Author(s): Delahunty, J. ; Gumbert, T.A. ; Hale, S.

Title: The biasing influence of linguistic variations in DNA profiling evidence

Journal: Australian Journal of Forensic Sciences

ISSN: 0045-0618 **Year:** 2014 **Pages:** 348 - 360

Volume: 46 **Issue:** 3

Abstract: Extensive psychological research has confirmed that probabilistic statistical information that is mathematically equivalent, if presented in different linguistic forms, is not psychologically equivalent. Commentators have argued that certain forms of evidence are more prejudicial than probative and should be excluded from trial. In a recent homicide case based on circumstantial evidence, the sole evidence linking a suspect to crime scene was mitochondrial DNA profiling evidence derived from a si ...

THE BIASING INFLUENCE OF LINGUISTIC VARIATIONS IN DNA PROFILING EVIDENCE

Jane Goodman-Delahunty,[1] Thea Gumbert[2] and Sandra Hale[3]

Abstract

Extensive psychological research has confirmed that probabilistic statistical information that is mathematically equivalent, if presented in different linguistic forms, is not psychologically equivalent. Commentators have argued that certain forms of evidence are more prejudicial than probative and should be excluded from trial. In a recent homicide case based on circumstantial evidence, the sole evidence linking a suspect to the crime scene was mitochondrial DNA profiling evidence derived from a single loose hair found at the crime scene. Variations in the form of the linguistic evidence presented by DNA forensic scientists and research about its differential impact were cited as the basis for appeal of the conviction to the Australian High Court. A review of the research revealed that presentation of statistical information using natural frequencies is a recommended best practice. Random match probabilities expressed as frequencies are less susceptible of misinterpretation by legal professionals and lay jurors, and lead to fewer convictions.

Corresponding author:

Professor Jane Goodman-Delahunty, JD, PhD
Australian Graduate School of Policing and Security
Faculty of Arts
Manly Campus
Manly NSW 1655
Tel: +612 9752 9017
Mob: 0421 879 411
Email: jdelahunty@csu.edu.au

[1] Research Professor, Charles Sturt University, School of Psychology and Australian Graduate School of Policing and Security, Manly Campus, Manly, New South Wales
[2] Research Assistant, Charles Sturt University, Manly Campus, Manly, New South Wales
[3] Professor of Interpreting and Translation, School of International Studies, University of New South Wales

Introduction

On 27 November, 2005, the body of Ms Sevda Bayrak was found near the front door of her apartment in North Parramatta, Sydney. She had been stabbed 34 times. The chief suspects in the case were her current partner, Erdal Tunc, and her ex-boyfriend, Yusuf Aytugrul. Whilst Mr Tunc admitted having spent the previous evening with Ms Bayrak in her apartment before leaving around 8pm, Mr Aytugrul maintained that he had had little or no recent contact with Ms Bayrak, and had never been inside her apartment, into which she had moved just four weeks earlier. Crime scene investigators found a loose hair under the fingernail of Ms Bayrak and submitted it for DNA testing. There was no other physical evidence, or indeed other evidence that directly placed Mr Aytugrul at the scene of the crime.

As it was not possible to conduct a nuclear DNA analysis upon the hair, a mitochondrial DNA analysis was performed. Mitochondrial DNA analysis has a much lower power of discrimination than does nuclear DNA (for example, the contributor's gender cannot be determined) and therefore does not yield such determinative match results. The DNA profile yielded by the analysis of the hair excluded Tunc, but was reported as a match for Aytugrul. The DNA profiling expert who conducted the test on the hair reported that the likelihood of a random match between the crime scene sample and someone in the general population was one in 1680 based on consultation of a database containing 4,839 individuals. This figure was generally accepted by other experts for both the prosecution and defence who testified at trial, and was also briefly expressed interpreted to mean that 99.9% of the population would not be expected to share the same DNA profile, and could be excluded as suspects (exclusion percentage) (*Aytugrul v The Queen*, 2010).

However, in consideration of the fact that the suspect, the deceased, and the bulk of her friends and relatives were of Turkish origin, the DNA profile match statistics were also interpreted for a sub-population of Turkish individuals. In these analyses, the prosecution and defence experts arrived at a "true match probability" between approximately one in 100 and one in 200, but agreed upon a "cautious upper limit" of one in 50 persons of Turkish origin (Respondent's submissions, *Aytugrul v The Queen*, 2011:s6.23). In all, in this case, a range of linguistic formulations were used by the expert witnesses to convey the incriminating value of the DNA evidence to the jury: both 1 in 1680 for the general population and a corresponding exclusion rate of 99.9%, and a random match probability of one in 50 for the Turkish sub-population.

Probative or prejudicial evidence?

The DNA profiling evidence comprised a significant component of the overall prosecution case against Mr Aytugrul on charges of homicide. Since there were no witnesses to the crime, the evidence was purely circumstantial. It included information that the defendant had refused to accept the termination of his romantic relationship with the deceased, two years prior, and that since then he had stalked, harassed and threatened her, with behaviours of this nature escalating in the weeks leading up to her death. Jurors were shown copies of text messages indicating that the defendant had been covertly observing the deceased late into the night, and messages he sent to her containing veiled threats to her physical safety. The accused had contacted Centrelink to report her for alleged welfare fraud. He also authored a poem which has been published in the local Turkish Australian newspaper entitled "I Can Not Give Up", containing the lines; "My heart burning for you like

mad…Even if you say give up, I can not give you up" (Respondent's Submissions, *Aytugrul v The Queen,* 2011: s6, para 6.48). The deceased's car had been vandalised on several occasions in the months prior to her death, and on one occasion, the brake lines were cut. Relatives of the deceased testified that she had reported being subjected to unannounced visits by the defendant and that she had feared for her safety after noticing unusual events which suggested he was trying to access her apartment at night, and that she had actually moved to her new apartment in an attempt to prevent further contact with him. The defendant persisted and eventually succeeded in locating her address through various individuals and government agencies. In the course of the criminal investigation, the defendant had lied to police interviewers about his relationship with the deceased, his actions and whereabouts in the days and weeks preceding her death.

In light of the multiple sources of circumstantial evidence, the extent to which the jury relied upon the DNA profiling evidence in determining to convict Yusuf Aytugrul of the murder of Sevda Bayrak is difficult to ascertain. The Crown contended that "The statistical [DNA] evidence was presented as but one part of the circumstantial case. Its significance lay in the fact that the suspect population was relatively small, namely, a person or persons who would kill Ms Bayrak for no reason" (Respondent's Submissions, *Aytugrul v The Queen,* 2011, para 6.43).

The conviction of Yusuf Aytugrul was appealed on the basis that the presentation of statistical information by the DNA profiling expert for the Crown, specifically the exclusion rate of 99.9% for the general population, was misleading to the jury and prejudicial to the defendant. In his submission to the High Court, the appellant's counsel argued that there was no need to introduce the exclusion percentage of 99.9%, as this was merely a different way of stating the random match probability of one in 1680, which had already been presented to the jury. Although it was conceded that these two figures were simply a mathematical inversion of each other, the appellant's submission emphasised that the two different statements could be psychologically interpreted and understood in very different ways. The submission contended that "The 99.9% figure invites a subconscious rounding up to 100%" and created "a residual risk of unfairness deriving from the subliminal impact" of this figure (Appellant's Submissions, *Aytugrul v The Queen*, 2011:s36, para 36). Moreover, the appellant pointed out that the judge had repeatedly, though inadvertently, expressed the figure as 99.99% in summing up the evidence to the jury, demonstrating that percentage figures with decimals inherently increase the risk of "inadvertent imprecision," (2011:s38). Finally, the appellant made reference to the "prosecutor's fallacy" meaning that jurors may have been prone to confuse the exclusion percentage of 99.9% with the likelihood 99.9% that the defendant was in fact the source of the DNA hair found at the crime scene, or that the jury would infer that this number represented the likelihood that the defendant was guilty of the crime charged.

Coulthard (2011) argued that jurors have great difficulty understanding likelihood ratios or probability figures presented by expert witnesses, especially in DNA evidence, and they will inevitably reach 'semantically encoded decisions'. Linguistic theories of pragmatics suggest that "the way a speaker chooses to encode some piece of information is in part a function of the speaker's evaluation of the hearer's knowledge as well as of the inferences the hearer can be expected to make" (Prideaux, 1991:117). In other words, even if the expert witness or the judge did not intend to deliberately mislead the jury, jurors will use their 'common sense' to make meaning of a technical term, inferring a lay understanding of the meaning of 99.9%. Research into juror comprehensibility of instructions and expert evidence (Charrow and Charrow, 1979; Dumas, 2000; Levi, 1993; Napier et al, n.d; Saxton, 1998) has

found that jurors rarely fully understand technical explanations, due to linguistic features such as the use of the passive voice, technical terminology, complex syntax, multiple negation and excessive embedding. These results tend to indicate that explanations by the expert witness of the exclusion percentage may not have been fully comprehended by the fact finders.

As the case of *Aytugrul* demonstrated, an increasing area of focus and contention in criminal cases in many jurisdictions internationally is the linguistic form of the presentation of DNA match statistics to triers of fact, whether they are judges or lay jurors. A substantial body of psychological research has demonstrated that variations in the presentation of probabilistic statistical evidence influence the extent to which triers of fact find the evidence persuasive, and that changes in the presentation format alone can lead to changes in the weight accorded to the inculpatory DNA evidence, and as a consequence, changes in conviction rates (Koehler, Chia & Lindsey, 1995; Koehler, 2001; Koehler, 2001b). These results are consistent with the large body of forensic linguistic research into the significance of language in the adversarial system, where lexical, syntactic and discoursal features found in questions can affect the answers elicited (Hale, 2004/2010; Harris, 1984; Loftus & Palmer, 1974; Woodbury, 1984), presentational style of answers can affect the fact finders' assessments of credibility (Conley and O'Barr, 1990; Coulthard & Johnson, 2007; O'Barr, 1982), and different linguistic features in jury instructions can affect juror comprehensibility (Charrow and Charrow, 1979; Levi, 1993; Napier et al, n.d; Saxton, 1998).

This article reviews recent research findings on the form of probabilistic statistical evidence, in particular their application in cases in which inculpatory DNA match statistics are presented. The legal issues that flow from the linguistic variations were illustrated in the foregoing Australian homicide case in which the potentially biasing or prejudicial effects of exclusion percentages presented by expert witnesses to jurors was the basis of the appeal (Appellant's Submissions, *Aytugrul v The Queen*, 2011).

The power of inculpatory DNA profiling evidence

The introduction by the prosecution of incriminating DNA profiling evidence in a criminal case is a significant predictor of the likelihood that a criminal case will proceed to trial and result in a conviction (Briody, 2004). An archival analysis of sexual assault cases revealed that convictions were 23-33 times more likely when DNA profiling evidence was admitted than in matched cases where no such evidence was introduced (Briody, 2003). Research conducted in several jurisdictions using controlled trial simulation experiments with mock-juries has further illuminated the powerful influence of DNA profiling evidence upon mock-juror judgments (Dartnall & Goodman-Delahunty, 2006; Goodman-Delahunty & Hewson, 2010a; 2010b, Lieberman et al., 2008). For instance, the mere presence of DNA profiling evidence in circumstantial cases in which no other direct evidence linked the suspect to the alleged crime produced conviction rates 15 times higher than that in a parallel control case without this evidence (Dartnall & Goodman-Delahunty, 2006, 2006). Similar results were obtained in a study in response to different case facts and different random match probabilities presented by a forensic scientist on the results of the DNA tests, i.e., the conviction rate of 23% when the DNA tests results were described as inconclusive increased to 59% when the expert presented results that the likelihood of a random match between the sample from the crime scene and the suspect was one in one billion (Goodman-Delahunty & Hewson, 2010). Jury-eligible participants in another study rated DNA evidence as more influential upon their verdicts than any other type of evidence, including a confession by the

suspect, eyewitness testimony, or other types of physical evidence such as blood typing or fingerprints (Lieberman et al., 2008). By varying whether the DNA profiling evidence was inculpatory or exculpatory in a simulated case using the same case facts, studies have shown that the presentation of inculpatory DNA evidence resulted in a 100% conviction rate whereas exonerating DNA evidence resulted in a 100% acquittal rate (Lieberman et al. 2008).

Legal and psychological researchers have concluded that jurors may attribute a "special aura of credibility" to forensic scientific evidence presented in the context of criminal trials (Schklar & Diamond, 1999:159). One difficulty is that jurors may be prone to misunderstandings about the types of inferences that can be drawn from DNA profiling evidence. For example, jurors often believe that DNA evidence can conclusively determine whether the suspect was the source of the DNA samples taken from the crime scene (i.e., the *source probability),* and whether or not the suspect, as opposed to another person, committed the crime (i.e., *guilt probability)* (Koehler, 1996). In fact, DNA profiling evidence does not allow these inferences. DNA profiling evidence relies upon a chain of inferences, whereby a reported match is inferred to be a true match, from which it is inferred that the suspect is the source of the DNA, that he was present and at the crime scene and is therefore guilty (Koehler et al. 1995). However, potential for error exists during each step of this chain. A reported match may not be a true match in the case of laboratory error, a true match may not mean that the suspect is the source if the match is adventitious or coincidental, the source of the DNA may not have been present at the crime scene in cases of contamination or evidence tampering, and even if they were present, may have left DNA traces in an innocent way, such as in cases where a victim and a suspect previously shared housing (Koehler et al., 1995). At best, DNA profiling evidence allows for estimation of the relative size of the population that may have contributed the DNA profile (Koehler, 1996; Twenty Scholars of Forensic Science, 2009).

DNA profiling evidence in particular is often inaccurately portrayed and regarded as an infallible procedure that is capable of uncovering absolute truth in criminal investigations. Several examples of criminal trials in the USA were identified in which jurors were encouraged to draw assumptions of DNA infallibility by forensic experts who explicitly instructed jurors that incorrect matches were an "impossible" result; that DNA analysis was "failsafe" or that "there is no way to get a false positive with this technology," (Koehler, Chia & Lindsey,1995:211). Another group of psychological researchers warned that "this type of scientific evidence may be so persuasive that its mere introduction in a criminal case is sufficient to seriously impede defence challenges," (Lieberman et al., 2008: 58). As a consequence, attention has shifted to the manner in which forensic scientists present the results of DNA profiling evidence in court, both orally and in written reports.

Linguistic research supports the assumption that the manner in which testimony is presented can greatly influence jurors' perceptions of the veracity of the evidence. O'Barr's (1982) seminal study, showed that those witnesses who used what was described as a 'powerful' speech style were rated as being more credible, more trustworthy and more intelligent than witnesses who used what was described as the 'powerless' style. The labels 'powerful' and 'powerless' were chosen to indicate the social status of the witnesses. Those who spoke in the powerful style were predominantly professionals from a high socio-economic background. In this case, the high status of the DNA evidence may be enhanced by the powerful speech style of the expert witness. Other studies have shown the effect of lexical choice on witness memory. The study by Loftus and Palmer (1974), for example, found that the connotations of different terms with the same denotations caused eye-witnesses to assess

events differently. In this study, subjects viewed a video of a car accident and then estimated the speed at which the cars had been travelling. Subjects who were asked "About how fast were the cars going when they *smashed* into each other?" estimated significantly higher speeds than did those who were asked the same question, but with the verb "smashed" replaced by "hit", "bumped", "collided" or "contacted". Moreover, when asked one week later whether they recalled seeing broken glass at the scene of the accident, subjects who had been asked about the speed of the "smash" were more likely to respond in the affirmative, even though no broken glass was present in the video. Similarly, the use of the definite article 'the' led to different recollections of the existence of certain articles in the witnessed scene than did the use of the indefinite article. These results indicated that subtle linguistic differences in the presentation of evidence impacted the way jurors comprehended and evaluated features of evidence.

Random match probabilities

The presentation by forensic scientists of the results of DNA profiling tests typically includes three elements: a description of the sources of the samples tested; an overview of the test procedures and outcomes, i.e., whether a "match" occurred; and some interpretation of the results. Most typically, the interpretation of the results entails presentation of a probability statistic to express the strength of the evidence in terms of the likelihood of obtaining these results by chance alone, or the probability of a random match between the sample found at the crime scene and the sample taken from the defendant. The random match probability (RMP) refers to the probability that a DNA sample from a person randomly drawn from a reference population, will match the DNA profile of the obtained sample. At trial, RMPs are often expressed in the order one in many millions, billions, or even trillions (Koehler et al. 1995), and the technology has advanced to the point that expressions of one in a septillion are feasible (Walsh, 2011). Despite widespread use of DNA profiling evidence in criminal forensic proceedings globally in the past two decades, no clear guidelines have emerged as to how RMPs are calculated and presented. "The courts do not have the technical guidance needed to make sensible and consistent judgments concerning the admissibility of RMPs," (Koehler et al. 1995:203). The shortcomings and limitations of DNA evidence, including the way that random match probabilities are calculated and presented to juries has been the subject of extensive psychological research given the serious consequences for the legal system of common misconceptions and misunderstandings of this evidence (Koehler, Chia & Lindsey, 1995).

The "Prosecutor's Fallacy"

Random match probabilities can be misleading to juries in terms of how they apply to a particular defendant. For example, upon hearing that a particular profile has a random match probability of 1 in 1 trillion, a juror may conclude that no other individual on earth could share that profile, when in fact the chance that others could share the profile is less than 1 in 200. For an RMP of 1 in 1 billion, the chance that others could have a matching profile shrinks to 99.6% (Koehler, 1996). The error of inverting the RMP and equating it with the culpability of the defendant is well-documented, and is commonly known as the "Prosecutor's Fallacy" (Twenty Scholars of Forensic Science, 2009) or the "inverse fallacy" (Koehler, 1997a).

Laypersons and legal professionals are at risk of conflating the random match statistics with issues of source probability, by reasoning, for example, that the RMP of 1 in 1

million equates to a 1 in 1 million chance that the true source of the DNA profile is someone other than the defendant (Koehler, 1997a). At times, the prosecutor's fallacy has been introduced in evidence by testifying DNA experts. For instance, in a US sexual assault case, a prosecution expert attested that an RMP of 1 in 3 million equated to a 0.000033% likelihood that the defendant was *not* the source of the DNA (20 Scholar of Forensic Science, 2009). The case was successfully appealed to the US Supreme Court (*McDaniel v. Brown*, 2010).

Potentially prejudicial statistical presentations

The contention by Koehler et al. (1995) has been that RMPs presented to jurors may be both irrelevant and potentially prejudicial, as they not only exploit human errors in mathematical reasoning, but fail to take into account laboratory error rates. Citing results from laboratory proficiency tests carried out in the USA, the authors noted that laboratory or human error occurs in 1-2% of DNA analyses, and are actually far more important to juror considerations of DNA match statistics than is the RMP itself, however this "lab error rate" (LE) is very rarely mentioned in court. In other words, mathematically and practically, the relevance of a large RMP of one in one billion is diagnostically small compared to a 2% LE: "as the discriminatory power of a profiling technique increases, errors caused by coincidental circumstance become increasingly insignificant to errors that arise for other reasons… [and] increases in discriminatory power no longer increase the diagnostic value of a reported match" (Koehler et al., 1995:216). Put another way; "if an error occurs, say, 1 time in 100, it makes no difference whether the DNA frequency statistic is 1 in a million or 1 in 57 billion. The chance that a reported DNA match is erroneous is about 1 in 100 regardless of how small the DNA frequency statistic becomes," (Koehler, 1997a:222). Research has shown that even if LE rates are presented to jurors, these values tend to be under-weighted due to flawed mathematical reasoning whereby the RMP and LE are combined by averaging, a strategy which "substantially over-weights the contribution of the smaller error source" (Koehler et al. 1995:212).

Jurors tend to over-weight RMPs because the very small match statistics tend to be vivid and memorable, and may therefore exert a strong influence upon verdicts. In a series of controlled studies, Koehler et al. (1995) examined the influence of RMPs, LE rates and aggregated error rates upon conviction rates in samples of students and jurors. Conviction rates were significantly higher when RMPs were provided, occurring at double to triple the rate than of decisions to convict in conditions where no RMP was presented. No significant differences in conviction rates emerged based on the presence or absence of LE rates in conjunction with the RMP, suggesting that jurors did indeed tend to over-weight the contribution of the RMP when left to their own devices to integrate expert evidence about combining the RMP and LE. However, conviction rates among participants who received an aggregated error rate reflecting the true value of the combined RMP and LE were significantly lower, and were similar to rates observed in conditions where no RMP was presented.

In light of these results, the researchers concluded that the probative value of RMPs may be outweighed by their prejudicial impact, and cautioned that "scientists should not be permitted to describe the significance of a reported DNA match in terms of RMPs or using vague comments about the improbability of laboratory error", as the evidentiary value of the RMP was "effectively trumped" by the LE (p.217). They further emphasised the importance of requiring DNA profiling experts to distinguish between "a match" and a "report of a match" by acknowledging the possibility of human or laboratory error because "the probative

value of a reported DNA match is more closely linked to the scientists' ability to avoid committing errors than it is to the tiny theoretical probabilities that forensic scientists testify to in open court" (Koehler, 1997a:222).

Exemplar cueing theory and the linguistic presentation of statistical evidence

A considerable body of research has demonstrated that the capacity of experts and non-experts to draw appropriate inferences from statistical information is influenced by the form and syntactical order of the presentation of the numerical information (Hoffrage, Lindsey, Hertwig & Gigerenzer, 2000; Koehler, 2001). For instance, the probability of a random match may be framed in terms of frequency (e.g. "1 in 1,000,000") or in terms of probability (e.g. ".000001"). Mathematically, the statistics can be presented as probabilities, percentages, or absolute frequencies, and the change in the form of the presentation does not modify the mathematical value of that information (Hoffrage et al., 2000). However, a change in the form of the presentation changes the psychological value and weight of that information as changes in the linguistic presentation stimulate different types of reasoning and different inferences. Statistics framed as natural frequencies encourage a broad or "outside" view of events, which are perceived as but one instance among a larger group of similar events. Statistics framed as probabilities encourage a narrower consideration of a single case in isolation (Kahneman & Lovallo, 1993). In the context of DNA profiling match statistics, an interpretation framed in the language of probability tends to be more persuasive and accorded more weight by the trier of fact because it highlights the suspect's chance of matching by coincidence alone. However, when framed in the language of frequencies (e.g. 1 in 1,000), the information is less persuasive and is accorded less weight as it draws attention to the fact that others in addition to the accused have a matching DNA profile (Koehler, 2001). In essence, the results of this body of research determined that jurors' ability to understand and apply expert evidence on DNA profiling statistics may be a function of the linguistic formulation or manner in which the expert presents the numerical information rather than the content of that information. In other words, linguistic variations in providing quantitative information had profound psychological but not mathematical consequences (Koehler, 2001b:497).

As a consequence, it is possible to shift the focus or "target" of a DNA match statistic, by varying the linguistic form of the presentation of the quantitative information either towards a single individual or towards a larger population. In a single-target approach, a DNA statistic is framed in terms of the likelihood that the profile of the suspect will match the sample from the crime scene by coincidence, if he or she is not the source of the DNA. Conversely, in a multi-target approach, the statistic will be framed in terms of the likelihood that the DNA profiles of others in a large, reference population will also match the crime scene sample, broadening the focus of considerations by finders of fact explicitly drawing attention towards other persons.

To further demonstrate the various ways that DNA match statistics may be manipulated to exert more or less incriminating influence on jurors, Koehler (2001) presented the following example of the same mathematical statistical information in four different ways, employing frequency versus probability statistics, and single-target versus multiple target approaches, as follows:

> *The probability that the suspect would match the blood drops if he were not the source is 0.1%;*

The frequency with which the suspect would match the blood drops if he were not the source is 1 in 1,000;

0.1% of the people in Houston who are not the source would also match the blood drops;

1 in 1,000 of the people in Houston who are not the source would also match the blood drops (Koehler, 2001:1288).

In the above example, the psychological impact of the first linguistic statement is the most inculpatory or incriminating to jurors, as it refers to a single target in terms of probability. The linguistic presentation in the fourth statement is the least inculpatory, as it refers to multiple targets and uses a natural frequency. The explanation provided by Koehler (2001; 2001b) to account for the psychological differences observed in the perceived inculpatory weight of the linguistic variations was "exemplar cueing theory." This theory posits that the more easily a trier of fact can imagine examples of others who match a target DNA profile by chance (exemplars), the less persuasive or incriminating they will find a DNA match statistic. Jurors are more likely to conclude that the DNA profile for the accused matches that obtained from the crime scene when the DNA profiling expert expresses the results of the DNA tests in terms of a single-target, probability-framed statistic than in terms of a multi-target, frequency-framed statistic.

Controlled experiments conducted with mock-jurors provided support for exemplar cueing theory. Koehler (2001b) presented mock-jurors with various manipulations of the statistical evidence presented in the context of the Clinton-Lewinsky scandal, and asked them to rate their certainty as to whether Clinton was the source of the DNA obtained from Lewinsky's dress. Whilst 28% of respondents in the single-target, probability condition were at least 99% certain that Clinton was the source, only 8% of respondents in the multi-target, frequency condition rated themselves as 99% certain. These results were replicated using a non-famous case example (Koehler, 2001b). In light of these findings, within the context of the adversary legal system, lawyers were advised to engage in strategic considerations regarding how best to frame DNA profiling match statistics to secure a persuasive advantage for their client. For instance, defence counsel may wish to use as large as possible a reference group, such as a large city, state, or nation, to draw jurors' attention to the potential of as many adventitious matches as possible, decreasing the inculpatory weight of the DNA evidence.

The findings by Koehler in studies conducted in North America were replicated in Germany by Hoffrage et al. (2000), generating additional support for exemplar cueing theory. By presenting two criminal case files or dossiers containing evidence of incriminating DNA profiling test results to legal professionals who were about to qualify as judges and to advanced law students, these researchers illustrated that it was not only lay jurors, jury eligible citizens or students who had difficulty in aggregating and understanding probability statistics. The legal professionals were asked to infer the probability of a true DNA match, given either probability- or frequency-based match statistics. In this study, the researchers examined both the accuracy of the inferences drawn, and the influence of the form of the presentation of the quantitative information on conviction rates. The accuracy of their estimates was significantly lower in the probability condition, with only 13% of legal professionals and less than 1% of students making a correct inference, compared to 68% and

44% respectively in the frequency condition. Probability-based evidence also significantly influenced judgments of guilt, with 45% of legal professionals and 55% of students rendering guilty verdicts in the probability condition, compared to 32% and 33% in the frequency condition (Hoffrage et al., 2000). These findings are significant in that they demonstrated the direction of the inferential errors, and the influence of linguistic formulations of statistical evidence upon ultimate verdicts: the significantly biasing effect of probability-based presentations was apparent in the higher proportion of guilty verdicts. Moreover, the fact that these effects were replicated in samples of legal professionals, including trainee judges, indicated that legal professionals were not better equipped than lay jurors to respond to this complex probabilistic evidence in the context of criminal trials.

In a parallel study, advanced medical students were given a series of diagnostic tasks involving the aggregation of statistical information about base rates, test sensitivity and false positive rates for various diseases, and asked to calculate the likelihood that a patient with a positive test would actually have the disease. The participants were significantly less likely to make the correct inference when the figures were expressed in terms of probability than in terms of frequency. Thus, legal professionals and medical professionals who must consider quantitative information in the context of criminal cases or in the process of rendering a medical diagnosis or decision were susceptible to the same cognitive biases. Hoffrage et al. concluded that "statistics expressed as natural frequencies improve the statistical thinking of experts and nonexperts alike" (2000:2261) and urged that, "when verdicts hinge on statistical evidence, understanding that evidence is crucial, and pursuing this simple method of fostering statistical insight could contribute to that goal" (2000:2262).

However, drawing on exemplar cueing theory, frequency-framed statistics can also be manipulated to appear more or less impressive to jurors. For example, Koehler (2001b) noted that numerators higher than one are more convincing to jurors, as they tend to invite exemplar-based thinking. Therefore, a figure of two in 200,000 was rated as less inculpatory than was a figure of one in 100,000, even though these figures are mathematically identical. Conversely, fractional numerators were viewed as more incriminating as they discouraged contemplation of other exemplars, with mock-jurors rating a match statistic of 0.1 in 100 as more inculpatory than the information presented as a frequency of one in 100,000, even though the former is objectively smaller. Results of this nature serve to emphasise the importance of the linguistic form of the expert evidence in promoting more accurate juror understanding of the statistical information.

In the appeal to the Australian High Court, counsel for the appellant Mr Aytugrul referred to exemplar cueing theory, contending that "by inviting the jury to focus on the percentage of people who were expected to be excluded from sharing the profile the trial judge was inadvertently diverting attention from the significant number of people who would be expected to share the profile" (Appellant's Submissions, *Aytugrul v The Queen*, 2011:s45, para 45). Upon review of this issue, the New South Wales Court of Criminal Appeals held that the fact that jurors may well have been able to convert an RMP into a percentage exclusion rate was irrelevant, as this possibility "does not inevitably lead to the conclusion that such a conversion should be invited or encouraged," (Appellant's Submissions, *Aytugrul v The Queen*, 2011:s47, para 47).

The influence of DNA match statistics in *Aytugrul v The Queen*

The submission by the appellant to the Australian High Court in the case *Aytugrul* sought a ruling regarding the linguistic form of the probabilistic expert evidence to establish a "clear and consistent approach using mathematical expressions which have been found to maximise juror comprehension and reduce the risk of fallacious reasoning," (Appellant's Submissions, *Aytugrul v The Queen*, Submissions of Appellant, 2011:s40, para 47). This point is important, as past research has shown a connection between how *persuasive* a particular statistical presentation is and the *accuracy of* comprehension of that statistic. For example, in the study of German legal professionals, the accuracy rates of their inferences increased and conviction rates decreased when the mathematically equivalent information was presented in natural frequencies than as an RMP (Hoffrage et al., 2000).

Indeed, the respondent's (Crown) submission emphasised that "Admissibility is not in an inverse ratio to persuasiveness" (2011:s6.30), and further submitted that the RMP of 1 in 1680 could have been presented as 10 in 16,800 and that the latter figure may have been "less persuasive because it suggested the existence of other possible matches but that [this] did not render the 1 in 1680 ratio inadmissible or unfair" (*Aytugrul v The Queen*, Submissions of Respondent, 2011:s6.30). The respondent's submission to the High Court argued that there was no appreciable risk of prejudice in providing the percentage exclusion rate alongside the RMP, as these were mathematically equivalent, and that "percentage quantifications in themselves do not have greater subliminal impact than random match probabilities or other statistical formulations. Their effect in any given case would depend largely on the figures themselves and on the other circumstances of the case" (2011:s6.15). The submission agreed that there were risks of prejudice in cases citing "overwhelmingly high" figures, such as in the case of *R v GK,* (in which the jury was presented with three formulations; a paternity index of 147,000 to 1, also expressed as 99.9993%, and 500 million times more likely that GK was the father than was another person)(Respondent's Submissions, *Aytugrul v The Queen*, 2011, para 6.4) but that "such concerns are very far removed from the present case which lies at the other end of the spectrum. In the present case, the random probability match ratios were quite low, as low as 1 in 50, and posed no risk that the jury might be misled into thinking that the ratios negated any possibility of matches other than the appellant" (Respondent's Submissions, *Aytugrul v R*, 2011:s6.12, para 6.12). Similarly, the risks of a possible subliminal impact in cases involving ratios in the hundreds of millions, or billions, were deemed inapplicable in the present case, where the RMP was as low as 1 in 50.

The respondent's submission further argued that the RMP figures presented by the various experts were not accompanied by any degree of certainty, as the jury was given a detailed explanation of the nature of the statistical evidence, as well as its limitations, and was expressly informed that the figures were estimates and averages. In addition, explicit attention had been drawn at trial to the possibility of adventitious matches both in the Turkish sub-population and the overall Australian population. The prosecution expert testified that there would be "a substantial number of Turkish origin individuals" who had this profile (Respondent's Submissions, *Aytugrul v The Queen*, 2011, para 6.18), including appellant's maternal relations over many generations, and persons whom he would not consider to be relatives. The jury's attention was specifically drawn to possibility of other matches using the analogy of a football stadium. They were informed that in a stadium of 16,000 people, 10 would be expected to match the profile, although the actual number of matches would vary on any given day (Respondent's Submissions, *Aytugrul v The Queen*, 2011, para 6.35). The Crown emphasised that the 99.9% exclusion ratio was only mentioned briefly during the expert testimony and played a very minor part in the overall case. Moreover, the judge had explicitly reminded the jury that the hair might have had nothing to do with the case, probing

"Did the hair come from the person who killed Sevda Bayrak? If it did not, it does not matter who it came from, and all of that evidence of mitochondrial DNA takes this case nowhere" (Respondent's Submissions, *Aytugrul v The Queen*, 2011, para 6.41).

Conclusion

In the appeal by Aytugrul against conviction, no legal challenges were raised in relation to the accuracy or reliability of the DNA analysis, nor the mathematical interpretation of the results. A major, substantive legal issue at stake surrounded the expression and conveyance of the DNA profiling test results to the jury, specifically whether the probative value of the exclusionary DNA match statistic of 99.9% was exceeded by its prejudicial effect upon the jury. In this case, the appeal was unanimously dismissed by the High Court, however the Court acknowledged that "the evidence that was given did not, and was not said to, establish that the mitochondrial DNA profile found in the hair definitely came from the appellant" (*Aytugrul v The Queen, 2012*, para 24) and that consequently there was no risk that jurors would subconsciously round up the figure of 99.9% to 100%. In other cases where a conviction relies more directly on DNA match statistics, it is foreseeable that the use of exclusion percentages rather than natural frequencies may be accepted as unfairly prejudicial.

Whatever the probative or prejudicial value of the exclusion percentage in the ***A****ytugrul* case, its central role as the basis for appealing the conviction urges consideration of the practical impact of exclusion percentages on the weight attributed to DNA profiling evidence at trial. Arguably, exclusion percentages provide little if any advantage above frequency-based statistics in terms of their mathematical accuracy, but carry an unacceptably high risk not only of miscomprehension that may prejudice the jury and lead to miscarriages of justice, but also of resultant, costly appeals. Given these risks, a more prudent position is to avoid the presentation of exclusion percentages to triers of fact, and to require expert witnesses and legal practitioners to present their evidence in the form of natural frequencies, which have been shown to enhance the understanding by jurors and legal professionals of statistical evidence and to minimise the risk of erroneous inferences and unwarranted convictions.

References

Aytugrul v The Queen (2012) HCA 15.

Aytugrul v The Queen (2010) 205 A Crim R 157.

Aytugrul v The Queen 2011, No S315 of 2011, Submissions of Appellant to the High Court.

Aytugrul v The Queen 2011, No S315 of 2011, Submissions of Respondent to the High Court.

Aytugrul v The Queen, No S315 of 2011, Hearing of 8 December 2011, [2011] HCATrans329.

Briody, M. (2003) The effects of DNA evidence on sexual offence cases in court. *Current Issues in Criminal Justice* 14: 159--181.

Briody, M. (2004) The effects of DNA evidence on homicide cases in court. *Australian and New Zealand Journal of Criminology* 37: 231--252.

Charrow, R., & Charrow, V. (1979) Making legal language understandable: a psycholinguistic study of jury instructions. *Columbia Law Review*, 79: 1306-1374.

Conley, J., & O'Barr, W. M. (1990) *Rules versus relationships. The ethnography of legal discourse* Chicago and London: The University of Chicago Press.

Coulthard, M. (2010) Experts and opinions. In M. Coulthard & A. Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (pp. 473-486). London and New York: Routledge.

Coulthard, M., & Johnson, A. (2007) *An introduction to forensic linguistics: Language in evidence*. London and New York: Routledge.

Dartnall, S., & Goodman-Delahunty, J. (2006). Enhancing juror understanding of probabilistic DNA evidence. *Australian Journal of Forensic Sciences*, 38:86-96.

Dumas, B. K. (2000). US pattern jury instructions: problems and proposals. *Forensic Linguistics, 7*(1), 49-71.

Findlay, M. (2008) Juror comprehension and the hard case- Making forensic evidence simple.*International Journal of Law, Crime and Justice* 36: 15-53.

Goodman-Delahunty, J., & Hewson, L. (2010a). Improving jury understanding and use of DNA evidence. *Trends & Issues in Crime and Criminal Justice,* 392:1-6.

Goodman-Delahunty, J., & Hewson, L. (2010b). *Improving jury understanding and use of DNA evidence.* Canberra: Australian Institute of Criminology Technical and Background Paper Series. Hale, S. (2004/2010). *The discourse of court interpreting. Discourse practices of the law, the witness and the interpreter*. Amsterdam and Philadelphia: John Benjamins.

Harris, S. (1984). Questions as a mode of control in magistrates' courts. *International Journal of the Sociology of Language, 49*, 5-28.

Hoffrage, U., Lindsey, S., Hertwig, R. and Gigerenzer, G. (2000) Communicating statistical information. *Science* 290: 2261--2262.

Kahneman, D. and Lovallo, D. (1993) Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science* 39: 17--31.

Koehler, J.J (1996) On conveying the probative value of DNA evidence: Frequencies, likelihood ratios and error rates. *University of Colorado Law Review* 67: 859-886.

Koehler, J.J. (1997a) One in millions, billions or trillions: Lessons from *People v. Collins* (1968) for *People v. Simpson* (1995). *Journal of Legal Education* 47: 214--223.

Koehler, J.J. (2001a) The psychology of numbers in the courtroom: How to make DNA match statistics seem impressive or insufficient. *Southern California Law Review* 74: 1275--1305.

Koehler, J.J. (2001b) When are people persuaded by DNA match statistics? *Law and Human Behavior*, 25: 493--513.

Koehler, J.J., Chia, A. and Lindsey, J.S. (1995) The random match probability (RMP) in DNA evidence: Irrelevant and prejudicial? *Jurimetrics* 35: 201--219.

Levi, J. (1993) Evaluating jury comprehension of the Illinois capital sentencing instructions. *American Speech*, 68: 20-49.

Lieberman, J.D., Carrell, C.A., Miethe, T.D. and Krauss, D.A. (2008) Gold versus platinum: Do jurors recognise the superiority and limitations of DNA evidence compared to other types of forensic evidence? *Psychology, Public Policy & Law* 14: 27--62.

Loftus, E.F., and Palmer, J.C. (1974). Reconstruction of automobile destruction: an example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behaviour* 14: 585-589.

McClellan, P. (2011, March). *Looking inside the jury room.* Paper presented at the meeting of the Law Society of New South Wales Young Lawyers 2011 Annual Criminal Law Seminar, Sydney, Australia.

McDaniel v. Brown, 130 S.Ct. 665 (2010).

Napier, J., Shearim, G., Rohan, M. & Spencer, D. (n.d). Comprehension of courtroom discourse: How do deaf and hearing people compare? Unpublished research report, Macquarie University.

O'Barr, W. M. (1982). *Linguistics evidence. Language, power, and strategy in the courtroom*. New York: Academic Press.

Prideaux, G. D. (1991) Syntactic form and textual rhetoric: The cognitive basis for certain pragmatic principles. *Journal of Pragmatics*, 16: 113-129.

Saxton, B. (1998). How well do jurors understand jury instructions? A field test using real juries and real trials in Wyoming. *Land and Water Literary Review*, 33: 59.

Schklar, J. and Diamond, S. (1999). Juror reactions to DNA evidence: Errors and expectancies. *Law and Human Behavior* 23: 159--184.

Twenty Scholars of Forensic Evidence. (2009). *Amicus curiae brief in support of Troy Brown, McDaniel v. Brown*, US Ninth Circuit Court of Appeals, No. 08-559.

Walsh, S. (2011, December) *DNA profiling: The form of the expert evidence.* Paper presentation at the Australian Academy of Forensic Sciences Conference on Impressions and Expressions: Expert Evidence in Reports & Courts, Sydney, Australia.

Woodbury, H. (1984). The strategic use of questions in court. *Semiotica, 48* (3/4): 197 228.