B. D. Gaudette, [1] *B.Sc.*

A Supplementary Discussion of Probabilities and Human Hair Comparisons

REFERENCE: Gaudette, B. D., "**A Supplementary Discussion of Probabilities and Human Hair Comparisons,**" *Journal of Forensic Sciences,* JFSCA, Vol. 27, No. 2, April 1982, pp. 279-289.

ABSTRACT: The author presents an expanded discussion of his work on the probability of errors in forensic science comparisons of human hair. The presentation includes a clarification of some areas of past work on this topic, a discussion of the concept of error probabilities as applied to associative evidence in general and to hair comparisons in particular, an explanation of factors that differentiate hair comparisons from other forms of associative evidence, and a direct response to criticisms of the author's past work.

KEYWORDS: criminalistics, hair, probability, human identification

The recent appearance in the literature of two papers [*1,2*], as well as personal communications with some forensic scientists, has indicated that the underlying principles and concepts of my three papers on probabilities and human hair comparison [*3-5*] are sometimes not well understood. Since the probability concepts involved are very subtle, further explanation seems to be required to ensure that all who use the results have a good understanding of their derivation and meaning. This paper will consist of four parts. First, I will attempt to clarify some areas from my papers that appear to cause misconceptions. Next, I will discuss the general concept of error probabilities in human hair comparison. Then the extra dimension that makes hair comparison different from most other forms of associative evidence will be discussed. In the fourth section, a direct reply will be made to the criticisms of my work expressed by Barnett and Ogle [*1*]. It is hoped that this paper will contribute to an understanding of the probabilities that relate to human hair comparison. A better understanding of the strength of hair comparison evidence should then lead to an increased use of this valuable type of evidence by forensic science laboratories and the courts.

Clarification of Some Common Misconceptions

First, because a card coding system was used as a part of the original research, some people feel that I am advocating card coding as a new method of hair comparison and that my results are only valid when card coding is used. In fact, for routine casework in the Royal Canadian Mounted Police laboratory system, hair comparisons are made directly with a comparison microscope and not by card coding. Similarly, all the important comparisons reported in my papers [*3-5*] were made directly with a comparison microscope. The card coding was just a way of quickly eliminating a large number of obviously dissimilar hairs (for ex-

Accepted for publication 3 Nov. 1981.
[1] Scientific advisor-hair and fibers, Science and Technology Advisory Group, Royal Canadian Mounted Police, Ottawa, Ontario, Canada.

ample, dark brown as compared to light brown) and as such was simply a labor-saving device. Since all hairs whose characteristics were close were compared directly by comparison microscope, the use of the card coding system should affect neither the validity of the probability results obtained nor their applicability to casework. While the card coding was carried out according to rules similar to those used in forensic hair comparison, a coding system would not likely be suitable for routine casework applications because the thought processes and pattern recognition involved in hair comparison are so complex and because any coding procedure is, of necessity, somewhat restrictive.

A second area requiring clarification concerns the independence assumption that was made in the probability calculations for my first paper [*3*]. It was assumed there that deliberately selecting approximately nine mutually dissimilar hairs to represent the hairs on the scalp of an individual made these nine sufficiently independent of each other that the same calculation could be made as if the nine hairs were randomly chosen from all the hairs in the population studied. Thus it was calculated that the probability P of misclassifying at least one pair of hairs $= 1 - (1 - p)^n$, where n was 9 and p was $9/366\,630 = 1/40\,737$, where p is the number of indistinguishable pairs of hairs divided by the population studied. This yielded a result of about 1 in 4500. Several readers thought that this independence assumption was somewhat suspect since the hairs were preselected in a nonrandom manner. However, it has been pointed out to me by several people[2] that even if this independence assumption is dispensed with, the conclusion and indeed the figure of 1 in 4500 remains the same. This results from the Bonferonnie Inequality where, if we let Ei represent the event that the single hair from person A matches with the ith hair from those chosen to represent the hairs on the scalp of person B, we have the inequality:

$$P\,(\text{at least one } Ei \text{ occurs among } E_1, \ldots, E_n) \leq P(E_1) + P(E_2) + \ldots + P(E_n)$$

If $n = 9$ and $p = 9/366\,630$, then $np = 1/4526$. This upper bound is thus only minutely larger than the figure derived by assuming independence. Similarly, for pubic hairs, where $n = 7.5$ and $p = 16/101\,368$, $np = 1/845$, again very close to the results quoted using the independence assumption. Thus, whether or not the independence assumption was valid, the results remain the same.

A third common misunderstanding of my papers arises from a confusion of the statistical concept of population with the lay concept that population refers only to people. In statistics, a population is defined as an aggregate or totality of elementary units, such as people or things, about which information is desired. A population can also be thought of as the aggregate of observations about items in a given situation, rather than the items themselves [*6*].

In my research, the population considered was not a population of people, but rather a population of hair comparisons. Using that population, the probability of *an event occurring* was determined. The event in question was that of a careful, experienced hair examiner concluding that the characteristics of a given hair were consistent with its having originated from a particular individual, when in fact it came from someone else.

Error Probabilities in Human Hair Comparison

If we ignore the possibility of inconclusive results, there are two possible conclusions that can arise from a hair comparison: (a) that the unknown hair is consistent with having originated from the person from whom the standard hair sample was obtained or (b) that the unknown hair is *not* consistent with such origin. Similarly, there are two possible states of nature or reality: (1) the hair originated from the same person as the standard hair sample or

[2]In particular, Malcolm Greig and Piet de Jong of the University of British Columbia and David Binder of Statistics Canada.

(2) the hair did not originate from the same person as the standard hair sample. If the true state of nature is (1) and we give conclusion (a), we are correct. Similarly if the true state of nature is (2) and we give conclusion (b), we are also correct. However, if the true state of nature is (1) and we give conclusion (b), we have committed an error. This type of error is known as a Type I error. If the true state of nature is (2) and we give conclusion (a), we have committed a Type II error. These concepts are illustrated in Table 1. The probability of a Type I error is represented by α and the probability of a Type II error is β.

In hair comparison in the forensic sciences, the consequences of a Type II error are much more serious than the consequences of a Type I error, since a Type II error would result in wrongly incriminating evidence being presented against a suspect. On the other hand, a Type I error would only result in no evidence from a hair comparison being presented against a guilty person, since negative hair comparisons do not necessarily imply innocence or lack of association or contact. It is because of the differences in consequences that the level of discrimination in hair comparison ("if in doubt—throw it out") is set so as to minimize the possibility of Type II errors.

There are two possible sources of Type II errors, examiner error and systematic or structural errors. In hair comparison, structural errors arise as a result of the fact that hair comparison is not at present a positive means of personal identification—it is possible for two different individuals to have hairs on their heads that match. With an experienced, careful examiner and a quality assurance system, the possibility of examiner errors should be negligible. In the discussion that follows, only structural Type II errors will be considered, and the use of the term "Type II errors" will imply structural Type II errors.

In assessing the value of associative evidence we need to know the extent to which the ratio $P(A)/P(N)$ has been changed by the evidence. The probability that two individuals (or one individual and a crime scene) were in association is $P(A)$ and the probability that they were not in association is $P(N)$. Given the associative evidence E, the ratio becomes $P(A|E)/P(N|E)$ and the extent of change in the ratio can be measured by

$$\frac{P(A|E)/P(N|E)}{P(A)/P(N)}$$

Probability theory tells us that

$$\frac{P(A|E)/P(N|E)}{P(A)/P(N)} = \frac{P(E|A)}{P(E|N)}$$

From Table 1 we see that $P(E|A)$ is the same as $1 - \alpha$ and $P(E|N)$ is β. Thus the value V of a particular type of evidence in establishing association[3] can be given by

$$V = \frac{P(A|E)/P(N|E)}{P(A)/P(N)} = \frac{P(E|A)}{P(E|N)} = \frac{1-\alpha}{\beta}$$

If α is small, then $(1 - \alpha)/\beta \approx 1/\beta$; by simply determining β, the probability of a Type II error, we can determine the value of a particular type of associative evidence.

Since hair comparison evidence has been used successfully for so many years by major crime detection laboratories, we would expect α, the incidence of Type I errors, to be small. Some experimental verification of this is given in my third paper [5], where in all incidences

[3]Note that this is the value in establishing association, not guilt, since assessments of guilt will depend not only on association but also on a number of other factors such as intent, consent, alternative explanations for association, and fine points of the law.

TABLE 1—*Types of errors.*

	Reality	
Conclusion	Hair Is from Individual	Hair Is Not from Individual
Match	Correct; $P = 1 - \alpha$	Type II error; $P = \beta$
Nonmatch	Type I error; $P = \alpha$	Correct; $P = 1 - \beta$

of both experiments described, no Type I errors were committed. Thus it appears reasonable to assume that the value of hair evidence can be estimated by $1/P$ (Type II error).

The Extra Dimension of Hair Comparison

Knowledge of the probability of Type II errors is important in all forms of associative evidence. In blood comparisons, for example, the probability of such an error occurring is directly related to the frequency of occurrence of the various system types (such as ABO or phosphoglucomutase) in a particular population of people. As shown in Table 2, a 2-by-2 matrix exists with three distinct questions that can be asked about Type II errors in blood comparisons. First, we can ask what is the probability that two *individuals* chosen at random will have blood that is indistinguishable by a certain system or combination of systems. This is related to the discriminating power [7]. Secondly, we can ask what is the probability that anyone else in the *population* will have blood that is indistinguishable from that of a particular *individual*. The third question is what is the probability that anyone in a particular *population* will have blood indistinguishable from that of anyone in another *population*.

With blood, all the blood from a particular individual will be of the same type in any given system, regardless of the part of the body from which it originates. With hair, however, the situation is quite different. Not only are there the obvious differences in hairs from different body areas, there is also the fact that all the hairs from a particular body area are not homogenous. Indeed, a considerable amount of intrapersonal variation and inhomogeneity exists. Adjacent hairs on a person's scalp, for example, can be gray, dark brown, and light brown.

Everyone's hair has a range of characteristics over the scalp (or pubic region or other area). This is what makes hair comparisons so difficult. But once we have made a successful hair comparison, this variation works *for* us to add an extra dimension that strengthens hair evidence. The probability of occurrence of a Type II error for hair comparisons is composed of two factors:

(a) The probability that another individual will have a hair somewhere on his or her scalp (or other area) that will match any one of the hairs from the accused; and

(b) The probability that *that particular hair* (out of all the hairs on the person's scalp) would be the one to be shed and subsequently be found at the crime scene.

TABLE 2—*Probability questions about Type II errors in blood comparisons.*

	Unit Comprising Known Sample	
Unit Compared	Individual	Population
Individual	Question 1	Question 2
Population	Question 2	Question 3

The second factor is not available for blood comparisons; for this reason, there is an extra dimension to hair comparison. This extra dimension is reflected in the fact that we now have a 3-by-3 matrix with six distinct questions that can be asked about the probability of Type II errors in human hair comparison. If *A* and *B* are two randomly chosen individuals, then these six questions can be phrased as follows:

Q (1). If one hair is chosen at random from person *A* and one hair is chosen at random from person *B*, what is the probability that these two hairs will match—that is, be consistent with each other?

Q (2). If one hair chosen at random from *A* is compared to a representative standard hair sample obtained from *B*, what is the probability that it will be consistent with having originated from *B*?

Q (3). If one hair chosen at random from *A* is compared to representative standard hair samples from a number of individuals other than *A* in a population sample, what is the probability that it will be consistent with having originated from at least one of those individuals?

Q (4). If a representative standard sample from *A* is compared to a representative standard sample from *B*, what is the probability that at least one pair of hairs from the different samples will match?

Q (5). If a representative standard hair sample from *A* is compared to representative standard hair samples from a number of individuals in a population sample, what is the probability that at least one hair from *A* will be consistent with at least one hair from at least one of these other hair samples?

Q (6). If all of the hairs from all of the standard hair samples from all of the individuals in a particular population sample are compared to all of the hairs from all of the standard hair samples from all of the individuals in another population sample, what is the probability that at least one pair of hairs from different individuals will match?

The following answers, which represent probability *estimates*, all assume that *A* and *B* and all the populations are white.

The probability sought in question 1 represents the probability of a Type II error occurring when a randomly chosen single *hair* is compared to another randomly chosen single *hair*. This probability was previously calculated to be about 1/40 700 for scalp hair [*3*] and 1/6336 for pubic hair [*4*].

Question 2 asks the probability of a Type II error occurring when a randomly chosen single *hair* is compared to a representative hair sample from a randomly chosen *individual*. This probability is about 1/4500 for scalp hair [*3*] and 1/800 for pubic hair [*4*], whether or not an independence assumption is made in the calculations.

Question 3 is concerned with the probability of a Type II error when a randomly chosen single *hair* is found to match at least one representative hair sample from the group of hair samples that comprise the *population sample*. The answer to this question will depend on the size of the population sample chosen. For the population sample of 100 individuals considered in my first paper [*3*], of the 861 hairs in the scalp hair sample, 15 matched at least one other hair. Therefore the estimated probability is 15/861, which is about 1 in 57. Similarly, from the population sample of 60 individuals in the pubic hair study, 26 of the 454 hairs were found to match at least one other hair. This gives an estimated probability of 26/454 or about 1 in 17.

In question 4, we are asking the probability of a Type II error with at least one pair of hairs when a representative standard hair sample from an *individual* is compared to a representative standard hair sample from another *individual*. This is the same as Barnett and Ogle's statement III [*1*]. Using my data [*3,4*], they have calculated this probability to be 9/4950 = 1/550 for scalp hair and 16/1770 = 1/111 for pubic hair.

Question 5 asks the probability of a Type II error occurring with at least one pair of hairs when a representative standard hair sample from an *individual* is compared to representative samples from the group of hair samples that comprise the *population sample*. This question is equivalent to a form of Barnett and Ogle's statement II [*1*]. As they have demonstrated, this probability is 13/100 = 1/7.7 for scalp hair. With pubic hairs 23 *individuals* (not 25 as stated by Barnett and Ogle [*1*]) out of 60 were found to have nonunique hair [*4*]. Thus this probability would be 23/60 = 1/2.6 for pubic hairs.

Question 6 deals with the probability of a Type II error occurring with at least one pair of hairs when representative hair samples from the group of hair samples that comprise a *population sample* are compared to representative hair samples from another group of hair samples that comprise another *population sample*. An alternate way of phrasing this question would be to ask whether the event ever occurs that two different individuals are found to have hairs that are macroscopically and microscopically consistent with each other. Provided that the population samples are large enough, the answer to this question is, of course, yes, based on most examiners' personal experience and common knowledge. This answer was also proven in my research [*3-5*]. Thus, this probability is 1, a certainty.

These six questions and their associated probabilities are shown in Table 3 for scalp hair and Table 4 for pubic hair.

We must now consider which of these six questions most accurately reflect the issues encountered in forensic science hair comparison. Question 6, while clearly relevant, is trivial, since its answer is common knowledge among forensic scientists. Question 5 would be an important question if hair, like blood, were homogenous throughout an individual. Hair does, however, exhibit a wide degree of intrapersonal variation. Also, in forensic analysis of hair, questioned exhibits generally consist of hairs that must be considered singly, not of whole hair samples from individuals. Thus question 5 is only of academic interest. Furthermore, since the answer to it is dependent on the size of population sample chosen, it is basically meaningless.

Question 4 is of interest to forensic hair examiners since it gives an estimate of the rate of occurrence of two individuals having overlapping hair standards. Again, however, since questioned exhibits consist generally of single hairs, this question does not reflect an accurate model of forensic hair comparisons.

Question 3 deals with single hairs as unknowns and as such is closer to the true situation in

TABLE 3—*Probability estimates of Type II errors in scalp hair comparisons among whites.*

	Unit Comprising Known Sample		
Unit Compared	Single Hair	Individual	Population Sample
Single hair	1 in 40 700 (Q1)	1 in 4500 (Q2)	1 in 57 (Q3)
Individual	1 in 4500 (Q2)	1 in 550 (Q4)	1 in 7.7 (Q5)
Population sample	1 in 57 (Q3)	1 in 7.7 (Q5)	1 in 1 (Q6)

TABLE 4—*Probability estimates of Type II errors in pubic hair comparisons among whites.*

	Unit Comprising Known Sample		
Unit Compared	Single Hair	Individual	Population Sample
Single hair	1 in 6336 (Q1)	1 in 800 (Q2)	1 in 17 (Q3)
Individual	1 in 800 (Q2)	1 in 111 (Q4)	1 in 2.6 (Q5)
Population sample	1 in 17 (Q3)	1 in 2.6 (Q5)	1 in 1 (Q6)

forensic hair comparisons. But in such comparisons we are concerned with the potential for erroneous matches to *particular* hair samples and not to *any* of a large group of hair samples. Thus this question is also only of academic interest. As with question 5, the answer to question 3 depends on the size of the population chosen, so it is basically meaningless. Questions 3 and 5 are included in this discussion simply for the purposes of instruction, completeness, and comparison.

The situation in question 1, where single hairs are compared to single hairs, also does not correspond to the actual situation in forensic hair comparisons since we do not compare unknown hairs to unknown hairs and standard samples generally consist of more than one hair.

Question 2, on the other hand, does ask about a realistic situation as far as hair comparison is concerned—the situation where a single unknown hair is compared to a known sample from a particular individual. Assume, for example, that a single unknown hair is found at a crime scene. Given a suspect who is known to be innocent and was never at the crime scene, the probability that the unknown hair would be found by an experienced hair examiner to be consistent with a scalp hair sample from the suspect would be 1/4500—that is, the probability of an accidental match would be 1/4500. Since the answer to question 6 tells us that hair is not, at present, a positive means of personal identification, the important question facing hair examiners concerns the probability of an accidental match; how reliable and valuable is hair evidence? This is the form in which my previous results [*3,4*] were phrased.

Some of the misunderstanding of my papers arises from confusing my results and the answer to question 2 with some of the other questions, and in wrongly thinking that some of these other questions are appropriate tests of the value of hair evidence.

Discussion of Barnett and Ogle's Paper

Barnett and Ogle [*1*] would appear to have two main criticisms of my work, namely "experimental bias" and "improper statistical treatment of the data."

With regards to the statistical treatments, they state that I have failed to relate probability calculations to the question posed. In particular, they state that I have asked their probability statement II but produced an answer to their probability statement IV. Let us examine this claim.

It should first be noted that Barnett and Ogle have failed to ask a complete set of questions. Of the matrix of six possible questions about Type II errors in hair comparison, Barnett and Ogle have directly considered only three, since their probability statement I concerns Type I errors. Actually, their probability statement II is ambiguous, since whenever they use the words *person, hair,* or *individual,* they have not specified whether they mean "one randomly chosen" or "any from anywhere in the sample." Thus, if one uses the appropriate combinations of modifiers, their statement II can be made to correspond to all six of the questions I have formulated. If we assume they mean what I shall term form IIa, "What is the probability that a *randomly chosen* person will have a *randomly chosen* hair that cannot be distinguished from one *randomly chosen* hair from another *randomly chosen* individual?" then it would correspond to my question 1. Similarly, the following alternate forms of their statement II would correspond to my questions 2 through 6, respectively:

Form IIb: What is the probability that a *randomly chosen* person will have a hair *anywhere on his or her scalp* (or other area) that cannot be distinguished from one *randomly chosen* hair from another *randomly chosen* individual? (Corresponds to question 2.)

Form IIc: What is the probability that a *randomly chosen* person will have a *randomly chosen* hair that cannot be distinguished from one hair from another individual *anywhere in the population sample*? (Corresponds to question 3.)

Form IId: What is the probability that a *randomly chosen* person will have a hair *any-*

where on his or her scalp (or other area) that cannot be distinguished from one hair *anywhere on the scalp* (or other area) of another *randomly chosen* individual? (Corresponds to question 4.)

Form IIe: What is the probability that a *randomly chosen* person will have a hair *anywhere on his or her scalp* (or other area) that cannot be distinguished from one hair from another individual *anywhere in the population sample*? (Corresponds to question 5.)

Form IIf: What is the probability that a person *anywhere in a population sample* will have a hair *anywhere on his or her scalp* (or other area) that cannot be distinguished from one hair from another individual *anywhere in another population sample*? (Corresponds to question 6.)

Since Barnett and Ogle's discussion of the answers to their statement II relates to form IIe of this question, I will assume that their statement II corresponds to my question 5. Table 5 gives the relationship of Barnett and Ogle's probability "statements" to the probability questions I have outlined in the previous section of this paper.

The question posed in my first paper [*3*] was this: "You have testified that the hair found at the scene of the crime is similar to those of the accused's scalp. What are the chances that it could have originated from someone else?" I submit that this question is analogous to my question 2 and that it is also analogous to form IIb of Barnett and Ogle's second statement. However, it *is* different from form IIe of their statement, which was what they discussed. Their assertion that I have answered the wrong question is also mistaken. My papers [*3,4*] posed my question 2 and answered my question 2; Barnett and Ogle in their criticism do not even directly consider question 2.

In addition, their statement that "criminalists are concerned with identifying and distinguishing people, not hair," indicates a failure to understand the extra dimension that distinguishes hair comparisons from most other forms of associative evidence, as well as a confusion of the lay concept of population with the statistical concept.

With regards to my experimental design, Barnett and Ogle would appear to have four main criticisms. First, they state:

> The bias in Gaudette and Keeping's data stems from their confusion of two distinctly different tasks: the task of discriminating between two (randomly selected) hairs and the task of correctly assigning an unknown hair to its true source.... Since there are always observable differences between any two hairs (even from the same individual), the experimental method of Gaudette and Keeping should have resulted in a high rate of success in distinguishing between any two hairs.

These statements indicate a lack of understanding of the process of hair comparison. When an unknown hair is compared to a standard sample, a conclusion that it is consistent with having originated from the same person as the standard sample requires that it:

(a) have characteristics that fall within the range of characteristics observable in the standard sample; and

(b) have no unexplained *forensically significant differences* from at least one hair present in that standard sample.

Requirement (a) is just a preliminary screening for requirement (b). Thus discriminating between hairs is a prerequisite task in assigning an unknown hair to a source. In my experiments I required that there be no major and no more than five minor characteristics dissimilar between hairs if they were to be determined to match. I was therefore looking for *forensically significant* differences, not *any* differences, as implied by Barnett and Ogle. I submit that my experiments were an accurate model of the process of hair comparison and that they do represent and apply to an "exhaustive search" process.

The second criticism given by Barnett and Ogle is of the use of "nonindividualizing fea-

TABLE 5—*Relationship of Gaudette's probability questions to Barnett and Ogle's probability statements.*

Gaudette	Barnett and Ogle
1	IV (and IIa)
2	not discussed (IIb)
3	not discussed (IIc)
4	III (and IId)
5	IIe
6	not discussed (IIf)

tures" in the comparison of hairs. This is also the criticism voiced by Craddock [*2*]. The features chosen for inclusion in my study were those features commonly used by forensic hair examiners. A further discussion and analysis of these features will be the topic of a future paper.

For instance, Barnett and Ogle criticize the inclusion of characteristics of the root. But clearly an anagen hair is different from a telogen hair not only just in the root but often above the root and throughout the rest of the hair. The inclusion of type of root in the characteristics ensures that hairs in different phases of the hair cycle will not be compared to each other.

They also criticize the inclusion of length and characteristics of the tip. Obviously the validity of these characteristics in casework will depend on the circumstances of the case. If a suspect is apprehended shortly after commission of a crime, these characteristics will be very valid. If it is several weeks before the suspect is arrested, they may be of little value. As pointed out in my third paper [*5*], my results are based on the presence of representative standard samples and the ability to use all the microscopic and macroscopic characteristics outlined. If circumstances dictate that all the characteristics may not be valid, the strength of the testimony should be adjusted accordingly. I consider that for the purposes of obtaining a first estimate of the probabilities of an accidental match, the inclusion of such characteristics was certainly valid.

Barnett and Ogle also assert (without referencing any source) that medullary index and medulla type "have not been shown to be related to differences between individuals." However, I submit that hair examiners generally agree that these characteristics are valid. They are, for example, included in an introductory manual used in FBI training courses [*8*].

Barnett and Ogle further criticize the use of cross sections as multiple measurements of the same characteristic. It is my experience that cross sections provide valuable additional information.

However, even if cross sections, medulla characteristics, and the like were totally useless in comparing hairs, the results would not be affected; it would merely mean that we were wasting time looking at these additional characteristics. This is because my method and calculations evaluate hair comparison as a system, not as the product of a number of characteristics. I am not saying, for example, that if 1 in 4 hairs have absent medullas, 1 in 5 are light brown, and 1 in 3 have large pigment granules, the chances of two hairs matching on these three characteristics would be 1 in 60. Again, I feel that Barnett and Ogle and Craddock's criticisms in this regard center on a failure to understand the underlying biological differences that lend an extra dimension to hair comparison and distinguish it from blood comparison and most other forms of associative evidence.

The third criticism of my method offered by Barnett and Ogle is that "the use of dissimilar hairs from each individual placed the examiner in the position of knowing in advance that any match between two hairs was erroneous." This criticism has some validity, although I believe that in practice the effect of this bias would have been slight. In performing the experiments I attempted to apply the same level of discrimination and the same criteria for elimination that I would use in casework, where it is not known in advance that any match between two hairs would be an incorrect one. If the effect of any such bias were major, no pairs of hairs would have been found to be similar, since as Barnett and Ogle themselves state "it is axiomatic that no two hairs are truly identical." In other words, if my level of discrimination were indeed too high as a result of bias, I would easily have been able to eliminate all of the hairs.

Barnett and Ogle also criticize my work for presenting no information on how well the technique works for identifying the correct individual, that is, for not determining the probability of Type I errors in forensic hair comparisons. As pointed out previously, as long as the probability of a Type I error is low, knowledge of the probability of occurrence of Type II errors alone will permit a good estimate of the value of hair evidence to be made. Thus there was no need to determine the exact probability of the occurrence of Type I errors. It was only necessary to show that this probability is low, which I did in my third paper [5]. The final criticism offered by Barnett and Ogle is:

> Gaudette and Keeping ... point out that different examiners get different results with the technique. This means that, even if the data are correct and usable, the numbers must be determined anew by each person who uses them.

Individual examiners did describe some characteristics differently [3] and even committed Type II errors with different hairs [5], but the total result is the same: the *number* of Type II errors committed by different examiners was constant. Thus the probabilities calculated remain valid, particularly since they are just estimates for average cases. The probability numbers do not need to be determined anew by each person who uses them, provided that the user understands them and they are in agreement with the user's personal observations and experience. If we had to personally redo all past research in order to use it, science would never advance.

In my first paper [3], I expressed my hope that it would stimulate other work and further discussion. Thus I was pleased to see the publication of Barnett and Ogle's paper [1]. As previously pointed out, they have made a number of misleading and erroneous statements as a result of their misunderstanding of the process of forensic hair comparison, my papers, and the application of probabilities. I also regret their use of overblown, inflammatory statements, such as "The errors in the derivation introduced a problem to the administration of justice greater than that which the experiments attempted to solve." However, if one can overlook these major faults in their paper, there still remain some valuable contributions to the field. In particular, they have presented a good discussion of associative evidence in general and have delineated the characteristics of good probability statements. They have also worked out probability calculations that provide the answers to my questions 4 and 5.

I am grateful to Barnett and Ogle for stimulating me to further thinking about my previous papers. As a result of this, I have improved my understanding of the subject and I hope I have been able to provide a better discussion of error probabilities in human hair comparison.

Acknowledgments

The advice and assistance of Dr. D. A. Binder of Statistics Canada in Ottawa, Ontario, and Dr. G. R. Carmody of the Department of Biology at Carleton University, also in Ottowa, are gratefully acknowledged.

References

[1] 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.

[2] Craddock, J. G., "A New Approach to the Microscopical Examination of Human Hair," Report 355, Home Office Central Research Establishment, Aldermaston, England, Aug. 1980, pp. 57-62.

[3] Gaudette, B. D. and Keeping, E. S., "An Attempt at Determining Probabilities in Human Scalp Hair Comparison," *Journal of Forensic Sciences*, Vol. 19, No. 3, July 1974, pp. 599-606.
[4] Gaudette, B. D., "Probabilities and Human Pubic Hair Comparisons," *Journal of Forensic Sciences*, Vol. 21, No. 3, July 1976, pp. 514-517.
[5] 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.
[6] Chou, Y. L., *Statistical Analysis*, Holt, Rinehart, and Winston, New York, 1975, p. 10.
[7] Smalldon, K. W. and Moffat, A. C., "The Calculation of Discriminating Power for a Series of Correlated Attributes," *Journal of the Forensic Science Society*, Vol. 13, No. 4, Oct. 1973, pp. 291-295.
[8] Hicks, J. W., "Microscopy of Hairs: A Practical Guide and Manual," Federal Bureau of Investigations, Washington, DC, 1977.

Address requests for reprints or additional information to
B. D. Gaudette
Science and Technology Advisory Group
L Directorate, CPS Bldg.
Royal Canadian Mounted Police
1200 Alta Vista Dr.
Ottawa, Ont. K1A 0R2 Canada