Statistics on Box Cutter Knife Samples

	Inconclusive Rate (%)	False Positive Rate (%)	False Negative Rate (%)	Accuracy (%)
GSR #1	0.00	0.00	0.00	100.00
GSR #2	0.00	0.00	1.00	99.50
GSR #3	0.00	0.00	0.00	100.00

Peer Review Analysis/ Difficult Samples

Only one duct tape sample was misidentified by one GSR. None of the samples were deemed difficult for this set.

Combined Performance Statistics

Exhibit 52 summarizes inconclusive rates for all separation methods and tape types across all GSRs. Exhibits 53 and 54 display the same data. The mean GSR inconclusive rates ranged from 0.00 to 5.17 percent with a standard deviation ranging from 0.00 to 5.51 percent. Across the hand torn tape, NGB, NPB, 3MGB, and 3MGG had the highest mean inconclusive rates and standard deviations. The Elmendorf torn 3MGG tape had the same mean inconclusive rate as the hand torn 3MGG tape, but a larger standard deviation. The box cutter knife cut 3MPG tape had the same mean inconclusive rate and standard deviation as the hand torn 3MPG tape, which was lower than the scissor cut 3MPG tape. In general, across all tape samples, GSR #1 had the most inconclusive reports GSR #3 had the fewest inconclusive reports.

Exhibit 52: Combined Inconclusive Rates (%)

Sample	GSR #1	GSR #2	GSR #3	Mean	SD
Hand Torn NGB	2.50	3.50	1.50	2.50	1.00
Hand Torn NGG	1.50	0.00	0.00	0.50	0.87
Hand Torn NPB	6.00	6.00	0.00	4.00	3.46
Hand Torn NPG	0.00	0.50	0.00	0.17	0.29
Hand Torn 3MGB	4.00	3.50	1.50	3.00	1.32
Hand Torn 3MGG	8.00	5.50	2.00	5.17	3.01
Hand Torn 3MPB	0.50	0.00	0.00	0.17	0.29
Hand Torn 3MPG	0.00	0.00	0.00	0.00	0.00
Elmendorf Torn 3MGG	11.50	2.50	1.51	5.17	5.51
Scissors Cut 3MPG	1.50	1.50	0.00	1.00	0.87
Box Cutter Cut 3MPG	0.00	0.00	0.00	0.00	0.00

Exhibit 53: Individual GSR Inconclusive Rates

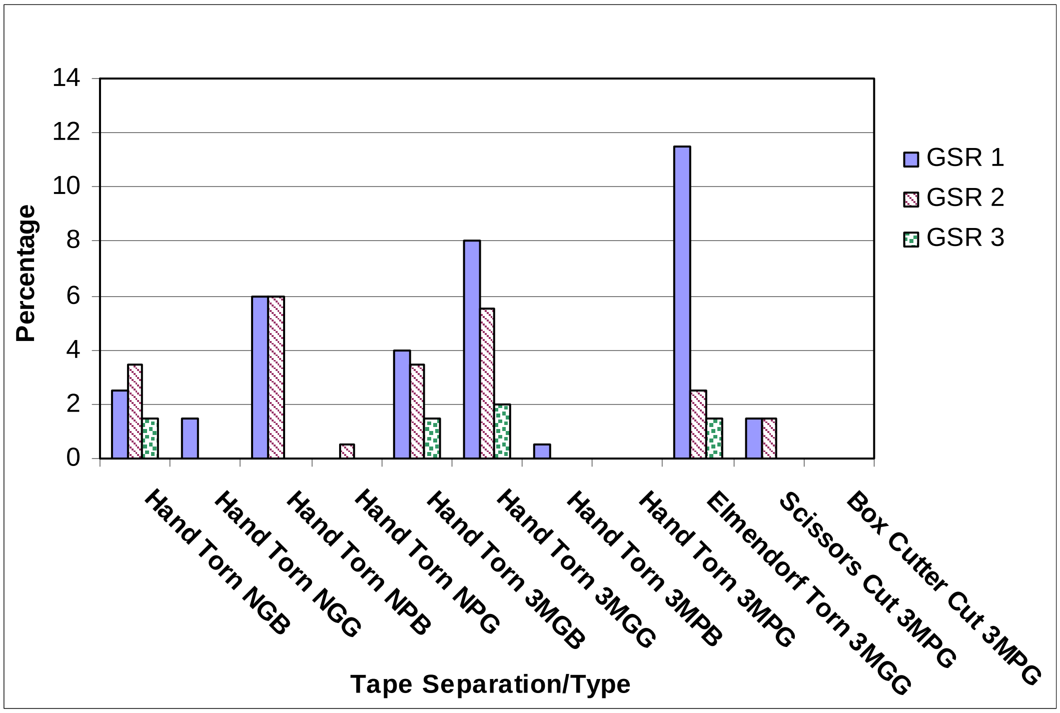


Exhibit 54: Mean GSR Inconclusive Rate

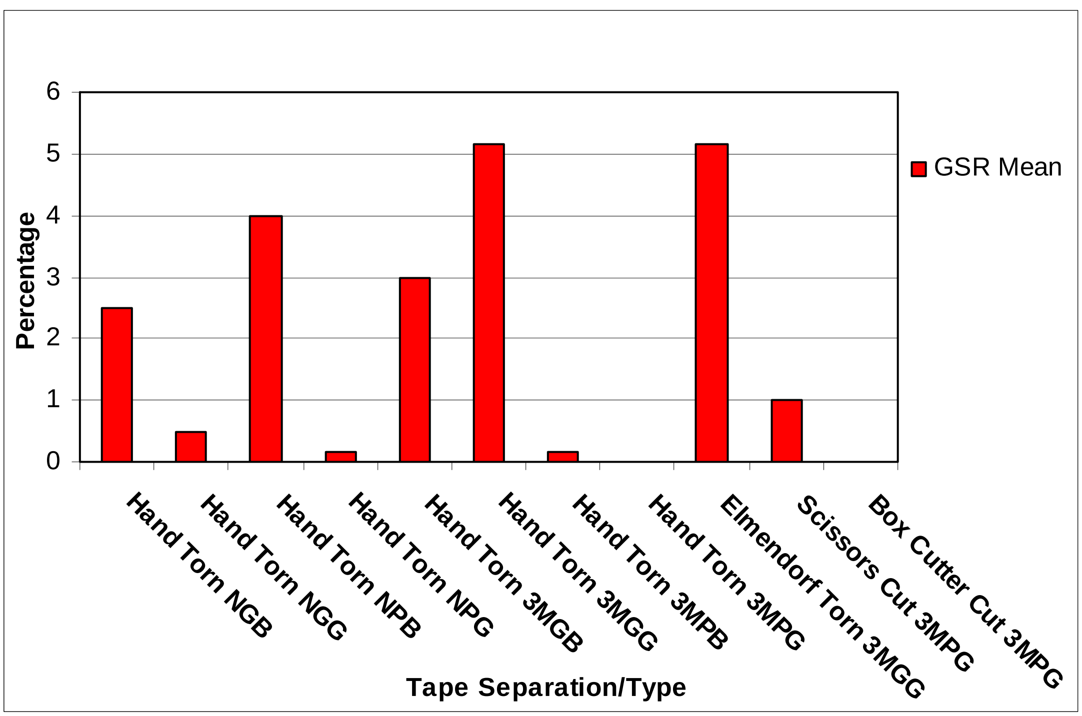


Exhibit 55 summarizes false-positive rates for all separation methods and tape types across all GSRs. Exhibits 56 and 57 display the same data. The mean GSR false-positive rates range from 0.00 to 3.33 percent with a standard deviation ranging from 0.00 to 4.16 percent. In the hand torn tape set, all false-positives were reported with tapes NPB, NPG, 3MGG, and 3MPG. NPG and 3MPG had the same mean false-positive rate, but 3MPG had a slightly higher standard deviation. The Elmendorf torn 3MGG tape had a higher mean false-positive rate and standard deviation than the hand torn 3MGG tape. The box cutter knife cut 3MPG tape had a lower mean false-positive rate and standard deviation than the hand torn 3MPG tape, and the hand torn 3MPG tape had a lower mean false-positive rate and standard deviation than the scissor cut 3MPG tape. In general, across all tape samples, GSR #2 had the most false-positive identifications and GSRs #1 had the least false-positive identifications.

Exhibit 55: Combined False-Positive Rates (%)

Sample	**GSR #1**	**GSR #2**	**GSR #3**	**Mean**	**SD**
Hand Torn NGB	0.00	0.00	0.00	0.00	0.00
Hand Torn NGG	0.00	0.00	0.00	0.00	0.00
Hand Torn NPB	0.00	0.00	1.00	0.33	0.58
Hand Torn NPG	1.00	0.00	1.00	0.67	0.58
Hand Torn 3MGB	0.00	0.00	0.00	0.00	0.00
Hand Torn 3MGG	0.00	0.00	1.03	0.34	0.60
Hand Torn 3MPB	0.00	0.00	0.00	0.00	0.00
Hand Torn 3MPG	0.00	0.00	2.00	0.67	1.15
Elmendorf Torn 3MGG	0.00	0.00	0.00	0.00	0.00
Scissors Cut 3MPG	0.00	8.00	2.00	3.33	4.16
Box Cutter Cut 3MPG	0.00	0.00	0.00	0.00	0.00

Exhibit 56: Individual GSR False-Positive Rates

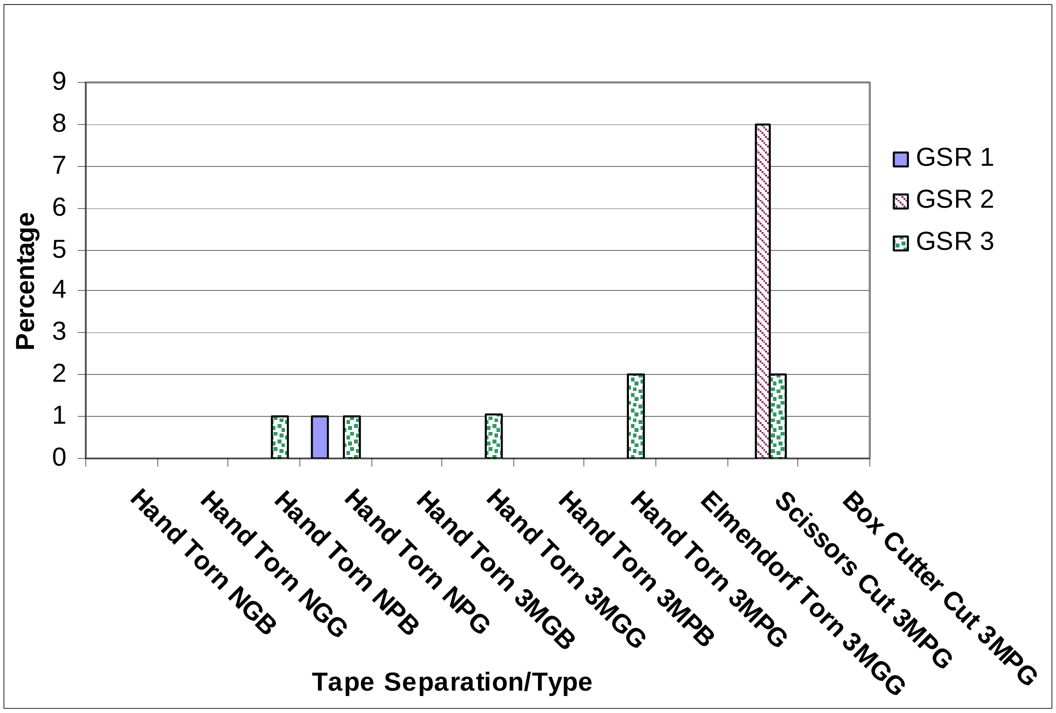


Exhibit 57: Mean GSR False-Positive Rates

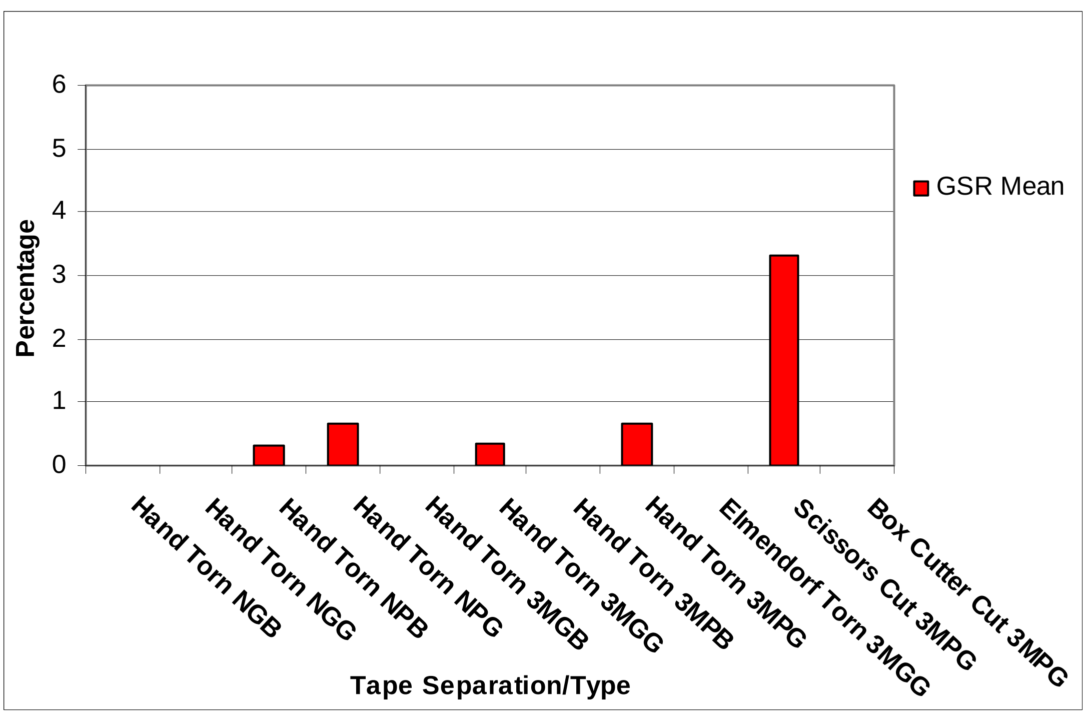


Exhibit 58 summarizes false-negative rates for all separation methods and tape types across all GSRs. Exhibit 59 and 60 display the same data. The mean GSR false-negative rates range from 0.00 to 2.67 percent, with a standard deviation ranging from 0.00 to 2.78 percent. In the hand torn tape set, NPB, 3MGB, and 3MGG had the highest mean false-negative rates and standard deviations. The Elmendorf torn 3MGG tape had a lower mean false-negative rate and standard deviation than the hand torn 3MGG tape. The box cutter knife cut 3MPG tape had the same mean inconclusive rate and standard deviation as the scissor cut 3MPG tape, which was higher than the hand torn 3MPG tape. In general, across all tape samples, GSR #3 had few or no false-negative identifications, while GSR #1 and GSR #2 had higher false-negative rates.

Exhibit 58: Combined False-Negative Rates (%)

Sample	GSR #1	GSR #2	GSR #3	Mean	SD
Hand Torn NGB	0.00	0.00	0.00	0.00	0.00
Hand Torn NGG	0.00	1.00	0.00	0.33	0.58
Hand Torn NPB	3.37	1.12	0.00	1.50	1.72
Hand Torn NPG	0.00	0.00	0.00	0.00	0.00
Hand Torn 3MGB	4.26	3.19	0.00	2.48	2.21
Hand Torn 3MGG	5.88	1.12	1.01	2.67	2.78
Hand Torn 3MPB	0.00	1.00	0.00	0.33	0.58
Hand Torn 3MPG	0.00	0.00	0.00	0.00	0.00
Elmendorf Torn 3MGG	1.27	1.05	1.03	1.12	0.13
Scissors Cut 3MPG	0.00	0.00	1.00	0.33	0.58
Box Cutter Cut 3MPG	0.00	1.00	0.00	0.33	0.58

Exhibit 59: Individual GSR False-Negative Rates

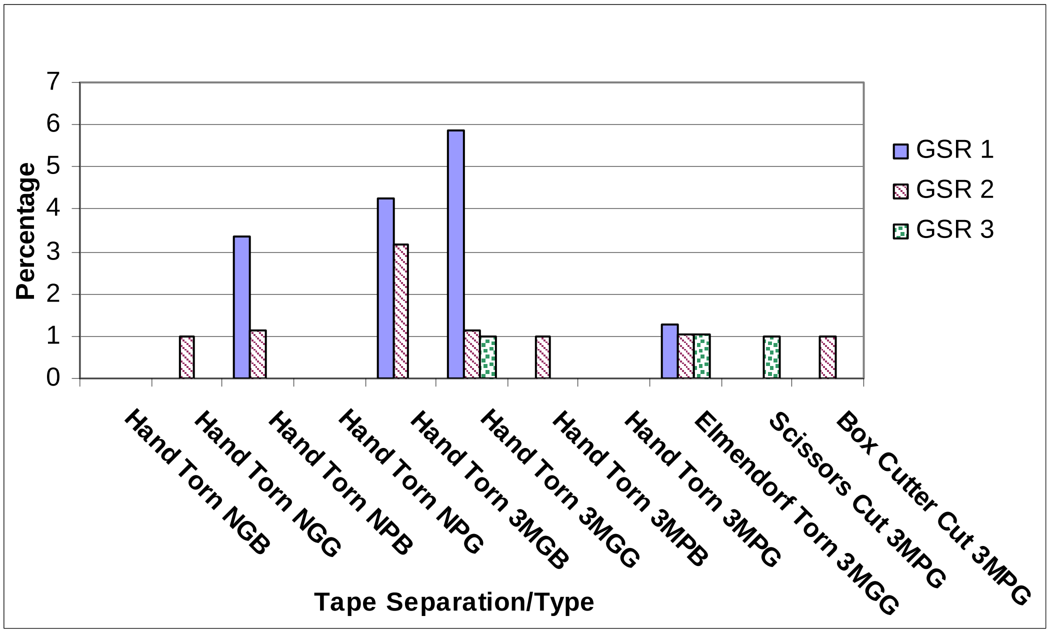


Exhibit 60: Mean GSR False-Negative Rates

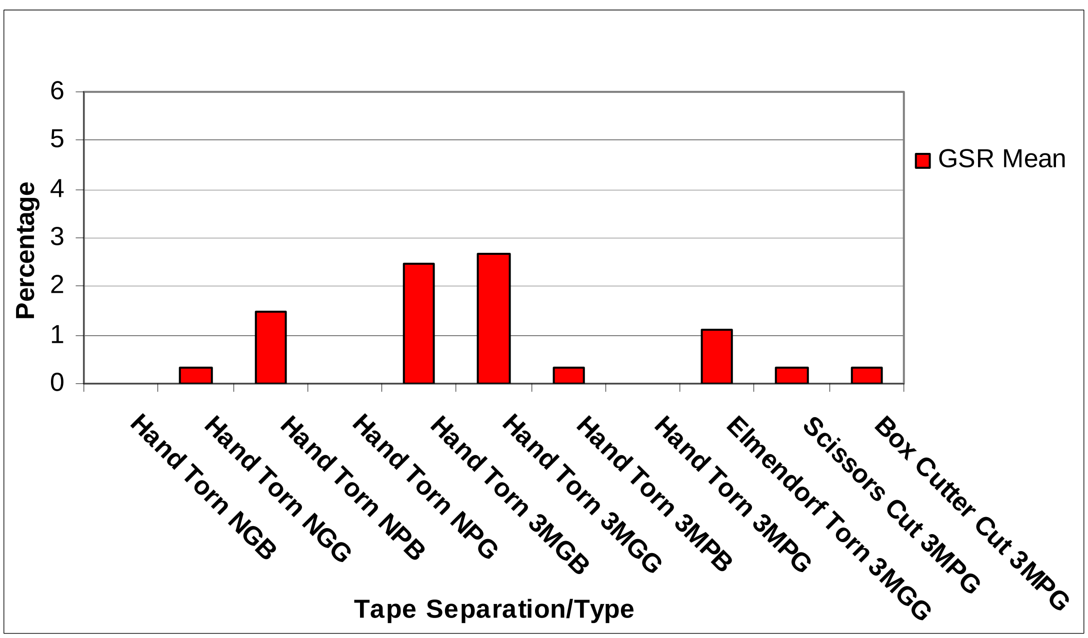


Exhibit 61 summarizes accuracy rates for all separation methods and tape types across all GSRs. Exhibits 62 and 63 display the same data. The mean GSR accuracy rates range from 98.15 to 100.00 percent with a standard deviation ranging from 0.00 to 2.05 percent. In the hand torn tape set, 3MGB and 3MGG were the only tape types to have a mean accuracy below 99.00 percent. 3MGB and 3MGG also had the largest standard deviation. The Elmendorf torn 3MGG tape had a higher mean accuracy rate and standard deviation than the hand torn 3MGG tape. The box cutter knife cut 3MPG tape had a higher mean accuracy rate and standard deviation than the hand torn 3MPG tape, and the hand torn 3MPG tape had a higher mean accuracy rate and standard deviation than the scissor cut 3MPG tape. In general, all three GSRs had high accuracy rates across all tape samples.

Exhibit 61: Combined Accuracy Rates (%)

Sample	GSR #1	GSR #2	GSR #3	Mean	SD
Hand Torn NGB	100.00	100.00	100.00	100.00	0.00
Hand Torn NGG	100.00	99.50	100.00	99.83	0.29
Hand Torn NPB	98.40	99.47	99.50	99.12	0.63
Hand Torn NPG	99.50	100.00	99.50	99.67	0.29
Hand Torn 3MGB	97.92	98.45	100.00	98.79	1.08
Hand Torn 3MGG	97.28	99.47	98.98	98.58	1.15
Hand Torn 3MPB	100.00	99.50	100.00	99.83	0.29
Hand Torn 3MPG	100.00	100.00	99.00	99.67	0.58
Elmendorf Torn 3MGG	99.44	99.49	99.49	99.47	0.03
Scissors Cut 3MPG	100.00	95.94	98.50	98.15	2.05
Box Cutter Cut 3MPG	100.00	99.50	100.00	99.83	0.29

Exhibit 62: Individual GSR Accuracy Rates

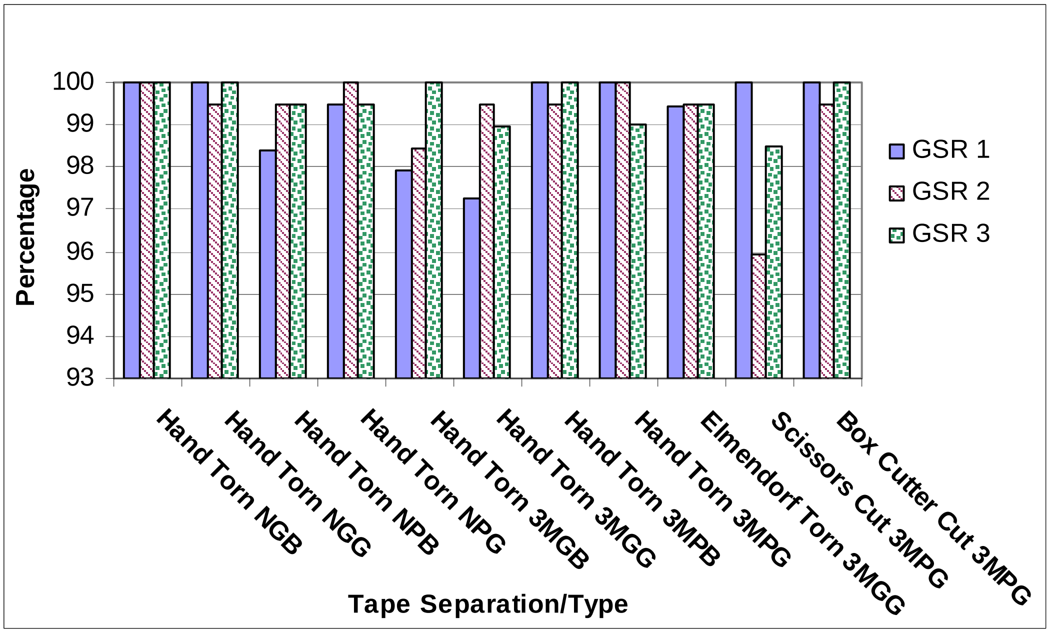
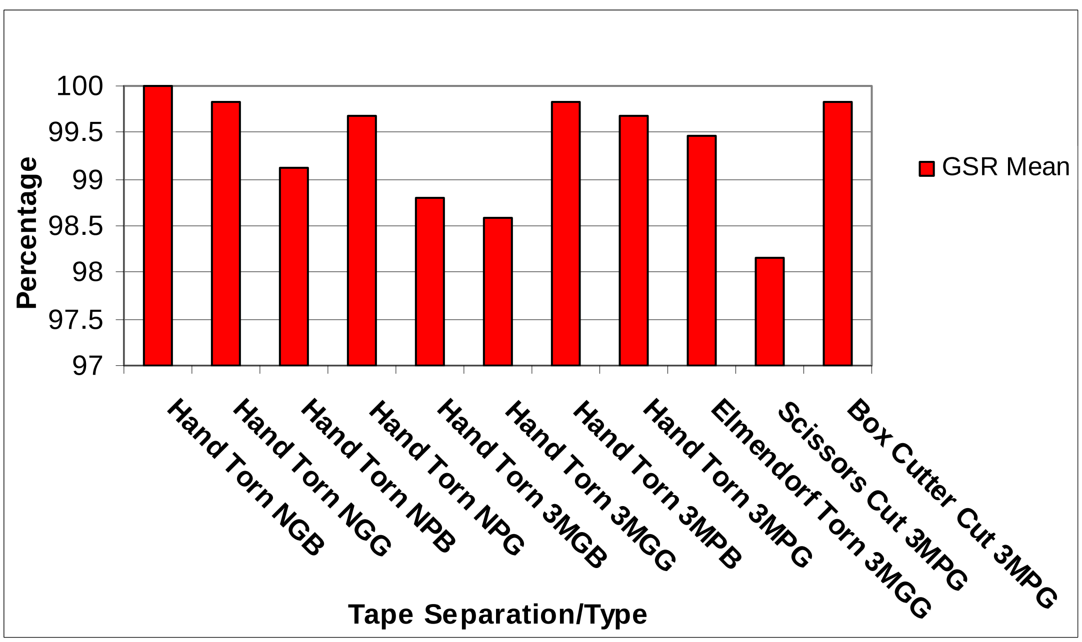


Exhibit 63: Mean GSR Accuracy Rates



Conclusions

Overall Performance

This study confirms that it is possible to use physical end matching to identify duct tape samples as matching or non-matching and that differences between analysts, brands, tape grades, tape color, and method of separation have varying contributions to misidentifications and inconclusive results. The mean GSR accuracy observed ranged from 98.58 to 100 percent for torn tape and from 98.15 to 99.83 percent for cut tape. The mean GSR false-positive rate ranged from 0.00 to 0.67 percent for torn tape and from 0.00 to 3.33 percent for cut tape. The mean GSR false-negative rate ranged from 0.00 to 2.67 percent for torn tape and was 0.33 percent for both types of cut tape.

Overall the three GSRs were consistent in their analyses. Given that an identification was made, all three GSRs showed an accuracy of 95.94 percent or higher for all tape types and separation methods. Error rates were low, and many of the errors that did occur were seen with similar tape types or on similar tape samples. In the study, two analysts were more likely to report inconclusive results as compared with the third analyst. Reporting inconclusive results on difficult samples likely reduced the false-positive error rate, but at the price of reducing the percentage of correctly identified true matches and possibly increasing the false-negative error rate. False-positive errors are more serious in forensic science than false-negative errors; therefore, reducing the number of false-positive errors is worth the expense of possibly increasing false-negative errors or reducing the number of correctly identified true matches.

This research focused on two common brands, Nashua and 3M. In this study, 3M tapes had a higher percentage of misidentifications and inconclusive reports than Nashua tapes, although the percentage of false-positive identifications was fairly even between the brands.

Overall results showed the 3M general grade tapes contributed to a majority of the inconclusive and false-negative identifications. As such, the combination of brand and grade likely interacted to affect an analyst's ability to correctly identify duct tape end matches.

General grade tapes had a higher percentage of overall misidentifications and inconclusive results than professional grade tapes. This is probably due to the structural differences between the two grades, where professional grade tapes are stronger and less prone to distorting or stretching upon separation. It is likely that the higher level of stretch and distortion during the separation of general grade tapes caused the increased percentage of false-negative identifications and inconclusive results, while the lack of stretch and distortion during separation of professional grade tapes caused the increased percentage of false-positive identifications. It is probable that tape grade affects an analyst's ability to conduct these examinations.

The total number of misidentifications in this study was similar for both tape colors. The black tapes gave a slightly higher percentage of false-negative identifications, while the gray tapes gave a higher rate of false-positive identifications. The black tapes gave a higher percentage of inconclusive results. The inconclusive results on black tapes were spread across the various brands and grades, while the majority of the inconclusive identifications of gray tapes were with 3MGG tape. The increase in inconclusive results with black tapes across all brands and grades may be occurring because the scrim is more difficult to examine in darker tapes, so if an analyst is suspicious of a match or non-match but cannot confirm scrim alignment they are likely to report an inconclusive result. In general, brand and grade of tape had more of an effect than color on an analyst's ability to correctly identify duct tape end matches.

When comparing separation methods, the Elmendorf torn tape results were compared with the 3M general grade gray hand torn tape results. Overall, the Elmendorf torn tape pairs and the hand torn tape pairs gave similar results. The hand torn tapes had a slightly higher percentage of false-positive and false-negative identifications, and the total amount of reported inconclusive results was equal for both separation methods. Consistent tearing conditions do not appear to have a significant affect on an analyst's ability to correctly analyze duct tape physical end matches.

The results for scissor cut tape and tape cut with a box cutter knife were compared with the 3M professional grade gray hand torn tape results. The scissor cut tape had a higher percentage of inconclusive results and false-positive identifications compared with the hand torn tapes, but the hand torn tapes had a slightly higher percentage of false-negative identifications. It should be noted that a majority of the false-positive identifications on the scissor cut tapes were made by a single analyst who reported no false-positive identifications with any other tape types or separation methods. The tapes cut with a box cutter knife had an equal number of inconclusive results compared with the hand torn tapes, but the hand torn tapes had a slightly higher percentage of both false-positive and false-negative identifications. In general, tape cut with scissors had a greater affect on an analyst's ability to correctly analyze duct tape physical end matches than hand torn tape. More research should be conducted in this area to confirm this conclusion considering the high percentage of false-positive identifications by a single analyst may have skewed the overall results. Tape cut with a box cutter knife did not appear to have a significant affect on an analyst's ability to correctly analyze duct tape physical end matches as compared with hand torn tape.

Peer Review

This study demonstrates that peer review is a necessary step in duct tape analysis and can greatly reduce the number of misidentifications made by analysts. With every separation method there were misidentifications made by one or more analysts, but more likely than not those misidentifications would have been noticed and corrected by the remaining analyst(s). Conducting independent, blind peer review is strongly encouraged on all duct tape samples undergoing physical end matching analysis, and it is recommended that a minimum of two trained analysts perform this peer review.

Implications for Policy and Practice

The following recommendations have been developed for analysts conducting duct tape end matching:

- Report inconclusive results on difficult to match samples to avoid false-positive results.
- Conduct examinations using procedures similar to the ones suggested here unless new procedures can be assessed and shown to provide higher accuracy.
- Use a light box and illuminated magnification lamp or stereo microscope to enhance end match analysis and minimize the need to remove adhesive from the tape for further scrim analysis.
- Training for duct tape analysts should include literature review, difficult practice sets, and completion of a competency exam (see appendix A) before analysts performing duct tape end matching on actual casework.

Limitations

There are many more brands available to consumers, and the brands available vary by region. The brands studied were not chosen at random, as a random sample of two brands is not sufficient to allow overly broad generalization. Also, a random sample of brands might not be representative if some brands have much larger market share. If the study had shown very little difference in error rates, the question would still be open as to the potential effects of brand. However, the fact that there were indications of an interaction of brand and grade on error rates is suggestive though not conclusive.

Similarly, industrial grade tape was not studied, nor the many other colors other than gray and black. It is noted that not all combinations of brand, grade, and color are available, further complicating study design and interpretation.

This study did not examine the effects of within-roll factors such as position within the roll, the side on which the tear was started, nor possible variations in identifying features within the roll.

The different separation methods were studied on two of the same rolls as that used in the hand torn sample. Also, the samples were matched in secondary designs from the primary hand torn design. This introduced a blocking factor of time that is confounded with separation method. In theory, this factor could also interact with the analyst factor, to further complicate interpretation.

Three analysts were studied. All analysts were GSRs in the same program, and all received the same training. This feature of the study helped control confounding many factors but may have limited the ability to generalize the results to the general population of analysts. Results found in this study probably represent a lower bound on the variation to be expected.

The sample size chosen was a trade-off between the large sample sizes needed to estimate small error rates and the small sample sizes needed to practically implement a useful design. With a sample size of 100, for cases where no errors were noted, the usual upper 95 percent confidence limit could be obtained as approximately 3 percent (Louis 1981). To accurately estimate smaller error rates requires much larger sample sizes.

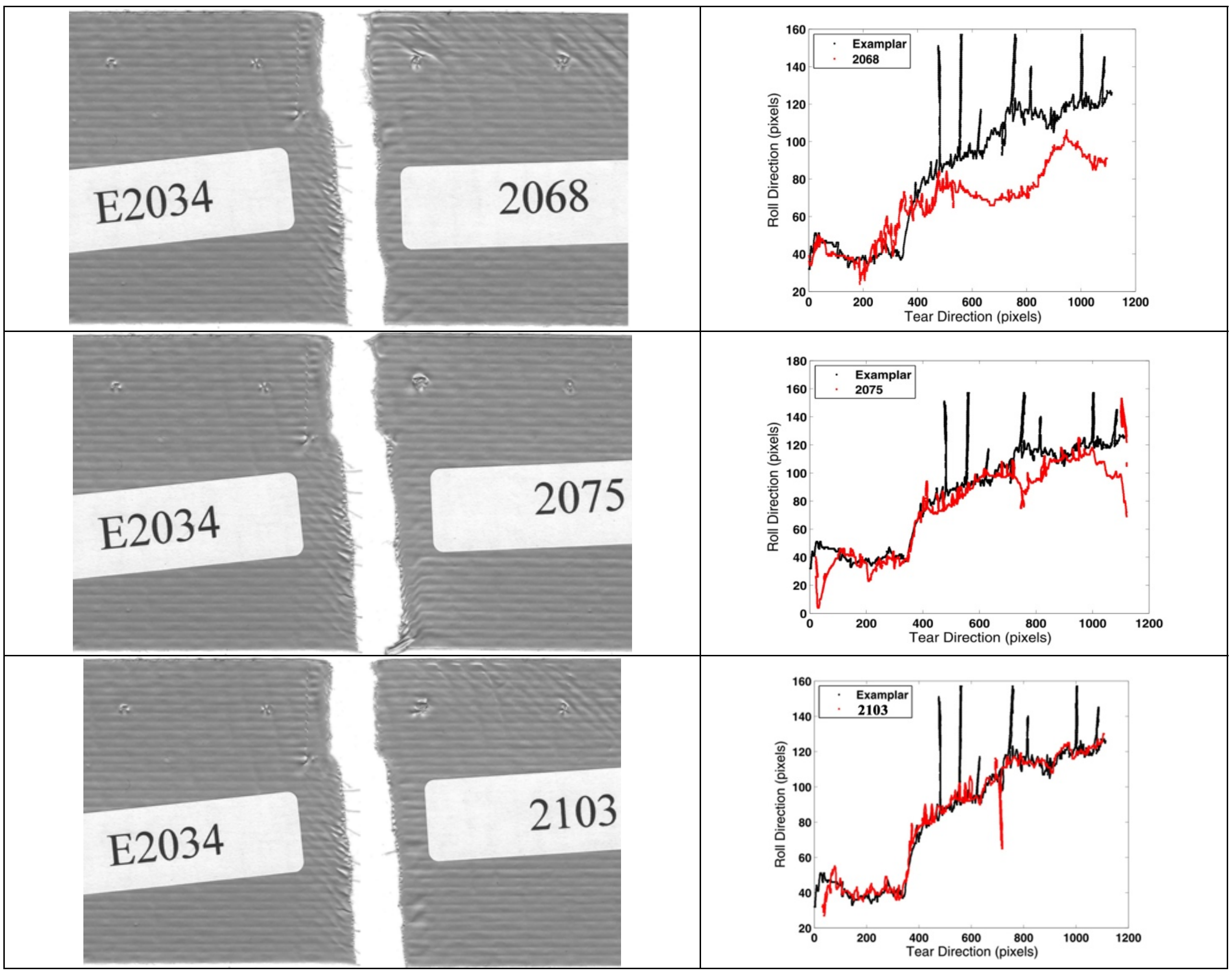
Formal statistical analysis was not performed due to limitations arising from the distribution of the data. In essence, the error rates were too low to permit straightforward application of statistical analysis methods such as linear models or generalized linear models without introducing a variety of assumptions about the results. Other informal analyses could be considered, but are expected to produce similar interpretations.

Implications for Further Research

It is recommended that studies be conducted using a larger number of analysts, preferably analysts from various backgrounds and experiences. It is also recommended that a larger population of duct tape brands be included in future studies in combination with various separation methods to fully evaluate which of the factors (brand, tape grade, or separation method) have the greatest effects on analyst accuracy and error rates. It is reasonable to restrict attention to a particular grade, though if other factors were considered they should at least be incorporated as nuisance factors. If main effects only are of interest, there are a variety of so-called screening designs that could be considered.

Finally, it is recommended that future studies be conducted on duct tape end matching using mathematical and computerized methods of analysis and that those results be compared with this study to determine the reliability and reproducibility of computerized analysis versus human analysis. This study provides an assessment of analyst performance: however, there still is the issue of being able to quantitatively characterize duct tape end matches. End matches and physical fractures in general do not lend themselves to discreet points of comparison as in the case of latent prints, nor can we use the consecutive matching striae (CMS) as used in toolmark/firearms identification. There is a high probability of a match when the CMS count exceeds four in multiple areas. However, the CMS method is not suitable for duct tape end matches. We recommend that future studies be conducted on duct tape end matching using quantitative image analysis software with a mathematical algorithm that can profile the tape ends. Using such an algorithm one could assess the results of this study to determine the reliability and reproducibility of computerized analysis versus human analysis. The main goal for such an automated analysis would be to profile the edges of torn or cut duct tape specimen thus providing a graphical method for further comparison. Once an analyst has this profile, automated comparisons can be performed objectively on the two ends of a duct tape using residuals analysis, or fitting errors. These residuals act as an observable estimate of the unobservable statistical error. In exhibit 64, Exemplar E2034 is compared with three possible candidate duct tapes. Associated with each image is a graphical depiction of the edge of the torn duct tape. This was accomplished by converting all tape images to a binary image and then converting the torn edge profile of these tape segments into graphical format suitable for mathematical comparison. In this case, specimens 2068, 2075, and 2103 are candidates for a match. The mathematical results illustrate that the best fitting match to the exemplar E2034 is unknown 2103. The spikes in these lower graphical figures are scrim yarns from E2034 and, for demonstration purposes, have been ignored in the calculations. In a real comparison situation, these yarns would be fitted on the reverse side of the corresponding duct tape specimens. The next phase would involve the residuals analysis, which would look at the sum of the squares of the observable error. We would expect that the true matching duct tape would have the lowest residuals error.

Exhibit 64: Illustration of the application of a software algorithm to process and convert the duct tape images.



References

Agron, N. and B. Schecter. "Physical Comparisons and Some Characteristics of Electrical Tape." *AFTE Journal* 18 (1986): 53─59.

Bradley M.J., R.L. Keagy, P.C. Lowe, M.P. Rickenback, D.M. Wright, and M.A. LeBeau. "A Validation Study for Duct Tape End Matches." *Journal of Forensic Sciences* 51(May 2006): 504-508.

Claytor L.K. and A.L. Davis. "A Validation of Fracture Matching Through the Microscopic Examination of the Fractured Surfaces of Hacksaw Blades." *AFTE Journal* 42(2010): 323-334.

Johnston, John. *Pressure Sensitive Adhesive Tapes: A Guide to their Function, Design, Manufacture, and Use*. Illinois: Pressure Sensitive Tape Council, 2003.

Louis, T.A. "Confidence intervals for a binomial parameter after observing no successes." *American Statistician* 35(1981): 154.

National Institute of Justice (n.d.). *Glossary*. In *Firearm Examiner Training*. Retrieved from http://www.ojp.usdoj.gov/nij/training/firearms-training/glossary.htm.

National Research Council (2009). *Strengthening Forensic Science in the United States: A Path Forward* (ISBN: 0-309-13131-6). Washington, DC: National Academies Press.

Smith, Jenny M. "Forensic Examination of Pressure Sensitive Tape," in *Forensic Analysis on the Cutting Edge: New Methods for Trace Evidence Analysis*. ed. Robert D. Blackledge. New Jersey: Wiley, 2007: 291-332.

Tsach T., S. Wiesner, Y. Shor. "Empirical Proof of Physical Match: Systematic Research With Tensile Machine". *Forensic Science International* 116(2007): 77-83.

Tulleners F. and K. Chan. "The Statistical Evaluation of Duct Tape Fracture Match as Physical Evidence." Paper presented to the American Academy of Forensic Sciences, Seattle, February 2010.

Dissemination of Research Findings

American Academy of Forensic Sciences 63rd Annual Scientific Meeting, February 21-26, 2011, Chicago, Illinois.

Journal of Forensic Sciences (Peer reviewed article in the to be submitted in winter 2011)

California Association of Criminalists Spring 2011 Seminar – Los Angeles area

NIJ Pattern Evidence Symposium, August 2010 "The Statistical Evaluation of Torn Duct Tape Matches"

Appendix A

Suggested Procedures for the Training and Analyses of Duct Tape End Matches

Purpose: A suggested training program focused on training the analyst to conduct end match comparisons of duct tape. This proposal wills briefly discus the various analytical methods, but most of the focus will be on the non-destructive end matching of torn duct tape specimens.

Introduction

Duct tape is a multi composition material consisting of a backing material, a fiber reinforcement layer, and an adhesive layers. Each of these component parts are manufactured in wide varieties. The end use application of the product or the consumer usage will dictate how the manufacturer will develop particular duct tape specimen. Due to it manufacturing process, there are many analytical schemes that can be used to identify a particular duct tape specimen. These involve a variety of instrumental techniques that can differentiate backing layers, color coating, adhesive coating, fibers, etc. However, the most significant finding is that of an end match. If all the visual characteristics are the same, then a successful end match should not require any further chemical or instrumental analysis.

Training Time

About 200 total hours should be allocated to this training. In order to avoid analyst fatigue, the actual tape comparison work should be limited to 10 hours a week

Training Outline

I. DUCT TAPE HISTORY AND USAGE
   a. Development of duct tape
   b. Current duct tape usage
   c. Duct tape types and purposes
   d. Manufacturing process
   e. Manufacturing sources

II. DUCT TAPE STRUCTURAL COMPOSITION
   a. Discuss the three basic layers
      i. Backing
      ii. Support
      iii. Adhesive
   b. Backing types
      i. Compositions
      ii. Colors
   c. Support types
      i. Scrim
         1. Structure and definition
         2. Fibers compositions
         3. Weave patterns
   d. Adhesive

        i. Adhesive types
        ii. Adhesive purposes

III. PHYSICAL MATCH INTERPRETATION
    a. What is a physical match?
        i. Is it absolute identity?
        ii. Sufficient features
    b. Are physical matches unique?
    c. Class characteristics description
        i. Discuss AFTE criteria
    d. Subclass characteristics description
        i. Discuss AFTE criteria
    e. Individual characteristics issues
        i. Discuss AFTE criteria
    f. Mode of separation
        i. Torn, distortion
        ii. Cut tape samples, tools
        iii. Intermediate separations
    g. Conclusions
        i. Match
        ii. Non-Match
        iii. Inconclusive
            1. Ranges

IV. LITERATURE REVIEW OF SELECTED DUCT TAPE ARTICLES
    a. Student to review pertinent journal articles that describe the various analytical procedures with a focus on the physical matching issues
    b. Student to summarize and discuss relevant issues of each article

V. DUCT TAPE EXAMINATIONS
    a. Overview of basic examination protocols
        i. Chain of custody issues
        ii. Documentation techniques
            1. Digital imaging
            2. Case notes
        iii. DNA evidence recovery
        iv. Latent fingerprint recovery
    b. Equipment requirements
        i. Magnification lamp
        ii. Stereo Microscope
        iii. Light box
    c. Visual examination notes
        i. Type of tape
        ii. Condition – distortion, damage, environmental changes
        iii. Color of backing
        iv. Adhesive color and texture

v. Width
vi. Surface texture
    1. Surface calendar marks
vii. Construction
    1. Scrim pattern and sequence
    2. Thickness
viii. Stereo Microscope comparisons
    1. Unusual inclusions
ix. Characteristics determination
    1. What are the class characteristics
    2. Any subclass characteristics
    3. Proposed individual characteristics

VI. DUCT TAPE END MATCH EXAMINATIONS
a. Review the class characteristics attributes
    i. Tape size, color, calendaring marks, scrim size, scrim pattern
b. Compare the end matching using the following equipment as required:
    i. Stereo Microscope
    ii. Illuminated magnifying lamp
    iii. Light box
c. Document the end match profile by commenting on the following factors:
    i. Scrim alignment
    ii. Angle of tear/cut, presence of gaps or overlap
    iii. Unique tear pattern
    iv. Overall percent of correspondence of matching area
    v. Matching of exposed yarn to the corresponding tape missing yarn in the scrim

VII. DUCT TAPE PRACTICAL TRAINING SET ONE
a. Inter-compare 20 envelopes containing four exemplar duct tape specimens to four specimens with random numbers. Irrespective of any physical end match the analyst will still be required to inter-compare all 16 specimens to each other. This equates to 320 comparisons
b. The specimens will be label E-001 to E-080. Their matching counter parts will have random numbers assigned using the RAN function from Excel. There will only be one area of match - right side of the Exemplar and left side of the random number specimen.
c. Report the results of the findings as one of the following: End match, Inconclusive, or Non match
d. Review with the trainer the detailed result of your finding and documentation
e. Re-analyze off-target values

VIII. DUCT TAPE PRACTICAL TRAINING SET TWO
a. Inter-compare 25 envelopes containing five exemplar duct tape specimens to five specimens with random numbers. Irrespective of any physical end match the analyst will still be required to inter-compare all 25 specimens to each other.

b. The specimens will be label E-001 to E-100. Their matching counter parts will have random number assigned using the RAN function from Excel. There will be four areas of match – left or right side of the Exemplar and left or right side of the random number specimen. This equates to 100 comparisons per envelope and 2500 comparisons total.
c. Report the results of the findings as one of the following: End match, Inconclusive, or Non match
d. Review with the trainer the detailed result of your finding and documentation
e. Re-analyze off-target values

IX. DUCT TAPE COMPETENCY TEST

a. This set contains 100 exemplar tape specimens and 100 randomly numbered duct tape specimens torn form duct tapes that has minimal stretching techniques.
b. Duct specimens are torn and placed with three pairs to a transparency sheet. The left portion of the tape specimens is numbers E-xxx and the right portion of the tape specimen is given a random number
c. The sheet with the duct tape is documented by a digital scanner or copy machine
d. Remove six matching random number tape specimens and replaced with six random number tape specimens that do not have a counterpart in the exemplars data set
e. All specimens have to be inter-compared. That is exemplar #001 will have to be compared with 100 random numbered specimens even if an end match is found half way thru the testing.
f. The total number of comparisons will be ~10,000
g. Expectations: when the analyst is at this stage of the training one should expect at least 98.5% accuracy for specimens.

Appendix B

Training Literature Review

Aziz, N., P.F. Greenwood, K. Grice, R.J. Watling, W. van Bronswijk. 2008 “Chemical Fingerprinting of Adhesive Tapes by GCMS Detection of Petroleum Hydrocarbon Products.” *Journal of Forensic Sciences.* 53(2008): 1130-1137.

Bakowski, N. L., E.C. Bender, and T.O. Munson. “Comparison and Identification of Adhesives Used in Improvised Explosive Devices by Pyrolysis-capillary Column Gas Chromatography-Mass Spectrometry.” *Journal of Analytical and Applied Pyrolysis*. 8 (1985):483-492.

Bradley M.J., J.M. Gauntt, A.H. Mehltretter, P.C. Lowe, D.M. Wright. “A Validation Study for Vinyl Electrical Tape End Matches.” *Journal of Forensic Sciences*. (2011). Retrieved March 20, 2011: DOI: 10.1111/j.1556-4029.2011.01736.x.

Bradley M.J., R.L. Keagy, P.C. Lowe, M.P. Rickenback, D.M. Wright, and M.A. LeBeau. “A Validation Study for Duct Tape End Matches.” *Journal of Forensic Sciences* 51(May 2006): 504-508.

Carter, James F., Polly L. Grundy, Jenny C. Hill, Neil C. Ronan, Emma L. Titterton, and Richard Sleeman. "Forensic Isotope Ratio Mass Spectrometry of Packaging Tapes." *Analyst*. 129(2004): 1206-1210.

Curry, C.J. “Pyrolysis-Mass Spectrometry Studies of Adhesives.” *Journal of Analytical and Applied Pyrolysis*. 11(1987): 213-225.

Dobney A.M., W. Wiarda, P. de Joode, G.J.Q. van der Peijl. “Elemental Comparison of Packaging Tapes Using HR ICP/MS.” *Problems of Forensic Sciences* 47(2001): 275-287.

Dobney A.M., W. Wiarda, P. de Joode, G.J.Q. van der Peijl. “Sector Field ICP-MS Applied to the Forensic Analysis of Commercially Available Adhesive Packaging Tapes.” *Journal of Analytical Atomic Spectroscopy*. 5(2002): 478–484.

Federal Bureau of Investigation Scientific Working Group on Materials Analysis (SWGMAT). *Guidelines for the Forensic Examinations of Pressure-Sensitive Tapes*. Washington, DC: U.S. Department of Justice, Federal Bureau of Investigation Forensic Science Communications. 2008.

Goodpaster, J.V., A. B. Sturdevant, K. L. Andrews, and L. Brun-Conti. "Identification and Comparison of Electrical Tapes Using Instrumental and Statistical Techniques: I. Microscopic Surface Texture and Elemental Composition." *Journal of Forensic Sciences.* 52(2007): 610-629.

Goodpaster, J.V., A. B. Sturdevant, K. L. Andrews, E. M. Briley, and L. Brun-Conti. "Identification and Comparison of Electrical Tapes Using Instrumental and Statistical

Techniques: II. Organic Composition of the Tape Backing and Adhesive." *Journal Forensic Sciences.* 54(2009): 328-338.

Hobbs A.L., J. Gauntt, R. Keagy, P.C. Lowe, D. Ward. "A New Approach for the Analysis of Duct Tape Backings." *Forensic Science Communications* 9(2007).

Johnston, John. "Adhesive Tapes as Trace Evidence. ASI," Paper presented to the Pressure Sensitive Tape Council Tech 30 Global Conference VI, Orlando, FL, 2007.

Johnston, John. *Pressure Sensitive Adhesive Tapes: A Guide to their Function, Design, Manufacture, and Use*. Illinois: Pressure Sensitive Tape Council, 2003.

Katterwe, Horst W., "Fracture Matching and Repetitive Experiments: A Contribution of Validation." *AFTE Journal.* 37(Summer 2005): 229-241.

Kumooka Y. "Discrimination of rubber-based pressure sensitive adhesives by size exclusion chromatography." *Forensic Science International.* 171(2007): 5-8.

Maynard, P., K. Gates, C. Roux, and C. Lennard. "Adhesive Tape Analysis: Establishing the Evidential Value of Specific Techniques." *Journal of Forensic Sciences* 46(2001): 280-287.

Merrill R. A. and E.G. Bartick. "Analysis of Pressure Sensitive Adhesive Tape: I. Evaluation of Infrared ATR Accessory Advances." *Journal of Forensic Sciences* 45(2000): 93-98.

Sakayanagi, M., Y. Konda, K. Watanabe, and Y. Harigaya. "Identification of Pressure-Sensitive Adhesive Polypropylene Tape. *Journal of Forensic Sciences.* 48(2003): 68-6.

Smith, Jenny M. "Forensic Examination of Pressure Sensitive Tape," in *Forensic Analysis on the Cutting Edge: New Methods for Trace Evidence Analysis*. ed. Robert D. Blackledge. New Jersey: Wiley, 2007: 291-332.

Tsach T., S. Wiesner, Y. Shor. "Empirical Proof of Physical Match: Systematic Research With Tensile Machine." *Forensic Science International.* 116(2007): 77-83.

Williams, E. R. and T.O. Munson. "The Comparison of Black Polyvinylchloride (PVC) Tapes by Pyrolysis Gas Chromatography." *Journal of Forensic Sciences.* 33(1988): 1163–1170.

Appendix C

Duct Tape Research Study

Exemplar # Duct Tape	Unknown Duct Tape #	Distance of Matching Areas	% Match	Match Category	Describe your reasoning

Appendix D

Date	Envelope number	Exemplar number	Evidential number	Overlap (Y/N)	Gap (Y/N)	Angle alignment (Y/N)	Scrim alignment (Y/N)	Match (Y/N)	Match length (mm)	Tape width (mm)	Match %	Notes:
	1											
	2											
	3											
	4											
	5											
	6											
	7											
	8											
	9											
	10											
	11											
	12											
	13											
	14											
	15											
	16											
	17											
	18											
	19											
	20											
	21											
	22											
	23											
	24											
	25											
	26											
	27											
	28											
	29											
	30											
	31											
	32											
	33											
	34											
	35											
	36											
	37											
	38											
	39											
	40											
	41											
	42											
	43											
	44											
	45											
	46											
	47											
	48											
	49											
	50											