

# ‘The Innocent v The Fickle Few’: How Jurors Understand Random-Match-Probabilities and Judges’ Directions when Reasoning about DNA and Refuting Evidence



\*Michelle Cowley

Royal Statistical Society, England

Submission: March 12, 2017; Published: March 31, 2017

\*Corresponding author: Michelle Cowley, Royal Statistical Society, Royal Statistical Society Fellow, Errol St., London, EC1Y 8LX, England, Tel: 00353894389634; Email: cowleym@eircom.net

## Abstract

DNA evidence is one of the most significant modern advances in the search for truth since the cross examination, but its format as a random-match-probability makes it difficult for people to assign an appropriate probative value [1]. While Frequentist theories propose that the presentation of the match as a frequency rather than a probability facilitates more accurate assessment [2], Exemplar-Cueing Theory predicts that the subjective weight assigned may be affected by the frequency or probability format, and how easily examples of the event, i.e., ‘exemplars’, are generated from linguistic cues that frame the match in light of further evidence [3]. This paper presents two juror research studies to examine the difficulties that jurors have in assigning appropriate probative value to DNA evidence when contradictory evidence is presented.

Study 1 showed that refuting evidence significantly reduced guilt judgments when exemplars were linguistically cued, even when the probability match and the refuting evidence had the same objective probative value. Moreover, qualitative reason for judgment responses revealed that interpreting refuting evidence was found to be complex and not necessarily reductive; refutation was found indicative of innocence or guilt depending on whether exemplars have been cued or not.

Study 2 showed that the introduction of judges’ directions to linguistically cue exemplars, did not increase the impact of refuting evidence beyond its objective probative value, but less guilty verdicts were returned when jurors were instructed to consider all possible explanations of the evidence. The results are discussed in light of contradictory frequentist and exemplar-cueing theoretical positions, and their real-world consequences.

**Keywords:** DNA Evidence; Random-Match-Probability; Juror Decision-Making; Exemplar-Cueing Theory; Frequentist Theories; Reasoning and Rationality

## Introduction

The power of DNA evidence in convicting the guilty and exonerating the innocent is well established [4,5]. The probative value of DNA evidence has been held up as a most valuable ‘Sword and Shield’ of criminal justice, and is now on a par with fingerprint evidence in identifying or eliminating suspects in a criminal trial [6]. DNA evidence is largely considered one of the greatest advances in the search for truth since the cross-examination [1], yet its presentation as a random-match-probability makes it one of the most complex sorts of evidence for people to assign appropriate probative value to [3,7]. Thus, if DNA matches are not 100% conclusive [8], and legal practitioners question whether everyday reasoners, such as jurors, can think about DNA evidence in terms of a random-match-probability in

a way that is consonant with the statistical reasoning essential for reasoning accurately about probabilistic information [9], then we must find out more about what factors affect rational juror reasoning about DNA evidence. Consider what happened in the media following the now well-known disappearance case of Madeleine McCann [10], the child who went missing from a holiday apartment in Portugal in 2007<sup>1</sup>. Although what happened has never been definitively discovered, the media speculation focusing on the case in those early days and weeks of the investigation took a turn for the worst by theorizing that the child’s parents killed her rather than more explicitly theorizing about the possibility of child abduction. Due to disproportionate focus on Madeleine’s DNA evidence found in the apartment

and hire car implying 100% conclusiveness, rather than a statistical match criterion between zero and one-hundred percent, the possibility that the child's parents were primarily suspect were found to be exaggerated. Moreover, the alternative possibility that the transport of Madeleine's belongings had led to subsequent cross-contamination also explained the presence of DNA evidence in the hire car, and was indicative of her parents' innocence regardless of how conclusive the match. In this case the headline language used by the media was key in implicating that the DNA evidence meant more than its objective probative value could logically prove. The negative consequences of exaggerating a powerful evidential paradigm such as DNA evidence through miscommunication are several fold including: the waste of investigatory time, the waste of policing resources, the facilitation of the guilty party at the expense of the innocent, invasion of privacy, and the placement of a considerable personal and procedural correction burden on individuals. How then can we be more careful in the communication of DNA evidence, whether in the media, in investigatory proceedings, or in the courtroom? We need to understand the communication format of DNA evidence in the first instance.

### Procedural issues in the presentation of DNA as statistical evidence

DNA is calculated probabilistically and it is presented as a random-match-probability (or 'random occurrence ratio', [6]). The more indicative of guilt the DNA evidence, the lower the chance that the match between a crime scene sample and the suspect is due to coincidence (i.e., the match is a low probability event). For example, when a random-match-probability is low and can be understood as more indicative of guilt than not, the chances that the suspect would match the DNA by chance would be one-in-one-million rather than one-in-ten-thousand. Thus, if there is a match between the DNA evidence found at the crime scene and the suspect's DNA sample, then the likelihood that coincidence is at play is less when the chance that the suspect's sample would match DNA evidence at the crime scene is one-in-one-million rather than one in ten thousand [11]. In the United Kingdom, for example, the presentation of DNA evidence during a trial requires that the details of a random-match-probability's calculation are made explicit to allow the defendant's legal team to critically evaluate its accuracy and relevance.

The Forensic Science Service provider may be required to make available the working or database(s) on which the calculations are based, and a pretrial review of the evidence and issues relevant to expert evidence regarding DNA is usual. While DNA evidence is expert evidence, and expert evidence comes

under the auspices of opinion evidence [12], forensic experts presenting DNA evidence are only to be directed to present the random-match-probability relevant to the trial, that is, the random-match probability and the frequency with which these matching DNA characteristics are to be found in the population at large. The information presented is strictly technical rather than opinion-specific. Technically then, there are several ways to challenge DNA evidence in the courtroom despite the DNA evidence indicating a low probability event including: challenging the expert level of the expert witness, challenging the chain-of-custody records for samples to rule out cross-contamination or tampering, examining how up to date the laboratory technology and protocols are and so on [13].

For instance the defence team can argue for admissibility disqualifications for breaches in chain-of-custody issues related to the collection, preservation, and laboratory handling of the samples - if there is a missing component in the travel log of chain-of-custody records the DNA evidence must be argued inadmissible. That said, from a psycho-linguistic perspective the expert's presentation of the random-match-probability and the frequency with which these matching DNA characteristics are likely to be found in the population at large is a two-part proposition to be understood by the jury. The first part is the likelihood that the DNA match from the crime scene to the suspect's DNA is by chance. The smaller the coincidence the more likely that the suspect's DNA is the DNA at the crime scene. The second part opens up the consideration of other possible matches within a population of people to which the suspect belongs. That is, the possibility that there could be alternative suspects is communicated. The adversarial nature of the courtroom demands that jurors consider the communication of alternative possibilities within opposing arguments about a suspect's guilt or innocence. Thus, manipulation of the language framing this proposition may be another way for advocacy to undermine the impact of DNA evidence.

### Communication issues in the understanding of DNA evidence

A body of experimental evidence in the psychology of law and human behavior has shown that people do have a tendency to judge the probative value of reported evidence by the ease with which the communication of a match brings to mind explanatory possibilities, regardless of frequency or probabilistic format [1-3,14]. Let us turn to examine these findings.

<sup>1</sup>DNA Database and Data Protection Published Transcript - Lord Hansard Text - House of Parliament UK (2003) <https://www.publications.parliament.uk/pa/ld200203/ldhansrd/vo030703/text/30703-12.htm>

<sup>1</sup>BBC News Online, 5th August 2008, <http://news.bbc.co.uk/1/hi/uk/7541810.stm>

<sup>1</sup>The Leveson Inquiry -<http://webarchive.nationalarchives.gov.uk/20140122145147/http://www.levesoninquiry.org.uk/>

## Frequentist theories and exemplar generation in human reasoning

Psychological studies of human reasoning about frequencies and probabilities show that people often have a fickle understanding of statistical information in risk or gambling contexts [15]. For example, when people are asked to choose which lottery with more examples of winning, even when the probability of winning is the exact same in each lottery (e.g., 10 winning jelly beans out of 100 vs. 1 jelly bean out of 10 - they tend to prefer the first option even though the chances are 1 in 10 in each; [16]), and people often choose to play a large lottery rather than a small lottery when the small lottery has a higher probability of choosing a winner [17,18].

Likewise, in clinical settings people tended to think that patients with a mental illness were more likely to pose a risk when released if the risk was presented as a frequency (e.g., 1 in 20) than when it was presented as a probability (.05) [2]. An explicit frightening image of patients was reported when the frequency of risk was presented. The explanation is that the phrase 'one in twenty' mentally ill patients provides a probability attached to a single person for which an image of an individual may be prompted more readily than when the probability value presents a statistic [19]. Therefore, when people consider DNA random-match-probabilities in their reasoning, the subjective weight that they assign to that evidence may be affected not only by the probability value, but by how easily exemplars are prompted (i.e., examples or alternative examples explaining the evidence), or are available to mind from memory when reasoning about the DNA match [20].

## Exemplar-Cueing Theory and linguistic cues in juror reasoning about DNA

When then are people convinced by the communication of DNA match statistics? Recent studies of juror reasoning suggest that jurors do not automatically assume that guilt is indicated when considering a low random-match-probability [1]. The same evidence may convince most jurors that the suspect is almost definitively the source of the DNA evidence found at a crime scene, or definitively not the source of the DNA evidence depending on how that same DNA evidence is presented [1]. Consider that laypeople have been found more impressed with DNA statistics when they are presented as likelihood ratios rather than frequencies [1]. Aside from studies showing that jurors underestimate the probative value of DNA evidence relative to theoretical Bayesian norms [21], studies examining some of the effects that changes in linguistic framing have on estimates of evidence strength, guilt ratings, and overall verdicts delivered found that linguistic cues were sufficient in making an impact in conditions under which the random-match-probabilities were statistically equivalent [3].

Koehler and Macchi [3] present a conceptual framework known as Exemplar-Cueing Theory to explain much of the effects

of linguistic framing on DNA random-match-probability in juror reasoning. They predict that where the probabilistic statistic identifies how prevalent a matching profile is in a given reference population, jurors will judge the probative value of the reported match by the ease with which they can imagine examples of others who would also match in light of the presentation format of the DNA match statistic. In a series of mock juror studies they presented mock jurors with a burglary scenario in which the suspect's DNA was subsequently presented as a DNA match during trial. The random-match-probability was presented as a low or less high random-match-probability, as a frequency or probability, and delivered in the presence of a linguistic cue (i.e., '...nonetheless') that was intended to prompt thinking of alternatives (i.e., alternative matching suspects). The linguistic cue 'nonetheless' led to jurors indicating concern about the possibility that the match was coincidental and they were less likely to believe that the defendant was the source of the DNA or guilty of the crime, when low-probability DNA matching evidence was presented. When evidence was not presented in an exemplar prompting manner with low-probability DNA matching evidence, jurors were apparently less concerned about the possibility of a coincidence, and therefore more likely to believe that the defendant was the source of the DNA and therefore guilty of the crime.

Using the same methodology, this paper explicates the role of exemplar cueing in two other key courtroom communication events in which DNA random-match-probabilities are presented in the presence or absence of exemplar prompting linguistic cues. In two studies the paper tests whether:

- i. evidence refuting that the suspect is the source of the DNA match, or
- ii. Judges' directions to 'consider all possible explanations of the evidence', either suppress or enhance the number of guilt verdicts, or ratings of guiltiness indicative of exemplar-cueing or lack thereof.

The prediction is that exemplar-cueing communication factors such as Judges' directions, to consider all possible explanations, or evidence refuting that the suspect is the source of the DNA evidence, decreases consideration of guiltiness, and that exemplar-suppressing communication factors such as the absence of exemplar cueing refuting evidence, or Judges' directions, will increase consideration of guiltiness.

## Study 1 - Multiple exemplar-cueing and refuting evidence

This study examined if the language used to frame the DNA random-match-probability cued the consideration of alternative explanations facilitating guilt reduction in the presence of refuting evidence. In line with Koehler & Macchi [3] this study tested whether the linguistic cue, 'nonetheless', which is known to cue the generation of alternative possibilities concordant

with innocence, was more effective in reducing guilt if refuting evidence was presented.

## Method

I. **Participants:** Sixty-four people took part. There were twenty-seven men and thirty-seven women. Their mean age was 29 years and their ages ranged from 18 to 65 years. They were jury eligible members of the public who volunteered to take part in response to Oxford newspaper advertisements. They received a token of £4 for taking part.

II. **Design & Procedure:** The study employed a 2 (linguistic cue: single exemplar, multiple exemplar) x 2 (refutation: present, absent) between-subjects design. The random-match-probability was a low random-match-probability and identical in each condition. A short scenario was adapted from a burglary case [3]. Participants were asked to consider the following scenario and answer questions on two following pages: 'A masked man bursts into a jewellery store and announces he intends to rob it. However, the owner of the store resists and pops the robber on the nose. DNA analysis of the blood reveals a match with a saliva sample taken from one of several suspects. The matching suspect is charged with attempted robbery and brought to trial'. Participants received a sentence in which the random-match-probability was presented in such a way as to cue a single exemplar (i.e., the suspect) or multiple exemplars (i.e., the suspect and alternative suspects). In the single exemplar conditions participants received the additional sentence 'You learn that the chance that the suspect would match the blood drop if he were not the source is 0.0001%' (i.e., single exemplar). In the multiple exemplar conditions the participants received the additional sentence 'You learn that the chance that the suspect would nonetheless match the blood drop if he were not the source is 0.0001%' (i.e., multiple exemplars).

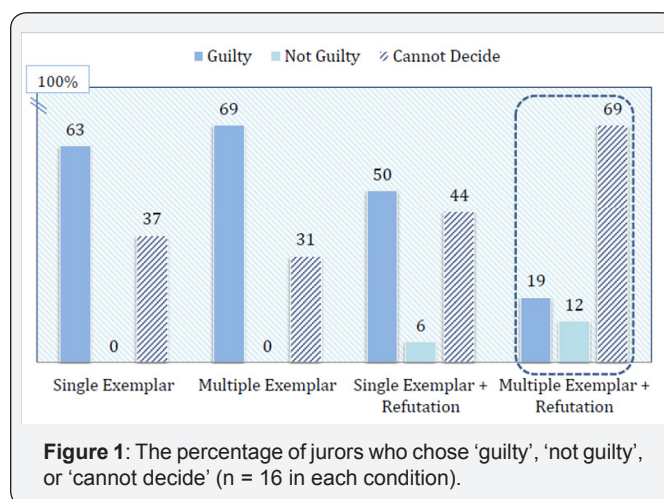
The difference between the single and multiple exemplar conditions is the presence of the alternative exemplar cueing word 'nonetheless'. Participants either received an additional sentence containing a piece of refuting evidence 'You learn that there is a record of the suspect's debit card being used at a petrol station on the other side of town, and the card was used at the same time as the attempted robbery' (i.e., refutation present), or they received no refuting evidence (i.e., refutation absent). In conditions in which there were two evidence sentences the order of evidence presentation was counterbalanced across conditions. Responses were required on rating scales (where 0 represented 'not at all strong', and where 10 represented 'extremely strong') to questions pertaining to how strong participants thought the DNA evidence against the suspect was, how strong they thought the overall evidence against the suspect was, and how guilty they thought the suspect was overall. Responses were required as percentages when asked what participants thought the probability was that the suspect was the source of the DNA evidence from the DNA evidence presented on the previous page and what they thought the probability was that the suspect

was the source of the DNA evidence from the overall evidence presented on the previous page.

A categorical response was required for participants' decisions about whether the suspect was guilty, not guilty, or whether they could not decide. Reason for choice responses were coded qualitatively. They were transcribed and segmented into individual components mentioning the DNA evidence or the refuting evidence. The number of each sort of evidence was calculated nominally for each participant. The evidence statements we were interested in corresponded to respondents' thinking about the inferences they drew related to two components of their thinking:

- whether the DNA evidence required corroboration to draw guilt conclusions,
- Whether evidence refuting the DNA match (when present) was indicative of innocence such as, suspects alternative to the present suspect that could match the DNA evidence.

## III. Results and discussion



Participants returned the verdict 'not guilty' (5%) significantly less than 'guilty' (50%) or 'cannot decide' (45%),  $\chi^2 = 23.844(2)$ ,  $p < .0005$  as Figure 1 shows. Jurors tended to choose 'guilt' more than 'cannot decide' or 'not guilty' in each condition, except where they were cued to consider multiple exemplars and the refutation. Overall, the pattern of verdicts tended towards 'guilty' when single exemplars (67%) rather than multiple exemplars were cued (33%,  $\chi^2 = 6.25(1)$ ,  $p < .02$ ). When examined individually per condition this trend of choosing guilty verdicts was not significant: when jurors were cued to consider the single exemplar they chose 'guilty' (63%) more often than 'cannot decide' (37%) and 'not guilty' (0%,  $\chi^2 = 1.000(1)$ ,  $p > .05$ ); when they were cued to consider multiple exemplars they chose 'guilty' (69%) more often than 'cannot decide' (31%) and 'not guilty' (0%,  $\chi^2 = 2.25(1)$ ,  $p > .05$ ); and when presented with a refutation in addition to the single exemplar they chose 'guilty' (50%) and 'cannot decide' (44%)

more often than 'not guilty' (6%,  $\chi^2 = 5.375(2)$ ,  $p > .05$ ). There was no significant tendency towards 'guilty' in the absence of the refutation (56%) than in its presence (44%,  $\chi^2 = 1.000(1)$ ,  $p > .05$ ). But when jurors are cued to consider multiple exemplars in the presence of a refutation, a pattern in favour of indecision rather than guilt occurs. Cueing jurors to consider alternative suspects to the defendant, increases the impact of refuting evidence. Jurors chose 'cannot decide' (69%) more often than 'guilty' (19%) and 'not guilty' (12%,  $\chi^2 = 9.125(2)$ ,  $p < .01$ ), as Figure 1 shows.

### Guilt ratings and overall ratings of evidence strength

Participants' ratings of guilt tended not to be significantly lower when they considered the refutation ( $M = 6.49$ ,  $MSE = .426$ ) than not ( $M = 7.5$ ,  $MSE = .432$ ,  $F(1,64) = 2.781(1)$ ,  $p > .05$ ), but their overall ratings of guilt were lowered by the cueing of multiple exemplars ( $M = 6.39$ ,  $MSE = .440$ ) rather than a single exemplar ( $M = 7.59$ ,  $MSE = .418$ ,  $F(1,64) = 3.895$ ,  $p = .05$ ). The effect of interaction between exemplar cueing and refutation on overall guilt ratings was not significant ( $F(1,64) = 1.08$ ,  $M = 6.99$ ,  $MSE = 6.06$ ,  $p > .05$ ), even though the presence of a multiple exemplar cue and a refutation led to a larger proportion of 'cannot decide' verdicts than any other condition. The result suggests that while alternative possibilities and refutations may interact to create doubtful verdicts in DNA reasoning, underlying guilt ratings may not necessarily reflect categorical decisions. Yet when there was both DNA and refuting evidence participants were asked for an overall rating of evidence strength against the suspect, in addition to an underlying rating of guilt. Overall ratings of evidence strength were significantly lower in the presence of a refutation when multiple exemplars were cued ( $M = 5.84$ ) than not ( $M = 7.34$ ,  $t = 2.381(30)$ ,  $p = .012$ ).

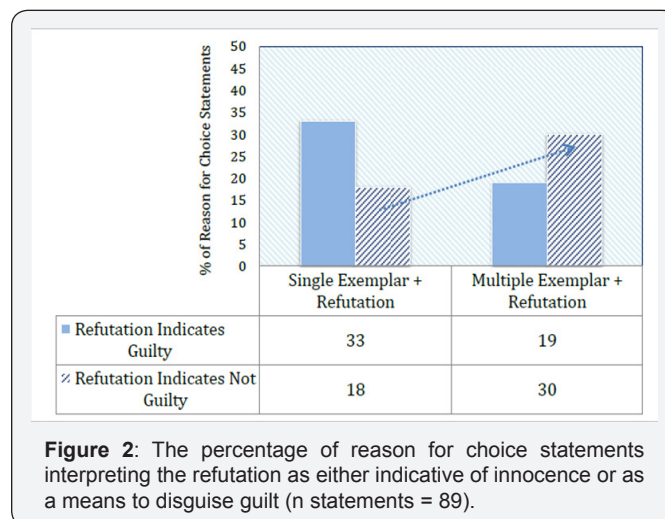
### DNA evidence strength and source probabilities

Participants' ratings of evidence strength specific to the DNA evidence were not significantly lower when multiple exemplars were cued ( $M = 6.75$ ,  $MSE = .4400$ ) than not ( $M = 7.64$ ,  $MSE = .418$ ,  $F(1,64) = 2.134$ ,  $p > .05$ ). The presence of the refutation did not significantly affect DNA evidence strength ratings ( $M = 6.76$ ,  $MSE = .432$ ) compared to refutation absence ( $M = 7.63$ ,  $MSE = .426$ ,  $F(1,64) = 2.089$ ,  $p > .05$ ), but there was an effect for the interaction between exemplar cue and refutation presence ( $M = 7.19$ ,  $MSE = .606$ ,  $F(1,64) = 4.858$ ,  $p < .05$ ), demonstrating that the refuting evidence tends to have a refuting evidence effect when alternative suspects are linguistically cued. Participants' recordings of what they thought the probability was that the suspect was the source of the DNA were lower in the presence of the refutation ( $M = 80.5\%$ ,  $MSE = 3.3\%$ ) than in its absence ( $M = 92.6\%$ ,  $MSE = 3.3\%$ ,  $F(1,64) = 6.093$ ,  $P = .01$ ). Cueing multiple exemplars significantly lowered the source probability ratings ( $M = 80.9\%$ ,  $MSE = 3.4\%$ ) than cueing a single exemplar ( $M = 92.3\%$ ,  $MSE = 3.2\%$ ,  $F(1,64) = 6.115$ ,  $p < .02$ ), and so there was no main effect for the interaction between exemplar cue

and refutation ( $M = 86.6\%$ ,  $MSE = 4.6\%$ ,  $F(1,64) = 1.314$ ,  $p > .05$ ). These results suggest that thinking about the evidence in terms of whether or not the suspect was a source of the DNA, is somewhat distinct from thinking about the evidence in terms of guilt, or innocence. Participants were asked what they thought the overall probability was that the suspect was the source of the DNA evidence given all of the evidence considered (i.e., refutation and DNA). The overall probability was not significantly lower in the presence of refutations when the multiple exemplars were cued (69.5%) than not (81.1%,  $t = 1.259(30)$ ,  $p > .05$ ).

### Reasons for choice: Refuting evidence and corroboration requirements

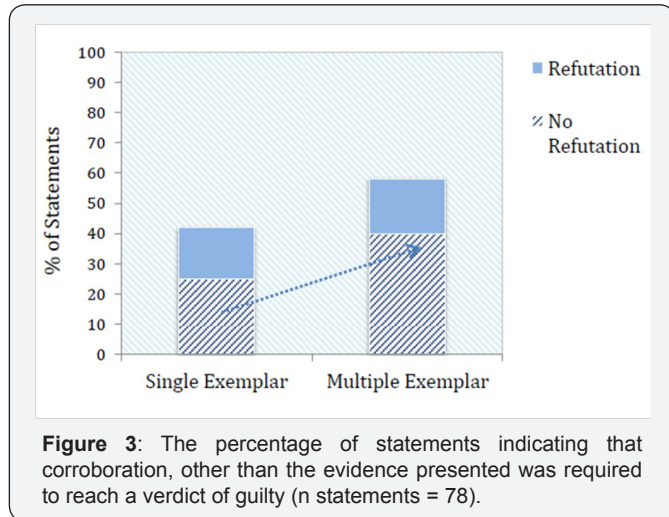
Participants were provided with space to record their reasons for choice for each question. A total of 89 evidence statements referring to the two refuting evidence conditions were generated. The refuting evidence was interpreted in one of two ways by participants: they either reasoned that the refutation indicated innocence by providing evidence of the suspect being in a different place at the time of the crime, or they reasoned that the refutation indicated guilt because the suspect could have either disguised his whereabouts by having an accomplice use the card, or could have had his card stolen (Figure 2). Overall multiple exemplar cueing produced refuting reasoning indicative of innocence (30%) rather than guilt (19%), more than single exemplar cueing (18% and 33% respectively,  $\chi^2 = 5.92(1)$ ,  $p < .02$ ). This result suggests that the language used to frame the random-match-probability may not only cue or suppress the consideration of alternative suspects, but may simultaneously affect the way that other evidence is interpreted.



**Figure 2:** The percentage of reason for choice statements interpreting the refutation as either indicative of innocence or as a means to disguise guilt (n statements = 89).

Even when presented with a refutation that logically should reduce the evidential weight apportioned to guilt, participants tended to at times interpret this refutation as additive evidence of guilt when alternative exemplars were not cued (Figure 3). A total of 78 statements were generated indicating that corroboration, other than the evidence presented, was required

to reach a verdict. More corroboratory requirement statements were generated when multiple exemplars were cued in the absence of other evidence (40%) more than any other condition (25%, 18%, and 17% for single exemplar without refutation, and multiple exemplar conditions respectively), suggesting that in the absence of evidence other than DNA, participants require corroboration when they are considering alternative suspects ( $\chi^2 = 9.875(2)$ ,  $p < .01$ ).



**Figure 3:** The percentage of statements indicating that corroboration, other than the evidence presented was required to reach a verdict of guilty (n statements = 78).

**Study 2 - Judges' directions**

This study examined if the presence or absence of a judge's direction to consider all possible explanations of the evidence affected how much the evidence refuting the DNA random-match-probability impacted guilt ratings, or cued the consideration of alternatives to guiltiness, in the absence of any exemplar-cueing linguistic framing.

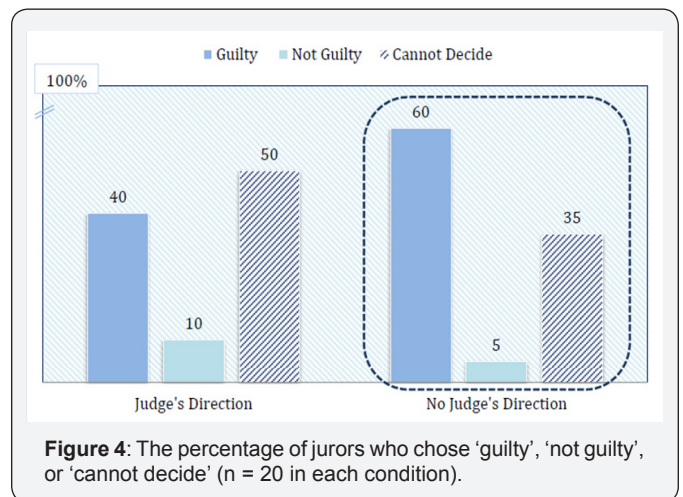
**Method**

I. **Participants:** Forty people took part. There were fourteen men and twenty-six women. Their mean age was 33 years and their ages ranged from 18 to 68 years. They were jury eligible members of the public who volunteered to take part in response to Oxford newspaper advertisements. They received a token of £4 for taking part.

II. **Design and procedure:** The study employed a 2 x (judge's direction present or judge's direction absent) x 1 (identical refutation with random-match-probability in each of the two conditions) design. The random-match-probability was identical in each condition, and the debit card refutation was present in every condition. The same short burglary scenario from Study 1 was used. Half of the participants were asked to consider the scenario following a judge's direction, and the other half without it. The judge's direction was given on a piece of paper inside an envelope when the participant had finished reading the scenario but before they answered the questions about guilt or innocence on the following pages. The instruction was worded: 'The judge has instructed that all possible

explanations of the evidence should be dutifully considered. If a guilty verdict is returned, then it should be beyond reasonable doubt'. Participants then answered the same questions about evidence and guilt as Study 1.

III. **Results and discussion:** While participants returned verdicts of not guilty (7.5%) less often than guilty (50%) or cannot decide (42.5%) regardless of whether there was a judge's direction to consider all possible explanations or not, they did so not significantly overall ( $\chi^2 = 1.64(2)$ ,  $p > .05$ ). Only when within-condition discretionary comparisons are made can we see that significantly more guilty verdicts are returned (60%), than not guilty (5%) or cannot decide (35%,  $\chi^2 = 9.1(2)$ ,  $p < .01$ ) in the absence of a judge's direction to consider all possible explanations. When the judge's direction is given there are less not guilty verdicts (10%) than guilty (40%) or cannot decide verdicts (50%), but only marginally so ( $\chi^2 = 5.2(2)$ ,  $p = .07$ ) as Figure 4 shows:



**Figure 4:** The percentage of jurors who chose 'guilty', 'not guilty', or 'cannot decide' (n = 20 in each condition).

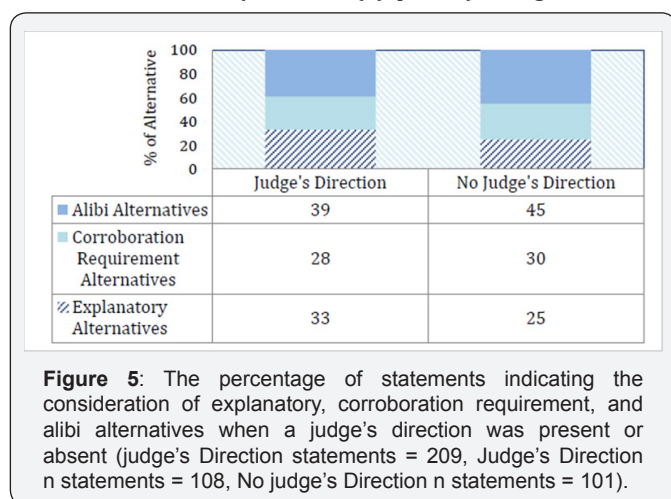
**DNA, source probability, and overall evidence strength ratings**

Participants' did not differ significantly in their ranking of: DNA evidence strength (Mean 8.29 v Mean 7.55,  $t = 1.150(35.1)$ ,  $p > .05$ ), the probability that the suspect was the source of the DNA evidence (Mean 91.89% v 85.4%,  $t = .819(35.3)$ ,  $p > .05$ ), the overall evidence strength (Mean 7.25 v 6.25,  $t = 1.355(37.4)$ ,  $p > .05$ ), or the overall rating of guilt (Mean 7.65 v 6.85,  $t = .992(38)$ ,  $p > .05$ ). Critically, only in the categorical outcomes corresponding to guilty, not guilty, or cannot decide are there significant differences in participants' juror reasoning. When judge's directions to consider all possible explanations of the evidence are absent, guilty verdicts are returned significantly more often. Despite these different overall outcomes, the rankings of DNA evidence strength, source probability, and overall evidence strength rankings do not differ significantly when the judge's direction is absent or present.

**Reasons for choice: Judges' directions as qualitative cues**

Participants were provided with space to record their reasons for choice for each question. A total of 209 evidence statements referring to alternative possible explanations were generated. These evidence statements fell into the three main categories of:

- i. Non-explicit alternative explanations of events such as 'perhaps the police confused his saliva with someone else's' (29%),
- ii. Corroboration requirement alternatives such as 'the evidence is not strong... we only have his DNA' (28.5%), and
- iii. References to alibi-related alternatives such as 'which could be used by an accomplice to establish an alibi' (42.5%), and these alternatives did not differ in their frequency between the absence or presence of a judge's direction or from one another ( $\chi^2 = 1.93(2)$ ,  $p > .05$ ), as Figure 5 shows:



Thus, judges' directions did not significantly cue exemplars relevant to the consideration of explanatory, corroborating, or alibi alternatives more often than when a judge's direction was absent. That an equivalent amount of alternatives were generated, whether the judge's direction was present or absent, shows that the generation of alternatives were neither prompted nor suppressed respective of a judge's direction.

## General discussion

The results corroborate the theory that communication factors conducive to prompting explanatory alternatives can affect the impact of low random-match-probability DNA evidence on returned verdicts, guilt, evidence strength, and source probability ratings by jurors [3], especially in light of directly refuting evidence. That said, jurors did chose the verdict 'guilty' more often than 'not guilty' or 'cannot decide' in both studies overall indicative of the powerful impact that DNA evidence has on juror reasoning. In study 1, jurors chose guilty significantly more than other verdicts unless both an exemplar-cueing linguistic frame and evidence refuting the presence of the suspect at the crime scene were present.

Overall ratings of guilt decreased when exemplar-cues were present regardless of the presence of refuting evidence. Ratings of total evidence strength and strength of DNA evidence singularly decreased when exemplar-cueing linguistic frames and refuting evidence were present. Jurors explained their verdicts and ratings by reasons indicative of innocence explanations in the presence of exemplar-cueing, and guilty explanations in its absence. In sum, when alternative explanations are prompted in tandem with refuting evidence, jurors are unwilling to convict. Juror statements relevant to corroboration showed that where low random-match-probability DNA evidence was the only evidence (i.e., there was no exemplar-cueing or other evidence), jurors persistently requested more evidence. The implication is that jurors understand how corroboration requirements are necessary even in cases where the DNA evidence is highly compelling. In study 2, ratings of guilt, evidence strength, and source probability did not differ between scenarios in which a judge's direction to 'consider all possible explanations of the evidence' was present or absent. However, when the direction was absent, significantly more 'guilty' than 'not guilty', or 'undecided' verdicts were returned. Thus, the judge's direction does play an important role according to our juror reasoning in cases requiring the consideration of low random-match-probability DNA evidence.

## Implications and future directions

Jurors in these studies behaved quite logically. In the spirit of it being preferable to let ten guilty men go free rather than convict one innocent, jurors understood to take account of refuting evidence, to consider alternative explanations more often than not, understood the need to request corroborating evidence, and demonstrated the necessity of the presence of a judge's direction. While the studies herein were not conducted in courtrooms, or in groups, the value of examining the individual reasoning and qualitative reasons have proved theoretically enriching and informative [22]. Despite the presentation of low random-match-probability evidence in all scenarios, jurors were reluctant to convict when presented with minimal contradictory evidence, or forewarned by judges' directions to take account of all the evidence. Future legal empirical studies employing large-scale multivariate mixed-method designs to study these judges' directions and refuting scenarios, in the presence of a range of high and low random-match-probabilities, larger sets of evidence (e.g., corroborating and refuting, interactive evidential argumentation etc.) and group decision-making are likely to yield further important future insights.

## Acknowledgment

This research was funded by an SLSA Small Grant awarded to Michelle Cowley. The preliminary results were presented at the Evidence & Proof One-Day Conference at the Institute of Criminology, University of Cambridge. Thank you to Samantha Lundrigan, Katrin Mueller-Johnson and Mandeep Dhama for

discussions of this work. Thank you to Avanti Perera and Paul Honey for their valuable research assistance.

## References

1. Koehler JJ (2001) The Psychology of Numbers in the Courtroom: How to Make DNA-Match Statistics Seem Impressive or Insufficient. *Southern California Law Review* 74: 1275-1306.
2. Slovic P, Monahan J, MacGregor DG (2000) Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law & Human Behavior* 24(3): 271-296.
3. Koehler JJ, Macchi L (2004) Thinking about low probability events: Exemplar cueing theory. *Psychological Science* 15(8): 540-546.
4. Nuffield Council on Bioethics (2007) *The forensic use of bioinformation: Ethical issues*. Cambridge Publishers, Cambridge, UK.
5. Kassin SM, Bogart D, Kerner J (2012) Confessions that corrupt: Evidence from the DNA Exoneration Case Files. *Psychological Science* 23(1): 41-45.
6. Semikhodskii A (2007) *Dealing with DNA evidence: A legal guide*. Routledge-Cavendish, USA.
7. Dawid AP (2001) *Bayes's Theorem and Weighing Evidence by Juries*. Bayes Theorem, Oxford University Press, UK.
8. McCartney C (2008) 'LCN DNA: proof beyond reasonable doubt?' *Nature Reviews Genetics* 9(5): 325.
9. Aitken C, Roberts P, Jackson G (2010) *Practitioner Guide 1: Fundamentals of probability and statistical evidence in criminal proceedings: Guidance for judges, lawyers, forensic scientists, and expert witnesses*. In *Communicating and interpreting statistical evidence in the administration of criminal justice Four-Part Series*. Royal Statistical Society, UK.
10. BBC News (2008) McCann DNA evidence 'exaggerated'. BBC News Channel Online News, UK.
11. Thompson WC, Schumann EL (1987) Interpretation of statistical evidence in criminal trials: The prosecutors' fallacy and the defense attorney's fallacy. *Law & Human Behavior* 11(3): 167-187.
12. Dwyer D (2003) The duties of expert witnesses of fact and opinion: R v Clark (Sally). *International Journal of Evidence and Proof* 7(4): 264-269.
13. Freckleton I, Selby H (2005) *Expert Evidence: Law, Practice, Procedure and Advocacy* (3rd edn), Lawbook Co, Australia.
14. Koehler JJ (1997) When are people persuaded by DNA match statistics? *Law & Human Behavior* 25: 493-513.
15. Chater N, Tenenbaum JB, Yuille A (2006) Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Science* 10(7): 287-291.
16. Windschitl P D, Wells G L (1996) Base rates do not constrain non-probability judgments. *Behavioral and Brain Sciences* 19: 40-41.
17. Denes Raj V, Epstein S (1994) Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality Theory & Social Psychology* 66(5): 819 - 829.
18. Pacini R, Epstein S (1999) The relation of rational and experiential information processing styles to personality, basis beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology* 76: 972-987.
19. Gigerenzer G, Hoffrage U (1995) How to improve Bayesian reasoning without instructions: Frequency formats. *Psychological Review* 102(4): 684-704.
20. Tversky A, Kahneman D (1973) Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5(2): 207-232.
21. Schklor J, Diamond SS (1999) Juror reactions to DNA evidence. *Law & Human Behaviour* 23(2): 159-184.
22. Devine DJ, Caughlin DE (2014) Do they matter? A meta-analytic investigation of individual characteristics and guilt judgments. *Psycholog, Public Policy and Law* 20(2): 109-134.



This work is licensed under Creative Commons Attribution 4.0 License  
DOI: [10.19080/JFSCI.2017.03.555601](https://doi.org/10.19080/JFSCI.2017.03.555601)

### Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats  
( Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission

<https://juniperpublishers.com/online-submission.php>