# Abstract

Due to inadequacies in historical hair examinations, the use and perceived value of microscopic hair evidence has reduced. Many reports have heavily criticised the use of pattern-based evidence including hair evidence as being unreliable, with police forces focussing on more individualising evidence. Hair evidence has been utilised in casework for many years due to its ability to transfer easily between individuals and crime scenes. Hair evidence can provide important information in criminal investigations, and in certain circumstances, aid in the identification of an individual, along with providing information to aid in the reconstruction of a crime scene. It has become apparent in recent studies, such as the 2012 FBI review, that hair evidence has been inappropriately utilised in case work and this, along with a preference for more discriminating methods such as DNA profiling, has led to both a lack of confidence in and reduction of use of hair evidence. The need for more standardised and objective methods is required along with improved methods for interpreting hair evidence. Since these reports, there has been no study investigating how these have affected the perceptions of hair evidence by forensic hair examiners or that has investigated what methods are being utilised to create a robust, objective and reliable interpretation of this evidence. This study aimed to assess the current similarities and differences in microscopic hair examinations and perceptions of its evidential value between countries, with a focus on the analysis and interpretation methods used by case work examiners and researchers.

This study will discuss the findings of a survey and interviews completed by hair examiners. Previous surveys have been conducted by Aitken & Robertson (1986) and Murphy (2013), however the current international status of microscopic hair examinations is unknown.

This study gathered responses across 9 countries. From this, it was identified that the microscopic examination of hair evidence is still valued highly by those practicing it. The amount of cases that an examiner has worked on that involve the examination of microscopic hair evidence is positively correlated to the perceived evidential value and inversely correlated to the value of morphological characteristics of hair. There are inconsistencies between the approaches used in interpreting this type of evidence. Subsequently, this study has generated an international understanding of the current status of hair examinations which was not previously known from an operational perspective.

## Key words

Microscopic Hair Evidence

Survey

Interpretation

International

# Introduction

Hair analysis in forensic investigations can provide a wealth of information and has been utilised in casework for many years (Robertson, 2017). Whether human or animal hair, this evidence type may act as associative evidence, be used for intelligence gathering, and also provide activity level information (Lmwinkelried,1982; Fallon, Stone and Petty,1985; Deedrick and Koch, 2004a). Hair examinations may take the form of both morphological analysis and DNA analysis, the former of which involves characterising the features of the hair such as pigment, medulla type and cuticle profile using microscopy techniques. By observing these features, conclusions based upon comparisons with known sources and published data help in determining possible source, including species and body area and provide descriptions of individuals for investigations where suspects have not been identified. (Robertson,1999; Deedrick and Koch,2004; SWGMAT,2005; ENFSI,2015).

Although this form of evidence has been proven useful in criminal investigations, the perceived importance and use of microscopic hair examinations as a stand-alone approach has greatly reduced in recent years. Upon the introduction of DNA analysis techniques, the use of pattern-based methods of identification have come under significant scrutiny, including the examination of morphological features of hair.

In 2012, the Department of Justice (DoJ) and Federal Bureau of Investigation (FBI) launched a review into historical cases where morphological examinations of hair evidence were conducted by the FBI and influenced a conviction. In 2015, it was revealed that in the 268 reports that were scrutinised that contained morphological hair examination, 257 contained erroneous statements; many of which were stating association between crime scene and suspect samples where there was not sufficient evidence to do so (FBI, 2015). A report by the ABS Group (2018), an independent subsidiary of its parent organization, American Bureau of Shipping (ABS), took over the FBI review and concluded that as of June 2018, there were errors in 856 of 1729 reports, 450 out of 484 transcripts and in 31 of 35 examiners testimony with 98% of errors in reports containing erroneous statements. Reasons for these errors may have been caused by failures in many aspects of the recovery, analysis and interpretation of the evidence but generally it is thought that these were primarily due to overstating the conclusions (Garrett and Neufeld,2009; Lee and Pagliaro,2016).

Prior to the FBI (2015) conclusions, a report by the National Academies of Science (NAS) (United States of America National Research Council of the National Academies, 2009), identified a series of factors that have led to unreliable conclusions for evidence that uses feature comparison approaches, including hair examination; these include; inadequate training and educational requirements, lack of standardised procedures and high-quality research in both the scientific theory and validity of methods and poor proficiency testing. These issues noted in the NAS report had previously been identified by Rowe in 2001, where he made suggestions on future developments required if microscopic hair comparisons were to survive future Daubert challenges after this approach was deemed as not suitable for admission of scientific evidence in federal courts upon a Daubert hearing in 1995 (*Williamson v. Reynolds*). These suggestions included the global use of proficiency testing, the further study of error rates and a greater development of scientific literature and data to support conclusions. Similar recommendations are still being made yet it is unclear as to how many analysts and laboratories have taken this onboard.

Following the FBI (2015) review and NAS report, several further studies and reports have critically assessed the failures that have taken place within forensic science techniques, many of which apply to hair evidence. In 2016, the PCAST report (The President’s Council of Advisors on Science and Technology, 2016) was released to address whether additional processes can be applied to forensic science methods to improve the validity of evidence. Within this report, it is claimed that reliability and validity studies from the 1970’s and 1980’s cannot be used to support microscopic hair analysis. Modern studies have investigated the effectiveness of microscopic studies of hair by using mitochondrial DNA (mtDNA) to confirm identification however significant flaws of misidentification were found (Houck and Budowle, 2002).

Murrie *et al* (2019) investigated error rates across a range of forensic science techniques and although this study does not assess error rates in hair examinations in particular, it is useful to note, that morphological hair examination is not alone in its lack of empirical studies into its reliability and validity of methods used and the presence of false positive associations.

The use of hair examinations in forensic casework is not a new debate, Taupin in 2004, evaluated its use historically and critiqued its place amongst more recent forensic techniques. Hair analysis, like many other forms of evidence that lacks an ability to positively identify an individual, suffers from a lack of published data on the commonality of microscopical features within and between individuals and disparity between methods used to help interpret the microscopical data along with a lack of objective methods for assessing these microscopical characteristics. The use of published datasets, collections of data, whether compiled in databases or otherwise, hair reference sets (both physical and digital) along with coding systems for categorising hair features may all aid in the interpretation of hair evidence and overcome some of the criticisms it has incurred. In this study, the authors use the definition of interpretation as described by Cook *et al* (1998) who states, “the drawing of rational and balanced inferences from observations, test results and measurements” and ‘interpretation methods’ as any resource or approach that aids in achieving the aims of Cook’s definition. Studies that attempt to improve the interpretation of hair evidence have been conducted over the years including investigations into inter and intra-variation (Jasuja & Minakshi, 2002; Sato and Seta, 1985) differences between hairs from different body areas (Tolgyesi, Coble, Fang & Kairinen, 1983) the development of datasets to allow for prevalence of hair types and features to be determined (Podolak and Blythe, 1985), the application of statistical approaches to hair examination data (Gaudette and Keeping, 1974; Gaudette, 1976; Wickenheiser and Hepworth, 1990) and the investigation of objective approaches to analyse hair features, such as colour (Verma *et.al.*, 2002; Birngruber, Ramsthaler and Verhoff, 2009; Vaughn, Oorschot and Baindur-Hudson, 2009; Brooks *et.al.* 2011). In addition to these studies, guidelines have been developed in an attempt to reduce the wide variability in analysis and interpretation methods seen between different analysts, institutions and countries. (SWGMAT, 2005; ENFSI, 2015).

It is clear that although the variables being chosen to be examined may have been refined over the years, and there was a push from the late 1980’s to early 2000’s to improve processes and increase knowledge in microscopic characteristics, this has only gone so far in establishing a unified approach to the analysis of hairs globally.

Albeit the criticisms and limitations, there are valid reasons to pursue the scientific study and use of microscopic hair evidence for casework. The use of preliminary microscopic examinations for the purposes of narrowing down a set of hairs for DNA analysis has been proven to be successful (Kowlowski et al, 2004). This same study also indicated its increased success rate in linking hairs to known sources compared to DNA profiling of the root material which is not always successful in obtaining a full DNA profile. Although advances in DNA phenotyping have led to improved prediction of hair colour and hair shape (wavy, straight, curly) (Pośpiech et al, 2018), morphological features can provide further intelligence information valuable to a case. By observing the root shape and growth stage, it can be identified whether a hair was forcibly removed or naturally shed and examining the morphological characteristics for any damage can indicate the nature of the offence such as observing thermal damage in hair exposed to flame or contact heat (Robertson, 1999). In instances where nuclear material is not sufficient and therefore DNA analysis not possible, these morphological features provide the main source of information. This is also true for organisations which do not have access or funds for DNA profiling of hairs and utilise the quicker and less expensive approach of microscopy analysis only.

Past surveys that investigated the use of microscopic hair examinations, are either now out of date or are focussed on one country. Aitken & Robertson (1986(a) and (b)) created and disseminated a survey in the 1980’s which focussed on the recording methods used in the analysis of hair samples. In 2013, Murphy created a survey investigating the analysis of hair evidence and this was disseminated to hair examiners in the USA only. These provide some information about the methods used but offers limited knowledge into examiner’s perceptions of the evidential value of hairs and the characteristics they deem to be of most use when coming to their conclusions. Any changes in these perceptions or methods used have also not been captured post the FBI (2015) and PCAST (The President’s Council of Advisors on Science and Technology, 2016) reports.

The aim of this paper is to assess the current similarities and differences in microscopic hair examinations and perceptions of its evidential value between countries, with a focus on the analysis and interpretation methods used by case work examiners and researchers. Current use of established approaches to improving the validity of any conclusions will be identified, including any newer methods being utilised. This is the first time a study has addressed the perceptions of evidential value of characteristics by forensic hair analysts and methods used to aid in interpretation at a global level. This is particularly pertinent given recent reports and global challenges in comparison-based evidence.

# Material and methods

A mixed method design was used in this study to allow for a wide global reach and maximising data collection. This was achieved via an online survey and then a more detailed exploration of some of the ideas and themes noted in the survey by the use of structured interviews with a sub-set of the participants.

### Participants

Participants were sourced by identifying laboratories or private companies that conduct microscopic hair examinations. Participants were recruited via an email containing the survey link. This link was also forwarded on to the members of the American Society of Trace Evidence Examiners (ASTEE). In total, 35 invitations were disseminated to laboratories and independent consultants across Europe, America, Africa, Asia and Australasia. It was decided that a targeted approach focussing on forensic hair examiners would be preferential over a publicly available survey. This ensured that views were from examiners only and not diluted by participants who have no experience in hair examination for casework. 58 responses were submitted to the survey from 9 countries; Australia (n= 2, 3.4%), Belgium (n= 1, 1.7%), Canada (n= 1, 1.7%), France (n= 1, 1.7%), Ireland (n= 1, 1.7%), New Zealand (n= 2, 3.4%), Sweden (n= 1, 1.7%), United Kingdom (UK) (n= 3, 5.2%) and United States of America (USA) (n= 46, 79.3%). 5.3% of participants were aged between 18-25 years old (n= 3), 35.1% were between 25 and 40 years old (n= 20), 49.1% were between 41 and 60 years old (n=28) and the remaining 10.5% were over the age of 61. Of the survey participants, 67% were females (n= 38) and 33% were male (n= 19). 77.6% (n= 45) of the participant group identified themselves as predominantly casework examiners, 20.7% (n= 12) stated they conducted both casework and research and only 1 (1.7%) participant stating that they only conducted research and development. There were 8.8% of participants stated they had between 0 and 5 years of experience working in forensic science (n= 5), 15.8% had 6-10 years (n= 9), 49.1% had 11-25 years (n= 28) and the remaining 26.3% had over 26 years of experience (n= 15). To identify what proportion of this time they worked in microscopic hair examinations, the number of years of experience in this particular discipline was questioned. From the participant group, 31.6% of participants had 0-5 years’ experience in hair examinations (n= 18), 14.0% had 6-10 years experience (n= 8), 36.8% had 11-25 years (n= 21) and the remaining 17.5% of participants had over 26 years of experience. A large amount of variation was seen in the length of the training period which ranged from 1 week to 7 years with 1 year, 6 months and 2 years being the most common (36.2%, 18.9% and 13.8% of participants respectively). There were two participants who stated that they had received either very brief to no training period at all. The approximate number of cases involving hair evidence the participants had worked on in their careers also varied dramatically. Of the 57 participants who conducted casework, 3.5% of participants having conducted less than 10 cases, 10.5% between 11 and 50, 3.5% between 51 and 100, 29.8% between 101 and 250 and 31.6% having conducted hair examinations in more than 250 cases. Of the participant group, 21.1% declined or were unable to answer this question.

Individuals who opted to leave contact details in the survey were contacted to identify if they were willing to participate in a semi-structured interview based on themes from the survey requiring further investigation. Six interviews were conducted with participants from Australia (n= 1, 16.7%), New Zealand (n= 1, 16.7%), Sweden (n= 1, 16.7%), UK (n= 1, 16.7%) and USA (n= 2, 33.3%). There were 4 out of 6 six participants who conducted both casework examinations and research into microscopic hair evidence (66.6%) and two worked on casework only (33.3%). All participants had worked on over 101 cases (66.7% between 101 and 250 cases and 33.3% on over 250) involving microscopic hair evidence. Half of the participants (50%) had gained accreditation that included the microscopic examination of hairs.

### Survey Procedure

The survey was created using Qualtrics software. The survey was conducted between last quarter of 2017 to first quarter 2018. Prior to completing the survey, participants were provided with a brief overview of the survey including the testing process, any risks and benefits of taking part and confidentiality statements. All data was submitted anonymously and is reported as such in this study. Participants were given the opportunity to refuse to answer any demographic questions and could opt out of open comment type questions but were required to answer all other questions to continue.

Full ethical approval was granted by the University’s ethical committee in relation to the appropriate guidelines.

### Survey Contents

The survey was split into eight parts in the following order; general demographic questions, perceptions of the evidential value of hair evidence, use of guidance manuals, types of hair examinations, use and value of morphological characteristics, interpretation methods, participation and frequency of proficiency testing, research and additional comments. Each of these parts asked a mix of qualitative and quantitative based questions (Likert scale). Where appropriate open questions were asked so as to allow the participant to illustrate their answers with examples and express their opinions in a non-coerced manner, which is a potential limitation of closed option questioning. A breakdown of the questions asked in each part can be seen in figure 1. Where Likert scales were used in questions, these were on a scale of 1-7. Three different scales were used and are noted on figure 1. Where the symbol Δ is indicated against a question, this represents a Likert scale where 1 = no value and 7 = extremely valuable, the symbol ° represents a Likert scale where 1 = strongly agree and 7 = strongly disagree and the symbol + represents a Likert scale where 1 = extremely useful and 7 = extremely useless. Skip logic was used so that those participants who conducted research in hair only (no casework) did not answer questions relating to decisions and interpretation methods used to create conclusions for the purposes of the court. These are indicated with a \* in figure 1.

Figure 1: Flow chart showing the structure of the survey questions. \* represents the set of questions not displayed to participants who indicated that they conducted research only as their main profession, Δ represents Likert scales with a numeric scale, º represents Likert scales showing level of agreement with a particular statement and the symbol + represents a Likert scale where 1 = extremely useful and 7 = extremely useless.

#### Survey Data analysis

Analysis of the survey results was carried out using IBM SPSS Statistics v.25, NVIVO v.11 and R v.3.4.2 software.

For questions that involved Likert scales (denoted by the symbol Δ, ° or + in figure 1), any general trends were identified via the generation of descriptive statistics including the mean, mode, standard deviation (SD) and percentages (%). Qualitative data was subjected to thematic analysis and word clouds were produced via frequency testing.

Following this, the data was then analysed for specific correlations between variables. The evidential value and usefulness of morphological characteristics were then compared by demographic groups. Principle component analysis (PCA) was carried out using R software in order to identify the common factors by reducing variables. Variables to be tested were chosen based on the previous data analysis results. In order to check the suitability of the data, the Kaiser-Meyer Olkin (KMO) test was applied (Hutcheson and Sofroniou, 1999). From the principal component analysis, factor maps illustrating loading of variables onto extracted factors were created and analyses of the individual responses on this feature-space were performed.

### Interviews

Interviews were carried out with a focus on exploring certain areas introduced in the survey including possible solutions to improving the limitations of hair evidence, these included; participant perceptions of the evidential value of hair evidence via case examples, factors affecting conclusions, methods used in interpretation and the effectiveness of these, knowledge and use of any information and data that indicated ‘commonality’ of hair characteristics or types and the use of grading systems for objective hair classifications, including their requirements for interpretation. A further breakdown of the questions can be seen in figure 2.

The interviews were conducted during 2019. Interviews were carried out using Skype or via telephone call, based on the participants preference, and were recorded using a dictation device. Interviewees were not supplied a list of questions but were told the key themes of the interview including the areas described above prior to the interview. Interviewees were given the opportunity to refuse to answer any questions and to provide as little or as much detail as they felt was suitable.

The audio recordings of each interview were then transcribed using online software (Otter.AI) and manually checked, and thematic analysis carried out to identify similar themes between interviewees.

Figure 2: Flow chart showing the structure of the interviews

# Results and Discussion

Fifty-eight respondents completed the survey. Respondents were predominantly individuals who conducted hair analysis in casework and research rather than solely for research purposes. As certain questions did not have a forced response, response rates differed for each question. Due to this, either number of responses are given or where percentage responses are provided the n values (total number of respondents to the question) have been stated. Of the 58 participants, 18 terminated the survey part way through therefore the completion rate of the survey is 40 out of 58. This is discussed further in the limitations section.

The results from both the survey and interview were collated into themes, these are: training activities, proficiency testing, occurrence of hair evidence in casework, perceptions of evidential value of hair evidence, analysis methods, perceived value of hair characteristics, interpretation of microscopic hair evidence, and proposed future improvements to hair examinations. These are summarised below.

### Training activities

When asked what activities had been completed as part of their training (both initial and ongoing), 48 participants responded describing a breadth of activities. The most common activities included completing a literature review of the subject area, conducting human hair comparisons and completion of competency tests. Figure 3 shows the full range of responses to this open text question.



Figure 3: Word cloud showing the range of training activities completed by participants.

Based on the frequency of responses seen in figure 3, it is apparent that training could be themed into passive and active types of learning. Passive learning is described as taking in information through activities such as reading, use of audio and visual aids, for example, use of photomicrographs, and attending lectures or demonstrations. These are important for continued professional development (CPD) so as to ensure the examiner is up-to-date on developments in hair research and new methods being developed. Some participants noted the topics of training they received rather than training activity type, which provided some insight into current areas of interest for CPD; these included animal hair examinations, hair disease, hair recovery and slide mounting, somatic and ethnic origin of hairs and scale casting. Active learning involves learning by participation and collaboration including discussions, teaching others and practical based activities. It has been suggested that this type of learning activity is more effective (Palloff and Pratt, 2005). There were a large range of active learning activities noted by participants, including moot courts, practical exercises, use of reference collections, supervised casework, observations of hair characteristics such as root growth stages and general microscope use. Assessment and test type activities also constituted a large proportion of what participants deemed as training. Proficiency tests, oral examinations, and practical examinations appeared to be a standard part of ongoing training, with only 11 of the 48 respondents to this question not specifically noting being part of any type of assessment or test (1 = UK, 1 = New Zealand, 1= Sweden and 8 = USA). The number of respondents who stated they partook in proficiency tests for this question is not indicative of the actual number who participate in proficiency tests, only those who perceive this as part of their training. To ascertain how many actually partake in proficiency or competency tests, a separate question was asked and described in the next section.

In the interviews, participants were asked if they had received any specific training for the interpretation of microscopic hair examinations, i.e. methods and resources to aid interpretation. All interviewees described this as being only basic and noted that this is a key aspect that needs to be improved. The interviewees also highlighted that many analysts rely solely on low-level microscopy and do not know how to conduct full microscopic examination of hairs including observations of all the microscopic features a hair may present. Training in these two areas could be a way forward to progress hair examinations, particularly using active type learning activities such as practical examination training and mock cases.

### Proficiency testing

As proficiency testing is a clear recommendation of key reports and all analysts should be part of some sort of periodic assessment (SWGMAT, 2005; United States of America National Research Council of the National Academies, 2009; ENFSI, 2015; The President’s Council of Advisors on Science and Technology, 2016) the survey participants were asked if they undertook proficiency testing and, if so, how often with 41 participants answering this question. The results indicate that proficiency testing is commonly undertaken by hair examiners with 90% (37 respondents) declaring that they do participate in proficiency testing of some kind. This is mostly completed on an annual basis (65% (24 respondents)). Nineteen percent (7 respondents) stated that proficiency testing occurs biennially and 16% stated it occurred triennially (6 respondents). The high proportion of analysts who undertake proficiency tests clearly is a positive step towards improvement in hair analysis and is likely to be linked with global recommendations and accreditation needs.

### Occurrence of hair evidence in casework

Although the use of trace evidence such as hairs has been questioned in terms of its validity over recent years (Taupin, 2004; United States of America National Research Council of the National Academies, 2009; The President’s Council of Advisors on Science and Technology, 2016), it appears that this has had minimal effect on the amount of cases that have had hair evidence submitted for analysis. Only 4.2% of participants (2 out of 48 participants who answered this question, one from Ireland and one from the USA) stated that they no longer received hair evidence in their casework submissions. The survey did not investigate any changes in number of submissions, therefore during the interviews, this was re-explored. The occurrence of microscopic hair evidence in casework has generally reduced in the last 10 years, with the most cases per year reported in this study as 10 and the lowest being just 1 case. Participants provided reasons for this reduction as the advancement of DNA profiling methods and the declining reputation of hair evidence. These comments support Taupin’s work in 2004 and it appears that hair evidence, albeit not completely removed from most examiner’s casework, has not recovered since this study.

### Perceptions of Evidential value of Hair Evidence

The evidential value of hairs in different crime type scenarios has been discussed frequently over the years, with hairs being criticised for their lack of individualising ability (Taupin, 2004; United States of America National Research Council of the National Academies, 2009; The President’s Council of Advisors on Science and Technology, 2016). Yet, there have been no studies that ask hair examiners for their perceptions of the evidential value of hair.

In this study, when participants were asked to score the evidential value of hair evidence, in a general sense and then specifically in major, serious and volume crimes the mean scores (n=47) were 4.6 (standard deviation (SD) = 1.42), 5.0 (SD=1.55), 5.1 (SD=1.44) and 4.2 (SD=1.56) respectively. The modal score fell at 5 or above for all categories (figure 4). In major and serious crimes, the value of hair evidence was scored a 6 by the majority therefore indicating its perceived importance in these types of investigations.

Six statements were presented to participants to provide an agreement score (1-7; where 1 = strongly agree and 7= strongly disagree), this can be seen in figure 5. The mean scores (n= 47), along with the SD values, are noted in brackets next to each statement, these were as follows:

* The microscopic examination of hair evidence is subjective (mean =2.9, SD = 1.60)
* The microscopic examination of hair evidence is time-consuming (mean = 2.5, SD = 1.35)
* The microscopic examination of hair evidence is cheap to perform (mean = 2.5, SD = 1.21)
* The microscopic examination of hair evidence is an unreliable method (mean = 5.5, SD =1.57)
* Microscopic methods should only be used as a screening tool prior to DNA analysis of hair evidence (mean = 4.5, SD =2.11)
* Experts should not make positive identifications from this type of evidence alone. (mean= 1.6, SD = 1.27)

Several observations can be made from the results of these statements from identifying the majority scores:

* This form of evidence is subjective and time-consuming and is cheap to perform
* The microscopic examination of hair evidence is a reliable method
* Microscopic methods should not be used only as a screening tool for DNA analysis suitability
* Experts should not make positive identifications from this type of evidence alone.

Additional benefits and limitations were raised by participants via the open text box. Participants noted that microscopic examinations of hair evidence are an investigative tool which can provide significant intelligence to investigative officers such as indicating the nature of the offence by observing the root area, i.e. whether it was forcibly removed. This method can swiftly eliminate samples for DNA testing therefore reducing the cost of the forensic examination and allowing these funds to be spent on testing additional evidence. Conversely, participants noted that hair evidence has limited specificity and cannot provide absolute identity, thus it needs to be interpreted with care.

Figure 4: Bar chart showing the scores given for the perceived evidential value of microscopic hair examinations generally, in major crimes, serious crimes and volume crimes (1 = no value and 7 = extremely valuable).

Figure 5: Bar chart showing the agreement scores given to a series of statements about hair evidence where 1 = strongly agree and 7= strongly disagree)

Further investigation into the perceptions of the evidential value of microscopic hair examinations were carried out in the interviews. A number of common themes emerged from these results:

* Microscopic hair examinations and DNA profiling should be used in conjunction with each other as part of the whole process of hair examinations and not in isolation from one another.
* Generally, examiners had mixed opinions on the criticism of microscopic hair evidence. They agreed that historically, bad reporting has sometimes taken place by examiners but disagreed with the testing of validity of such methods.
* Differences could be seen in the type of crimes where examiners would perform a microscopic examination of hair samples with some only using this for the most serious cases whilst others would apply this method to all crime types providing an adequate reference sample was collected.

Demographic data collected from the survey (amount of hair cases, experience in forensic science, experience in hair examinations and age) were compared against the participants’ perceptions of evidential value of hair evidence (scored 1-7) generally and particularly in major, serious and volume crimes using principle component analysis (PCA). In addition to the evidential value perceptions, the agreement scores for the statements of possible benefits and limitations of hair evidence were also compared to the demographic data.

The KMO test produced a value of 0.7 identifying the data as being suitable for PCA (Kaiser, 1974; Cerny and Kaiser, 1977). Figure 6 shows the scree plot showing the eigenvalues for the extracted principal components. Component 1 accounts for 51.2% of the variance in the data and the point of inflection suggests that components higher than this are not significant. Despite this, it should be noted that components 2 and 3 are greater than one, indicating significance (Field *et al*, 2012). Selection of number of significant components was therefore performed using parallel analysis. PCA of random data of the same dimensions as the questionnaire responses was performed. The scree plot for this data was compared with the questionnaire responses, represented by the broken line in figure 6. This comparison indicates that components 2 – 10 are not significantly better than noise.

The evidential value scores for all crime types are loaded heavily onto component 1 with loadings between 0.84 and 0.91 (table 1). It appears that dimension 1 is likely to represent a participant’s overall confidence in hair evidence. The PCA shows that evidential value (for all crime types) and confidence statements about hair evidence apart from whether a ‘positive ID should not be made’, ‘hair evidence is subjective’ and ‘hair analysis is time consuming’ are loaded onto principal component 1. PC 1 was analysed with demographic data and it was found that that the number of cases that an examiner has completed in hair examinations is positively correlated with PC1 scores. This suggests that having a greater number of cases involving hair evidence in an examiner’s case history is likely to increase their perception of evidential value in major, serious and volume crimes (R2 = 0.3, p = 0.00017). This was found to be independent of the number of years’ experience in forensic science (p = 0.63) and number of years’ experience working in hair analysis (p = 0.19). Additionally, the more cases that an examiner has worked on, the less they agree that hair evidence is unreliable and should be used as a screening test only.



Figure 6: Scree plot for the PCA of evidential value scores

Table 1: Loadings onto PC1 for evidential value scores and hair evidence statements

|  |  |  |
| --- | --- | --- |
| Scoring | Question | Loading onto PC1 |
| Evidential value scores | General evidential value | 0.90 |
| Evidential value in major crimes | 0.91 |
| Evidential value in serious crimes | 0.91 |
| Evidential value in volume crimes | 0.84 |
| Agreement scores | The microscopic examination of hair evidence is subjective | 0.48 |
| The microscopic examination of hair evidence is time-consuming | 0.54 |
| The microscopic examination of hair evidence is cheap to perform | -0.28 |
| The microscopic examination of hair evidence is an unreliable method | 0.81 |
| Microscopic methods should only be used as a screening tool prior to DNA analysis of hair evidence | 0.76 |
| Experts should not make positive identifications from this type of evidence alone | 0.31 |

### Analysis Methods and Perceived Value of Hair Characteristics

A significant proportion of responses (30 responses out of 45 participants who answered question) indicated that a framework of guidance was used however, a quarter of responses stated that they were unsure. Internal standard operating procedures (SOPs) and the Forensic Human Hair Guidelines by SWGMAT (2005) are a common source of guidance (28 responses) however ENFSI’s Best Practice Manual for the Microscopic Examination and Comparison of Human and Animal Hair (ENFSI, 2015) is less frequently used even though it is the most recent publication (9 responses). Additional guidance documents included within responses are;

* Microscopy of Hair – A Practical Guide and Manual (Hicks, 1977) (1 response)
* Atlas of Human Hair Microscopic Characteristics (Ogle and Fox, 1999) (1 response)
* Forensic Examination of Human Hair (Robertson, 1999) (1 response)

In addition to the above, 34 participants also noted that they used internal SOPs to support their analysis. When asked to select the types of examinations that are conducted, participants were provided with pre-determined options that they could select and then add in any other type of examination, these selections included; human hair examination, animal species identification, questioned vs known comparison, DNA testing suitability, racial determination, somatic determination and presence of damage, disease and alterations). Human hair identifications were the most common examination type (44 responses out of 44 participants who answered this question) and racial determination was the least common (31 responses). Figure 7 shows the responses to examination types conducted.

Participants were asked to identify techniques used in the analysis and comparison of hairs in casework (n = 45) and also any techniques they have used in hair research (n = 11). Stereomicroscopy (43 responses for casework and 9 in research), transmitted light microscopy (36 for casework and 8 in research), and comparison microscopy (42 for casework and 7 in research) are methods heavily relied upon by examiners in both casework and research. Scanning electron microscopy (SEM) for casework is rarely used (2 for casework, 3 in research). This is understandable as many examiners may not have access to SEM due to the cost of the equipment compared to other microscope set-ups. Additional methods used in casework were confocal microscopy (1 response), DNA analysis (2 responses), polarised light microscopy (4 responses), reflected high magnification microscopy (1 response), scale casting (2 responses) and visual analysis (4 responses). In research, participants stated that other methods used include confocal microscopy (1 response), DNA analysis (2 responses), oil immersion microscopy (1 response), transmitted electron microscopy (1 response) and visual analysis (1 response).

It was identified in the interviews that all participants (n = 6) create case notes from each examination however the level of detail and manner of recording differs. Only three individuals stated that they use a hair examination form with pre-determined characteristic types. One participant stated that they used the form provided in the appendices of the book ‘Forensic Examination of Hair’ (Robertson, 1999). The three remaining participants stated that they create case notes which include the drawing of sketches and taking photographs of samples.

Participants were asked to note from a list of morphological characteristics (see table 2 for list) which they observed when examining hair evidence. The number of participants who chose each characteristic can be seen in table 2. Although this was not an exhaustive list, it represents commonly observed characteristics for human and animal hair. Morphological characteristics that were least used by participants were shaft profile (31 responses), pigment granule shape (29 responses), scale profile (24 responses), medulla index (21 responses) and scale count (4 responses). The mean scores assigned to the value of each morphological characteristic are shown in table 1. The latter two of these are mainly used for animal hair observations and due to the majority of comparisons in casework being mainly human hair, this would explain why fewer analysts observe these as standard practice.

In addition to which characteristics hair examiners use, their perceptions on the usefulness of these characteristics were also investigated. The weighting of importance against each characteristic when coming to a conclusion in casework is something that can occur both intrinsically and in a holistic manner. Previously, the usefulness of many microscopic characteristics has not been investigated, partly due to the complexity of assessing this. The participants were asked to provide a usefulness score (1 = extremely useful to 7 = extremely useless). Mean and SD values for each of these can be found in table 2. The characteristics deemed most valuable were colour (mean score = 1.6), presence of artificial treatment (mean score = 1.8), presence of disease (mean score = 2.4) and root growth stage (mean score = 2.1) with the least valuable characteristics perceived as scale count (mean score = 4.6) and scale profile (mean score = 3.2). The full usefulness values for each characteristic can be found in figure 8 and table 1. Further comments stated that the usefulness of characteristics is dependent on whether you are examining human or animal hair. For example, scale type and medulla index were noted as useful characteristics when carrying out species determination on animal hair. It was also specified that not one characteristic is useful when considered alone.

As part of the interviews, participants were asked if any characteristics would increase the evidential value of hair evidence if present. Consistency was seen between participants as unusual hair lengths, dye patterns, pigmentation features and certain diseases and damage variations were deemed factors that could increase the evidential value of hair evidence in a case.

Demographic data collected from the survey (amount of hair cases, experience in forensic science, experience in hair examinations and age; these were noted in pre-defined number ranges) were compared against the participants perceptions of values of morphological characteristics shown in table 2 (scored 1-7).

Table 2: Table showing the total number of examiners using each characteristic, the mean and standard deviation of the value scores assigned and the codes used to define morphological characteristics of hair in PCA

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Characteristics** | **Total number of participants who indicated that they use this characteristic when analyzing hairs (n = 45)** | **Usefulness Mean scores (scored 1-7)** | **Usefulness****Standard deviation** | **Code given for PCA** |
| Colour  | 44 | 1.6 | 1.06 | VCOL |
| Cross-sectional shape  | 36 | 2.6 | 1.48 | VXS |
| Cuticle thickness | 38 | 2.9 | 1.32 | VCutThk |
| Hair width | 40 | 2.5 | 1.32 | VHW |
| Length | 44 | 2.4 | 1.19 | VL |
| Medulla distribution | 42 | 2.8 | 1.38 | VMedD |
| Medulla index | 21 | 3.3 | 1.51 | VMI |
| Medulla type | 43 | 2.3 | 1.16 | VMedT |
| Pigment aggregate size | 39 | 2.4 | 1.52 | VPAS |
| Pigment density | 42 | 2.0 | 1.26 | VPDe |
| Pigment distribution | 43 | 2.0 | 1.25 | VPDi |
| Pigment granule shape | 29 | 3.0 | 1.45 | VPGS |
| Presences of artificial treatment | 44 | 1.8 | 1.16 | VATr |
| Presence of cortical fusi | 38 | 3.1 | 1.54 | VCF |
| Presence of damage | 41 | 2.5 | 1.42 | VDam |
| Presence of disease | 33 | 2.4 | 1.57 | VDis |
| Presence of ovoid bodies | 38 | 3.0 | 1.42 | VOB |
| Root growth stage | 42 | 2.1 | 1.32 | VRST |
| Root shape | 41 | 2.3 | 1.30 | VRSH |
| Scale count | 4 | 4.6 | 1.54 | VSC |
| Scale pattern type | 35 | 2.8 | 1.82 | VScPa |
| Scale profile | 24 | 3.2 | 1.61 | VScPr |
| Shaft profile | 31 | 2.9 | 1.47 | VShPr |
| Tip shape | 43 | 2.6 | 1.24 | VTS |

Figure 7: Participants responses to types of hair examination they conduct

Figure 8: The percentage use of morphological characteristics against perceived usefulness scores

Before performing PCA on the responses, a KMO test was performed. The KMO test produced a value of 0.7 demonstrating that the data were suitable for PCA. Figure 9 shows the scree plot produced from the principal component analysis of the usefulness ratings. Parallel analysis was performed which, in conjunction with the point of inflection, suggests a single factor model. The factor map provided in figure 10 shows the distribution of loadings onto dimensions 1 and 2. Component 1 explains 51% of the variance in the data. Correlation of demographic data, as shown by the blue loading vectors in figure 10, show that component 1 is inversely correlated to small effect (R2 = 0.22, p = 0.001) with number of cases. This suggests that the number of cases is inversely associated with responses given to value of morphological characteristics. Therefore, it can be concluded that the more examinations that an examiner has worked on, the more value is placed on the value of morphological characteristics in general. Using the threshold of r > 0.6 as defined by Cohen as ‘Large effect’ (Cohen 1988), it can be identified that the following factors directly correlate to the number of cases that an examiner has worked on; colour (0.8), cross-sectional shape (0.6), cuticle thickness (0.7), hair width (0.8), length (0.7), medulla distribution (0.8), medulla index (0.6), medulla type (0.7), pigment aggregate size (0.8), pigment density (0.9), pigment distribution (0.8), pigment granule shape (0.7), artificial treatment (0.8), cortical fusi (0.7), damage (0.7), disease (0.8), ovoid bodies (0.8), scale profile (0.7), shaft profile (0.8) and tip shape (0.8).



Figure 9: Scree plot of eigenvalues for the PCA analysis of the demographic information against the value of morphological characteristics in hair



Figure 10: Variables factor map for the PCA analysis of the demographic information against the value of morphological characteristics in hair

Some comparisons can be made between previous surveys conducted by Aitken and Robertson (1986(a) and (b)), Murphy (2013) and the present survey.

The survey conducted by Aitken and Robertson (1986(a) and (b)), investigated the perceived evidential value of characteristics in human hair including length, medulla index, scale count and shaft width, by ranking on a scale from 1 – 5 with 1 representing “not at all useful” and 5 representing “very useful.” Differences can be observed between the scores assigned to all of these characteristics from the 1986 survey and the present survey. Previously, little value was placed onto medulla index, whereas this was rated as moderately useful by the examiners in the present survey. Scale count was scored as not at all useful in the previous survey however this has changed to a mean sore of 4 which translates as neutral in usefulness. Shaft width and length both scored a ‘3’ in Aitken and Robertson’s survey therefore sits in a neutral position however these ranked as moderately useful in this survey. These differences show that there has been a slight shift in the perception of the value of certain characteristics. The differences observed could be attributed to the fact that the present study did not differentiate between human and animal examination whereas the previous study focused on human hair examination. Although animal hair still does not constitute the majority of hair types encountered in casework, research into non-human hair as evidence has increased since 1986, with studies investigating transfer (D’Andrea *et. al,* 1998), persistence (Boehme *et al*, 2009) and the creation of guides for animal hair analysis (Deedrick and Koch, 2004b). This may have led to a greater awareness of characteristics used for the comparison of these hairs. It is not possible to identify any changes in perceived value of other characteristics as only four characteristics were measured in the 1986 survey, but it could be hypothesised that there is likely to be a change in perceptions for other characteristics used for animal ID and those that have incurred further research in the subsequent years.

Murphy (2013) enquired about what type of examinations were performed on hair samples and the use of particular characteristics in microscopic examinations in a survey in 2013. Where these characteristics overlap with this study, the percentage number of participants for both the Murphy (2013) survey and current study are compared in Table 3; please note that actual number of participants were not stated in the Murphy 2013 study, therefore % number of participants have been compared. A decrease was seen in individuals conducting examinations which determine racial (82% to 69%) and somatic origin (99% to 87%) and suitability for DNA analysis (99% to 96%). The most dramatic reduction was seen in racial origin determinations with a decrease from 82% to 69% in this study. A reduction in use was seen across all characteristics however this was most significant in pigment granule shape and shaft profile, where percentage number of participants changed from 100% to 64% and 69% respectively. This comparison should be taken with caution as the previous survey only gained data from examiners from the USA which may contribute to the differences seen.

Table 3: Table comparing the percentage use of morphological characteristics between the results of the Murphy (2013) survey and the current survey

|  |  |  |
| --- | --- | --- |
| **Morphological characteristics** | **Percentage (%) of participants from the Murphy (2013) survey using each characteristic**  | **Percentage (%) of participants from the current survey using each characteristic.** **n=45** |
| Artificial treatment | 100 | 98 |
| Colour | 100 | 98 |
| Cortex – cortical fusi | 100 (Combined percentage recorded) | 84 |
| Cortex – ovoid bodies | 84 |
| Cuticle thickness | 100 | 84 |
| Medulla type | 100 | 96 |
| Pigment aggregate size | 91 | 87 |
| Pigment density | 97 | 93 |
| Pigment distribution | 97 | 96 |
| Pigment granule shape | 100 | 64 |
| Presence of damage | 97 | 91 |
| Presence of disease | 79 | 73 |
| Root growth stage | 97 | 93 |
| Shaft profile | 100 | 69 |
| Tip shape | 100 | 96 |

It is interesting to compare the current use and perceived value of morphological hair characteristics by examiners in this study to research that has assessed the usefulness of these characteristics by some means, for example, by their ability to discriminate between individuals or body areas through intra and inter-variation determination.

Scale count was the least used and least valued characteristic perceived by examiners in this study. The value of scale counts in hair analysis was identified by Gamble and Kirk (1941) who examined scale counts on human scalp hairs. From this it was identified that there is little intra variation on an individual’s head but inter-variation between individuals was seen. These observations were criticised by Beeman (1942) who later stated that scale counts are not representative of an individual and these do not show enough differences between individuals to act in a discriminatory fashion. Sato and Seta (1985) also deemed scale count as being one of the least useful characteristics to use for discriminatory purposes due to only small amounts of intra and inter variation being observed. Porter and Fouweather (1975) stated that although little value can be placed on the use of scale counts in human hair comparisons, these are useful in species identification. The value in species identification was also stated for the medulla index which was the second least used characteristic by examiners in this study. Research by Kshirsagar, Singh and Fulari (2009), who investigated the medulla index of various species of animals, found that this characteristic was useful in discriminating between human and animal hair.

The cross-sectional shape of hairs and scale pattern was used by 80% and 78% of examiners respectively in this study. Tolgyesi *et.al.* (1983) found that these characteristics were useful in discriminating between hairs of different regions of the body when studying variation in different somatic regions which may account for the percentage of users in this study. The 1985 study by Sato and Seta, mentioned previously, concluded that because cross-sectional shape shows high intra-variation and scale pattern shows both little intra and inter variation, these are some of the least useful characteristics for discriminatory purposes in human hair identification.

The usefulness of hair colour and the presence of artificial treatments in hair examinations, seen in this study, is also supported by Porter and Fouweather (1975) who state that these can be important characteristics to be used in identifying individuals from their hair. Sato and Seta (1985) also stated that colour is a valuable characteristic in human hair comparison along with pigment density and distribution, tip shape and hair length. These characteristics are all used by over 90% of examiners in the present study and are perceived as valuable.

### Interpretation of microscopic hair evidence

Participants were asked firstly whether they interpret microscopic hair data and then the methods they utilised for aiding the interpretation of hair evidence. A set of predetermined choices were not used for this last question as it was felt that this would influence answers and artificially increase the number of methods actually in use; the limitations of doing this are discussed briefly in the limitations section of this paper.

Forty-four participants answered the question as to whether they interpret microscopic hair data in some form. Just over half (34 responses, 59%) of survey participants stated that they carried out interpretation on microscopic hair examination data, which implies that either many of the analysts do not have to make conclusions on the results they gather or did not understand the question; this is discussed further in the limitations of the study. When identifying the types of approaches and methods used to help in interpreting data, 23 participants responded with a breadth of different answers. There is a large amount of disparity between the methods used by participants with no common method used by examiners. Some examples given include: use of characteristics only (8 responses), verification by other examiners (2 responses), verbal scale (2 responses), no formal method used (1 response). This variation was also seen when asked about the conclusions given after an examination. When asked whether weight was assigned to morphological characteristics (n = 32), Fifty-three percentage (17 responses) of participants to whom this question was applicable indicated that they never assign weight with 16% (5 responses) stating rarely, 9% (3 responses) sometimes, 13% (4 responses) often and the remaining 9% (3 responses) always taking it into account. Ninety-one percent (29 out of 32 responses) of participants took intra-variation into account to some level with 69% (22 responses) stating always, 13% (4 responses) often, 9% (3 responses) sometimes. The remaining 9% (3 responses) of participants stated that they never take intra-variation into account. Commonality of features was considered to some level by 81% of participants (n = 32). This value is broken down into the individual categories: always (15 responses, 47%), often (5 responses, 15%), sometimes (5 responses, 16%), rarely (1 response, 3%) and never (6 responses, 19%). In a follow up question in the interviews, participants agreed that variation in hair samples is more problematic in certain hair types such as colourless or heavily pigmented hairs and that major consideration needs to be applied to the reference samples taken from individuals to ensure that this covers the full range of characteristics present in an individual’s hairs.

The interpretation of hair evidence was the main theme investigated further in the interviews. When the 6 participants were asked what interpretation methods they used for hair evidence, variation was still apparent between the responses given. Some examiners stated that they only holistically compared the range of characteristics present in the questioned hairs to the range present in the known hairs, with no apparent quantification, whilst others added a verbal scale of support to this. No examiner in these interviews used a statistical approach on the data and it is believed that a meaningful method of applying statistics to microscopic hair evidence cannot occur. Participants stated reasons such as “numbers cannot be applied to microscopic features” (1 participant), “characteristics are a form of continuous variation” (1 participant) and that there are “too many variables to consider” (1 participant).

Databases containing population data or transfer and persistence data are not readily available for examiners. However, the general consensus is that these types of databases would be useful when conducting the interpretation of casework examinations. The creation of a grading method (i.e. a standardised numeric scale or predetermined descriptive categories that could describe hair features more objectively) was deemed useful for casework, training and the development of databases. Interviewees stated that this would have to be applied to individual characteristics but would prove to be difficult for certain characteristics such as colour.

Previous research has been carried out into developing new approaches to interpretation methods such as the adaptation of automated digital methods (Podolak and Blythe, 1985; Verma *et.al.,* 2002; Birngruber, Ramsthaler and Verhoff, 2009; Vaughn, Oorschot and Baindur-Hudson, 2009; Brooks *et.al.* 2011) and the application of statistical approaches to hair comparison data (Gaudette and Keeping, 1974; Gaudette, 1976). Interestingly, these published approaches appear to have not translated to actual use in casework, at least for the participants in this study. When exploring this further with interviewees, the use of statistics to interpret hair evidence was stated by only 1 survey participant (Participant in New Zealand) and automated digital methods were not used by any participants in this study.

### Improvements for hair examinations

Participants were asked for their thoughts on current research and whether they had any suggestions to how research may improve hair analysis in the future. Participants indicated areas that they thought should be focussed on for further development of the field, these included investigating the efficiency of microscopic examinations with DNA methods, re-assessing racial characteristics, transfer and persistence type studies, black box studies and improved methods for reporting of hair evidence. In addition to this, any research that aided training generally in hair analysis was welcomed. Three statements were provided with participants asked to state their level of agreement which each (n = 40). These statements and the mean, SD and breakdown of scores for each are:

* ‘There is enough literature available in relation to the morphological examination of hair evidence’ (mean = 2.9, SD = 1.19). The breakdown of each of score for this statement is strongly agree (score 1) = 3, agree (score 2) = 12, somewhat agree (score 3) = 18, neither agree nor disagree (score 4) = 2, somewhat disagree (score 5) = 3, disagree (score 6) = 2, strongly disagree (score 7) = 0.
* ‘Recent failings have led to a reduction in value of this form of evidence, therefore reducing the need for further research’ (mean = 5.0, SD = 2.00). The breakdown of each of score for this statement is strongly agree (score 1) = 4, agree (score 2) = 2, somewhat agree (score 3) = 4, neither agree nor disagree (score 4) = 4, somewhat disagree (score 5) = 5, disagree (score 6) = 10, strongly disagree (score 7) = 11.
* Not enough resources are available to allow further research to be conducted (mean = 4.0, SD = 1.65). The breakdown of each of score for this statement is strongly agree (score 1) = 3, agree (score 2) = 5, somewhat agree (score 3) = 6, neither agree nor disagree (score 4) = 14, somewhat disagree (score 5) = 5, disagree (score 6) = 3, strongly disagree (score 7) = 4.

Areas for improvement in hair examinations was further explored in the interviews. All interviewees acknowledged the need for a better method of analysis and interpretation of microscopic hair evidence. When asked how they would like to improve hair examinations, other than the previously discussed methods, the answers provided were more varied. It was stated by a proportion of examiners (2 participants) that they would like methods to be more time efficient and for a collaborative approach to be implemented for data sharing, e.g. co-creation of databases for interpretation purposes.

Some similarities can be drawn between the recommendations from examiners and from official reports from NAS (United States of America National Research Council of the National Academies, 2009) and PCAST (The President’s Council of Advisors on Science and Technology, 2016). A collaborative approach was also recommended by both reports with both the NAS and PCAST report noting that this would help in the development and advancement of methods and the PCAST report adding that the sharing of databases between laboratories and external agencies would enable research into the reliability of subjective methods. The use of black-box studies was suggested by the PCAST report to evaluate and report on the validity of feature comparison methods. Additionally, both reports stated that the introduction of uniform terminology when reporting results would improve feature-comparison methods and that the accuracy, reliability, and validity of said methods should be tested. It was recommended in the NAS report that all laboratories should achieve accreditation and individuals should gain certification. Clearly, many of the participants in the study agree and support the recommendations made in these global reports but there appears to be limited moves towards making significant changes in procedure and integrating new approaches to achieve these recommendations at this time. This may be partly due to lack of funding to create these resources and a lack of time available that could be dedicated on research as casework will take precedence in analysts’ workloads.

### Limitations of Study

There are a number of limitations that have been identified when considering the results of this study. During data collection, an international perspective of the status of microscopic hair examinations was sought. However, participants only came forward from countries within America, Australasia and Europe. The responses discussed within this paper may not reflect the methods used by examiners in Africa or Asia. A large proportion of participants stated their country of residence as USA (46 participants, 79.3%). Therefore, this could also affect the representation of the results as this could be skewed towards practices predominantly used in the USA. When observing whether there were any significant differences between countries in both PCAs, it was seen that the sample size from countries other than the USA were too small for reliable comparison. As this survey was aimed at hair analysts and designed to preclude individuals who were not/have not been active in forensic hair examinations, this naturally meant sample number is small. The number of hair examinations, and therefore hair analysts, has decreased since the advent of DNA analysis and this will mean there is a smaller pool of individuals available for surveying but also to conduct the research required to meet recommendations by the NAS and PCAST report.

It was found that there was a large drop-off rate at the start of the survey section asking about interpretation methods. The completion rate of surveys typically reduces as the number of questions increases (Fan and Yan, 2010). However, this sudden drop off rate could be due to a number of reasons. This survey required participants to answer the question “Do you interpret microscopic hair examination data?”. As it was not possible to skip this question, this design may have been responsible for a disproportionately high number of aborted surveys. The wording of the interpretation questions may have also resulted an increase in drop-off rates. As seen from the survey results and stated in literature (United States of America National Research Council of the National Academies, 2009), there is a lack of standardised terminology used in hair examinations, particularly with the interpretation and reporting of conclusions. Due to this, how a survey participant translates the term ‘interpretation’ may differ or not be understood at all. It was noted by some participants that they were unsure as what was meant by ‘interpretation’. To ensure this was further understood and explored, the discussion of interpretation methods was included as a main theme for the interviews.

The survey and interviews collected responses from only casework examiners or researchers who carry out microscopic examinations of hair, therefore the results from this study cannot be generalised to forensic scientists who do not conduct this type of examination. Due to this, the measure of perceived value of hair evidence may be different between different forensic experts from different areas and be influenced by factors such as the actual number of court appearances a participant has undertaken and the amount of court feedback a participant has obtained during their career. If non-hair analysts were to provide their perceptions of hair evidence, the results may be different due to general bias as to the value of the evidence they have expertise in. Bias may also be present in the types of people who generally complete questionnaires and interviews. There is a tendency for those with either strong positive or negative opinions to take part in these methods of research which can skew the results to represent extreme views. (McLeod, 2018).

This survey, in its current design, did not allow for participants to specifically indicate whether each morphological characteristic was evidentially valuable for human, animal or both types of hair. By not asking participants to reflect on the value of each characteristic for human and animal hair examinations separately, it was not always clear as to whether participants would value these differently depending on whether they were from human or animal origin. Some participants did state that the value of these would be different depending on what type of examination was being conducted, with scale type and medulla index provided as examples of this, however no values were included to suggest how these features would differ. Due to this, a further study investigating the perceived value of morphological characteristics in human hair and animal hair examinations separately would be useful.

# Conclusions

Microscopic hair evidence has been under scrutiny for over 25 years, with many limitations acknowledged and ways forward recommended. In light of more recent reports and reviews highlighting the potential for erroneous conclusions in hair examinations, it may be believed that methods for interpretation and evaluation have moved forwards and now utilise more objective approaches and empirical data. This study surveyed 58 hair examiners from 9 countries to ascertain the current perceptions in evidential value of hair and its microscopic characteristics and to identify methods used to aid in the interpretation of hair evidence and explore some of the factors that may affect this. Although it appears that certain recommendations made in previous studies and reports (Rowe, 2001; Taupin, 2004; United States of America National Research Council of the National Academies, 2009; The President’s Council of Advisors on Science and Technology, 2016) have occurred including the extensive use of proficiency testing and use of accepted guidelines and protocols to ensure greater standardised practice between examiners, there still appears to be a wide variation in approaches used to aid interpretation with many still relying solely on personal experience rather than also using empirical data, research-informed decisions as to weighting of characteristics and less subjective methods. To conclude, the evidential value of hair evidence is still perceived highly by examiners. PCA of perceptions of evidential value suggest one latent dimension of ‘overall confidence in evidential value of hairs’ in the responses given. Similarly, PCA of ratings of usefulness of individual hair attributes can also be thought of as a single dimension of ‘usefulness of hair evidence attributes’. The extent to which examiners attribute evidential value is correlated with the number of cases in which hair evidence was used in an examiner’s case history. Despite this, there is a lack of standardised approach to interpretation in the form of the methods used and consideration of interpretation factors. Limitations of the survey and interviews are that although multiple countries have been sampled, no participants from Africa, Asia or South America were represented therefore there is not quite the full international perspective gained. Survey bias may also be present in the results due to a tendency for individuals who have strong opinions (positive or negative) to participate in surveys and interviews.

## Reference List

ABS Group (2018) *Root and Cultural Cause Analysis of Report and Testimony Errors by FBI MHCA Examiners.* Virginia: ABS Group.

AITKEN, C.G.G. & ROBERTSON, J. (1986a) The Value of Microscopic Features in the Examination of Human Head Hairs: Statistical Analysis of Questionnaire Returns. *Journal of Forensic Sciences.* 31. (2). Pp.546-562.

AITKEN, C.G.G. & ROBERTSON, J. (1986b) The Value of Microscopic Features in the Examination of Human Head Hairs: Analysis of Comments Contained in Questionnaire Returns. *Journal of Forensic Sciences.* 31. (2). Pp.563 - 573.

BEEMAN, J. (1942) The Scale Count of Human Hair. *Journal of Criminal Law and Criminology.* 32. [5]. Pp. 572 – 574.

BIRNGRUBER, C, RAMSTHALER. F, VERHOFF. M.A (2009) The Color(s) of Human Hair—Forensic hair analysis with SpectraCube®. *Forensic Science International*. 185. (Issues 1–3). Pp. 19 - 23. doi: 10.1016/j.forsciint.2008.12.018.

Boehmea, A., Brooksa, E., McNaught, I & Robertson., (2009) The Persistence of Animal Hairs in a Forensic Context, *Australian Journal of Forensic Sciences,* 41 (2).Pp 99-112.

BROOKS, E., COMBER, B., MCNAUGHT, I. & ROBERTSON, J. (2011) Digital Imaging and Image Analysis applied to Numerical applications in Forensic Hair Examination. *Science and Justice*. 51. (1, March). Pp. 28 - 37. doi: 10.1016/j.scijus.2010.06.008.

CERNY, C.A., KAISER, H.F. (1977), A Study of a Measure of Sampling Adequacy for Factor-Analytic Correlation Matrices. *Multivariate Behavioural Research,* 12 (1), P43-47.

COHEN, J. (1988) Statistical Power in the Behavioural Sciences. Lawrence Erlbaum Associates, New York, NY USA

COOK, R., EVETT, I.W., JACKSON, G., JONES, P.J., LAMBERT, J.A. (1998) A Model for Case Assessment and Interpretation. *Science & Justice.* 38. Pp. 151 – 156. Doi: 10.1016/S1355-0306(98)72099-4.

D’Andrea F, Fridez F, Coquoz R. (1998) Preliminary Experiments on the Transfer of Animal Hair During Simulated Criminal Behaviour. *Journal of Forensic Sciences*. 43 (6). Pp 1257–1258.

DEEDRICK, D.W. & KOCH, S.L. (2004a). Microscopy of hair part I: A practical guide and manual for human hairs. *Forensic Science Communications.* [Online] 6. (1) Available from: <https://www.researchgate.net/publication/318753210_Microscopy_of_Hair_Part_1_A_Practical_Guide_and_Manual_for_Human_Hairs>.[Accessed: 30/10/2019]

DEEDRICK, D.W. & KOCH, S.L. (2004b) Microscopy of Hair Part II: A Practical Guide and Manual for Animal Hairs. *Forensic Science Communications.* [Online] 6. (3)

EUROPEAN NETWORK OF FORENSIC SCIENCE INSTITUTES. (2015) Best Practice Manual for the Microscopic Examination and Comparison of Human and Animal Hair.

FALLON, T.C., STONE, I.C. & PETTY, C.S. (1985) *Hairs on Victims Hands: Value of Examination.* In: Proceedings of the International Symposium on Forensic Hair Comparisons; June 25-27; Quantico (VA). Washington, D.C.: U.S. Department of Justice, 145.

FAN, W., YAN, Z. (2010) Factors Affecting Response Rates of the Web Survey: A Systematic Review. *Computers in Human Behaviour.* 26. Pp. 132 – 139. Doi: 10.1016/j.chb.2009.10.015

FBI. (2015) *FBI Testimony on Microscopic Hair Analysis Contained Errors in at Least 90 Percent of Cases in Ongoing Review.* [Online]. Available from: https://www.fbi.gov/news/pressrel/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review. [Accessed: 25/10/19].

FIELD, A., MILES, J., FIELD, Z. (2012) *Discovering Statistics Using R,* Los Angeles, Sage.

GAMBLE, L.H., KIRK, P.L. (1941) Human Hair Studies: II Scale Counts. *Journal of Criminal Law and Criminology.* 31. (5). Pp. 627 – 636.

GARRETT, B.L., NEUFELD, P.J. (2009) Invalid Forensic Science Testimony and Wrongful Convictions. *Virginia Law Review.*  95. (1). Pp. 1-97.

GAUDETTE, B.D., KEEPING, S. (1974) An Attempt at Determining Probabilities in Human Scalp Hair Comparison. *Journal of Forensic Sciences.* 19. (3). Pp. 599 – 606.

GAUDETTE, B.D. (1976) Probabilities and Human Pubic Hair Comparisons. *Journal of Forensic Sciences.* 21. (3). Pp. 514 – 517.

HICKS, J. (1977) Microscopy of Hair – A Practical Guide and Manual. Washington D.C.; FBI

HOUCK, M.M., BUDOWLE, B. (2002) Correlation of Microscopic and Mitochondrial DNA Hair Comparisons. *Journal of Forensic Sciences.* 47. (5). Pp. 964 – 967.

HUTCHESON, G., SOFRONIOU, N. (1999) The Multivariate Social Scientist: Introductory Statistics using Generalized Linear Models. California, USA: Sage Publication.

IMWINKELRIED, E.J. (1982) Forensic Hair Analysis: The Case Against the Underemployment of Scientific Evidence. *Washington and Lee Law Review.* 39. (1). Pp.41-67.

KAISER. H. (1974) An Index of Factor Simplicity, *Psychometrika,* 39, Pp31-36.

KSHIRSAGAR, S.V., SINGH, B., FULARI, S.P. (2009) Comparative Study of Human and Animal Hair in Relation with Diameter and Medullary Index. *Journal of Forensic Medicine and Pathology.* 2. (3). Pp. 105 – 108.

JASUJA, O.P., MINAKSHI. (2002) A Study of Variations in Some Morphological Features of Human Hair. *Journal of Punjab Academy of Forensic Medicine & Toxicology.* 2. Pp. 1 - 6.

KOLOWSKI, J.C., PETRACO, N., WALLACE, M.M., DE FOREST, P.R., PRINZ, M. (2004) A Comparison Study of Hair Examination Methodologies. *Journal of Forensic Science.* 49. (6). Pp. 1253 – 1255.

LEE, H.C., PAGLIARO, E.M. (2016) Is Hair Reliable Forensic Evidence? *Journal of Forensic Pathology.* 1. (1). Pp. 1.

MCLEOD, S. A. (2018). *Questionnaire*. [Online]. Available from <https://www.simplypsychology.org/questionnaires.html>. [Accessed: 21/11/2019].

MURPHY, N.E. (2013) *Variation between Laboratory Procedures for the Microscopic Examination of Human Hair.* A Thesis submitted in partial fulfilment of the requirements for the degree of Master of Science. Boston: Boston University.

MURRIE, D.C., GARDNERA, B.O., KELLEYA, S., DROR, I.E. (2019) Perceptions and Estimates of Error Rates in Forensic science: A Survey of Forensic Analysts. *Forensic Science International*. 302. Pp. 1 - , Pp1-8. Doi: 10.1016/j.forsciint.2019.109887

OGLE, R.R. & FOX, M. (1999) Atlas of Human Hair Microscopic Characteristics. Boca Raton: CRC Press LLC.

PALLOFF, R. M., & PRATT, K. (2005). *Collaborating online*. San Francisco, CA: Jossey-Bass.

PODOLAK, A.G. & BLYTHE, C.E. (1985) *A Study of the Feasibility of establishing a Computer Data Bank for Hair Characterization using Standard Descriptive Criteria.* In: Proceedings of the International Symposium on Forensic Hair Comparisons; 1985 June 25-27; Quantico (VA). Washington, D.C.: U.S. Department of Justice, 149.

POŚPIECH, E., CHEN. Y., KUKLA-BARTOSZEK, M., BRESLIN, K., ALIFERI, A., ANDERSEN, J.D., BALLARD, D., CHAITANYA, L., FREIRE-ARADAS, A.D, VAN DER GAAG, K.J., GIRÓN-SANTAMARÍA, L., GROSS, T.E., GYSI, M., HUBER, G., MOSQUERA-MIGUEL, A., MURALIDHARAN, C., SKOWRON, M., CARRACEDO, Á., HAAS, C., MORLING, N., PARSON, W., PHILLIPS, C., SCHNEIDER, P.M., SIJEN, T., SYNDERCOMBE-COURT, D., VENNEMANN, M., WU, S., XU, S., JIN, L., WANG, S., ZHU, G., MARTIN, N.G., MEDLAND, S.E., BRANICKI, W., WALSH, S., LIU, F., KAYSER, M. (2018) Towards broadening Forensic DNA Phenotyping beyond pigmentation: Improving the prediction of head hair shape from DNA*.*  *Forensic Science International: Genetics*. (37). Pp. 241 - 251. doi: 10.1016/j.fsigen.2018.08.017.

ROBERTSON, J. (1999) Forensic Examination of Hair. London: Taylor & Francis.

ROBERTSON, J. (2017) Managing the Forensic Examination of Human Hairs in Contemporary Forensic Practice. *Australian Journal of Forensic Sciences.* 49. (3). Pp. 239 - 260. Doi: 10.1080/00450618.2017.1279838

ROWE, W.F. (2001) The Current Status of Microscopical Hair Comparisons, Mini-Review. *The Scientific World*. (1). Pp. 868 – 878. Doi: 10.1100/tsw.2001.358

SATO, H. & SETA, S. (1985) *An Appraisal of the Use of Macroscopic and Microscopic Data in Japanese Head Hair Comparisons.* In: Proceedings of the International Symposium on Forensic Hair Comparisons; 1985 June 25-27; Quantico (VA). Washington, D.C.: U.S. Department of Justice, 143.

SCIENTIFIC WORKING GROUP ON MATERIALS ANALYSIS. (2005) Forensic Human Hair Examination Guidelines. *Forensic Science Communications.* 7. (2).

TAUPIN, J.M. (2004) Forensic hair morphology comparison - a dying art or junk science? *Science and Justice*. (44). Pp. 95 – 100. Doi: 10.1016/S1355-0306(04)71695-0

TOLGYESI, E., COBLE, D.W., FANG, F.S. & KAIRINEN, E.O. (1983) A Comparative Study of Beard and Scalp Hair. *Journal of the Society of Cosmetic Chemists.* 34. (November). Pp. 361 - 382.

UNITED STATES OF AMERICA. NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES. (2009) *Strengthening Forensic Science in the United States: A Path Forward.* Washington D.C: The National Academies Press.

UNITED STATES OF AMERICA. THE PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY. (2016) *Report to the President Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods.* Washington D.C: PCAST.

VAUGHN, M.R., OORSCHOT, R.A.H. & BAINDUR-HUDSON, S. (2009) A Comparison of Hair Colour Measurement by Digital Image Analysis with Reflective Spectrophotometry. *Forensic Science International.* 183. (1-3, January). Pp.97-101. doi: 10.1016/j.forsciint.2008.11.002.

VERMA, M.S., PRATT, L., GANESH, C. & MEDINA, C. (2002) Hair-MAP: A Prototype Automated System for Forensic Hair Comparison and Analysis. *Forensic Science International.* 129. (3, October). Pp .168-186.

WICKENHEISER, R. A., HEPWORTH, D. G. (1990) Further evaluation of probabilities in human scalp hair comparisons. *Journal of Forensic Sciences*. 35. Pp. 1323 - 1329.

Williamson v. Reynolds [1995] 904 F. Supp. 1529 (E.D. Okla.) [Online] Available from: <https://law.justia.com/cases/federal/district-courts/FSupp/904/1529/1763070/> [Accessed 12/11/19]