

The Washington Post

The paradox of unanimity

By **David Post** February 8, 2016

In ancient Jewish law, a unanimous guilty verdict in a capital case resulted in an acquittal of the defendant.

The voluminous commentary on Jewish law contains a number of explanations for this paradoxical rule: that the court has an obligation to seek *some* “merit” in the accused and to seek to preserve his/her life, and unanimity suggests that it failed to fulfill that obligation; or that the sin of someone against whom the evidence was so compelling was so great that forcing the accused to life under its burden was a harsher punishment than death itself. (A good discussion of this issue can be found [here](#).)

But a [fascinating paper](#) about to appear in the Proceedings of the Royal Academy A (Math & Physics) suggests that there may be some additional wisdom behind the rule. (The math, I admit, is a bit beyond me, but [there’s an excellent non-technical summary](#) by Lisa Zyga available here at [phys.org](#).) The authors (Lachlan Gunn of Australia’s University of Adelaide and colleagues) put forth a formal model to explain what they call the “paradox of unanimity” — situations in which unanimity among observers or decision-makers indicates, by a kind of “this is too good to be true” reasoning, that some sort of systematic bias, or other failure in the process, is responsible for the result.

The intuition behind this is related to the problem of the biased coin. Imagine a coin-flipping exercise, where you have some reason to believe that the coin *might* be rigged. If you flip the coin over and over and it keeps coming up “heads,” you start to have increasing confidence that it *is* rigged, increasing with each additional “agreement” (i.e., each flip that comes up heads).

Now consider the case of an “identity parade,” in which witnesses to a crime are asked to pick out the perpetrator from a lineup. Ordinarily, if the first, and the second, and the third . . . all pick out the same individual, we would view each identification as providing additional confirmation that the identification was correct.

What the authors show, however, is that these “increasing confirmatory identifications in a police lineup or identity parade can, under certain conditions, *reduce* our confidence that a perpetrator has been correctly identified.”

What conditions produce this peculiar result? There has to be some *a priori* probability — even a *very* small one — that the lineup process that was used is systematically biased, i.e., designed to produce a particular result, intentionally or not. In that case, each additional confirmation actually *decreases* the likelihood that the person identified was in fact the perpetrator (by increasing the likelihood that the system was rigged).

Imagine that as a court case drags on, witness after witness is called. Let us suppose thirteen witnesses have testified to having seen the defendant commit the crime. Witnesses may be notoriously unreliable, but the sheer magnitude of the testimony is apparently overwhelming. Anyone can make a misidentification but intuition tells us that, with each additional witness in agreement, the chance of them all being incorrect will approach zero. Thus one might naively believe that the weight of as many as thirteen unanimous confirmations leaves us beyond reasonable doubt.

However, this is not necessarily the case and more confirmations can surprisingly disimprove our confidence that the defendant has been correctly identified as the perpetrator. . . .

The numbers are pretty staggering. With even a really small probability — 1 in 10,000, say — that the lineups are systematically rigged against the suspect, the probability that the identified suspect is guilty starts to decrease with the fifth positive ID, and the probability of guilt is actually lower with 10 positive identifications than with three!

And if the probability of systematic bias is higher than that — a not-unreasonable one in a hundred, say — even three positive identifications becomes suspicious, and 10 positive identifications reduces the likelihood that the suspect was actually guilty to around 50/50. With that 1/100 chance that the system’s been rigged, it becomes literally impossible, *no matter how many eye witnesses agree*, to conclude that the probability of the suspect’s guilt is 95 percent or more (which is often used as the threshold for “reasonable doubt”).

It’s pretty disturbing, actually. The results they obtain do, it is true, assume that the bias in question — what they call the system’s “hidden failure state” — is a substantial one; in their calculations, they assume that the bias, if it is indeed present, causes witnesses to choose a particular suspect 90 percent of the time.

But still. As every television viewer in America surely knows, there are any number of ways in which lineups or other ID procedures can be tuned to systematically produce a particular result, and the notion that results of a “poll” can become less reliable as more and more people agree unanimously on the outcome, is a bit unsettling.

And the implications may be quite far-reaching.

The paradox of unanimity has many other applications beyond the legal arena. One important one that the researchers discuss in their paper is cryptography. Data is often encrypted by verifying that some gigantic number provided by an adversary is prime or composite. One way to do this is to repeat a probabilistic test called the Rabin-Miller test until the probability that it mistakes a composite as prime is extremely low: a probability of 2^{-128} is typically considered acceptable.

The systemic failure that occurs in this situation is computer failure. Most people never consider the possibility that a stray cosmic ray may flip a bit that in turn causes the test to accept a composite number as a prime. After all, the probability for such an event occurring is extremely low, approximately 10^{-13} per month. But the important thing is that it's greater than 2^{-128} , so even though the failure rate is so tiny, it dominates over the desired level of security. . . .

The recent Volkswagen scandal is [another] good example. The company fraudulently programmed a computer chip to run the engine in a mode that minimized diesel fuel emissions during emission tests. But in reality, the emissions did not meet standards when the cars were running on the road. The low emissions were too consistent and ‘too good to be true.’ The emissions team that outed Volkswagen initially got suspicious when they found that emissions were almost at the same level whether a car was new or five years old! The consistency betrayed the systemic bias introduced by the nefarious computer chip.

[And] a famous case where overwhelming evidence was ‘too good to be true’ occurred in the 1993-2008 period. Police in Europe found the same female DNA in about 15 crime scenes across France, Germany, and Austria. This mysterious killer was dubbed the Phantom of Heilbronn and the police never found her. The DNA evidence was consistent and overwhelming, yet it was wrong. It turned out to be a systemic error. The cotton swabs used to collect the DNA samples were accidentally contaminated, by the same lady, in the factory that made the swabs.

Food for thought for sure.

David G. Post taught intellectual property and Internet law at Temple and Georgetown Law Schools, and is the author of *In Search of Jefferson's Moose: Notes on the State of Cyberspace* (Oxford). He is currently is a Fellow at the Center for Democracy and Technology, and an Adjunct Scholar at the Cato Institute.