Ray A. Wickenheiser,[1] *B.Sc., Hons. and David G. Hepworth,*[2] *B.Sc., Hons.*

Further Evaluation of Probabilities in Human Scalp Hair Comparisons

REFERENCE: Wickenheiser, R. A. and Hepworth, D. G., **"Further Evaluation of Probabilities in Human Scalp Hair Comparisons,"** *Journal of Forensic Sciences,* JFSCA. Vol. 35, No. 6, Nov. 1990, pp. 1323–1329.

ABSTRACT: Placing value on associative hair evidence is an integral part of court presentation. A modified repeat of the hair probability study by Gaudette and Keeping has been undertaken, with steps taken to remedy shortcomings of the original work. The results of this study demonstrate that, with the application of rigid selection criteria, the frequency of coincidental matches in forensic science hair comparisons is low. It also demonstrates that routine hair classification is not feasible, because of inconsistency in examiner discrimination. The macroscopic selection of 5 to 13 mutually dissimilar hairs has been shown to be frequently unrepresentative of the microscopic range of features present in a known hair sample.

KEYWORDS: forensic science, criminalistics, hair

Hair evidence has been shown to be statistically good evidence, as was first demonstrated by Gaudette and Keeping [*1*] and then further shown by Gaudette [*2*] and by Strauss [*3*]. The original study, "An Attempt at Determining Probabilities in Human Scalp Hair Comparisons" [*1*], was the first of a series of papers placing a statistical value on hair evidence. Representative scalp hairs were selected from 100 individuals, microscopic features were coded on punch cards, and the cards were sorted to eliminate unnecessary comparisons of obviously dissimilar hairs. Similarly coded hairs were compared one to one microscopically, which resulted in 9 pairs of hairs being declared microscopically indistinguishable.

Based on the number of individuals involved (100), the total number of hairs examined (861), the number of theoretical examinations performed (366 630), and the number of incorrect associations made (9), a number of probabilities were set forward. The most notable was, "It is estimated that if one human scalp hair found at the scene of a crime is indistinguishable from at least one of a group of nine dissimilar hairs taken from a given source, the probability that it could have originated from another source is small, about 1 in 4500."[3]

A critical paper, "Probabilities and Human Hair Comparisons," by Barnett and Ogle [*4*], raised four concerns: the validity of the application of findings to life situations, possible examiner bias, the use of non-individualizing features in the sorting procedure,

Received for publication 2 Oct. 1989; accepted for publication 20 Nov. 1989.

[1]Forensic science laboratory specialist, Hair and Fibre Section, RCMP Forensic Laboratory, Regina, Saskatchewan, Canada.

[2]Head, Hair and Fibre Section, RCMP Forensic Laboratory, Regina, Saskatchewan, Canada.

[3]Ref *1*, page 605.

and the statistical treatment of data. In our view, the Gaudette rebuttal [5] adequately dealt with three of the four criticisms. The only criticism not fully addressed by Gaudette was that of examiner bias. Despite efforts to apply the same level of discrimination and the same criteria for elimination used in routine casework, the examiner knew any hairs deemed indistinguishable would not be correctly associated. In our view, examiner bias could have been eliminated through the addition of an unknown number of potentially matching hairs. Participation of more than one examiner working independently would also provide a comparison of individual levels of discrimination in hair comparison as well as increase the validity of the findings.

Despite the criticism of Barnett and Ogle [4], the rebuttal by Gaudette [5], and the lack of agreement as to the value of hair evidence [6], there has been little additional study in this area. Since an integral part of the court presentation of associative evidence involves placing a value on findings, further evaluation of probabilities in human scalp hair comparisons was warranted. A repeat of the Gaudette and Keeping [1] study was undertaken, with steps taken to remedy shortcomings of the original study.

Method

Representative hair samples of 100 or more scalp hairs were obtained from 97 different Caucasian individuals, including a number of closely related individuals from several generations. From 5 to 13 mutually dissimilar hairs were macroscopically chosen to represent the range of characteristics demonstrated by each known hair sample. An independent party randomly numbered the mutually dissimilar hairs and, at their discretion, added a number of additional hairs (53) randomly chosen from the original known samples of the 97 individuals. Therefore, several duplicate hairs, which could potentially match the mutually dissimilar hairs yet also have originated from the same individual, were included in the study. In this way, the two participating examiners would be unbiased while comparing hairs. That is, when they encountered a situation in which two hairs were microscopically indistinguishable, they would not know whether the two hairs they were comparing were from the same source or whether they were dealing with a coincidental match. In total, 930 hairs were measured, mounted on numbered glass slides in Fisher Permount, and examined on a Leitz Orthulux II bright-field comparison microscope at objective magnifications of 100 and 240 power.

In order to determine the number of matching pairs of hairs present within the 930-hair group, it was necessary to compare each of the 930 hairs to each one of the 929 remaining hairs. In other words, Hair 1 was compared with Hairs 2 through 930 inclusive, and so on. As a consequence of this process, a total of $(930 \times 929)/2$, or 431 985, hair comparisons were necessary. Since it was impractical to perform such a large number of one-to-one comparisons (many of which represented quick eliminations of obviously dissimilar hairs) a system of sorting the hairs was devised.

The method used was somewhat similar in principle to that outlined by Gaudette and Keeping [1], except that a personal computer database was utilized in lieu of punch cards. The sole purpose of the sorting procedure was to eliminate unnecessary one-to-one microscopic comparisons of obviously dissimilar hairs.

As the sorting procedure was developed, two potential difficulties surfaced. First, if the sorting procedure was too stringent, hairs that were similar could be eliminated from the final one-to-one microscopic comparisons. Second, if the sorting procedure was too general, the number of hairs included for final one-to-one microscopic comparison could be so large that completion of the study would not be possible within a reasonable time frame. Therefore, it was necessary to design a sorting procedure general enough to include all similar hairs, yet stringent enough to exclude obviously dissimilar hairs.

Toward this end, the examiners selected 14 characteristics they routinely utilize in describing human scalp hairs. These characteristics were separated into two categories,

major and secondary. Based on the examiners' experience, major characteristics were identified as being those features that, if not identical in two hairs, would mark them as obviously dissimilar. Secondary characteristics were assigned to include features that could vary slightly in "indistinguishable" hairs or be more prone to suffer variations in the initial classification. The terminology used to describe the variations within each characteristic was based on familiarity and common use among the examiners, and it was found to cover the general range of characteristics exhibited in scalp hairs adequately.

The use of familiar characteristics and terminology allowed the examiners to describe each hair in a manner similar to that used routinely for years, and was done in order to draw from their experience to obtain maximum accuracy. It may be argued that the characteristics chosen are features that do not permit individualization of hair. However, failure to include more specific individualizing characteristics would not increase the rate of errors, but would only increase the number of one-to-one comparisons, thereby requiring more time to complete the study.

For the purposes of this study, both examiners individually classified each of the 930 hairs on the basis of their microscopic features, as outlined in Fig. 1. Each hair was classified into one (or more, as was often necessary) of the major color groups. After color classification, a single-digit code number was assigned for each of the remaining 13 major and secondary characteristics. As a result, each of the 930 hairs, when fully classified, was represented by a color and a 13-digit classification number. Hairs that contained several different features or colors were given more than one classification number or color, as necessary, to ensure that all features were represented. The features selected to be recorded in the major and secondary characteristic categories were purposely chosen to be as general as possible to prevent the final sorting from being too stringent.

When all 930 hairs were classified, the color and classification number for each hair was entered into a computer sort utility, and the 13-digit classification numbers were sorted in ascending numerical order according to color block. For example, all white hairs were sorted as a group, as were those classified as yellow, dark brown, and so on. A computer printout for each color block was obtained. Each printout listed, in ascending order, the classification number for all hairs included in the color block, as well as their individual identification numbers (that is, Nos. 1 through 930). Each printout was further sorted according to final selection criteria, as outlined in Table 1. All hairs within each color block that met the final selection criteria were identified for manual one-to-one microscopic comparisons.

The selection criteria rules were developed on the basis of examiner experience. It was agreed that, in the microscopic comparison of hairs, most experienced examiners would consider hairs obviously different if there were differences in color or in the major characteristics listed in Fig. 1. In the case of secondary characteristics (Fig. 1), however, it was felt that slight differences in certain features would not necessarily constitute grounds for elimination, unless there were several differences evident. Therefore, it was decided that more than four differences in the category of secondary characteristics would constitute grounds for elimination. More rigid rules were not applied in order to prevent potentially similar hairs from being excluded by the computer sort.

After each of the two examiners had independently examined, classified, sorted, and prepared final comparison lists, the one-to-one microscopic comparisons were conducted. The hair pairs were highly scrutinized for significant differences, and in each case where two hairs were selected as being a one-to-one match, they were essentially indistinguishable upon microscopic comparison. The one-to-one comparisons were basically a search for significant differences. This was particularly true for "featureless hairs," since they exhibit a limited number of microscopic features for comparison. In each case where two hairs were determined to be a one-to-one match, they shared all characteristics in common, including the variation of these characteristics from root to tip.

HAIR CLASSIFICATION

Colour:	White	Light Brown	Dark Brown	Red
	Yellow	Medium Brown	Red Brown	Black

MAJOR CHARACTERISTICS

Treatment

None	0
Dyed	1
Bleached	2
Curled/Permed	3
Other	4

Texture

Smooth	1
Streaky	2
Very Streaky	3
Granular	4

Pigment Distribution

Absent	0
Uniform	1
Peripheral	2
About Medulla	3
Unilateral	4
Clusters	5

Medulla

Absent	0
Continuous/Opaque	1
Continuous/Translucent	2
Continuous/Opaque/Translucent	3
Interrupted/Opaque	4
Interrupted/Translucent	5
Interrupted/Opaque/Translucent	6
Fragmentary/Opaque	7
Fragmentary/Translucent	8
Fragmentary/Opaque/Translucent	9

Medulary Index

Non Applicable	0
<0.250	1
>0.250	2

Maximum Diameter

Fine <0.04mm	1
Med 0.04-0.08mm	2
Coarse >0.08mm	3

Cortical Fusi

Absent	0
Present - Root	1
Present - Shaft	2

SECONDARY CHARACTERISTICS

Length

<2.5 cm	1
2.5 - 7.5 cm	2
7.5 - 15 cm	3
15 - 30 cm	4
>30 cm	5

Cuticular Margin

Smooth	1
Slightly Serrated	2
Serrated	3

Tip

Natural Taper	1
Recent Cut	2
Rounded	3
Rounded Frayed	4
Split	5
Crushed	6

Pigment Density

Absent	0
Sparse	1
Light	2
Medium	3
Heavy	4

Pigment Size

Absent	0
Fine	1
Medium	2
Large	3

Shaft

Constant	1
Slight/Smooth/Variation	2
Wide/Smooth/Variation	3
Slight/Abrupt/Variation	4
Wide/Abrupt/Variation	5

FIG. 1—*Microscopic hair classification categories and codes.*

TABLE 1—*Selection criteria for one-to-one comparisons.*

Color must agree
Major characteristics must agree (i.e., the first 7 digits of the classification number must be the same)
Length—variance of 1 allowed
Cuticular margin—variance of 1 allowed
Tip—variance of 1 allowed if 2, 3, 4, 5, or 6
Pigment density—variance of 1 allowed if 1, 2, 3, or 4
Pigment size—variance of 1 allowed if 1, 2, or 3
Shaft—variance of 1 allowed if 1; variance of 2 allowed if 2, 3, 4, or 5
No more than four (4) differences allowed in the category of secondary characteristics

Results

As a result of the individual computer sorts, Examiner 1 conducted 749 one-to-one microscopic comparisons, and Examiner 2 conducted 2006 comparisons.

The results of these comparisons are shown in Table 2. Examiner 1 found seven pairs of hairs that were microscopically indistinguishable, and Examiner 2 found six pairs. Two pairs of hairs were found in common by both examiners. In every case where a one-to-one match was found, the source of both hairs was identical. No incorrect associations were made by either examiner.

Discussion

The results of this study raise some important questions. First, why did Examiner 1 conduct only 749 one-to-one comparisons, while Examiner 2 performed 2006? Comparison of the individual classification numbers assigned to each hair by Examiners 1 and 2 indicate that the difference in final one-to-one comparisons was a direct result of individual variation in the original classification of the hairs. In many cases, the examiners classified the same hairs differently. A wider usage of descriptive terms by Examiner 1 produced hair classification numbers spread over a larger range, thereby resulting in fewer final comparisons. This was amplified by the fact that hairs sharing common major characteristics were grouped together. When the hair types were common (for example, light brown, untreated, streaky texture, peripherally pigmented, non-medullated, medium diameter, 2.5 to 7.5 cm in length), the size of the group was quite large. The addition of 1 more hair to a group of 25 hairs with the same major characteristic classification code could conceivably lead to 25 additional one-to-one comparisons.

The second major question arising from the results is why did the two examiners choose only 2 pairs of one-to-one matches in common? If all of the 13 chosen matches were indeed one-to-one matches and a proper sort was performed, the findings should have been more similar.

In an attempt to determine the reason for the discrepancy, each examiner reviewed the other's one-to-one matches. In all cases, they concluded that, at the very least, these hairs should have been identified as potential matches by the sorting procedure and compared one to one microscopically. Close examination of the sorting procedure revealed that each examiner had considerable day-to-day variation in hair feature classification. Changes in how the individual hairs were classified over time led to elimination by the sorting procedure, particularly in the areas of texture and pigment distribution.

TABLE 2—*Comparison of the results of one-to-one matches by the two examiners.*[a]

Examiner 1		Examiner 2	
1:1 Matches	Known Source	1:1 Matches	Known Source
Hair 99 vs. 314	both from S12	Hair 76 vs. 96	both from S15
Hair 195 vs. 813	both from S42	Hair 94[c] vs. 796	both from S50
Hair 202 vs. 807	both from S87	Hair 322 vs. 478[d]	both from S96
Hair 322 vs. 478[b]	both from S96	Hair 594 vs. 880[d]	both from S96
Hair 406 vs. 675	both from S24	Hair 690 vs. 719	both from S85
Hair 594 vs. 880[b]	both from S96	Hair 806 vs. 837	both from S64
Hair 822 vs. 823	both from S21		

[a]Note—a 1:1 match implies that no significant microscopic differences were found between the hairs in question.

[b]Same findings as Examiner 2.

[c]Hair number 94 was one of 53 duplicate hairs.

[d]Same findings as Examiner 1.

In many cases, hairs that were generally similar were not compared, as the resulting inconsistent classification numbers caused them to be eliminated by the computer sort. This finding is in agreement with that of Podolak and Blythe [7], who found that an individual's ability to classify hair features consistently varies over time. The end result was that the two examiners did not identify the same pairs of hairs for final one-to-one comparison and, with the exception of two pairs of hairs, identified different pairs as being microscopically indistinguishable. In hindsight, it is evident that, had the selection criteria for one-to-one comparison (Table 1) allowed some variation in texture and pigment distribution, it is probable that the two examiners would have identified more of the same pairs of hairs for final one-to-one comparisons.

The third question raised concerns the duplicate samples. Why was only 1 of the 53 duplicate hairs included in the final one-to-one matches? One would expect more of the remaining 52 duplicates to be paired with their respective known samples. Each of the 53 duplicate hairs was compared with the 5 to 13 hairs originally selected to represent their known samples of origin. It was found that 38 of the 53 duplicates had no hair in the known sample that was microscopically similar. That is, none of the 38 duplicate hairs should have been identified for one-to-one comparison as a result of the computer sort. Therefore, it appears that the initial macroscopic selection of 5 to 13 hairs to represent the range of a known sample has been demonstrated to be inadequate by itself to represent a known sample microscopically. A microscopic review of an entire standard is necessary to ensure that all features present in the standard are represented.

Several of the remaining 15 duplicate hairs were identified for one-to-one comparison by the computer sort. As a result of the sort, Examiner 1 identified 3 hairs for one-to-one microscopic comparison and eliminated them all by narrow margins. Examiner 2 selected 4 hairs for one-to-one microscopic comparison, eliminated 3, and identified 1 (the fourth) as being microscopically indistinguishable from another hair. None of the duplicate hairs compared by Examiner 1 and 2 were the same. The fact that the remaining 12 (Examiner 1) and 11 (Examiner 2) duplicate hairs were eliminated by the initial computer sort is further evidence of individual variation in the initial classification of the hairs.

The two examiners working independently did not make any incorrect associations during the course of this study. Based on the large number of potential one-to-one comparisons (431 985) conducted, it can be concluded that, if a one-to-one match is insisted upon, there is a very low incidence of error. This is true even if one argues that several comparisons were not conducted because of potential errors in the initial sort of the hairs. In cases of this type, it is far more likely that the error committed will be that of incorrect elimination rather than of incorrect association.

It should also be noted that, in this study, no particular efforts were made to include or exclude "featureless hairs." Several such hairs were encountered and, in one case, two hairs determined to be microscopically indistinguishable were in fact hairs that most examiners would classify as being "featureless." It is important, however, that "featureless" hairs be subjected to a high level of scrutiny because of their inherent lack of microscopic features for comparison purposes.

A final question raised by this study relates to normal forensic science examinations. How do these findings relate to routine forensic hair comparisons? Despite the results of this study, based on personal experience and the work of others [1,2], the authors concede that a one-to-one microscopic hair match is not a means of 100% positive identification. While it is possible that a forensically significant matching hair could have originated from another individual, certain conditions must be met. First, an individual (A) must have had access to the crime scene within a time frame reasonably associated with the crime. Second, that individual (A) must have at least one hair on his or her head that is microscopically indistinguishable from at least one hair of the known sample obtained from the wrongly accused person (B). Finally, the significant hair must have

been shed when the individual (A) had access to the crime scene. A parallel argument can be used in the case of secondary transfer, except that, in secondary transfer, the individual (A) with access to the crime scene must have on his or her person a hair that is microscopically indistinguishable from at least one hair from the known sample taken from the wrongly accused person (B). Again, the appropriate hair must be left at the crime scene by secondary transfer. Given these circumstances and assuming a one-to-one match exists, this study demonstrates that the probability of incorrect association in routine forensic hair comparison is remote.

Conclusions

The following conclusions were drawn:

1. If two hairs are microscopically indistinguishable from root to tip and thus show no significant differences (that is, there is a one-to-one match), the probability of incorrect association is remote.
2. The classification of hair varied between examiners.
3. The classification of hair by both examiners was inconsistent over time because of variations in discrimination. The sorting procedure used was therefore susceptible to error.
4. Macroscopic selection of 5 to 13 mutually dissimilar hairs was frequently unrepresentative of the microscopic range of features present in the known samples. Experimental work aimed at determining the optimum composition of a representative known hair sample is warranted.

Acknowledgments

The assistance of C/M G. D. E. Wilson of the Royal Canadian Mounted Police Forensic Laboratory in Regina, Saskatchewan, is gratefully acknowledged.

References

[1] Gaudette, B. D. and Keeping, E. S., "An Attempt at Determining Probabilities in Human Scalp Hair Comparisons," *Journal of Forensic Sciences*, Vol. 19, No. 3, July 1974, pp. 599–606.
[2] Gaudette, B. D., "Some Further Thoughts on Probabilities and Human Hair Comparisons," *Journal of Forensic Sciences*, Vol. 23, No. 4, Oct. 1978, pp. 758–763.
[3] Strauss, M. "Forensic Characterization of Human Hair I," *The Microscope*, Vol. 31, No. 1, 1983, pp. 15–29.
[4] Barnett, P. D. and Ogle, R. R., "Probabilities and Human Hair Comparison," *Journal of Forensic Sciences*, Vol. 27, No. 2, April 1982, pp. 272–278.
[5] Gaudette, B. D., "A Supplementary Discussion of Probabilities and Human Hair Comparisons," *Journal of Forensic Sciences*, Vol. 27, No. 2, April 1982, pp. 279–289.
[6] Hicks, J. W., Deadman, H. A., Deforest, P. R., Gaudette, B. D., Roe, G. M., and Wittig, M., "Evidential Value of Hair Examinations," Report on Panel Discussion, *Proceedings*, International Symposium on Forensic Hair Comparisons, FBI Academy, Quantico, VA, 25–27 June 1985, pp. 193–211.
[7] Podolak, A. G. and Blythe, C. E., "A Study of the Feasibility of Establishing a Computer Data Bank for Hair Characterization Using Standard Descriptive Criteria," *Proceedings*, International Symposium on Forensic Hair Comparisons, FBI Academy, Quantico, VA, 25–27 June 1985, p. 149.

Address requests for reprints or additional information to
D. G. Hepworth
Hair and Fibre Section
RCMP Forensic Laboratory
P.O. Box 6500
Regina, Sask.
Canada S4P 3J7