

Statistical Evidence: An Investigation of Its Nature and Its Usage in the Criminal Context

Abstract

In *US vs. Shonubi*, the prosecution relied on impressively accurate statistics to prove the amount of drugs Shonubi carried. However, the appellate court quashed the sentence because it was not based on “specific” evidence.

But is there any *real* difference between statistical and individualistic evidence or is it merely a matter of presentation? And even if this difference is real, why should it matter for the criminal context?

In this paper, I argue that there is a real difference *in quality*. While the mainstream holds that “all evidence is probabilistic in a sense”, I argue that it is not in the *same* sense. There are different kinds of uncertainty involved and each is *irreducible* to the other.

I then argue that this qualitative difference has important implications for the criminal context. The debate about the usage of statistical evidence in criminal courts is in fact about the proper balance between two principles in tension: accuracy and personal responsibility. I then suggest a procedure that mitigates between them.

I take my theory to practice, and analyse three real cases, through which I deal with three residuals issues: statistical evidence used for non-convicting purposes, DNA evidence, and statistical evidence submitted by the defence.

Introduction

Some have said they have discovered the truth, some have asserted that it cannot be apprehended, and others are still investigating.

*Sextus Empiricus*¹

This paper investigates the issue of statistical evidence from an uncustomary angle: the nature of statistical *information* rather than the *inference*.

The application of statistics in Law of Evidence is everything but a neglected subject².

Yet, one important question seems to be neglected, namely whether there is any

¹ Sextus Empiricus, *Outlines of Scepticism*, Julia Annas and Jonathan Barnes trans. (Cambridge, 1994), book 1, section I, §1.

² “The New Evidence Scholarship” confronts the attempts to model the legal notion of “proof” and “reasoning” with statistical tools. See Richard Lempert, “The New Evidence Scholarship: Analyzing the

difference *in quality* between statistical and individualistic evidence, or whether any such difference is only an illusion created by the precision involved. It is important to note that this is a fundamental and non-legal question. However, this question immediately leads to a legal one: what the existence/absence of such quality implicates on the legal issue in question.

Although the overwhelming support that the quantitative approach seems to prevail³, the aim of this paper is to suggest an account for a *qualitative* difference, and then examine its implications on a specific legal question: the *admissibility* of statistical evidence in *criminal* context.

However, as a result of the shift in focus from inference to information, two interesting and important questions remain beyond the scope of this paper: (1) how to *infer* from statistical evidence⁴; and (2) how to *combine* it with conventional one⁵.

Process of Proof”, 66 *Boston University Law Review* 439, 440; William Twining, “The Boston Symposium: A Comment”, 66 *Boston University Law Review* 391, 391. Other scholars focused on the appropriateness of such evidence and reasoning in the legal context. See Laurence H. Tribe, “Trials by Mathematics: Precision and Ritual in the Legal Process”, 84 *Harvard Law Review* 1329 (1971); Charles R. Nesson, “Reasonable Doubt and Permissive Inferences: The Value of Complexity”, 92 *Harvard Law Review* 1187 (1979). Another debate developed around Jonathan Cohen’s challenging paradox of the gatecrasher, questioning what inference can be drawn from statistical evidence on a particular case. See L. Jonathan Cohen, *The Probable and the Provable* (Clarendon Press, Oxford), 1977, pg 74; David Kaye, “The Paradox of the Gatecrasher and Other Stories”, 1979 *Arizona Student Law Journal* 101; Mark Colyvan, Helen M. Regan and Scott Ferson, “Is it a Crime to Belong to a Reference Class?”, 9 *The Journal of Political Philosophy* 168 (2001); Peter Tillers, “Introduction: Three Original Contributions to Three Important Problems in the Law of Evidence”, 18 *Cardozo Law Review* 1875 (1997) [Tillers I]. Of course, this survey is scarcely the tip of an iceberg.

³ See section I, under “The Quantitative Approach”.

⁴ For summaries on inference and the extensive legal debate around Bayes Theorem, see Paul Roberts & Adrian Zuckerman, *Criminal Evidence* (Oxford University Press, 2004), 123-132; Ian H. Dennis, *The Law of Evidence* (Sweet & Maxwell, 2nd ed, 2002), 110-116.

⁵ See the particular debate around *Adams*. See *R. V. Adams* (No. 1) [1996] 2 Cr App R 369, CA [Adams I]; *R. V. Adams* (No. 2) [1998] 1 Cr App R 377, CA [Adams II]; Philip A. Dawid, “Statistics and the Law”, in *Evidence* (Karin Tybjerg, John Swenson-Wright, and Andrew Bell eds., Cambridge University Press, to appear) [Dawid I] 9-14.

The outline of my argument is as follows. The qualitative difference between statistical and individualistic evidence lies in the way the information is delivered: while the information conveyed by individualistic evidence is delivered on *one* individual, the information of the statistical evidence must always be *mediated by the individual's membership in a certain group*. Therefore, in addition to the uncertainty regarding the *reliability* of the evidence from which both types of evidence suffer, this move from the group to the individual creates *an additional* kind of uncertainty regarding the *applicability* of the information on the individual. I argue that this kind of uncertainty exists *only* in statistical evidence and is *irreducible* to the first kind.

When considering the legal question of admissibility in the criminal context, the debate about statistical evidence is yet another manifestation of the fundamental tension between consequentialist and deontological justifications for criminal law. A pure consequentialist approach that aims to maximise *accuracy* should not be bothered by this qualitative difference and, thus, admit any statistical evidence that increases accuracy. In contrast, a pure deontological approach should be loyal to the principle of *personal responsibility*, and, thus, render any statistical evidence *inadmissible*.

My suggestion for mitigating between the two extremes is to overcome the lack of individualisation with a requirement for corroboration with additional *individualised* evidence. However, as long as such evidence is unavailable, statistical evidence should be inadmissible in the criminal context.

I end by exemplifying the clarification that my account can provide and analyse a few real cases that address some specific satellite issues: non-convicting statistical evidence, DNA evidence, and statistical evidence submitted by the defence

It is important for me to emphasise that my main aim in this paper is not to provide a knock-out answers but rather to clarify what the debate on statistical evidence is really about. Thus, my goal is to promote a better *understanding* of statistical evidence and, hence, my suggested procedure should be regarded only as one possibility to mitigate the tension between the conflicting principles.

The paper consists of two parts. Part one is a fundamental investigation of the nature of statistical evidence. It is divided into three sections. (I) criticises the existing accounts. (II) puts forward my own account for qualitative difference. (III) tackles some potential objections from the quantitative mainstream.

Part Two examines the legal implications that this qualitative difference has. (IV) examines this difference in light of the tension between the consequentialist and deontological justifications. (V) details my suggestion to mitigate between them and (VI), through a number of real cases, analyses the residual issues mentioned above.

Part One: An Investigation of the Nature of Statistical Evidence

Since I use the terms ‘statistical’ and ‘individualistic’ without supplying any definitions beforehand, a preliminary methodological clarification is required. I do not assume that there is a *real* difference between them, but only that some types of evidence are *considered* statistical and others ‘individualistic’. Whether this distinction is sustainable is the aim of my investigation. Thus, I assume only a vague understanding of what these terms mean.

Section I: The Insufficiency of the Existing Approaches

A novel yet comfortable way to approach this investigation is by using the dichotomy *qualitative/quantitative*. A *qualitative* approach insists on certain categorical differences between statistical and individualistic evidence while a *quantitative* approach argues that in a *final account* there is *no real* difference between them: while both suffer from uncertainty, statistical evidence courageously admits its precise level.

There is an important implication for this dispute: if the difference is in quality, then these two types of evidence *cannot* be reduced to the *same scale*. However, if the difference is merely a matter of quantity (or presentation), then there is one scale for measuring both.

Before turning to survey the existing literature, one note of caution is required. Perhaps due to the dominance of the quantitative approach, I had difficulties in locating many consistent and detailed accounts for *either* side. I collected some opinions from footnotes and artificially constructed others. Thus, my survey should not create the wrong

impression that the debate about the nature of statistical *information* is a focused and lively one.

The Quantitative Approach

The Approach

The quantitative approach is not obliged to renounce the intuition that there is a difference between statistical and individualistic evidence. Rather, it draws it as a matter of *presentation* rather than a real difference in *quality*. Thus, the argument goes, although any evidence suffers from an inherent uncertainty that gives rise to a risk of error, statistical evidence “makes the risk of error explicit”⁶. The legal debate then turns to the question of whether this difference in presentation should have any legal consequences⁷.

The quantitative approach seems to be the dominant one as it wins supporters even amongst the leaders of the camp that strongly *objects* to the usage of statistical evidence in the criminal court. As Tribe’s recognition can testify, eminent scholars can sometimes have deep disagreements regarding statistical evidence and yet share the basic standpoint that any evidence is statistical⁸. Shaviro and Koehler assert that “all evidence is probabilistic”⁹. While disagreeing with them on other matters¹⁰, Allen notes that he has

⁶ Koehler and Shaviro (n 34), 253.

⁷ See Section IV and *supra* 70 in particular.

⁸ “*all* factual evidence is ultimately ‘statistical’ and all legal proof ultimately ‘probabilistic’”, Tribe (n 2), note 2.

⁹ Koehler and Shaviro (n 34), 251.

¹⁰ Ronald J. Allen, “On the Significance of Batting Averages and Strikeout Totals: A Clarification of the Naked Statistical Evidence Debate, the Meaning of Evidence, and the Requirement of Proof beyond a Reasonable Doubt” 65 *Tulane Law Review* 1093 [Allen I]; and see Shaviro’s response to these disagreements, Daniel N. Shaviro, “Response to Professor Allen”, 65 *Tulane Law Review* 1111.

“some difficulty with the proposition that there is an important qualitative difference between “statistical” evidence and other kinds of evidence, for...all evidence is “probabilistic” in a nontrivial sense”¹¹.

The American judiciary system contains similar acknowledgments. “All evidence is probabilistic— statistical evidence merely explicitly so”¹² or “Statistical evidence is merely probabilistic evidence coded in numbers rather than words”¹³ are just a few examples¹⁴. Even eyewitness testimony is conceived as merely probabilistic¹⁵.

Criticism

The role that the quantitative approach seems to play in this debate was to *assert* the lack of a qualitative difference. It assumes that the burden of proof lies on the rival camp to come up with a consistent and coherent definition for such a qualitative difference. However, it does not strike me as evident why this debate has to begin with the hypothesis that such a qualitative difference does not exist.

Nevertheless, the main difficulty with this assertion is that it asserts much more than it pretends to: not only any evidence is ‘statistical’ in *a* sense, but also it is statistical *in the same sense*. The assertion is that there is *only a single kind of uncertainty*, from which all evidence suffers. Hence, to prove its case, the quantitative approach has to show that

¹¹ Allen I (ibid), note 10.

¹² *Riordan v. Kempiners*, 831 F.2d 690 (7th Cir. 1987), 698.

¹³ *U.S. v. Veysey*, 334 F.3d 600 (7th Cir. 2003), 606.

¹⁴ For more, see *Howard v. Wal-Mart Stores, Inc.*, 160 F.3d 358 (7th Cir. 1998), 360; *Milam v. State Farm Mutual Automobile Ins. Co.*, 972 F.2d 166, 170 (7th Cir.1992); 170; *United States v. Chaidez*, 919 F.2d 1193, 1200 (7th Cir.1990), 1200.

¹⁵ *Branion v. Gramly*, 855 F.2d 1256 (7th Cir.1988), 1264.

“uncertainty” is homogenous concept that cannot be divided into different kinds. This strong proposition cannot simply be asserted but requires a positive account in addition to just challenging rival qualitative suggestions.

The Qualitative Approach

What motivates a qualitative approach is the strong intuition that individualistic evidence such as eyewitness has a particularistic/specific/direct quality that statistical evidence lacks. The challenge it faces is not only to convince that the difference is real rather than a matter of presentation, but mainly to properly define this quality. However, I argue that none of the existing accounts is satisfactory.

Assertion: some simply asserts the existence of such qualitative difference. Judge Newman, for instance, in his judgement in *Shonubi*, requires evidence to be “specific” to the issue in question.¹⁶ However, as justifiably criticised, there is no available definition for this quality of “specificity”, neither in Newman’s verdict nor in the federal rules.¹⁷

Presentation: others hang the difference on difficulties in presenting and relying on statistics. Schmalbeck, for example, mentions that statistical evidence “is often

¹⁶ *United States v. Shonubi*, 998 F.2d 84 (2d Cir, 1993) [Shonubi II], 89.

¹⁷ Judge Weinstein of the lower instance, *United States v. Shonubi*, 962 F. Supp. 370, 375 (E.D.N.Y. 1997) [Shonubi V], 375. For detailed criticism, see Peter Tillers, “If Wishes Were Horses: Discursive Comments on Attempts to Prevent Individuals from Being Unfairly Burdened by their Reference Classes”, forthcoming in *Law, Probability and Risk* [Tillers II].

treacherous”¹⁸. However, this recognition is perfectly consistent with the quantitative approach, as it deals with the presentation rather than the substance.

Causation: a tempting attempt is Thomson’s detailed account using causation. Thomson explains the intuition by stating that individualistic evidence is intuitively conceived as such because it “is in an appropriate way casually connected with the (punitive) fact that the defendant caused the harm”¹⁹. The alleged fact that individualistic evidence aims to prove has *appropriate* causal connection to *both* the harmful result *and* the existence of evidence. However, statistical evidence does not keep the same causal connection as this evidence “neither supplies nor suggests anything which might have been a common cause *both* of those facts available to us *and* of the (putative) fact”²⁰.

Thomson's account seems to be inspired by Goldman’s causal theory of knowledge, and, unfortunately, suffers from similar problems²¹. First, Thomson does not clarify what distinguishes an *appropriate* causal connection from an inappropriate one²². Second, Thomson entangles the discussion instead of clarifying it, as she replaces

¹⁸ Richard Shmalbeck, “The Trouble with Statistical Evidence”, 49 *Law and Contemporary Problems* 221 (1986), 230. Schmalbeck suggests three kinds of “problems” from which statistical evidence suffer, but I focused only on the third. The first “problem”, ‘confidence level’, was part of my account as well. The second is that statistical evidence presents information “to one element in a string of events”. I failed to understand why this is applied only to statistical evidence.

¹⁹ Judith J. Thomson, “Liability and Individualized Evidence” 49 *Law and Contemporary Problems* 199, 203. Thomson analyses a civil case, but her account for the ‘individualistic’ quality is general.

²⁰ Thomson (n 19), 204.

²¹ Alvin I. Goldman, “A Causal Theory of Knowing”, *The Journal of Philosophy* 64 (1967), 357-372. For difficulties in this general theory, see Roderick M. Chisholm, “The indispensability of internal justification”, *Synthese* 74 (1988), 285-296, §5; Keith Lehrer, “The coherence theory of knowledge”, *Philosophical Topics* 14 (1986), 5-25, §12; Richard Foley, “What’s wrong with Reliablism?”, *The Monist* 68 (1985), 188-202, §II. For detailed criticism of Thomson’s specific account, see Schmalbeck (n 18).

²² Schmalbeck (n 18), 232.

individualisation with another notion which is not less complex, if not much more so: *causation*.

It seems that another attempt, Wright's one, suffers from the same problem. Wright classifies evidence into four types: "individualistic" evidence, *ex post* casual probabilities, *ex ante* casual probabilities, and "naked statistics"²³. In general, "individualistic" evidence shows that a certain casual generalisation has *instantiated* in a particular case that in front of a court. However, not only does the entanglement in the notion of causation haunt this account too, but I also doubt if replacing one term ("individualisation") with another ("instantiation") provides any clarification.²⁴

Luck: another suggestion can be extracted from Thomson's account: when one relies on statistical evidence, even if one is correct, it is merely *due to luck*²⁵. If I seek a justification for my belief that a certain event will occur, but I know in advance that there is a chance (as small as it may be) that it will not occur, I cannot claim I have a *justification* for this belief²⁶. Thus, the quality that statistical evidence lacks is the ability to supply a proper epistemic justification.

²³ Richard W. Wright, "Causation, Responsibility, Risk, Probability, Naked Statistics, and Proof: Pruning the Bramble Bush by Clarifying the Concepts", 73 *Iowa Law Review* 1000 (1988), 1050.

²⁴ Koehler also criticises Wright's account for confusing between probity and policy considerations. Jonathan Koehler, "The Probity/Policy Distinction in the Statistical Evidence Debate", 66 *Tul. L. Rev.* 141, 148 [Koehler I].

²⁵ Thomson uses this to answer a different question: why the Law should treat such evidence differently. However, if this quality is used to justify the different treatment, it can also be used to identify which evidence belongs to the group that deserves it.

²⁶ Thomson (n 19), 208.

However, the notion of 'luck' in Thompson's account is not less problematic than that of causation as it relies on a strong metaphysical commitment to chance regarding *past* events.²⁷ In the legal context, most litigated events belong to the past and, thus, accepting the notion of chance regarding *future* events is *insufficient*. One can easily argue that the correctness of a decision based on statistical evidence was not determined by luck, as by the time of taking the decision, no luck was involved. Rather, what prevents the correctness to be guaranteed is only *lack of information whether this luck was realised or not*.

What might seem as a sophistic quibble is in fact lethal to the qualitative difference Thomson tries to establish. What is, then, the difference between the shortage of information in the question of whether the eyewitness lies and in the question of whether a past luck was realised or not?

Weight: Stein argues that only case-specific evidence supply *enough weight* to support fact-finding in a specific case.²⁸ Stein distinguishes between two meanings of “probable”: the first is “calculated chances that probability estimates express” and the second is “evidential support that probability estimates rest upon”.²⁹ This distinction leads to a criterion of *weight* according to which “some probability estimates both can and should be perceived as more evidenced, and, correspondingly, as weightier than others”.³⁰ Stein

²⁷ For more about chance and propensity theory, see Donald Gillies, *Philosophical Theories of Probability* (Routledge, Oxon), 2000, section 6; Ian Hacking, *The Taming of Chance* (Cambridge University Press, 1990).

²⁸ Alex Stein, *Foundations of Evidence Law* (Oxford University Press, 2005), 80-91[Stein I]. It is unclear to me whether Stein goes as far as arguing that statistical evidence has *no weight* at all, or whether it has much less weight than case-specific evidence. I chose the latter understanding, as it is less radical. However, my following criticism would apply on the former understanding too.

²⁹ *Ibid*, 81.

³⁰ *Ibid*, 82.

concludes that “This criterion introduces the demand for case-specificity in attaching probability estimates to individual events”.³¹

Stein’s account suffers from two main problems. The first is that this criterion of weight is insufficient to support Stein’s conclusion. One can easily accept this criterion, and yet argue that it distinguishes between well-informed decisions and less-informed ones, regardless the question whether the information is case-specific or not. For example, in Shonubi’s case, the prosecution compiled statistics about drug smugglers from the same country of origin, entering to the US through the same airport in the same year using the same method of smuggling.³² This is not case-specific evidence as there was nothing specific in this information regarding Shonubi’s case, and he was not even included in this statistics. However, it does not strike me as evident that this impressively accurate information fails to give enough evidential support to the probability estimates.

The second problem is even more crucial, in my view: Stein’s account only replaces one question with another. The question on the table was whether there is a qualitative difference between statistical and individualistic evidence. Stein’s answer is that they differ in weight because case-specific evidence gives a better (“weightier”) support to the factual findings (“probability estimates”).

However, this begs the question *why* case-specific evidence is weightier than statistical one *by definition*. One has to explain what case-specific evidence *has* that statistical evidence *does not have* that makes case-specific evidence weightier. Responding that

³¹ Ibid.

³² *United States v. Shonubi*, 895 F.Supp. 460 (E.D.N.Y.) [Shonubi III].

case-specific evidence gives better evidential support because the question itself is case-specific will not do, because this assumes precisely what has to be proven. The question is what quality, if any, case-specific evidence has that makes them more appropriate to answer case-specific questions. This question is the one needed an answer and, yet, was not answered by Stein's account.

Section II: My Account for the Qualitative Difference

The Difference: Particular vs. Group-Mediated

My account for the difference between individualistic and statistical evidence stands on the way each type of evidence delivers its information: statistical evidence delivers it *via* a group while individualistic evidence delivers it on one individual. Statistical evidence teaches the fact-finder something about a certain group, and any conclusion she arrives at regarding an individual member is based on his membership in this group. Individualistic evidence, in contrast, requires no such group as it applies only on a single individual. Thus, the information conveyed in the statistical evidence is *group-mediated*.

Furthermore, this move from the group to the individual involves an additional kind of uncertainty from which only statistical evidence suffers. It is uncontroversial to admit that any evidence suffers from uncertainty. Indeed, any evidence suffers from uncertainty regarding its *reliability*, namely the correspondence between the facts *in reality* and the facts *alleged by the evidence*. When rival scholars agree that any evidence is statistical³³, they probably mean that any evidence is exposed to a risk of error in its attempt to represent the actual facts correctly³⁴.

However, statistical evidence suffers from *additional* type of uncertainty regarding its *applicability*, namely how the information it delivers on the group relates to a particular

³³ See the text around footnotes 8 and 11.

³⁴ See in particular Jonathan J. Koehler and Daniel N. Shaviro, "Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods" 75 *Cornell Law Review* 247 (1990) [Koehler II].

member³⁵. Even if the evidence is perfectly reliable, and, thus, fully corresponds with reality, a troubling question remains: what *support* this evidence gives to any conclusion about the individual member in question. Therefore, even when both parties agree that the statistical evidence is true, they can still dispute the conclusion that should be derived from it.

This applicability problem can be separated into two different questions: first, amongst the *sub-groups* that construct the group, to which one this particular individual belongs? Second, amongst the infinite groups to which this individual belongs, which one is the most appropriate to use.³⁶

These questions are illustrated in the following example. For statistical evidence, take a variation on the Gatecrasher paradox³⁷: 1,000 people watched a rodeo but only 100 tickets were sold. The organisers sue Humpty, a random watcher, and want to prove that Humpty gatecrashed with the evidence that 90% of the watchers did not pay³⁸. For individualistic evidence, take Alice, an eyewitness, who testifies to seeing Humpty sneak through the fence.

This example illustrates the qualitative difference. The organisers' statistical evidence delivers information on the group of watchers. Any conclusion on Humpty depends on his membership in this group. In contrast, Alice's testimony delivers information only on Humpty, and requires no group to which Humpty should belong.

³⁵ My distinction is based on Peter F. Strawson, *Introduction to Logical Theory* (Methuen & Co. Ltd, London), 1952, 233-262. In a legal landscape, a similar idea can be found in Shmalback (n 18), 225-228.

³⁶ This question is the well-known "problem of the reference class". For a clear description, see Gillies (n 27), 119.

³⁷ Cohen (n 2), 74.

³⁸ The legal context is intentionally vague. I assume (1) proving lack of payment is an essential element of the claim, whatever it would be; (2) there are grounds for such a claim.

As a result, the organisers' statistical evidence suffers from the additional kind of uncertainty regarding its *applicability* on Humpty, in addition to the uncertainty regarding its reliability. The evidence might be intentionally forged (the cashier stole the money of 900 tickets and covers himself with this story) or innocently inaccurate (they miscounted the revenue). Such questions are directed against the *reliability* of the evidence as they undermine the correspondence between what the evidence claims to have happened and what actually happened.

However, even if the reliability of this evidence is perfect, its *applicability* is still an issue, because it raises the two questions mentioned above.³⁹ The first question is whether Humpty belong to the sub-group of the 900 who saved themselves the trouble of paying, or to the sub-group of the 100 innocent payers? The second question is regarding the appropriateness of this statistics. Assuming there are two additional statistics: (1) 20% of the watchers of rodeos in general gatecrash, and (2) 80% of all egg-shaped creatures gatecrash. Which statistics is the most appropriate to apply on Humpty? Should we corroborate all of them or use only the most relevant one? And what should we do when the only statistics available to us is the less relevant one?

In contrast, Alice's testimony suffers from only one kind of uncertainty. Alice might intentionally lie (for she bears a grudge for Humpty), mistakenly misidentify him (for she confused him with a white rabbit) or simply incredible in general (due to the delusions she frequently has). However, all these questions go to the *reliability* of her evidence. Once her testimony is considered reliable, there can be no additional uncertainty about its applicability to Humpty because no group was involved to begin with.

³⁹ See text around n 36.

My argument is that the problem illustrated in this example is an *inherent* problem with *any* statistical evidence due to a fundamental problem concerning probabilities in general. Probabilities can be divided into *definite and indefinite*⁴⁰. Definite probabilities deal with a *particular* event while indefinite ones concern a property or a class. Statistical evidence is usually information about frequencies: within a certain group, a certain proportion of members have a certain feature⁴¹. This can only supply *indefinite* probabilities about a property⁴². However, what the court needs is *definite* probabilities regarding the *particular case*. Moving from indefinite probabilities to definite ones is considered very difficult regardless of one's preferred interpretation of probability⁴³.

A Method of Identifying Statistical Evidence

A simple test can be proposed to identify whether given evidence is statistical or individualistic. In statistical evidence, the same information can be applied to *any* member of the group *in precisely the same way*.

This bowls down to two conditions:

- (1) Precisely the same *information* can be used for each member AND
- (2) The exact same *conclusion* can be drawn on each member.

⁴⁰ John L. Pollock & Joseph Cruz, *Contemporary Theories of Knowledge* (2nd ed, Rowan & Littlefield, 1989), 93.

⁴¹ For a detailed account of the Frequency Theory, see Gillies (n 27), section 5.

⁴² Pollock (n 40), 93.

⁴³ For the full argument, see Pollock (n 40), 92-111. Pollock shows that *neither* the subjective, indefinite physical nor propensity theories can support such a move. He deals with definite “epistemic probabilities” required for justification, but his discussion can be applied to the definite probabilities courts need for a decision. For similar conclusions, see also Roberts & Zuckerman (n 4), 120-121.

Applying individualistic evidence to a different individual requires, in contrast, either (1) adapting its content or (2) changing the conclusion that this evidence draws.

For example, in order to use the organisers' statistical evidence on Hatter, another random watcher, nothing has to be changed in the evidence itself and the same conclusion that could be drawn on Humpty can be drawn on Hatter too. In contrast, in order to use Alice's testimony on Hatter, either (1) its content has to be changed to include him (so she has to see Hatter instead or in addition to Humpty) or (2) the conclusion should be different (so the fact that Humpty was seen sneaking *slightly reduces* the probability that Hatter sneaked too).

Pseudo-Statistical Evidence

In order to sharpen the definition of statistical evidence even further, I would like to address two examples for what should *not* be considered as statistical evidence, although it might be conceived as such at a first glance.

Group vs. List

Consider the following example: Guinevere testifies she saw Arthur, Lancelot, Galahad and Percival sneaking through the fence. Allegedly, this evidence can be applied to each of them in precisely the same way, meaning that the same information can be used to draw the same conclusion regarding each member. Thus, the information delivered by

this evidence is group-mediated. However, it would be completely counter-intuitive to classify such evidence as ‘statistical evidence’⁴⁴ ...

The fallacy here is confusing a *list* of separated pieces of evidence with a *group*. Considering Guinevere’s testimony as a single piece of evidence which refers to a group of individuals is a misleading analysis of the situation. Eyewitness testimony can contain many factual claims, i.e. can be divided into *several pieces of evidence*. Thus, Guinevere’s testimony can be divided into *four* factual claims; each of which constitutes a separate piece of evidence against different individual.

Identifying whether evidence is dividable is simple: one should ask the hypothetical question of whether the evidence can be true regarding only *part* of the members. If so, it is *dividable list*. In Guinevere’s testimony, she may have been right about Lancelot, but misidentified Galahad. In contrast, in statistical evidence, if the organisers miscounted the number of watchers, it will have the *same* effect on the conclusion for *each* of the members.

Statistical Evidence vs. Logical Necessity

Assume that, instead of presenting evidence that only 100 tickets sold, the rodeo organisers now present evidence that *not a single ticket was sold*.

This issue does not seem to have a significant practical importance and I found it hard to find any realistic examples. However, it raises interesting analytical questions: should

⁴⁴ Because I doubt if intuition and statistics are a good match, I avoided relying on intuitions when putting forward my account. Yet, I use intuitions to support *counter-arguments* against my account. However, if one denies the intuition, the denial undermines the objection rather than my account...

such evidence be considered as statistical?⁴⁵ Allegedly, the answer is a simple yes, as the information it delivers on Humpty is again only via his membership in the group of watchers. Using the suggested test, this evidence can be used against any of the watchers precisely in the same way.

However, I argue that this evidence should not be considered as statistical evidence, at least for legal purposes⁴⁶. The reason is because there is no uncertainty regarding the applicability on the individual member. Of course, like any other evidence, there is an uncertainty regarding its reliability. However, once the fact-finder is satisfied with its reliability, the conclusion that Humpty did not pay is *logically necessitated* from the information that no single ticket was sold and that Humpty watched this rodeo. The form of argument here is *deductive* rather than inductive, as it is logically impossible for the conclusion to be false if the premises are true.⁴⁷

However, the question remains why evidence that is group-mediated needs also to suffer from this additional kind of uncertainty in order to be counted as “statistical”? Why this additional type of uncertainty is a necessary condition for evidence to be classified as “statistical”?

The answer derived from the legal context in which this question rises. One of the unique features of the *adversarial* legal process is that the court decides only on matters which

⁴⁵ I owe this important objection to Professor Dawid.

⁴⁶ Whether other disciplines, such as statistics or philosophy, should regard such evidence as statistical or not is outside of my scope.

⁴⁷ For a clear description of the deductive/inductive distinction, see Roberts & Zuckerman (n 4)112-113.

are *in dispute between the parties*⁴⁸. This is a result of the parties' control over the process⁴⁹ and the orientation of adversary systems toward conflict-resolution rather than truth-finding⁵⁰. Although it might seem counterintuitive to a non-legal observer, certain prepositions can be regarded by the court as “true” even without any evidence supporting them but only as a result of parties' agreement.

Because the applicability of the information on the individual is secured by a logical deduction, *by definition* no dispute can rise between the parties about the applicability of this information on the individual. The only possible dispute is regarding its reliability, but in that respect, this evidence is not different from any other individualistic evidence.

Section III: Potential Quantitative Objections to My Qualitative Account

It is possible to identify at least two potential lines of quantitative attack on my account. The first argues that the two types of uncertainty, reliability and applicability, can be reduced into a single scale of uncertainty. The second line argues that all evidence is group-mediated, and, therefore, individualistic evidence does not exist.⁵¹

⁴⁸ Roberts & Zuckerman (n 4), 48-49.

⁴⁹ Ibid.

⁵⁰ Mirjan Damaška, “Evidentiary Barriers to Conviction and Two Models of Criminal Procedure: A Comparative Study”, 121 *University of Pennsylvania Law Review* 506 (1973) [Damaška-I]; Mirjan Damaška, “Structures of Authority and Comparative Criminal Procedure” 84 *Yale Law Journal* 480 (1975) [Damaška-II].

⁵¹ Tillers raises a similar objection, see Tillers II (n 17). However, his critique is against the “specific” evidence requirement, so I had to adjust it slightly in order to run it against my account too.

(A) The Two Types of Uncertainties are Reducible to One.

The quantitative camp can respond to my account by arguing that the two types of uncertainty I described, reliability and applicability, *are reducible to a single scale of uncertainty*. For instance, simple multiplication can be used: assuming that the probability that organisers' statistical evidence is true equals to 0.7 (reliability)⁵² and according to this evidence, if 900 out of 1,000 did not pay, the probability of Humpty avoided paying equals to 0.9 (applicability). Thus, the probability that the evidence is correct AND that Humpty did not pay is 0.63 (0.9×0.7)⁵³.

However, although such a multiplication of probabilities is supported by every interpretation of probability⁵⁴, it is allowed only when the probabilities are *independent*: when the first probability is insensitive to any changes in the second one⁵⁵. Hence, in order to be independent, the probability of the applicability should remain unchanged under any change in the probability of the reliability of the evidence. Thus, increasing its reliability should *never* change the amount of uncertainty of its applicability.

However, this seems to be a very strong assumption. If any assumption has to be made, it is more plausible to assume that the two types of uncertainty are *dependent* even in a *very weak* connection. As even a very weak connection is sufficient to break the assumption of

⁵²The above text assumes that such quantification of reliability is (1) possible (2) normatively desirable. However, both assumptions were strongly contested, and for good reasons, I think. See Roberts & Zuckerman (n 4), 125; Mark Kaplan, *Decision Theory as Philosophy* (Cambridge, 1996), 23-31.

⁵³ It is also contentious if and how this number can be translated into a legal decision (of guilt, for instance). See Roberts & Zuckerman (n 4), 127-128.

⁵⁴ Terence Anderson & William Twining, *Analysis of Evidence: How to Do Things with Facts* (Weodenfeld and Nicolson, London), 1991, 416.

⁵⁵ For formal definition of independence, see Gillies (n 27), 75.

independence on which this reduction stands. Hence, it is really questionable whether an assumption of independence between reliability and applicability can be easily made.

Moreover, even if independence holds, this reduction leads to problematic results as it grants the same amount of uncertainty to two fundamentally different situations⁵⁶. Compare these two extremes: (1) highly reliable evidence (0.99) shows that the chance that Humpy avoided paying is 0.1; and (2) immensely unreliable evidence (0.1), which shows that this chance is 0.99. Surprisingly, the reduction equips the two with the same degree of uncertainty. But can it seriously be argued that these two pieces of evidence suffer from precisely the same uncertainty?

A quantitative approach can reply that multiplication is not the correct method to reduce these two types into one. Another possibility is to set a certain threshold for reliability and admit only evidence that is above this threshold. A rough example for such a model is admitting only research that was conducted in the proper scientific standard (and, hence, is highly reliable), and then focusing on its applicability to the specific case.

However, such a model does not undermine my qualitative account but, rather, strengthens it. While I treat reliability as a continuous spectrum, this model treats it as a binary value: evidence can be considered reliable (whether its reliability is above the threshold) or not. However, this model preserves the qualitative difference because each type of uncertainty receives a *different treatment* instead of being reduced into one.

⁵⁶ This argument is based on Kaplan's argument against Orthodox Bayesianism, Kaplan (n 52), 26.

According to my qualitative account, these two types are irreducible to one scale not merely because the lack of a reduction procedure. Rather, they are irreducible because there is a difference in quality between statistical and individualised evidence.

(B) Individualistic Evidence does not exist

Another objection argues that *all* evidence is group-mediated, even if the reliance of some evidence on a group is much more implicit. Any individualistic evidence also relies on *properties* of the individual rather than referring only to the individual him/herself. For instance, Alice's perception was not of "Humpty" but of some properties that Humpty has (she saw someone short, chubby, etc). She then concluded she saw *him*. However, this is an *opinion* rather than an actual report on what she saw⁵⁷. Therefore, even eyewitness testimony is actually group-mediated evidence, and, hence individualistic evidence simply does not exist.⁵⁸

It is questionable whether this is a plausible or even sustainable account of how our cognitive face recognition really works⁵⁹, but for the sake of the argument I would assume it does.

My account has a pragmatic solution to such difficulties.⁶⁰ My question is only *how many* individuals the evidence is applicable precisely in the same way. If the answer is

⁵⁷ For illuminating criticism on the fact/opinion distinction, see Roberts & Zuckerman (n 4), 143-146.

⁵⁸ Note that I do not argue that eyewitness testimony is *always* individualistic evidence. Eyewitness can also be group-mediated evidence (an eyewitness who claims to see 9 unidentified people sneaks into the rodeo while there were 10 people inside in total). However, the objection here is that eyewitness is *never* individualistic.

⁵⁹ For a thorough review on the subject of face recognition, see Mark H. Johnson, "Subcortical Face Processing", *Nature Reviews Neuroscience* 6, 766-774 (2005). This review surveys the evidence that supports the existence of a fast pathway for face detection that can produce a face-selective response *before* early visual cortical areas are activated.

one, this evidence can be considered individualistic – if it is more than one, than the evidence is group-mediated.

Now in order to undermine my method, the objection has to convince that *no single piece of evidence* can be applied to *only one individual*. Alternatively, it has to show that *any evidence* can be applied to a group of individuals without changing its informative content or the drawn conclusion.

Both directions are hopeless, as most evidence cannot be simply pulled out from one case and be applied to a different individual *as is*, without first changing either content or conclusion. Indeed, many types of evidence fit my definition of individualistic evidence. In addition to the eyewitness testimony analysed above, a few more examples can be listed: the individual's medical records; the accused's confession; previous convictions (when admissible); etc. The variety of examples reinforces the stance that individualistic evidence is a viable classification.

Summary

My account of the qualitative difference between statistical and individualistic evidence lies in the way each delivers its information on the individual. While individualistic evidence refers only to one individual, statistical evidence refers to a group, and a move is required from the group to the individual member. However, this move suffers from an additional and irreducible kind of uncertainty. I also accompanied my qualitative account with simple tests to identify what evidence is statistical and to distinguish it from pseudo-statistical evidence.

⁶⁰ Compare to *Gayle's* attempt to maintain the description/identification distinction, *R. v. Gayle* [1999] 2Cr App R 130.

However, the fact that the difference is in quality does not automatically imply that individualistic evidence is better than statistical one. For instance, eyewitnesses can be very unreliable or inaccurate⁶¹. Hence, I now proceed to analyse the implications of this qualitative difference on the legal question whether statistical evidence should be admissible in the criminal court.

⁶¹ For the case against eyewitness testimony, see Roberts & Zuckerman (n 4), 218-221.

Part II: Investigating Whether the Qualitative Difference should have Legal Implications

After accounting for the existence of a qualitative difference between statistical and individualised evidence, the obvious normative question is which implications, if any, this difference should have in the legal context.

I focus only in one issue: the admissibility of statistical evidence in the criminal context. My method is to identify which basic principles are in tension in this issue, and to reveal what commitment is required, and which implications should follow when one wants to argue in favour or against the admissibility of statistical evidence in the criminal context.

Section IV: The Legal Implications of the Group Mediation

In various legal fields, it is common to distinguish between *efficiency* justifications and *fairness* ones.⁶² While this might not be an exhaustive distinction, I think it supplies an adequate analytical tool.⁶³

When justifying the criminal process, this efficiency/fairness distinction is best articulated, in my opinion, as the tension between two types of justifications for

⁶² This distinction probably relies on the ethical distinction between consequentialist and deontological theories. Williams tracks down the sources of this distinction in Kant and Bentham. Bernard Williams, "Ethics" in *Philosophy: A Guide Through The Subject*, 1st volume, 545-583 (Oxford University Press, 2nd ed, 2000), 551-556

⁶³ A strategy similar to mine is employed in many legal fields. See, for instance, Harel for criminal law, Alon Harel, "Efficiency and Fairness in Criminal Law: The Case for a Criminal Law Principle of Comparative Fault", 82 *California Law Review* 1181 (1994); Coffee for class actions, John C. Coffee, "The Regulation of Entrepreneurial Litigation: Balancing Fairness and Efficiency in the Large Class Action", 54 *University of Chicago Law Review* 877 (1987); Anderson for company law, Alison Grey Anderson, "Conflicts of Interest: Efficiency, Fairness and Corporate Structure", 24 *UCLA Law Review* 738 (1977-1978); Nelson for tort law, William Nelson, "From Fairness to Efficiency: The Transformation of Tort Law in New York, 1920-1980", 47 *Buffalo Law Review* 117 (1999); Krawiec for equities law, Kimberly D. Krawiec, "Fairness, Efficiency, and Insider Trading: Deconstructing the Coin to the Realm in the Information Age", 95 *Northwestern University Law Review* 443 (2000-2001).

punishment: consequentialists and deontologicals.⁶⁴ Consequentialist justifications evaluate the criminal process by its good (or bad) consequences, mainly its efficiency in *crime reduction*. Deontological justifications are diverse, but the common ground is their emphasis on *the rights of the individual*, rather than the aggregative “common good” that haunts the consequentialists. The notion of fairness, thus, is manifested in the constraints these rights put on the measurements taken to reduce crime.

Before turning to extract the attitude that each justification should have towards statistical evidence, one important note of caution is necessary. I am well aware that I use radically naïve version of either approach (and of this distinction in general). Hence, I doubt if either of them are actually held by someone, as some may try combining or balancing between the two edges, or at least enhancing one of them to immune it from some obvious criticism. However, for my purposes, these simple versions are sufficient. I use the purest variations intentionally, in order to make the tension between the two approaches most evident.

⁶⁴ Robert & Zuckerman (n 4), 10-16. There are several more ways to describe this tension. One way is to pose “Due Process Model” versus “Crime Control Model”. See Damaška I (n 50), Damaška II (n 50); Alex Stein, “Political Analysis of Procedural Law, *A Review Article*”, 51 *Modern Law Review* 659 (1988) [Stein II], 660. While pure crime control is occupied with the procedural efficiency of the system, due process insists on fairness and dignity values to which the individual is entitled even at the expense of the system's efficiency., See also Robert S. Summers, “Evaluating and Improving Legal Processes A Plea for Process Values”, 60 *Cornell Law Review* (1974). Another way, rooted in the utilitarian tradition, is to pose accurate fact-finding as the ultimate goal of this system, and procedural fairness as one of its constraints. See William L. Twining, *Rethinking Evidence* (Oxford, 1990), 71-78. However, it is nontrivial for such utilitarian approach to account for a notion of fairness that cannot, in a final account, be reduced to an efficiency consideration. For broader criticism, see Williams (n 62), 552-553. I chose to use the deontological/consequentialist articulation for its resemblance to the ethical one, see *ibid*, 551.

A. The Consequentialist Approach

In this paper, I use the notion of “accuracy” as a *continuum* measurement of *aggregative* success in fact-finding of a given legal system. “Continuum” because the realistic recognition that “determination of facts always entails risk of error”.⁶⁵ Thus, a legal system can never be fully accurate, but rather *more or less* accurate than an alternative system. “Aggregative” because consequentialists value *aggregative* utility in general.⁶⁶ Hence, I use accuracy here to assess the quality of fact-finding of a system rather than its quality in a particular decision.⁶⁷

Accuracy in fact-finding promotes crime reduction in at least two ways. First, better accuracy can reduce the number of guilty people who were mistakenly acquitted. As fewer criminals will be able to evade the law, crime reduction will benefit from a better power of deterrence and more criminals being neutralised by imprisonment or other means. Second, better accuracy can reduce the number of innocents that are mistakenly convicted. Although avoiding wrongful convictions is usually the flagship of the deontological theories, it can promote crime reduction as well. Convicting an innocent person means that the real criminal walks free whilst the law enforcement authorities consider this case closed. Therefore, aiming to diminish wrongful convictions can be also accommodated by consequentialist theories and better accuracy can promote crime reduction regardless of the change in the number of convictions.⁶⁸

⁶⁵ Stein I (n 28), 12.

⁶⁶ Williams (n 62), 552.

⁶⁷ I do not argue that this is the only possible usage of the word accuracy, since in other contexts it may be used to assess a particular decision too. For example, a single instance of scientific measurement in the lab may be described as “*more* accurate” than the same experiment performed in my backyard.

⁶⁸ Note that I do not argue that accuracy can be reduced to crime reduction or vice versa, just that there is a positive correlation between them so promoting accuracy promote crime reduction too. For more detailed account on the connection between accuracy and consequentialism, see, for instance, Stein I (n 28), 12-15.

Koehler and Shaviro argue that insofar as accuracy is concerned, statistical evidence should be admitted⁶⁹. This move was challenged intensively, but the stronger objections seem to be those which emphasise other principles that should override accuracy and proclaim statistical evidence inadmissible *even if* they increase accuracy.⁷⁰ Therefore, it seems safe to accept that admitting *any* reliable statistical evidence promotes accuracy.⁷¹ This accuracy, in turn, promotes crime reduction, which in a final account, promotes the consequentialist justification for the criminal process. Thus, pure consequentialist justification should welcome the admissibility of reliable statistical information.⁷²

B. The Deontological Approach

These justifications usually focus on two strata: the individual and the special nature of the criminal proceedings⁷³. The latter is unique by being the only process in which the modern state *morally judges* the individual. There is a strong sense of *blame* in the

⁶⁹ Koehler & Shaviro (n 34). Note that their argument apply on single-case accuracy too.

⁷⁰ Tribe's case against statistical evidence is that they "bring this probabilistic element of inference to explicit attention in a quantified way", Tribe (n 2), note 2. He aims to show that "much turns on whether such explicit attention is attempted", Ibid. For that purpose, he brings detailed and strong policy considerations for why such quantification is not welcome. Nesson's argument is that statistical evidence reduces the "criminal cases to a simplified assessment of..."chances of guilt". Nesson (n 2), 1192. This simplicity leads to loss of public trust in the jury, since this trust is based on their perception and intuition, while "the coldly statistical case gives them no opportunity to exercise that perception or intuition". Ibid, 1196. This simplicity also ruins the public trust in the presumption of innocence, as prior to statistical evidence, this presumption had a vague notion to which every individual could be attached. statistical evidence gives this presumption a precise definition that not every one agrees with. Ibid, 1197. However, even if one fully accepts either of these arguments, the conclusion is not that 'admitting statistical evidence fails to promote accuracy', but that 'although admitting statistical evidence can promote accuracy, there are other considerations which outweigh its contribution to accuracy'. Put it bluntly, they argue that the contribution statistical evidence has to accuracy does not worth its price.

⁷¹ I added the constraint of "reliable" because unreliable statistical evidence would probably be rejected by any approach, as it dubious how it promotes either accuracy or fairness. Anyway, such unreliable statistical evidence is a residual issue as the core of this debate is the statistical evidence that is produced under very high standards. See, for instance, the statistics produced in Shonubi (n 32).

⁷² See also Stein who concludes that "Under the accuracy-maximising criterion, any evidence that pertains to a trial issue is generally admissible", Stein I (n 28), 15.

⁷³ Dennis (n 4), 42-43, Roberts & Zuckerman (n 4), 12-14.

criminal conviction, whose intention is to cause the convicted to feel guilt and remorse and the victim to feel that “justice has been done”. The conviction carries not only a sanction, but a social *censure*, which lasts long after the sanction ends. This uniqueness can explain why wrongly convicted criminals insist on “clearing their names” and ascribing such attention to the conviction itself, rather than to the sanction⁷⁴.

This justification appears in many debates. For instance, by giving the moral basis for the determination that convicting an innocent person is much worse than acquitting a guilty one, it establishes the hallmark of criminal law of evidence: the presumption of innocence⁷⁵. It also appears in the debate about due procedure and proportionality of punishment⁷⁶.

I, on the other hand, want to illuminate another aspect of the deontological justification: *the individualistic character of the moral responsibility*. The modern notion of moral responsibility is strongly committed to the idea of *personal* responsibility. This justification takes for granted, as it should, that one should be blamed and punished only for actions that one did oneself. The moral condemnation with which the modern state censures the criminal is so deeply immersed in the notion of personal responsibility, that it seems too obvious to be even mentioned⁷⁷.

⁷⁴ This can also explain why their families continue to struggle for clearing their names even after their deaths. See, for instance, *R. v. Derek William Bentley (Deceased)* [2001] 1 Cr App R 307, CA.

⁷⁵ Roberts & Zuckerman (n 4), 14, 19. For this asymmetry in general, see Stein I (n 28), 15-16.

⁷⁶ Andrew von Hirsch, "Proportionality in the Philosophy of Punishment: From 'Why Punish' to 'How Much'", 25 *Israel Law Review* (1991) 549 [Hirsch I]; Andrew Ashworth, *Sentencing & Criminal Justice* (Weidenfeld & Nicolson, 1992), section 4; Andrew von Hirsch, *Censure and Sanctions*, (Oxford University Press, 1993) [Hirsch II], sections 2, 4-5.

⁷⁷ Roberts and Zuckerman, for instance, do not include this notion in their extensive theoretical overview; see Roberts & Zuckerman (n 4), 18-25.

I argue that precisely this obvious element of personal responsibility is exactly what is threatened when statistical evidence is admitted. The same logic that abolished *collective* punishment is valid here. No individual should be punished only because he lives in the same neighbourhood in which the criminal lives too. This individual did nothing wrong and belonging to the group that lives in this neighbourhood is not a crime on its own.

Similarly, an individual should not be punished for belonging to *any* group, but only for actions he did himself⁷⁸. Yet, statistical evidence requires us to draw a conclusion on the accused based on his membership in a certain group. This is an inevitable feature of any statistical evidence, as group-mediation what qualitatively differs it from particular evidence. Thus, whenever the evidence attempts to deliver its information on a particular individual, some group must always be used as an intermediary.

My argument also explains what stands behind the intuitive recoiling from basing a criminal conviction *solely* on statistical evidence, as even the most enthusiastic supporters of statistical evidence try to avoid awarding a verdict based solely on statistical evidence.⁷⁹

My argument can be extended further, as using statistical evidence creates a new type of *discrimination*. As noted in the previous section, statistical evidence can be identified by the fact that it can deliver its information about any individual member in the group in the

⁷⁸ For a similar observation see Adrian Zuckerman, "Law, Fact or Justice?", 66 B.U. L. REV. 487-498. For a detailed discussion on whether belonging to a group can be a crime, see Colyvan's argument (n 2) and Tillers' response (n 17).

⁷⁹ See Kaye's explanations why statistical evidence is insufficient *even in civil litigation*, Kaye (n 2), 106-108. Brook criticises Kaye's account for its insufficiency, James Brook, "The Use of Statistical Evidence of Identification in Civil Litigation: Well-Worn Hypotheticals, Real Cases, and Controversy", 29 St. Louis University Law Journal 293 (1984-1985), 323-329. However, Brook himself implicitly agrees that statistical evidence is insufficient for conviction when he relies on the criminal/civil distinction in explaining why "trial by mathematics has gotten such bad press", *ibid*, 309-310.

precise same way. If the statistical evidence is all what the prosecution has, every member in this group is equally suspect. However, the prosecution wants to use this evidence only against one specific accused. What is the justification for picking him, and not any other member in the group?

The feeling of unfairness should be galling. I suspect that some might find Nesson's prisoners example as hard to identify with, so let us assume that the Inland Revenue has accurate statistics according to which 98% of the legal practitioners defraud in their tax report (e.g. because they have lower litigation costs). Criminal proceedings are being initiated against a few of them. However, one would expect the state to interfere with her life and drag her into the intimidating criminal process only when the state has evidence *against her personally*. Once again, belonging to the group of legal practitioners is not a criminal offence on its own (although some might gladly argue that it should be...).

Note that my argument should be distinguished from other strong and popular arguments against the usage of statistical evidence in the criminal context. The first is that even if statistical evidence should be admissible, a conviction should not be based solely on statistical evidence.⁸⁰ The second argument reveals that a conviction relying on statistical evidence requires a quantification of the notion of "proof beyond reasonable doubt", in order to know when the statistical evidence figures are sufficient for conviction. However, "proof beyond reasonable doubt" cannot, or should not, be translated into a numerical value.⁸¹ The last argument is based on the "problem of the reference class". As

⁸⁰ See next section for a detailed comparison between this approach and mine.

⁸¹ Craig R. Callen, "Adjudication and the Appearance of Statistical Evidence" 65 *Tulane Law Review* 457 (1991); Tribe (n 2). See also Roberts' interesting proposal for understanding this notion as a procedural requirement rather than a threshold, numerical or not. Roberts & Zuckerman (n 4), 366-373.

noted above, this problem is one of the causes for the uncertainty regarding applicability.⁸² In contrast, the argument here focuses on the difficulty to choose the appropriate group: since the accused is a member in an infinite number of groups, the question is which group is the most appropriate one to use and what should be done if the information available is only on another group which is less than optimal?⁸³ However, this argument has more practical flavour, as it is willing to assume that such evidence is acceptable to begin with. It focuses on the difficulty *finding* the *right* group rather than the difficulty *applying* group-mediated information *in general*, which is what my argument establishes when identifying the qualitative difference created, inter alia, by the problem of reference class.

Summary

This section reveals the basic tension lies beneath the question about the admissibility of statistical evidence. A pure consequentialist approach should support the admissibility of any accurate statistical evidence, while a pure deontological approach should render such evidence inadmissible because of its infringement of the principle of personal responsibility. As noted above, both approaches were presented in their purest version.⁸⁴ The obvious question, then, is how these two extremes should be mitigated. This is the subject of the next section.

⁸² Text around supra 36.

⁸³ See again the text around supra 36, and especially Colyvan (n 2)

⁸⁴ See text after supra 64.

Section V: A Suggestion for Mitigation between the Approaches

In my view, both solutions of the pure approaches are unsustainable and too radical. On the one hand, the pure consequentialist solution to unrestrictedly admit all statistical evidence is unattractive even to the most supportive advocates of statistical evidence.⁸⁵

The principle of the personal responsibility cannot be completely ignored and, hence, some limitations should be put on the usage of statistical evidence in criminal trials.

On the other hand, the pure deontological solution to always exclude statistical evidence is not very attractive either, as statistical information can contribute valuable information to the fact-finder. Moreover, excluding such evidence might be unnecessary when it is clear from other evidence why this particular individual that was marked from the group.

My suggestion

My suggestion is that statistical evidence should be admissible only if it is *corroborated* with *additional* evidence. This additional evidence should satisfy two conditions:

1. Be *individualistic* rather than statistical.
2. Refers to *the same material fact* that is in dispute.

Note that according to these conditions, the same piece of individualistic evidence can be used to admit *more than one* piece of statistical evidence, as long as that each piece of statistical evidence refers to the same material fact as the individualistic one.

The Individualisation Condition

⁸⁵ As described in supra 79.

I argue that this additional piece of evidence *must be individualised* evidence rather than statistical one. What was missing in the original statistical evidence is the individualisation. The additional evidence is required to *justify why this particular individual was marked from the whole group*. Therefore, this marking cannot be justified with a membership in yet another group because it does not provide the missing individualisation. Moreover, justifying one piece of statistical evidence with another one quickly leads to *an infinite regress*. The same need for justification exists about the additional piece of statistical evidence, so there a third piece of evidence is required, and so forth. Hence, only individualised evidence can both justify the marking and avoid the need for yet another piece of evidence.

The Same Issue Condition

The additional individualised evidence must also be relevant to the *same material fact that is in dispute*.⁸⁶ The rationale beyond this requirement is that the prosecution has to overcome the lack of individualisation that is created by the reliance of statistical evidence for proving a specific material fact. Thus, the individualised evidence needs to justify the usage of statistical information when proving *this material fact*.

For example, individualised evidence proving that Humpty watched the rodeo is not enough, because it refers to another separate factual question: did he watch the rodeo. It is separate because it can be answered without answering the factual question in dispute: did he pay the admission for it.

⁸⁶ What counts as the *same* material fact is a nontrivial question. However, it is common need for any requirement of corroboration and is not unique to my account. Therefore, this question is outside of the scope of this paper. For the notion of corroboration in general, see Roberts & Zuckerman (n 4), chapter 10. For the notion of evidential relevance in general, see *ibid*, 98-108.

Advantages

Although my suggestion does not eliminate the deontological objection to statistical evidence altogether, it blunts its sting. Part of the deontological concern is around the question “amongst all of the group members, why this particular individual was brought to trial”? My suggestion addresses this concern directly by supplying an explanation to this “marking” by requiring additional *individualistic* evidence to be corroborated with the statistical one.

From the consequentialist point of view, my suggestion provides a solution to the science-fictionist nightmare with which consequentialism can be easily blamed. Some might intuitively object a world in which people are sent to jail immediately after the state would be able to prove that these people match certain statistics. My approach guarantees that the state would be able to justify not only why fact-finding in these cases is more accurate than in a system that excludes statistical evidence altogether, but also why *these particular individuals* were put to trail. This should take away the intuition that it was merely a “trial by mathematics”.⁸⁷

Objections

Objections can be heard from both sides. A pure deontological approach might object to the admissibility of statistical evidence even when it is corroborated with additional

⁸⁷ This term was coined by Tribe (n 2).

individualised evidence. In contrast, a pure consequentialist approach might object the requirement for such additional evidence and argue for unrestricted admissibility.

Objecting to the Admissibility of Statistical Evidence even when Additional Evidence Exists

A pure deontological approach might raise three potential objections. The first points to the immense difficulty to combine statistical and individualised evidence⁸⁸ and argues that the lack of a combining procedure should entail the inadmissibility of such evidence. However, this argument is a hard one to run, as it is also far from clear how to combine two individualised pieces of evidence.⁸⁹ Yet, nobody argues that this should be a reason for always excluding them...

A second objection might argue that my approach violates the presumption of innocence, since the prosecution uses evidence that might still be in dispute as a springboard to import other evidence that would otherwise be excluded.

The answer to this objection is that if the additional evidence is beaten, the statistical evidence should automatically follow. Different pieces of evidence are sometimes interconnected, so when one falls, others might fall too. In this respect, there is nothing unusual in statistical evidence.⁹⁰

⁸⁸ Some aspects of this difficulty are discussed as part of the extensive debate about the usage of Bayes Theorem for legal purposes. For an excellent summary, see Roberts & Zuckerman (n 4), 123-132.

⁸⁹ This issue is also discussed in relation to Bayes Theorem, *ibid*. It also appears in relation to corroboration in general, see Roberts & Zuckerman (n 4), chapter 10. For an historical summary, see Anderson & Twining (n 54), 333-339.

⁹⁰ One can counter my answer by questioning how practical it is for a jury to be able to ignore evidence already presented to them, even if in theory they should. But there are many procedural measurements for overcoming such concerns and there is nothing unique to statistical evidence about it. For instance, if the statistical evidence is crucial to the prosecution case, the defence can reply with “no case to answer”, which will be assessed by the judge rather than the jury, see Dennis (n 4), 95. If the case does proceed, the judge

A third objection can argue that the additional evidence is *doubly weighted*: first as evidence on its own merits and, second, by rendering another piece of evidence admissible. Thus, it contributes to the prosecution case two different arguments, while in a final account it is only a single piece of evidence.

I believe that if this objection is appealing, it is mainly because it is confusing. It is true that the same alleged fact appears in two different arguments, but there is nothing with it. For a double countability argument to work, the same evidence should not only appear twice, but also should be *weighted* twice. However, the weight in the second stage is obtained *only from the content of the statistical evidence*.⁹¹ The individualised evidence is used only to admit this statistical evidence that would otherwise be excluded.

One can counter-argue that even if the additional evidence it is not counted twice, the existence of other evidence (statistical evidence) regarding the same factual question can strengthen the first one (individualised evidence). Thus, the individualised evidence might be outweighed only because the statistical evidence was added to it.

I am not sure that that this is a problem at all, as I do not see what is normatively wrong with the idea that two pieces of evidence reinforce each other, in a manner coined as “the sum is bigger than its parts”. However, even if one considers this to be a problem, I only argue that in this respect, there is no difference between (1) the case of one piece of statistical evidence and another piece of individualised evidence and (2) the case of two pieces of individualised evidence.

can guide the jury to ignore this evidence in his summary, see Roberts & Zuckerman (n 4), 77-82. The jurors might disregard his guidance, but this concern exists for any guidance given by the judge, so, again, this is not a unique problem for my suggestion.

⁹¹ But see Stein who argues that statistical evidence has no weight for the specific case, Stein I (n 28) 81-90. My criticism on his approach is brought in the text near supra 28.

Consequentialist Objections to the Requirement of Additional Evidence

Three additional objections can be raised from consequentialist point of view.⁹² The first lies on the intuition that “the police does not randomly pick people to arrest and the prosecution does not randomly prosecute; thus there is probably a good reason why this individual is here in court”. This intuition seems to follow the idiom of “there is no smoke without a fire”.

However, this misleading intuition should be dropped the sooner the better. If the police and prosecution have such a good reason, they should present it as evidence in court and let the jury examine it with scrutiny. Moreover, this intuition seems to contradict the presumption of innocence, as instead of beginning the trial with the presumption that the accused is innocent, it starts with the presumption that “he is probably here for a good reason”, and implicitly places the burden on the accused to convince the jury that it is solely smoke that the prosecution case contains rather than a real fire.

The second objection is to deny the existence of any discrimination by claiming that this evidence would be used against any member of this group that would be put on trial. Therefore, there is no discrimination here, as any member of this group will “enjoy” the same treatment under the same circumstances.

This objection seems to miss the point. The discrimination is not between one accused and another, but between those members who were brought to trial and those who were

⁹² Note that I already discussed some quantitative objections in section III, when defending my qualitative account from potential quantitative attacks. The objections here, however, are those who possible *even* if my qualitative difference is admitted.

not. For example, if only five out of the thousand watchers can be found, the discrimination is not between the first accused and the other four, but between those five and the rest of the rodeo watchers.

The last objection seems to be the most significant one. One can argue that there are other principles that override, or at least limit, the principle of personal responsibility. For some other good reasons, we might be willing to sacrifice personal responsibility and admit statistical evidence *even without* additional individualised evidence.⁹³

The important thing to notice is that such an objection lifts the debate about the admissibility of statistical evidence into a different realm. The consequentialist objections in the first part argued that there is no difference in quality between statistical and individualised evidence.⁹⁴ The invited conclusion was that if there is no real difference, they should also be treated in the same way by the law, as long as the difficulties in presentation are overcome.

However, the current objection admits that the issue at stakes is much more substantial than mere presentation difficulties. It transforms the debate into a principal debate about the appropriate balance between different and contradicting principles. Such an upgrade is most welcomed, I think, as this should be the realm in which the statistical evidence debate is to be conducted. If this is the only plausible objection that the quantitative camp is left with, I consider the main goal of this paper as being achieved.

⁹³ Some caution is required when such overriding principles would be suggested, as the question is whether such principles completely override the principle of personal responsibility or only limit it, and if so to what extent.

⁹⁴ See section III for these objections.

Other Potential Alternatives

Even if the above objections can be overcome, one can suggest other alternatives to my suggestion for mitigating between the extremes. I compare my suggestion to two alternatives in particular. The first is to accommodate the difficulty that statistical evidence raises through weight instead of inadmissibility. The second is disallowing *conviction* to be based solely on statistical evidence.

Weight Reduction instead of Inadmissibility

The alternative here is to accommodate the lack of individualisation by admitting the statistical evidence and then reducing its weight instead of rendering it inadmissible.⁹⁵

This alternative might be conceived as more pragmatic and moderate than my suggestion that completely denies the information from the fact-finder.

However, I argue that this alternative cannot stand for both theoretical and practical reasons. First, this suggestion seeks to reduce two qualitatively-different pieces of evidence into the same homogenous scale (called “weight” this time). Therefore, this suggestion is merely a concrete instantiation of the general quantitative objection seeking to reduce the two kinds of uncertainty into one, and my answer to that objection is relevant here too.⁹⁶

⁹⁵ I intentionally kept the notion of “weight” as vague as possible, because I want this alternative to cover any suggestion that includes an adjustment of the probative role that such evidence would have in court instead of rendering it inadmissible and outside of the fact-finding process altogether, as I suggest.

⁹⁶ See text after *supra* 51.

Second, this suggestion seems to address the wrong problem. The need for mitigation does not rise from the concern that this evidence lacks the appropriate weight,⁹⁷ but from the principle of personal responsibility dictating that an individual should not be blamed for belonging to a certain group.⁹⁸ Sometimes, statistical evidence can contribute much more valuable information to the fact-finder than many other individualistic ones. Therefore, reducing the weight of the statistical evidence fails to solve the right problem because the problem with statistical evidence is its lack of individualisation rather than its lack of weight.

Third, this alternative can also be *unjustifiably stricter* than my approach. If there is additional evidence in the case, my approach will give the statistical evidence its full weight, while this alternative would *always* reduce the weight. This weight reduction is inconsistent with the pragmatic pretence that this alternative claims for. Moreover, it is unjustified, because once the lack of individualisation was overcome, what is the justification to continue and reduce weight?

Last, this alternative suffers from practical problems too. It is far from clear what the method with which the weight should be adjusted is. Put it bluntly, how much weight should the fact-finder reduce? Like in the general quantitative objection, the practicalities here just reflect the fundamental theoretical difficulties that are resulted from the attempt to merge two qualitatively-different uncertainties into a homogenous scale.

No conviction should be based solely on statistical evidence

⁹⁷ See also my response to Stein, text around supra 28.

⁹⁸ See text after supra 76.

This alternative must share an important feature with my suggestion while it can differ on another. The common feature is the requirement for additional evidence. If, according to this alternative, statistical evidence alone is insufficient for conviction, then it practically means that for conviction to be reached, additional piece of evidence is required. However, a potential dispute between my suggestion and this alternative is about the nature of the additional evidence which is required.

I set two conditions for such additional evidence: (1) individualisation and (2) same material fact. An alternative can reject either one or both of these conditions. However, out of the three permutations possible, only one is really defensible. If this alternative stands on rejection of the individualisation condition, it would fall into an infinite regress described above.⁹⁹ Therefore, the only condition which would be plausible to relax is the second one.

Obviously, the additional evidence should not go to an issue that is in agreement, otherwise the requirement will be easily nullified (as even the accused's name can serve as such individualised evidence). But an alternative can be suggested so the requirement would be satisfied with individualised evidence that goes to *any* issue in dispute, not necessarily the *same* one for which the statistical evidence was brought. This is the only plausible suggestion I could reconstruct from such an alternative.

The distance between my suggestion and this alternative is very narrow to begin with.¹⁰⁰ Nevertheless, I think this alternative should also be rejected as it fails to explain the marking of the member from the *actual* group that was used. The rationale beyond my suggestion is that the additional evidence should explain why the prosecution chose to

⁹⁹ See text before supra 86.

¹⁰⁰ There is also no dispute between my suggestion and this alternative regarding statistical evidence using non-convicting purposes. See my analysis of *Shonubi* in Section VI.

focus on this particular member and not on any other members of this group. For such an explanation to be relevant, I argue that it has to refer to the same factual question.

Also, since I could not find any hypothetical or actual example that would illustrate the difference between my suggestion and this alternative, I am not sure that this difference has any practical significance.

Section VI: Investigating Real Cases

After facing various objections and alternatives, my suggestion can be now taken to practice and help solving a few real cases. I use these cases not only for illustration, but also to extend the scope of my analysis to some residual issues: non-convicting statistical evidence, DNA evidence, and defence statistical evidence.

Non-Convicting Evidence & the Case of Shonubi

So far, I dealt only with statistical evidence concerning conviction. Shonubi's case presents an excellent opportunity to broaden the scope to statistical evidence used for non-convicting purposes.

Shonubi is a Nigerian citizen and resident of US who was arrested for heroin smuggling using the method of swallowing heroin-filled balloons. He was caught with 427.4g in his stomach. Apparently, this was not his first trip, and he had made at least seven other trips. This was in agreement between the parties, but the amount of drugs in these previous trips was in dispute, as it had a significant influence on his sentence. In *Shonubi I*,¹⁰¹ Judge Weinstein simply multiplied the amount found on the last trip by eight, which was more than the 3000g needed to award Shonubi with the longest sentence possible for drug-smuggling. In *Shonubi II*,¹⁰² the Second Circuit rejected this method due to a lack of "specific evidence", and returned this case to Judge Weinstein, who in *Shonubi III*¹⁰³ allowed the admissibility of the statistics produced by the prosecution about 117 other (1)

¹⁰¹ *United States v. Shonubi*, 802 F. Supp. 859 (E.D.N.Y., 1992) [Shonubi I].

¹⁰² *Supra* 16.

¹⁰³ *Supra* 32.

Nigerian drug-smugglers (2) caught in the same airport (JFK), (3) around the same period of time (up to a year before Shonubi's arrest) and (4) used the same method of balloon swallowing.¹⁰⁴ Using this statistical evidence and other evidence from an expert economist on the patterns of heroin smuggling and trafficking, Judge Weinstein ruled that it is more probable than not that the total amount of drugs was between 1000g and 3000g, which awarded Shonubi with the second longest sentence. Defence appealed again and Shonubi IV¹⁰⁵ rejected this method as well due to a lack of "direct" evidence regarding the amount of heroin Shonubi actually carried. In Shonubi V, Weinstein reacted by commenting that "what the [federal] Rules [of Evidence] do not do is categorise, for exclusionary purposes, forms of relevant evidence as direct, circumstantial, and "specific".¹⁰⁶

I want to use my analysis to try to clarify this discussion, mainly regarding the prosecution's statistics, which is typical statistical evidence. From an accuracy perspective, it is hard to even understand how such valuable and precise information can be ignored, as Tillers' passionate reaction exemplifies.¹⁰⁷

Nevertheless, from a deontological perspective, perhaps what motivated the Second Circuit's search for "direct"/"specific" evidence is the need to establish Shonubi's personal responsibility with evidence that does not suffer from lack of individualisation as the prosecution statistics does.

¹⁰⁴ One should acknowledge the impressive efforts made to minimise the problem of the reference class, *supra* 36.

¹⁰⁵ *United States v. Shonubi*, 103 F.3d 1085 (2d Cir, 1997) [Shonubi IV].

¹⁰⁶ Shonubi V (n 17), 375. For a more detailed description of this case, see Tillers I (n 2), 1879-1883

¹⁰⁷ Tillers I (n 2).

I suggest understanding the debate between the two American instances as a principal disagreement on the proper balance between the two justifications. Thus, I would not rush to proclaim that the Second Circuit judges “still may not have a grasp of some basic characteristics of probabilistic and statistical methods and arguments”¹⁰⁸. Although one can argue that the balance they advocate is wrong, this stance should be understood as a legitimate approach rather than a mere misunderstanding of statistical methods and arguments.

Nevertheless, a certain question casts some doubts on the Second Circuit decision: *should the balance between accuracy and personal responsibility be the same in matters of sentence as in matters of conviction?* The personal responsibility argument is driven by the high price of mistakenly convicting an innocent person. However, using such evidence for sentencing crimes in which *the accused admits* is a different story, as the price of a mistake in such cases is almost incomparable to the one of wrong conviction. Proposing the proper balance between accuracy and personal responsibility regarding sentencing in general is beyond my scope, but, clearly, the weight given to personal responsibility in sentencing should be much lower than in convicting.

DNA Evidence & the Case of Denis Adams

What is “DNA Evidence”?¹⁰⁹

¹⁰⁸ Tillers I (n 2), 1879.

¹⁰⁹ Jeffreys, who invented DNA profiling, gives a scientific description in Alec J. Jeffreys, Victoria Wilson & Swee Lay Thein, “Hypervariable ‘Minisatellite’ Regions in Human DNA”, 314 *Nature*, 67-73 (1985) [Jeffreys I] and up-to-date on current techniques in Alec J. Jeffreys and Keiji Tamaki, “Human Tandem Sequences in Forensic DNA Typing” *Legal Medicine* xx, 1-7 (2005) [Jeffreys II]. See also Dawid I (n 5) 9;

Two samples of DNA are compared: the first is taken from the crime scene and the second from the accused. The comparison does not include the whole genome, since a full comparison is extremely laborious and expensive¹¹⁰. Instead, the expert focuses only on those parts of the DNA sequence that are known to be distinctive amongst human beings.

However, finding a match between the samples is insufficient to prove that the DNA in the crime scene *originated* from the accused. Although only identical twins can have precisely the same DNA, those parts of the sequence that were compared can be identical in more than one individual.¹¹¹ Thus, the expert has to calculate the probability for such a match if these two samples did not originate from the same person, i.e. if it was someone else who left this sample in the crime scene but there is a random match between this anonymous person and the accused.¹¹² For this, she uses a database to assess the frequency of those parts in the general population.¹¹³ The numbers she reaches can be impressively low and 1 in a billion is no longer a science fiction¹¹⁴.

Thus, evidence about DNA evidence consists of two components: (1) whether a match between these parts was found and (2) what the probability is for such a match to be randomly found if the two samples originated from two different individuals.

¹¹⁰ Most of the DNA sequence is less variable between different humans. Therefore, the basis for matching two samples of DNA is those parts that are prone to mutation and are, thus, hypervariable. Jeffreys I (n 109).

¹¹¹ Jeffreys II (n 109).

¹¹² Note that this is *not* the probability of guilt. Presenting it as a probability of guilt was entitled as “the prosecutor fallacy” and explained in *R. v. Doherty & Adams* [1997] 1 Cr App R 369, 372-373; Dawid I (n 5), 9-10.

¹¹³ Philip A. Dawid, “Bayes’s Theorem and Weighing Evidence by Jury”, *Proceedings of the British Academy* 113, 71-90 (British Academy, 2002) [Dawid II], 86-88.

¹¹⁴ Dawid I (n 5), 9.

Analysis

While DNA match is usually referred as a single piece of evidence¹¹⁵, it is, in my stance, a *list* of separate pieces of *statistical evidence*. It comprises from a *series of matches* between those parts of the sequence that tend to vary the most amongst human beings.

Why “list”: DNA evidence is a list rather than a single piece of evidence because it is *dividable*.¹¹⁶ Like any other human activity, the risk of error exists even in the most careful scientific method. Hence, when examining a specific part of the sequence, the expert might mistakenly consider it as matching the other sample, while in reality they do not.¹¹⁷ This stands for the uncertainty regarding the reliability from which any evidence suffers. What is important for my purpose is that the expert can be wrong about one part while be right about the others. Therefore, according to the test laid above¹¹⁸, this is a *list* of separated evidence rather than a group-mediated one.

Why “of statistical evidence”: each match of a single part of the sequence is statistical rather than individualistic because it applies on more than one individual in precisely the same way.¹¹⁹ In order to use this evidence against another member of this group (those individuals whose DNA match this part too), there is no need to change either the content of the evidence or the conclusion derived from it. The fact that this is group-mediated evidence is best expressed by the fact that the expert uses a database to estimate how

¹¹⁵ See David I (n 5), Dennis I (n 5), Dennis II (n 5).

¹¹⁶ For the criterion of divisibility, see text around n 44.

¹¹⁷ Refer to the disturbing criticism of Professor Donnelly for the defence on the method employed by Dr Harris, including, for instance, using a DNA band that was “so weak as to necessitate Dr Harris highlighting it”. Adams II (n 5), 379.

¹¹⁸ Text around n 44.

¹¹⁹ For elaboration on these conditions, see text around n 43.

many people have this match. Therefore, this evidence suffers *also* from uncertainty regarding its applicability.

The number: The final result of “1 in a billion” is based on *corroboration* of these pieces of evidence into a single number using statistical methods. This number is not consensual as this calculation can be in dispute amongst different experts.¹²⁰ Therefore, this impressive number should not be regarded as inevitable property of the DNA evidence, but rather as an expert opinion suggesting how to interpret this evidence.

Hence, in my stance, DNA evidence is statistical evidence that differs qualitatively from individualistic one and deserves the same treatment as any other statistical evidence. I exemplify my stance through the case of Denis Adams.

Denis Adams

Denis Adams was convicted for rape based on DNA evidence. However, not only did his girlfriend gave him an alibi by testifying that he was with her on that night, but, more importantly, the victim herself also did not recognise him¹²¹. The case visited four

¹²⁰ In *Adams*, Dr Harris for the Crown argued for 1 in 200 millions while Professor Donnelly for the defence argued for 1 in 2 million, Adams I (n 5), 470.

¹²¹ The description she gave of the attacker was not very consistent, though. On one occasion she said he was aged 40-42; while on another he was 20-25. Adams II (n 5), 378. Nevertheless, the crux of this issue is whether a DNA match can be overcome by other evidence or is it convulsive evidence. If *Adams* does not provide the appropriate background, one should replace this case with any hypothetical scenario in which the defence has other excellent evidence against the DNA match, and then investigate whether it would be sufficient to overcome the DNA evidence.

instances and attracted scholarly attention mainly due to a unique attempt to guide the jury through the process of applying the Bayes Theorem.¹²²

In contrast, I want to focus on the *admissibility* of the DNA evidence in this case. The prosecution offered no other evidence besides the DNA evidence and it remained unclear how Adams was picked. This case was the first case in which CPS based the entire case solely on DNA evidence,¹²³ but this seems to be an increasing trend nowadays.¹²⁴ This poses DNA evidence as conclusive evidence such that it is hard to imagine what evidence the defence can provide to overcome it.¹²⁵

Each of the pure approaches leads to a different conclusion in this question of admissibility. According to a pure consequentialist approach, DNA evidence should always be admissible. However, according to a pure deontological approach, the counter-intuitive result is that DNA should stop being used in courts *at all*.

According to my suggestion, the solution is more plausible. If the prosecution cannot explain why it picked Adams, this evidence should be *inadmissible*. This explanation has to be in a form of additional evidence which is both individualised and goes to the issue in dispute (i.e. whether it was Adams who had intercourse the victim). Since one piece of individualised evidence can be corroborated with more than one piece of statistical

¹²² For a clear description of Bayes Theorem and its problematic implementation in *Adams*, see Dennis (n 4), 110-116. See also Dawid I (n 5), 9-14; Roberts & Zuckerman (n 4), 123-132.

¹²³ *Adams I* (n 5), 469.

¹²⁴ *Dawid II* (n 113), 79.

¹²⁵ For a general argument against the concept of conclusive evidence, see Unger: "being absolutely sure" is *always* fundamentally wrong as one is entitled to ignore *any* new evidence whatsoever. Peter Unger, "An Argument for Scepticism", 1 *Philosophical Exchange* 4 (1974).

evidence, once such an additional evidence is provided, the whole *list* of matches becomes admissible. As for Adams, I argue that not only should Dennis Adams have been acquitted unless the prosecution could supply additional evidence, but also that the trend of prosecuting solely on the base of DNA match should cease.

Defence statistical evidence & the Case of Sally Clark

So far, the statistical evidence I dealt with submitted only by the prosecution's, yet the defence may also wish to submit statistical evidence. I examine this issue by looking at the well-known case of *Sally Clark* from another perspective. Clark was convicted for murdering her own two babies in two separate events. The attention was attracted mainly to Professor Meadow's proclamation that the chance of two natural deaths in the same family is 1-per-73 million¹²⁶ and it was criticised on two grounds. First, Meadow unjustifiably asserted that the two events are statistically independent.¹²⁷ Second, even if correct, this number has no meaning without *a comparison* to the probability of the alternative hypothesis of a mother will kill her own two babies. The probability of this alternative was estimated to be above 1-per-2.2 *billion*.¹²⁸ However, the court dismissed this evidence by noting that “legally speaking, the exercise is not realistic”.¹²⁹

¹²⁶ Dawid I (n 5); *R. v Sally Clark (No. 1)* 2000 WL 1421196, §108.

¹²⁷ Dawid I (n 5), 3-4; Stephen Watkins, “Conviction by Mathematical Error?”, 320 *British Medical Journal* 2 (2000).

¹²⁸ Dawid I (n 5), 4;

¹²⁹ Clark (n 126), §160. I think these statistics should have been dismissed but for a different reason. The statistics is based on the number of previous cases in which mothered murdered their babies. Thus, it has to rely on the rate of convictions. Such evidence aims to prove a legal question (whether Clark was a murderer) rather than a factual question (whether she physically smothered her babies). However, evidence should be brought only to prove questions of fact. For the basic distinction between questions of fact and questions of law, see Dennis (n 4), 91-92.

The question is, therefore, what should be the appropriate balance between consequentialist and deontological approaches when the statistical evidence is *submitted by the defence*. A second question is whether my suggestion for the requirement of additional individualised evidence should apply *also* to defence statistical evidence.

To deal with these questions, I differentiate between statistical evidence brought *to establish a reasonable doubt* in the prosecution's case, and statistical evidence brought to *prove* something in balance of probabilities (e.g. insanity¹³⁰). Obviously, the question which burden of proof is required depends not only on the issue in dispute but also on the rules of the specific legal system in which the case is being managed.

Defence Statistical Evidence for Establishing a Reasonable Doubt

According to a pure consequentialist approach, reliable defence statistical evidence should be treated precisely the same as prosecution one. Accuracy requires that *any* information that is factually relevant to the question in dispute will be taken into account. Thus, there is no reason to admit only prosecution statistical evidence and ignore relevant information from the defence. Any other conclusion will unfairly present the advocates of the accuracy argument as merely trying to raise the rate of convictions.

However, according to a pure deontological approach, the fact that the statistical evidence came from the defence side neutralises the tension with the principle of personal responsibility. This principle yields that the social and moral censure should not be attached to a person based only on group-mediated evidence. However, it does *not* yield

¹³⁰ Proving insanity is one of the exceptions to “proof beyond reasonable doubt”, *Woolmington v. DPP* [1935] AC 462, 481-482.

that when this individual tries to defend himself from the charges raised against him, he cannot use such information himself. When he challenges the plausibility of the prosecution's arguments or argues for the plausibility of other innocent scenarios, he should be able to back these up with *any* relevant information, including statistical evidence. Hence, a pure deontological approach should advocate free admissibility to *any* reliable defence statistical evidence.

This conclusion leads to what seems as a counter-intuitive result. The pure deontological approach discriminates between prosecution statistical evidence (inadmissible) and defence one (admissible). This conclusion seems to lead to an unfair advantage for the accused. If one accepts the deontological approach, so the objection goes, defendants will start flooding the courts with endless statistics about how unlikely it is for someone like them to be guilty. The prosecution, on the other hand, will remain with its hands tied behind the back, as its statistical evidence ought to be backed up with individualised evidence.

But I think that such an inequality is not necessarily a bad thing. The notion of “Equality of Arms” does not imply equal evidential rules for prosecution and defence, but that each party should have a reasonable opportunity to present its case.¹³¹ There are other evidential rules that discriminate between prosecution and defence evidence.¹³² Therefore, reciprocity is *not* a principle in criminal law and the defence is often privileged because the acknowledgement of the bigger damage from convincing an

¹³¹ Dennis (n 4), 39.

¹³² For instance, PACE §78(1) applies only to prosecution evidence, Dennis (n 4), 84.

innocent individual comparing to acquitting a guilty one.¹³³ My approach is consistent with this preference.

To summarise, although the consequentialist and deontological approaches are in dispute regarding prosecution statistical evidence, they seem to agree that defence statistical evidence for establishing a reasonable doubt should always be admitted. Since both approaches are in agreement, there is no need for mitigation between them, and, thus, no need for any additional evidence.

Defence Statistical Evidence for Proving Something

Assume that instead of an expert psychiatrist that personally examined him, the accused wants to submit statistics about the rate of insanity amongst people in a position similar to his. While this might sound like a bad joke, I find it hard for a consequentialist approach to resist such evidence, as this statistics can be very valuable information when reaching a decision about the accused's sanity.

An instant consequentialist reply can be based on “the best evidence principle”¹³⁴ and argue that such statistics should be dismissed because the accused can supply better evidence: a report of a psychiatrist who examined him personally. But this answer puts the consequentialist approach in a fairly uncomfortable corner. First, it has to admit that individualised psychiatrist evidence is better *in principle* than any statistical information. Alternatively, if it denies this inherent preference for the individualised evidence, it has to

¹³³ Roberts & Zuckerman (n 4), 19.

¹³⁴ Alex Stein, “The Refoundation of Evidence Law”, 9 Canadian Journal of Law and Jurisprudence (1996) [Stein III], 339; Dale Nance, “The Best Evidence Principle”, 73 Iowa Law Review 227 (1988), note 244.

come up with plausible criteria how to evaluate each type of evidence and choose between them in each particular case. Second, if the accused justifiably cannot get individualised evidence, should this approach deny him justice although it considers this evidence as factually relevant, or should it develop a set of rules when statistical evidence should be allowed to substitute for individualised evidence? Indeed, it seems to me like an uncomfortable position to be in.

The deontological approach is in a much more comfortable position. Personal responsibility is not the only *individualistic* notion in criminal law. Unlike *justifications* such as self defence, insanity is considered an *exemption* from criminal responsibility because it is based on the specific circumstances of the *individual*.¹³⁵ If the accused wants to enjoy this exemption, he has to prove that he *personally* suffers from the unique circumstances that entitle him this exemption.

Therefore, statistical evidence cannot be used unless, similarly to prosecution statistical evidence, the accused can justify the marking with additional individualised evidence. For instance, if the prosecution challenges the expert psychiatrist from the defence, the accused can try to strengthen his defence with relevant statistics.

Note that such statistics may have already reached the court via explanations, references to professional literature and background experience that is included in the report of the forensic psychiatric when justifying her conclusion. My account clarifies, however, why

¹³⁵ This argument is based on the justification/exemption distinction, see Roberts & Zuckerman (n 4), 376. For more detailed analysis of this distinction, see George P. Fletcher, *Rethinking Criminal Law* (Little, Brown, 1978), §§10.1 and §§10.3; J. C. Smith, *Justification and Excuse in the Criminal Law* (Stevens, 1989), in particular sections 1 and 4.

such statistical information raises no difficulty when it is integrated in a report on an accused person.

This type of evidence is also a good example that deontological approach toward statistical evidence is *not always pro-accused*. In this issue, the consequentialist approach allows the accused to bypass the need to be examined by an expert psychiatric about his specific condition, while deontological insistence on individualistic evidence plays against the accused. This illustrates that the debate between the justifications cannot be reduced to terms of pro/cons-accused.

Conclusion

This paper proposed a new method of analysis for a well-worn subject.

First, the focus of my investigation differs from what the literature has dealt with so far. Instead of focusing on questions such as the nature of the statistical *inference* and its suitability to law, I focused my investigation in the nature of statistical *information*.

My impression is that so far this issue have fallen between the chairs as it is beside the main interests of the different disciplines: the legal discipline focused mainly on the question of what the appropriate legal treatment for such information is; statistics develops techniques to gather, analyse and present such information; and philosophy is occupied with the different interpretations of the notion of probability and how to account for inductive logic. Thus, the properties of the statistical information are seemed to be neglected.

Second, I argued for a qualitative difference between statistical evidence and individualistic evidence. My approach goes against the dominant quantitative approach that seems to be accepted by both sides of the legal debates about law & probability.

Third, I returned to the legal context and unveiled the basic principles stands beyond the debate about admissibility of statistical evidence. I showed that this debate reflects a deeper tension between consequentialist and deontological justifications.

Forth, I put forward a suggestion how these contradicting justifications can be mitigated with a requirement for additional evidence that is individualised and goes to the same issue. I defended my suggestion from possible attacks from the pure approaches as well as from other alternatives for mitigation.

I did not end my investigation without first taking my account into a practical test of analysing some real cases. I used these cases to expand my account into other types of statistical evidence: non-convicting, DNA, and defence.

Although the scope of my paper was narrowed to a single legal question (admissibility in the criminal context), I believe that the foundations I laid in this investigation can be used to approach many other legal questions. However, even if this paper fails to put forward a convincing case for my stance on the nature of statistical information, I hope it at least promotes the need for a further investigation...