

If Wishes Were Horses: Discursive Comments on Attempts to Prevent Individuals from
Being Unfairly Burdened by their Reference Classes*

© 2005 Peter Tillers

1

In *United States v. Shonubi* (*Shonubi IV*¹) an eminent appellate judge, Jon O. Newman, who spoke for a three-judge panel of the US Court of Appeals for the Second Circuit, struck down an attempt (*Shonubi III*²) by an equally- eminent trial judge, Jack B. Weinstein, to use some statistics about the behavior of 117 Nigerian drug smugglers as one³ basis for his assessment of the amount of cocaine that one drug smuggler – Charles O. Shonubi, a resident of the United States but a Nigerian national – imported into the United States during eight round-trips between New York’s JFK Airport and Nigeria.

Shonubi was detained and then arrested at New York’s JFK Airport on December 10, 1991, at the end of one of eight drug-smuggling trips that he was thought to have taken. This particular trip turned out to be Shonubi’s last trans-Atlantic smuggling venture for some time to come: although it is not quite correct to say that government

* Professor of Law, Cardozo School of Law, Yeshiva University (New York).

I am grateful for the helpful comments by Craig Callen.

As this paper was going to press, the US Supreme Court ruled (not unexpectedly, see n. 12, below) that important aspects of the current US Sentencing Guidelines are unconstitutional. See *United States v. Booker*, 2005 U.S. LEXIS 628 (Jan. 12, 2005). However, as the argument of this paper shows, the issues presented by *Shonubi V* are not limited to the question of the types of evidence that may properly be considered in federal sentencing proceedings. For example, the logic of *Shonubi V* applies to trials, and the notion of “specific evidence” for sentencing purposes is directly pertinent to current controversies about racial profiling by actors such as law enforcement agencies.

¹ 103 F.3d 1085 (2nd Cir., 1997).

² *United States v. Shonubi*, 895 F.Supp. 460 (E.D.N.Y.)

³ Judge Weinstein did not rely exclusively on overtly-statistical evidence. For example, Weinstein made inferential hay out of the financial motivations and reckonings of drug smugglers. Judge Weinstein also heard Charles O. Shonubi’s testimony, and Judge Weinstein drew some inferences about Mr. Shonubi based on Mr. Shonubi’s testimony and Mr. Shonubi’s demeanor.

officials caught Shonubi red-handed, it might readily be said that on December 10, 1991, government officials caught him stomach-filled. In the investigation following Shonubi's detention at JFK Airport on December 10, US Customs officials discovered that Shonubi's intestinal tract contained 103 balloons filled with a paste that had been mixed with heroin. Government officials later determined⁴ that Shonubi's intestinal tract contained a total of 427.4 grams of heroin.⁵

Although Shonubi was eventually caught, he was not a complete failure as a drug smuggler. As noted, Shonubi had managed to conduct at least seven other drug-smuggling trips without being detected. Only Shonubi's apprehension at the end of his very last smuggling trip led to the discovery that Shonubi had made at least seven other smuggling runs between New York's JFK Airport and Nigeria. But Shonubi's early success as a drug smuggler gave rise to a difficulty. Because Shonubi evaded detection until his last known smuggling trip, authorities had not found heroin in Shonubi's possession (or in his stomach) during any of Shonubi's previous drug-smuggling escapades. But once Shonubi was convicted for importing heroin during his eighth drug-smuggling run to New York, the question of how much heroin Shonubi imported during his earlier trips became important. Under the applicable sentencing guidelines the severity of Shonubi's punishment depended in part on the aggregate weight in grams of the heroin that Shonubi imported during *all eight* of his known drug-

⁴ The statement in the text is true only in a metaphorical sense. The amount said to have been "found" in Shonubi's possession was itself an *estimate*, an estimate based on testing of the contents of just four (4) of the 103 balloons found in Shonubi's intestinal tract.

⁵ In common parlance, Shonubi was a "mule" who illegally imported heroin by using a modus operandi in wide currency among mules whose smuggling trips originated in Nigeria. Smugglers using this m.o. fill balloons with a paste mixed with heroin and they then swallow such balloons shortly before embarking on a trip for the United States by plane. If a mule survives the trans-Atlantic trip – if none of the balloons burst while in the mule's intestinal tract – the normal workings of the digestive and intestinal tract eventually lead to the expulsion of the heroin-filled balloons from the mule's body. See David Margolick, "At the Bar; For Lawyers and Judges in the Fight against Drugs, the Repugnant becomes Run of the Mill," *New York Times* (National Desk; Law Page), Section B, p. 7, column 5 (July 24, 1992). See also, e.g., Mike Claffey, "After 9/11 lull, heroin, Ecstasy busts soar," *Daily News, Domestic News* (May 15, 2002) (reporting, inter alia, that a 12-year-old boy from Nigeria ... got through customs at Kennedy with 87 bags of heroin in his stomach"; but then he got sick and was arrested).

smuggling trips from Nigeria. For example, if Shonubi was found to have imported 1000 grams of heroin or more, and not just 427.4 grams, his sentence would have to be substantially “enhanced,” and if it was found that he imported 3000 grams of heroin or more, his sentence would have to be enhanced yet more.

Stated roughly, the question facing the sentencing judge was the following: “What was the aggregate amount of heroin (in grams) that Shonubi imported during all eight drug-smuggling trips that Shonubi is believed to have made?” Stated differently and somewhat more precisely, the question before the sentencing court was this: “Given that the evidence convincingly establishes the amount of heroin Shonubi imported during his eighth trip, how much heroin did Shonubi import in the seven other smuggling trips that the evidence shows that Shonubi very probably took?” The most precise (if somewhat misleading⁶) description of the problem facing the court is this:

Given that the evidence convincingly establishes that Shonubi imported 427.4⁷ grams during his eighth and final trip and given that the only legally-material break points for sentencing purposes are 1000 and 3000 grams, respectively, the questions facing the sentencing judge are (i) whether Shonubi imported at total of least 572.6 grams during the other seven drug-smuggling trips that the evidence shows he very probably made and (ii) whether Shonubi imported a total of at least 2572.6 grams in those first seven trips.”

⁶ See n. 7, *infra*.

⁷ The attentive reader will note that is quite improbable that *any* evidence that could have been presented at the sentencing proceeding would have convinced Judge Weinstein that it was more probable than not that (i) Shonubi imported 427.4 grams rather than 427.3 grams (or 427.3 grams or less); (ii) Shonubi imported 427.4 grams rather than 427.5 grams (or 427.5 grams or more); (iii) Shonubi imported 427.4 grams rather than either (a) 427.3 grams or less or (b) 427.5 grams or more; or, even, (iv) Shonubi imported 427.4 grams or more rather than 427.3 grams or less. Judge Weinstein’s estimate of the aggregate weight of the heroin that Shonubi imported might have seemed more plausible to the Second Circuit Court of Appeals if Judge Weinstein expressed his estimates with less discrete numbers, numbers that more plausibly expressed what Judge Weinstein actually came to believe.

Judge Jack B. Weinstein was the sentencing judge. He concluded that to determine how much heroin Shonubi imported, he, Judge Weinstein, could and should consider statistics compiled by the US Customs Service about the amounts of heroin imported by similar smugglers (Nigerian nationals) using similar smuggling methods (swallowing of balloons stuffed with heroin) in similar circumstances and places (smuggling trips originating in Nigeria and terminating at JFK Airport in New York) in the same period of time (from September 1, 1990 through December 10, 1991) in which Shonubi carried out his eight known drug-smuggling runs via JFK Airport. By doing this, Judge Weinstein implicitly concluded that a prior ruling by the Second Circuit Court of Appeals in an earlier phase of the same case (Shonubi II⁸) did not bar him from considering the statistics compiled by the US Customs Service.

After considering the Customs Service statistics, a report by two court-appointed experts,⁹ and some other evidence,¹⁰ Judge Weinstein further ruled that the evidence presented at the sentencing hearing demonstrated that the aggregate amount of heroin that Shonubi had imported during his eight known drug-smuggling trips was between 1000 and 3000 grams. Stated differently, Judge Weinstein found that the evidence presented at the sentencing hearing established practically to a certainty that Shonubi imported 427.4 grams on his eighth and final trip, and he found that other evidence presented at the sentencing proceeding – some of it statistical and some, non-statistical – established that it was more probable than not that Shonubi had imported at least 572.6 grams in all seven of his remaining known drug-smuggling trips. Judge Weinstein wrote a long opinion – Shonubi III – setting forth the justification for this

⁸ United States v. Shonubi, 998 F.2d 84 (2d Cir, 1993) (Oakes, Newman & Cardamone, JJ.)

⁹ One of those experts was David A. Schum. The other (putative) expert was this writer. These two court-appointed experts submitted a joint report. (This joint report took the form of an affidavit.)

¹⁰ The judicial opinions resulting from Shonubi's trial and sentencing – there are five of them! – are now generally viewed as opinions that deal with the use of statistical evidence in courtroom proceedings. However, the evidence that Judge Weinstein actually considered in the proceeding discussed in Shonubi IV was not limited to statistical evidence. See n. 3, *supra*.

factual finding and “explaining how a sentencing judge – and a trier of fact generally – reaches a decision.”

Charles Shonubi appealed his sentence – for the second time.¹¹ The US Court of Appeals for the Second Circuit *again* concluded – in Shonubi IV – that Judge Weinstein had erred *again*. The Second Circuit did *not* repeat its earlier suggestion that there had to be “direct” evidence of the amount of heroin that Shonubi imported in his various smuggling trips; indeed, in Shonubi IV the Second Circuit expressly disavowed any reliance on a “direct evidence” requirement. Instead, the Second Circuit Court of Appeals expanded and expressly reaffirmed its earlier suggestion (in Shonubi II) that a finding of the amount of heroin that Shonubi imported in his trips had to be supported by “specific evidence.” Furthermore, Judge Newman (speaking for a three-judge panel of the Second Circuit) ruled that the Customs Service statistics about the quantities of heroin imported by other balloon-swallowing Nigerian drug smugglers did not constitute such “specific evidence” of the amount of heroin that Shonubi had imported during his first seven known drug-smuggling trips.¹²

¹¹ In the first sentencing proceeding Judge Weinstein found that “the defendant did import at least 3419.2 grams of heroin” in his eight drug-smuggling trips taken together. *United States v. Shonubi*, 802 F. Supp. 859, 864 (E.D.N.Y., 1992) (Shonubi I). To reach the conclusion that Shonubi had imported 3419.2 grams Judge Weinstein accepted the federal probation service’s recommendation that he multiply by eight (8) the amount of heroin that Shonubi was found to have been carrying at the time of his arrest. Shonubi appealed his sentence. The Court of Appeals for the Second Circuit, in Shonubi II, vacated the sentence imposed by Judge Weinstein and it remanded the matter to Judge Weinstein for re-sentencing. *United States v. Shonubi*, 998 F.2d 84 (2d Cir, 1993) (Oakes, Newman & Cardamone, JJ.) (Shonubi II).

¹² The implications of the Second Circuit’s “specific evidence” doctrine reach beyond sentencing proceedings because the rules governing the admissibility of evidence at trials on the merits are at least as strict as and ordinarily much more strict than the admissibility rules applicable in sentencing proceedings. Furthermore, as a result of recent US Supreme Court decisions (see *Apprendi v. New Jersey*, 530 U.S. 466 (2000) and *Blakely v. Washington*, – U.S. –, 124 S. Ct. 2531, 159 L. Ed. 2d 403, 2004 U.S. LEXIS 4573 (2004)), many issues that previously would have been resolved by the trial judge in a sentencing proceeding will now have to be resolved by the jury. Hence, if the misbegotten specific evidence requirement survives at all, there is ample reason to think that the Second Circuit’s specific evidence requirement will survive whatever form the revamping of the allocation of judge-jury factfinding responsibilities in criminal cases ultimately takes.

Judge Weinstein responded to Shonubi IV with a biting opinion: Shonubi V.¹³ He accused the Second Circuit of resurrecting the medieval-canon law theory of proof, a theory that graded different categories or forms of proof. Judge Weinstein's criticism of the Second Circuit's specific evidence doctrine was ironic as well as biting. Judge Weinstein knew, and he knew that Judge Newman knew, that the medieval-canon system of proof used *numbers* to characterize the quality of different types of evidence and proof. Judge Weinstein effectively charged that Judge Newman effectively favors numerical grading of evidence and proof!¹⁴

I responded to Shonubi IV with a critique of my own.¹⁵ My paper – hereafter “Critique” – focused on Judge Newman's talk about the need for “specific evidence.” I argued that Judge Newman's specific evidence requirement is practically incoherent: I argued that the attempted distinction between specific and non-specific evidence is almost unintelligible. I further argued that the logic of the specific evidence requirement – if taken seriously – leads to the conclusion that statistical evidence should always be inadmissible, a conclusion that is wildly inconsistent with the Second Circuit's general position on the use of statistical evidence in the courtroom. Finally, I argued that there is no meaningful way of limiting the reach of the specific evidence requirement to statistics and statistical argument – this was an easy argument to make because there is,

¹³ *United States v. Shonubi*, 962 F. Supp. 370 (E.D.N.Y., 1997).

¹⁴ This charge, apart from being mildly titillating, is slightly hyperbolic. But I suspect that Judge Weinstein's riposte to Judge Newman was also meant to communicate a subtle and important insight: when talk turns to evidence, inference, and proof, it is far more difficult than some people imagine to avoid quantitative expressions and numbers. Judge Weinstein's reply to Judge Newman was perhaps designed to make the point that the real issue facing judges and lawmakers is not *whether* to allow use numbers or quantitative expressions when talking about or presenting evidence in the courtroom – but *which sorts of quantitative expressions are most suitable* in this courtroom proceeding or that.

¹⁵ P. Tillers, “Introduction: Three Original Contributions to Three Important Problems in the Law of Evidence,” 18 *Cardozo L. Rev.* 1875 (1997). The pertinent portion of that article is available online (bereft of most citations) as “*United States v. Shonubi: A Statistical Oddity*” at <http://tillers.net/shonubi.html>.

after all, no principled way to limit the reach of an unintelligible requirement – and I argued that the Second Circuit’s misguided doctrine about evidentiary specificity threatens to disfigure, not just the law’s treatment of statistics and statistical argument, but also the law’s treatment of ordinary evidence and inference in the courtroom.

On the last point I was a bit of a Chicken Little. As it happens, the Second Circuit’s view about the importance of evidentiary specificity in sentencing proceedings has won neither a wide judicial following¹⁶ nor widespread admiration in the academy. Now, however, three academics have stepped forward to defend the Second Circuit’s specific evidence rule; Mark Colyvan, Helen M. Regan, and Scott Ferson did so in their deliciously-entitled paper *Is It a Crime to Belong to a Reference Class?*¹⁷

¹⁶ Very few federal courts outside of the Second Circuit have even mentioned the Second Circuit’s specific evidence requirement and, as of December 18, 2004, no federal court outside of the Second Circuit had endorsed the Second Circuit’s rule. Cf. *United States v. Mora-Higuera*, 269 F.3d 905 (8th Cir., 2001) (referring to the “Second Circuit’s *more rigorous* “specific evidence” standard”)(emphasis added). Only one state court has even mentioned *Shonubi IV* and that court did so for an unrelated point. See *Guzman v. St. Francis Hospital, Inc.*, 2001 Wis. App. 21, 240 Wis. 2d 559 (2000).

Judge Francis M. Allegra of the US Court of Federal Claims saw the Second Circuit’s specific evidence requirement as a condemnation of the use of “naked statistical evidence,” statistical evidence unaccompanied by non-statistical evidence. *Hart v. Secretary of HHS*, 60 Fed. Cl. 598, 2004 U.S. Claims LEXIS 123 (2004) (“[Statistical] evidence is useful only when properly combined with more conventional evidence, so that probabilistic statistics constitute but a single link in a longer chain of proof. See *United States v. Shonubi*, 103 F.3d 1085, 1091-92 (2d Cir. 1997).”).

¹⁷ Mark Colyvan, Helen M. Regan, and Scott Ferson, “Is It a Crime to Belong to a Reference Class?,” 9 *Journal of Political Philosophy* 168-181 (2001), *reprinted in*, Henry E. Kyburg, Jr. & Mariam Thalos, eds., *PROBABILITY IS THE VERY GUIDE OF LIFE* 331-347 (2003).

It is not premature to assert that the answer to the question that Colyvan, Ferson, and Regan pose in the title of their paper is, in some sense, “surely not.” The qualifier “in some sense” is necessary because much depends on how one can become a member of a reference class. For example, if “murderers” is a reference class and if a person can become a member of the reference class “murderers” by committing a murder, the correct answer to the question posed by Colyvan and his colleagues in the title of their paper is that it *is* possible for “mere membership” in a reference class to be a crime.

In their paper¹⁸ Colyvan, Regan, and Ferson defend the specific evidence requirement that the Second Circuit created in Shonubi IV. As the title of their paper suggests, Colyvan, Regan and Ferson argue that the Second Circuit's requirement of evidentiary specificity makes good sense. They further argue that my characterization of the evidentiary specificity doctrine as nonsense is itself nonsense. Here I will say a few words about these claims and about one or two related issues.

3

Colyvan, Regan, and Ferson say that it is important to consider the Second Circuit's evidentiary specificity requirement from the standpoint of the "reference class problem."¹⁹ They chide discussants for their (alleged) failure to discuss the implications

¹⁸ Sometimes I will refer to the authors as "CR & F." I use the acronym "CRF Paper" to refer to their paper and citations to specific pages in the paper refer to the book edited by Kyburg (see n. 17, *supra*).

¹⁹ The reference class problem can be described in various ways. Stated loosely and conceived broadly, the reference class problem is a variant of the following question: Assuming that experience, observations of regularities in nature, or associations among events generate knowledge – or a basis for judgment – about the probability of unknown or uncertain events, and given that any particular event has a large or infinite number of attributes – that is, given that any particular event belongs to a large or infinite number of sets, classifications, or categories of events – which category or categories of events of that large or infinite set of categories should serve as the basis for a conclusion or judgment about the probability of that particular event? The "problem" here is that (i) different reference classes seem to generate – or push toward – different probabilities, (ii) many reference classes – perhaps an infinite number – are available for consideration, and (iii) it is difficult to justify the choice of any particular reference class or any particular combination of reference classes. Branden Fitelson, Alan Hájek, and Ned Hall provide the following elegant statement of the reference class problem: "Any version of frequentism faces the notorious *reference class problem*. Any event, in all its detail, occurs exactly once, so if non-trivial frequencies are to be associated with it, it must be regarded as a token of a more general event type, whose instances constitute its reference class. However, there are indefinitely many ways of typing a given event. This would not be a problem if its relative frequency was the same in each reference class, or if one such class stood out as natural or privileged. The problem gains teeth to the extent that various competing reference classes have equal claim to determining the probability and that they yield different relative frequencies for the event." B. Fitelson, A. Hájek & N. Hall "Probability," p. 6, <http://fitelson.org/probability.pdf>, *forthcoming in* ROUTLEDGE ENCYCLOPEDIA OF PHILOSOPHY OF SCIENCE (Jessica Pfeiffer, Sherri Rausch & Sahotra Sarkar, eds., Routledge, 2005).

of the reference class problem for the question of the advisability of the Second Circuit's rule requiring specific evidence in sentencing proceedings. However, Colyvan, Regan, and Ferson do not make it entirely clear whether they think the judges who considered the specific evidence requirement also erred by failing to discuss the reference class problem.

Colyvan, Regan, and Ferson do not unambiguously assert that the correct solution to the reference class problem shows that the specific evidence requirement is well-advised. One might reasonably think that this must have been the thesis that Colyvan, Regan, and Ferson were advancing. The difficulty with this interpretation of the authors' intentions is that Colyvan, Regan, and Ferson expressly say they offer no solution to the reference class problem. (I am inclined to agree that they have not offered a solution.) They only want to establish, they say, that it is a good idea to discuss the reference class problem.

I am not against discussion of the reference class problem. I am in favor of such discussion.²⁰ But if Colyvan, Regan, and Ferson really mean to say (as they say they do)

²⁰I discussed the reference class problem at considerable length in my revision of the first volume of John Henry Wigmore's treatise on the law of evidence. See, e.g., 1A WIGMORE ON EVIDENCE §37.6 & 37.7, at 1046-1095 (*passim*), esp. §37.6 n. 8, at 1050-1054 (P. Tillers. Rev., 1983).

The joint testimony that David Schum and I submitted (in the form of a joint affidavit) to Judge Weinstein discussed some of the issues that were presented by the reference class that emerged from the statistics collected by the U.S. Customs Service.

Colyvan, Regan, and Ferson suggest that I did not discuss the reference class problem in my published critique of Shonubi IV. CR & F do hedge a bit: in n. 9 of the CRF Paper they acknowledge that I "actually" did discuss the reference class problem. But they then go on to say that in that paper I didn't say much about the reference class problem. I beg to differ. While it is true that I did not attempt to present – in that paper – a *comprehensive* discussion of "the" reference class problem, I certainly *meant* to discuss *some* of the possible problems with the use of some kinds of reference classes in some situations. And I think I did in fact do so. (The expression "reference class" appears six times in the online version of this paper.) Moreover – and more important – there is a real sense in which my previously-published critique of Shonubi IV is *nothing except* a discussion of the appropriate limits on the use of reference classes. (It may also be pertinent to note that a Venn Diagram adorns the title page of the online version of my critique of the Second Circuit's opinion in Shonubi IV. This Venn Diagram was an effort – apparently unsuccessful – to inject a bit of humor into discussion of the

that their discussion of the reference class problem does not solve the reference class problem, I do not quite understand why they think that their (inconclusive) discussion of that problem shows that the specific evidence requirement makes sense. The answer to this little riddle may be that, despite their avowals to the contrary, Colyvan, Regan, and Ferson *do* think they have offered a solution to the reference class problem. At one point in their argument they seem to prefer a causality-based explanation of the efficacy of reference classes in statistical argument. At another point they seem to suggest that a special form of abduction – inference-to-the-best-explanation theory – is the solution. Perhaps Colyvan, Regan, and Ferson think they have demonstrated the superiority of one or another approach to the reference class problem. I confess I am not sure which solution they favor. But it hardly matters: most informed observers would agree that it will take more than a small handful of pages of discussion to vindicate a claim that one approach or another is the proper solution to the reference class problem.²¹

4

I re-learned the meaning of “orthogonal” only a few years ago. (I must have known the meaning of the word when I studied plane geometry and solid geometry when I was in junior high school.) As a result of my recent re-education, I am now in a position to say that one of the arguments that Colyvan, Regan, and Ferson make about evidentiary specificity is orthogonal to mine – that, in other words, the argument in my critique of Shonubi IV and one of the arguments made by Colyvan, Regan, and Ferson are like the

reference class problem. I concluded my critique of Shonubi IV by saying, “I very much hope that the Second Circuit's opinion in Shonubi IV is not a reference class or a part of a reference class that should serve as a basis for predictions about how other courts (or the Second Circuit itself) will deal with statistical evidence and statistical inference in the future. In short, I fervently hope that the Second Circuit's opinion in Shonubi IV is thoroughly unrepresentative and atypical.” The reference class problem was very much on my mind – and, I think, palpable in my paper.)

²¹ One also needs to know what a “solution” to “the” reference class problem would do. For example, must an acceptable solution generate an autonomous procedure for wresting probabilities about individual events out of the appropriate reference class or combination of reference classes? Or might a “solution” merely tell us how human beings (often) make probabilistic sense out of experience – without providing a replica of or substitute for human judgment?

two fabled ships that pass each other unnoticed in the night. In particular, Colyvan, Regan, and Ferson argue in part that the class of “Nigerian balloon-swallowing heroin smugglers arrested at JFK Airport from September 1, 1990, through December 10, 1991,” is an inadequate basis for a valid statistical argument about the amount of heroin that Charles O. Shonubi – a Nigerian balloon-swallowing heroin smuggler arrested at JFK Airport – imported into the United States in seven drug smuggling trips that he was believed to have taken from September 1, 1990, until his arrest at the end of his eighth trip on December 10, 1991. CR & F may or may not be right about this. But for present purposes whether they are right or wrong on this point does not matter: while the question of the sufficiency or insufficiency of the particular reference class mentioned above for valid statistical argument about the amount of heroin imported by Shonubi is clearly important for an assessment of the sentence that Judge Weinstein imposed on Shonubi in the second sentencing proceeding (which was the sentencing proceeding discussed in Shonubi III), the question of the sufficiency or insufficiency of that particular reference class for that particular purpose is immaterial to the argument I made in my Critique. I did *not* argue that the statistics that Judge Weinstein used in Shonubi III provide a sufficient basis for inferences amount the amount of heroin that Charles O. Shonubi imported. I also did *not* argue that Judge Weinstein used appropriate statistical methods to draw inferences from the statistics that were available to him. (I thought that I should leave such questions to people who are better qualified to handle them.) In general, my Critique did not defend any particular statistical methodology. Conversely, the argument in my Critique did not depend on my view of the validity or invalidity of the statistics or statistical methods used by Judge Weinstein. Bluntly stated, my criticisms of the Second Circuit’s specific evidence requirement stand even if Judge Weinstein used truly rotten statistics and truly rotten statistical methods.

5

My Critique advanced the following three important propositions:

- (i) it is *not* always inferentially illegitimate to base inferences about the behavior of an individual on the behavior of other people;

(ii) if group-to-individual inference is a vice, this vice is not obviated by avoidance of formal statistical methods; and

(iii) the Second Circuit's requirement that statistics be limited to the behavior of the individual about whom inferences are to be drawn either is incoherent or the logic of the requirement altogether bars the use of statistical evidence to show individual human behavior.

Colyvan, Regan, and Ferson seem to respond to these three claims in the following way:

CR & F directly challenge proposition (i).

CR & F ignore proposition (ii); they do not address proposition (ii).

CR & F reject the thesis that the logic of the Second Circuit's specific evidence requirement entails the conclusion the use of any statistical evidence to show human behavior is inappropriate; but they reject this thesis *on the ground that proposition (i) is false*; and they maintain that it is logically consistent to disallow the use of statistics about groups to draw inferences about the behavior of individuals but allow the use of statistics about an individual to show the behavior of the *same individual*.

If the above summary is correct, the central point that divides me from my three critics is a difference of opinion about the validity or invalidity of using statistics about groups of individuals to support or draw inferences about individuals. The discussion in this paper therefore focuses on this disagreement.

6

There is an elephant in the room. Judge Newman and Colyvan, Regan, and Ferson – all of them – studiously ignore the elephant. The elephant is a question that usually plays a

large and *visible* role in controversies about the influence of non-individuated classifications on the distribution of official penalties and government burdens. In debates about the role of non-individuated classifications in such contexts – such as in debates about the propriety of racial “profiling” – the question usually arises (i) whether non-individuated classifications should be spurned because the use of such classifications is epistemically unwarranted or (ii) whether they should be spurned, not because they generate an unacceptable frequency of errors about factual hypotheses, but on some other ground – for example, on the ground that the use of non-individuated classifications to determine the incidence of official penalties and the distribution of social burdens is *unjust* or *socially harmful*.

One can imagine reasons why Judge Newman might have wanted to ignore the possibility that inferential, epistemic, or statistical considerations do not explain or justify the specific evidence requirement. For example, since even unelected judges may wish to avoid public anger and scorn, a judge might prefer to condemn a certain type of evidence on the ground that such evidence lacks probative value rather than on the ground that evidence should be ignored to avert the social harm or offense that its use would occasion even if the spurned evidence has probative force. So one can understand why judges might not want to make the true basis of their condemnation of some forms of evidence fully transparent. But it is harder to understand why Colyvan, Regan, and Ferson chose to ignore the possibility that the Second Circuit’s specific evidence requirement is best explained or justified by considerations of justice or social policy rather than by epistemic, inferential, or statistical principles.

The reference class that occasioned the birth of the Second Circuit’s specific evidence requirement was, roughly, “balloon-swallowing Nigerian heroin importers arrested at JFK Airport from September 1, 1990, through December 10, 1991.” Nationality is an ingredient of this reference class – and, since almost all Nigerians are Black, perhaps it is fair to say that in some sense race is also effectively an ingredient of this particular reference class. Given the turbulent history of race relations in the United States and given the current widespread legal and social hostility toward racial prejudice and discrimination, a massive amount of intelligence is not required to imagine that

considerations of social policy or justice rather than epistemic, inferential, or statistical considerations might induce a judge or a court to object to the reference class of “**Nigerian** balloon-swallowing heroin importers arrested at JFK Airport from September 1, 1990, through December 10, 1991.”

It is true that the Second Circuit’s specific evidence requirement does not expressly advert to the danger or dangers of the use of race or nationality to support inference about individuals. But it does not follow that explanations for the specific evidence requirement that are not rooted in epistemic, inferential, or statistical considerations must be ruled out. The facial neutrality of the phrase “specific evidence” with respect to race and nationality may mean only that it is necessary to consider possible non-epistemic explanations and justifications that are not tightly tied to nationality or race. Judge Newman’s brief discussion of his short list of examples of evidence that is specific in his eyes and examples of evidence that is not²² suggests that the specific evidence requirement rests on a normative antipathy to inferences about individuals that rest on evidence about groups of individuals.²³ In short, it does not take much imagination to suppose that the specific evidence requirement has its roots in Judge Newman’s (and, by extension, the Second Circuit’s) intuition, feeling, or sense that it is morally, ethically, and socially odious allow the use in the courtroom and in legal proceedings (particularly in criminal proceedings) of forms of evidence that allow inferences about individuals to be drawn that cause individuals to suffer legal burdens or penalties that those individuals do not deserve to have inflicted upon them because of their individual choices and behavior. There is ample reason to suspect that Judge Newman just did not

²² For example, Judge Newman wrote, “The ‘specific evidence’ we required to prove a relevant-conduct quantity [sic] of drugs for purposes of enhancing a sentence must be evidence that points specifically to a drug quantity for which the defendant is responsible. By mentioning ‘drug records’ and ‘admissions’ as examples of specific evidence we thought it reasonably clear that we were referring to the defendant — *his* admissions and records of *his* drug transactions. And by ‘live testimony’ we were referring to testimony about *his* drug transactions.” Shonubi IV, 103 F.3d at 1089-1090. He also wrote, “The statistical and economic analyses relate to drug trafficking generally and not to Shonubi specifically.” *Id.* at 1091.

²³ Colyvan, Regan, and Ferson also interpret the Second Circuit’s prohibition of non-specific evidence as a requirement that only the behavior of the affected individual be used to support inferences about that individual. CRF Paper at 339.

think that it would be *fair* to let the extent of the punishment inflicted on the *individual* Charles O. Shonubi depend even in part on the behavior of a group of other people, which in this instance consisted of a small sub-group of (other) Nigerian nationals.

So why did Judge Newman choose not to talk about the elephant in the room, the possibility that a rationale rooted in consideration of justice or social policy rather than in epistemic, inferential, or statistical considerations explains and justifies the specific evidence requirement?

I do not know the answer. And surely there were multiple reasons²⁴ for his choice to proceed as he did with his opinion in *Shonubi IV*. But I have a hunch that one particular consideration played a role.

It is possible that Judge Newman sensed that an extra-epistemic explanation for the specific evidence requirement just would not pan out. The road to hell is paved with

²⁴ In one part of *Shonubi IV* Judge Newman clearly expresses his concern about a sentencing scheme that allows (and often mandates) substantial enhancement of the punishment of a convicted defendant upon a showing by a mere preponderance of the evidence that factors allowing or requiring enhancement are actually present, and Judge Newman expressly states that this concern is part of the reason for the creation of the specific evidence requirement. See *Shonubi IV* 103 F.3d at 1089 (“A guideline system that prescribes punishment for unconvicted conduct at the same level of severity as convicted conduct obviously obliges courts to proceed carefully in determining the standards for establishing whether the relevant conduct has been proven. We have recognized the need for such care with regard to the basic issue of the degree of the burden of proof. Thus, though the Sentencing Commission has favored the preponderance-of-the-evidence standard for resolving all disputed fact issues at sentencing ..., we have ruled that a more rigorous standard should be used in determining disputed aspects of relevant conduct where such conduct, if proven, will significantly enhance a sentence...”; “A similar concern guided our decision on the prior appeal in this case. Aware of the consequences of a relevant conduct finding as to drug quantities, we invoked the rule from prior case law of our Circuit that, we observed, ‘uniformly requires specific evidence--e.g., drug records, admissions or live testimony--to calculate drug quantities for sentencing purposes.’”)

Compare David T. Wasserman, “The Morality of Statistical Proof and the Risk of Mistaken Liability,” 13 *Cardozo L. Rev.* 935 (1991)(it is morally proper to impose less exacting evidentiary standards for judgments about the appropriate extent of liability than for judgments about the existence or non-existence of liability; but it is morally improper to use non-individuated “base rates” on the question of the existence or non-existence of liability).

good intentions. If Judge Newman thought about non-epistemic explanations and justifications at all – and even if he just let his feelings run his brain – Judge Newman must have immediately understood that a conclusion that the specific evidence requirement is rooted in moral or ethical intuitions or in a sense of justice rather than in epistemic, inferential, or statistical considerations would not avert the need for further argument about the wisdom or unwisdom of a requirement of evidentiary specificity of the sort the Second Circuit had in mind. For example, if some sort of moral intuition or moral precept is the true basis of the requirement of evidentiary specificity, it would still be necessary to think carefully about two additional questions. First, there are effective and ineffective ways of realizing ideals. One general question is whether a specific evidence requirement does the job required of it. Second, if a moral or ethical precept, principle, or intuition is the basis of the specificity requirement, it is still necessary to settle questions about priorities. Should society give absolute priority to the principle that individuals should never suffer adverse legal or social consequences because of inferences about them that are false because of group-to-individual inferences? Is it sometimes morally or ethically permissible to inflict harm on individuals for the sake of acquiring accurate factual knowledge? If so, are group-to-individual inferences sometimes permitted if they advance the search for truth more than they harm it – even if by allowing group-to-individual inferences we sometimes harm individuals who have done nothing to deserve – by reason of their individual misconduct – the harm that they are made to suffer?

7

The Second Circuit's suggestion that non-specific statistical evidence is inferentially illegitimate is *intelligible* – but just barely so! – only if it is taken to stand for the proposition that statistical inference about an individual's unknown behavior can be supported only by statistics about that same individual. Colyvan, Regan, and Ferson turn this intimation by the Second Circuit into an explicit assertion: reference classes that are offered to show the behavior of an individual on a particular occasion must be limited to statistics that describe the behavior of the *same individual* on other occasions.

But if this is the thesis that the Second Circuit meant to advance, the Court embraced a wildly-implausible thesis.

I note first that a great deal of law on the books demonstrates that group behavior can be used as evidence in a wide variety of ways to show the behavior of an individual.²⁵

Furthermore, if statistical inferences about the behavior of an individual may rest only on statistics about that same individual, the following strange propositions seem to follow:

- (i) if we have not seen X in action but we learn that X was violently attacked by Y, we cannot infer that the chances that X attacked Y have increased;

²⁵ A handful of examples should suffice:

(i) Federal Rule of Evidence 406 permits the use of the routine practice of an organization to show that acts in conformity with that practice took place. For example, such routine practice is admissible to show that a particular letter was mailed. There is no requirement that evidence of such routine practice show the prior or routine practice of the particular person who allegedly mailed the letter on the particular occasion in question. It is enough that letters were routinely handled in a certain way by mail handlers in the organization and that the particular person who may have mailed the letter was a mail handler.

(ii) Evidence that a gang (a group of persons) exhibits animus toward the police is admissible to show that a particular member of the gang felt or exhibited animus toward the police. *Barnes v. City of New York*, 296 A.D.2d 330; 745 N.Y.S.2d 20 (2002). I have no doubt that an individual's membership in the Ku Klux Klan is admissible to show that individual's racial attitudes.

(iii) Surely I need not cite cases to show that it is legally permissible to show that some individual probably speaks Italian by showing that that same individual was raised by Italian-speaking parents in "Little Italy" in New York City or the North End (once a predominantly Italian neighborhood) in Boston.

These are three examples of situations in which the behavior of other people is relevant to how a particular individual behaves. Surely I need not multiply such examples. (It does not follow, of course, that group behavior always has substantial probative force on the behavior of a particular individual or even that group behavior is always relevant to a question about an individual's behavior. The only question at the moment is whether "non-specific" evidence *ever* has probative value for a question about individual behavior. The law says that sometimes such evidence is relevant, probative, and admissible.)

(ii) if we have not seen X in action but we learn that Z divorced him, we cannot infer that the chances that X physically attacked Z have increased;

(iii) if we have not seen X in action but we learn that X belongs to a neo-Nazi party, we cannot infer that the chances, in our minds, of an attack by X on a Jewish person have increased; and

(iv) if we have not seen X in action but we learn or believe that X is 6'10" inches tall and weighs 300 lbs., rather than, as we had earlier supposed, 4'10" tall and weighed 85 lbs, we cannot infer that the chances (in our minds) are now greater that he would swallow 200 balloons filled with 2000 grams of paste and heroin.

You (a hypothetical interlocutor) might reply:

Dear friend, I will disregard the fact that situation (iv) is not a case in which the person making an inference about an individual bases his inference on the evidence he has about other people. That aside, these are not really instances in which one has no specific evidence about X. For example, in situation (iv) we know that X weighs 300 lbs and in situation (i) we know that X was attacked by Y. That is specific information about X.

But if you (my interlocutor) were to say that to me, I might counter thus:

My dear fellow, you cannot distinguish the *Shonubi* situation on that basis. For, you see, in *Shonubi* the fact finder (the sentencing judge) also had specific information about Charles O. Shonubi: the trial judge knew, for example, that Shonubi was Nigerian, that Shonubi had been arrested, that Shonubi was a heroin smuggler, and that Shonubi swallowed balloons filled with a heroin-larded paste . The real question, you see, is, given the specific knowledge we have of the specific behavior and attributes of a specific person, what further inferences can we draw about some proposition of interest about that person given our knowledge of matters other than the proposition of interest, given our

knowledge of matters such as the behavior of similar persons or entities under similar circumstances or our knowledge of the same person under different but similar circumstances. You see, my friend, whenever we use our knowledge of other entities under similar circumstance or our knowledge of the same entity under different but similar circumstances, we are using our knowledge of matters in the world other than the evidence we have about this specific entity in this specific situation. Indeed, without such extra-specific, or “beyond-specific” knowledge – I refuse to call it “non-specific” or “unspecific,” since to do so begs the question under discussion here – we cannot possibly say that we are making observations of similar entities – and we cannot even possibly ever say that in any instance we are making observations of the “same” entity under similar circumstances! For example, without extra-specific knowledge we cannot say that the entity we saw yesterday and the entity that seems to be the same entity whose actions or attributes today are in question is or is not materially different in some pertinent respect from the entity that we thought we saw yesterday.

Oh, and by the way, in situation (iv) there *is* evidence of an entity other than the individual about whom an inference is in play. The other entity is 4’10” tall and weighs 85 lbs. More generally speaking, situation (iv) illustrates the essential role that general knowledge of the world and of the actors that populate the world plays in the drawing of any inference about any particular individual human being. Inferential use of knowledge of the behavior or attributes of particular groups in the world is just one species, or example, of the role that knowledge of the world plays in the drawing of inferences about an individual human being.

If you – my hypothetical interlocutor – were to tell me that I should not allowed base any inferences about the individual X by considering the manner in which the properties that are instantiated in X influence or are associated with some “other behavior” or with “matters other than X” – in short, if you were to tell me that I ought not consider behavior or matters other than the behavior or matter that is the ultimate subject of my inferential interest – I would have say to you that you have effectively

precluded the use of any evidence whatever. I would say to you that in your eyes the only evidence that is legitimate evidence is only the particular matter or act – the “specific” matter or act – that stands in question. It is not helpful to be told that the only valid evidence of what happened is what happened.

If I said such things to you – to my hypothetical protagonist – , you would be compelled to recognize (or so I imagine) the force of my argument. You might then decide to hedge your original position by telling me that what you really meant to say is that, yes, I can (and must) make use of knowledge that reaches beyond the specific situation that is the object of my inferential interest but that I can base inferences about X *only on my observations of the behavior and attributes of X*. But if you were to say that to me, I would say this to you:

Under the limitations and strictures you would put upon me, I would not be able to tell what I was seeing when I was watching X on *any occasion* – under the stricture you would impose on me I can not make any meaningful observations of X – because to know what I am seeing when I watch X I need to have knowledge about many things and matters other than X. So you see, the demand that I base my conclusions about X solely on my knowledge of X alone just takes us round and round and leads us ... nowhere. It is in the nature of evidence, I say – I have said this before, and now I say it again – it is in the nature of evidence to reach beyond itself and give us a basis for making guesses, drawing inferences, and making conclusions about things previously unknown or previously unobserved. You would destroy the very possibility of evidence were you to restrict evidence to those matters that are based solely on knowledge of the entity about whom guesses are to be made and inferences to be drawn.

It is true that many inferences about individual behavior based on statistics about group behavior or about other people are just rotten inferences. But it hardly follows that all inferences about individuals based on statistics about groups are inferentially worthless.

Numerous examples can be cited that strongly suggest that information about the behavior or characteristics of groups *at least occasionally* supports *to some degree* inferences about individuals.

That we are in some sense sometimes (always?) the prisoners of the attributes and behavior of groups of other persons is, of course, an uncomfortable thought.

In *LIFE AND FATE* Vasily Grossman wrote:

... There is a terrible similarity between the principles of Fascism and those of contemporary physics.

Fascism has rejected the concept of a separate individuality, the concept of 'a man', and operates only with vast aggregates. Contemporary physics speaks of the greater or lesser probability of occurrences within this or that aggregate of individual particles. And are not the terrible mechanics of Fascism founded on the principle of quantum mechanics, of political probability?

Fascism arrived at the idea of the liquidation of entire strata of the population, of entire nations and races, on the grounds that there was a greater probability of overt or covert opposition among these groupings than among others: the mechanics of probabilities and of human aggregates.²⁶

It is possible that the kind of liberal and humane sentiment that is found in this powerful passage also animates or colors the thinking of Judge Newman – and also that of Colyvan, Regan, and Ferson – about the importance of “specific evidence” and the dangers of “non-specific” evidence. If so, I applaud them all – Newman, Colyvan, Regan, and Ferson – for their humane spirit and for their liberal sentiments. Nonetheless, I cannot applaud the suggestion of these caring and well-meaning people that inferences based on non-specific evidence are always *inferentially unsound*.

²⁶ Vasily Grossman, *LIFE AND FATE* c. 19 p. 94 (trans. Robert Chandler 1985).

It is true that great social harms can ensue if some types of *valid* inferences are given free rein in law and society. But to say that the use by or in society of some kinds of inferences can cause social harm is to say one thing, and to say that some inferences are inferentially unsound or worthless is to say something quite different.

As already suggested here, the controversy involving Mr. Charles O. Shonubi has clear affinities with current controversies about racial "profiling."²⁷ The question CR & F pose about the possible criminality of just belonging to a reference class is related to controversies about matters such as DWB – the “crime” of driving while Black.²⁸ While I certainly do not propose to criminalize being Black (or, for that matter, being Latvian²⁹), driving by Black people, or any similar condition or conduct, I do wish to note that the question of whether it is *possible* to draw inferences about individuals without burdening individuals with adverse inferences resulting at least in part from such individuals' inherited or unchosen attributes is both an old question and a hard one. For example, the extended debates about “naked statistical evidence” were in part (sometimes implicitly, sometimes explicitly) about this question.³⁰ For my own part, I am not persuaded that it is possible to avoid saddling people with disabilities resulting from situations that individuals happen to find themselves in, through no fault or doing

²⁷ See, e.g., Sharon Davies, “Profiling Terror,” 1 Ohio St. J. Crim. L. 45 (2003); Frederick Schauer, *PROFILES, PROBABILITIES, AND STEREOTYPES* (2003).

²⁸ One notorious example of racially invidious traffic enforcement involved state troopers on the New Jersey Turnpike. See, e.g., Richard L. Jones, “New Jersey Prosecutors Cite Racial Profiling in Dismissal of 86 Criminal Cases,” *New York Times* (Late Edition) Section B, Column 1, Metropolitan Desk. p. 4 (April 20, 2002).

²⁹ My family has some experience with an older form of “profiling.” The 1941 NKVD warrant for the arrest of one of my Latvian grandmothers stipulates that the ground or cause for my grandmother’s arrest is “Wife of banker.” The stated ground for the arrest of my grandfather is “Banker and member of Farmer’s Party.” My grandmother and grandfather were deported to Siberia in cattle cars and by barge. Reports suggest that they both died of starvation and for lack of medicine. This was tough medicine for being a banker and a member of an agrarian political party and for being a wife of a banker.

³⁰ See, e.g., Lea Brilmayer & Lewis Kornhauser, *Review: Quantitative Methods and Legal Decisions*, 46 U. Chi. L. Rev. 116, 124 (1978); David T. Wasserman, “The Morality of Statistical Proof and the Risk of Mistaken Liability,” 13 *Cardozo L. Rev.* 935 (1991).

of their own. I believe even in the sphere of inference there is more than a grain of truth in the old adage that life is unfair.

Valid inference can be unfair.

Law and society, of course, can take steps to mitigate the unfairness of life. Law and society can also take steps to mitigate the unfairness of inference, including the unfairness even of some valid inference. When cognition becomes a social process and a social activity, and not merely a private, or solitary, activity, the rules about thinking and deliberation must change. When thoughts about evidence have social, material, and legal consequences, it is important to consider the cost of even *accurate* inferential processes. However, given the social character and consequences of some inference, it is *equally* important to consider the cost of *inaccurate* inference. It is important to remember that there is no free lunch – which in this context means that a choice to forfeit the making of valid inferences (e.g., by precluding the use of probative evidence) always exacts a price.

I think it follows that in a democratic society – in a society in which authority depends on the consent of the governed – when a court orders that a particular form of evidence or inference not be used, the members of such society have a right to be informed of the precise basis of such a prohibition – they have a right to know whether use of that type of evidence or inference is barred by the judiciary because it is thought (by the judiciary) that using such evidence causes social mischief or whether use of that kind of evidence or inference is precluded because it is thought that the use of that kind of evidence or inference increases the risk or frequency of factual error in legal proceedings.

A portion of a nursery rhyme proclaims:

If wishes were horses,
Beggars would ride.

In its *Shonubi* opinions the Second Circuit was maddeningly unclear about the ground or grounds for its objection to non-specific statistical evidence and inference. In particular, the court of appeals never clearly stated whether it saw non-specific statistical evidence as objectionable because it thought that such evidence is socially harmful or whether the court regarded non-specific statistical evidence as objectionable only because the court thought that such evidence is bereft of inferential force.

I cannot bring myself to believe that the genuinely-eminently three-judge panel that produced Shonubi IV did not appreciate the obvious and real difference between these two kinds of possible objections to non-specific evidence and inference. I am therefore forced to conclude that the panel *deliberately* sought to obscure the very real -- and *palpable* -- distinction between these two types of reasons for characterizing non-specific evidence as objectionable.

Why did the Judge Newman (who spoke for the three-judge panel and for US Court of Appeals for the Second Circuit) engage in such obscurantism?

I am not certain of the answer. (I have no inside knowledge.) But I do have a *suspicion*.

It is a fair guess that the tendency of the Second Circuit's doctrine about evidentiary specificity is to make some evidence and inference in some legal proceedings work the way that the Second Circuit court would *prefer* that evidence and inference work, to make evidence and inference work the way that the Second Circuit *wishes* evidence and inference would work. In particular, it is a good guess that Judge Newman and his colleagues have a strong normative preference for forms of evidence in legal proceedings that do not result in the infliction of penalties on individuals who, given the applicable norms, do not deserve to be penalized. Perhaps it is this sort of general moral preference or precept that explains the Judge Newman's embrace of a requirement of evidentiary specificity, and perhaps the influence of this unspoken moral preference or precept explains the brusqueness of Judge Newman's explanation of and justification for the specific evidence requirement in sentencing proceedings -- and, perhaps, in the future, in other proceedings as well.

I think that any such moral yearning for perfectly equitable and just evidence and inference is a pipe dream. And I wonder if such a pipe dream effectively advances the moral and social ends that the Second Circuit Court of Appeals or Judge Newman may desire. A more hard-headed approach might work better. A more bare-knuckled view – and a better view – of the problem of non-specific evidence would insist on the following points:

- (i) in the realm of evidence and inference – just as in other realms of social and political activity – there is always a potential conflict between the way that we would like things to be and the way that things are; and
- (ii) (a) it is therefore sometimes necessary to sacrifice the pursuit of the truth for the sake of doing good; and, conversely, (b) it is sometimes important to swallow unpleasant consequences in order to advance the search for the truth.

On this more tough-minded (and tragic) view of the human environment, it is no use pretending that conflicts between consequences and inferences do not exist. To try to imagine a world in which conflicts between the search for the truth, on the one hand, and the search for social value, on the other hand, do not occur is to engage in an attempt to live in a Panglossian world of evidence and inference.

In a Panglossian world valid inferences would always have beneficial consequences and inferences that produce unpleasant consequences would always be inferentially invalid. It is possible that this sort of Panglossian vision of inference is the sort of vision that the Second Circuit would like the public to embrace. Perhaps it thinks that delusions – or social myths – sometimes foster social peace and harmony?

Or perhaps the Second Circuit is just a little weak in the knees. Perhaps it just does not want to believe or acknowledge that even in the case of inference human life and existence inevitably have some tragic ingredients.

It is simply untrue that individuals need not be saddled with evidentiary signposts that they themselves did not implant in their environment.

Individuals are burdened with a vast multitude of evidentiary signs that are generated by the cosmos and by other people, rather than by the individuals about whom inferences are to be weighed.

Hence:

Any inferential strategy that purports to abjure all inferences about an individual that rest on events or conditions that are not generated by the decisions or behavior of that individual is bankrupt.

Therefore:

If statistical reasoning based on reference classes is ever to work, such reasoning can work only if resort is made to reference classes that consist at least in part of events that are *not* generated by the choices and behavior of the individual about whom inferences are under consideration.

Remember:

If turnips were watches,
I would wear one by my side.

Furthermore:

Wishing alone does not make a thing so.

The specific evidence requirement embraced by the Second Circuit may be an example of tacit “as if” thinking, somewhat analogous to the kind of “as if” philosophizing that is

found in one strand of neo-Kantian theorizing.³¹ But if the specific evidence requirement is an example of a strategy of wishing that things were so-and-so and then *saying* that things are as one wishes they were even while *knowing* that things are not and cannot be so-and-so, the Second Circuit might have done better if it had made any such “as if” philosophy that was tacitly at work in its thinking the *express* basis of its decision in Shonubi IV.

But perhaps it is just as well that the Second Circuit did not attempt to formulate some sort of “as if” justification for the specific evidence requirement that it promulgated. In today’s world it is vain, or pointless, to pretend that individuals are not burdened by reference classes; and it is vain to pretend that legal proceedings can be so designed such that evidence that is not specific to an individual plays no role in the drawing of inferences about that individual.

An “as if” strategy is a response to a perceived necessity, or a less-than-ideal but ineluctable reality. We should draw the lessons of prior experiences with such a strategy for dealing with unpleasant realities. For example, consider Kant’s monumental effort. Kant struggled with the reality and force of necessity. Kant saw the problem of necessity as the problem of inexorable laws of causation. What was to be done about such necessity? Kant’s solution was to affirm the need to believe that the realm of causation is the realm of freedom, to somehow have faith that the laws of nature are in reality the laws of freedom, or the emanations of free and spontaneous beings.³² Such a response may have been satisfactory to Kant, to some of his philosophical successors, and to some mystics and German Pietists. But this sort of faith-based solution is not, I think, the kind of terminus toward which either the Second Circuit Court of Appeals or Collyan, Regan, and Ferson imagined that they were careering.

³¹ Hans Vaihinger, *DIE PHILOSOPHIE DES ALS-OB. SYSTEM DER THEORETISCHEN, PRAKTISCHEN UND RELIGIÖSEN FIKTIONEN DER MENSCHHEIT* (1911).

³² See, e.g., Kant, *FOUNDATIONS OF THE METAPHYSICS OF MORALS* (1785).

But if necessity in inference – the unpleasant realities of the process of inference – cannot be done away with by an act of will or faith, it is a mistake to suppose that any individual can escape the inferential arrows that the natural and human evidential environment inevitably throws her way; and it is then an error for an individual to think he, she, or any individual can escape being subjected to inferences that depend on observations and judgments about the behavior and attributes of other human beings. Not in a pig's eye! No one can altogether avoid signs, signposts, and evidentiary hints that the operations of the world and other people have created. It is the failure to appreciate this basic point that makes the position of Judge Newman, Colyvan, Regan, and Ferson unsatisfactory and, ultimately, untrue.