Transforming Justice With Knowledge

By Lawrence W. Sherman[1]

High Sheriff of London's Talk, 10th March, 2014

Fishmongers' Hall, London

High Sheriff: D.C.F. Jones

Moderator: The Rt. Hon. The Lord Judge PC QC

Amidst the drumbeat of daily scandals and criticisms of British justice, my Talk tonight is intended as a message of hope. My message is that prospects for transforming justice with knowledge have never been better.

It is worth noting, for example, that one our distinguished Moderator's predecessors as Lord Chief Justice faced some difficulty in persuading King James I that knowledge of the law was a necessary qualification for acting as a judge. Lord Chief Justice Sir Edward Coke felt compelled to tell King James I that even the King himself could not sit as a Judge without "cognizance" of the laws of England. This statement managed to keep King James off the bench—although it may have been the cause of Sir Edward's subsequent dismissal. At least this principle seems no longer in dispute.

The main reason for predicting that knowledge can transform justice in *our* time comes from the digital revolution. The ease of data collection and analysis has made new knowledge about justice much cheaper and faster to generate. What I mean by knowledge is thus a bit sharper than what Chief Justice Coke meant. My hope for justice is the same as Home Secretary R.A. Butler had when he encouraged Isaac Wolfson to endow an Institute of Criminology at Cambridge in 1959: the hope that justice can reduce HARM by the systematic application of SCIENCE to the prevention of crime. In the years since then, the capacity for science to reduce harm, including UNFAIRNESS, has been vastly multiplied by faster and cheaper computations.

Take the example of the 2011 (US) National Academy of Sciences' publication of a study of parole decisions by Israeli judges. This analysis of over 1,000 decisions to release or detain prisoners serving long sentences found that at the outset of each hearing session, the petitioners had a 65% chance of being released. The data then showed a steady decline in those odds over the next several hours of each session, with the last petitioner heard getting a near-zero chance of release. After each break for morning coffee or lunch, the odds shot right back up to 65%. There were no differences in seriousness or other case characteristics between earlier and later cases in each session. Nobel Prize-winner Daniel Kahneman interpreted the findings as evidence of "cognitive exhaustion" by the judges, and implied that prison release turned on a lottery of judicial mood.

No sooner had that study been published, however, than my fellow criminologists at the University of Maryland asked the Maryland Parole Board to allow a replication of the Israeli study. By the summer of 2012, we had completed a similar analysis of over 1,000 Maryland cases. The Maryland study found NO difference in the odds of petitioners winning parole based on their order of appearance. But it also revealed a key difference between Maryland and Israeli procedures. In Israel, the judges did not receive the convicts' case files until the hearing began. In

[1] Wolfson Professor of Criminology and Director, Institute of Criminology, University of Cambridge; Fellow, Darwin College, Cambridge; Distinguished University Professor, University of Maryland.

Maryland, the judges could read all the case files days in advance, and they had 300% more time allocated for each decision in Maryland than in Israel.

Kahneman may still be right: time-pressured decisions without preparation may quickly exhaust open-mindedness in justice. But replication in science is also right: the Maryland study helped to specify the conditions under which judges do or do not change their patterns of decision-making over the course of a day.

Computers have fostered even greater progress in other areas of justice. These areas can be generally described as *targeting, testing, and tracking* decisions to best reduce HARM in society. By "harm" I mean the damage humans cause to each other by breaking lawful rules, including rules of fairness for justice. For criminal conduct, our knowledge about harm reduction can be quickly enhanced by adopting the new Canadian system of measuring crime by an index of crime severity.

Since 2009, Canada has avoided using the US-UK method of counting crimes as if all crimes are created equal. Instead, Canada publishes a comparative *index* of crime severity over time and across locations, adding up the number of days in prison that would be assigned to each reported crime if all the criminals had been caught and sentenced. If we use that index here, for example, it shows that crime severity has indeed dropped since 2002, but by substantially less than the official crime count has dropped. Whatever defects this severity index—or what I propose to call the "Crime Harm Index"—may have, it is a vast improvement on a system that gives equal weight to a homicide and a failure to pay for a BBC licence.

Indeed, it was computerized knowledge that showed BBC licence cases reportedly comprise 12% of Magistrates' Court business. That knowledge may have helped mobilize the recent proposal to de-criminalize such cases—which brings us back to the three "Ts" of targeting, testing and tracking for harm reduction.

By knowledge for "Targeting," I mean systematic analysis of the distribution of harm across all the units of work for each justice system. For police, it means knowledge that 5% of street addresses produce over half of all crime, places that we now call "hot spots," which police are increasing targeting.

For prosecutors, we know how to precisely forecast the chances that each defendant in each case under review will be charged with murder or rape in the next two years, based upon algorithms from 100,000 PNC records we have analyzed at Cambridge. This might mean targeting higher priority for prosecuting dangerous offenders and lower priority to safer ones, as a metric for whether it is in the "public interest" to prosecute.

For Police and Crime Commissioners, targeting could be based on knowledge that restorative justice conferences offer far greater benefits to female victims of serious crime than to male victims, primarily because women suffer much higher levels of post-traumatic stress symptoms than men do from such crimes as burglary and robbery.

Such targeting is closely linked to "testing," which we can define as statistically controlled comparisons of two or more options in taking action against harm. For police, the testing now includes over 20 randomized controlled trials of increased police patrols in hot spots of crime, most of which show substantial reductions of crime resulting from small increases in patrol. Police testing also includes evidence from ten controlled trials that, on average, there are 27% fewer reconvictions of consenting offenders who are randomly assigned to attend police-led restorative justice conferences: meeting with consenting victims and their families to acknowledge the harm the offenders have caused, and where possible to repair that harm.

Knowledge has little benefit if it is ignored. That is why the third "T," for "tracking," is so important. For all our advances in medical knowledge, we still don't know how to get doctors to wash their hands after touching one patient and before touching another. This life-saving procedure has been part of medical knowledge for almost two centuries. The US estimates it could prevent 100,000 hospital deaths per year, if doctors would only do it.

Until new testing shows us how to change doctors' behaviour, we could target this cause of death by tracking the compliance of each physician. I do not know of any hospitals that have actually done this. But I do know of police agencies that are starting to track the amount of patrol time actually delivered to each crime hot spot. By using GPS tracking devices on police cars or radios, weekly reports on patrol delivery are now being fed back to police in Cambridge-led experiments.

The key question is whether knowledge from such Tracking will change police practices. The answer depends as much on leadership and other factors as it does on having the knowledge itself. Here is where we must ***ack***nowledge that knowledge alone isn't everything, and is hardly sufficient to improve justice.

Not sufficient...but absolutely necessary. No matter how skilled our leaders, they need knowledge to decide best where to lead us.

In the long run, knowledge can carry the seeds of its own flowering. In the short run, it may have the effect of a lighted match falling on wet tinder, a social context that either resists or ignores the new facts. But when the social context changes, a revival or rediscovery of useful knowledge can lead to rapid changes in policy.

This pattern is clear in the story of how British sailors came to be called "Limeys." A ship's doctor named James Lind treated sailors dying of scurvy on a voyage in 1747. He conducted the first recorded clinical trial, a small experiment with powerful results. Patients given a syrup of juices from limes and other citrus fruits made miraculous recoveries, while other treatments failed miserably. Lind wrote up his results in 1753. But his evidence ran against the prevailing theories of the day, and was ignored until 1794, the year Lind died. Only 40 years after the knowledge arrived was citrus juice issued to all sailors on a single naval vessel, and not until the Napoleonic Wars of the 1800s was juice put in general use throughout the Navy.

This story helps us to reflect on the difference between systematic knowledge, ignorance and incompetence. Incompetence can be defined as failure to use available knowledge, while ignorance can be described as not knowing what knowledge is available. Social knowledge—or ignorance—can be said to exist when at least some in the society do or do not know the answer to a question, such as how to prevent Alzheimer's Disease. Individual knowledge, by contrast, is the basis of all professions, as Lord Chief Justice Coke insisted, requiring that individuals master the recognized body of knowledge that is essential for competent practice—even if, as ship's Doctor James Lind found, professionally-accepted "knowledge" is wrong.

Knowledge can thus be defined as the opposite of ignorance—perhaps too easily, since the methods of science are always provisional. There are many things some people think we know—such as the human causes of global warming, or the existence of Angels in America—that others will hotly dispute. Thus to label as "incompetent" those who refuse to apply what others call "knowledge" may be to deny the never-ending nature of the discovery of new knowledge.

Yet what can we say about doctors who will not wash their hands? Are they sincerely doubting the evidence of the harm they cause? Or can we agree that after 170 years of consistent evidence, they have no reasonable basis for disagreement—that they are, in fact, incompetent?

This question is the crucible of what is now called evidence-based policy. When governments make policy by intuitive feelings that run against clear evidence, they may be acting normally—especially where the evidence is new, shocking, and runs against prevailing theory. That was the story of scurvy and James Lind.

My own crucible is about arrests for domestic common assault, for which my first experiment in Minneapolis became the basis for international claims to "evidence-based policy." In the first controlled experiment in the use of arrest, Police Chief Anthony Bouza and I found that arrest caused the lowest rate of repeat assaults over the next six months in comparison to a police warning or police sending the suspect out of the premises. The effect of this knowledge was like a match falling on the tinder of accumulated frustration that nothing was being done about widespread domestic assault. Most US states immediately made arrest mandatory after our experiment was reported, even though we concluded that more tests were needed before we had sound knowledge.

By 1992, four more independent—and larger--experiments had found that the short-term effects of arrest varied substantially by whether the suspect had a job. Arrest worked for suspects who had a job to lose. For those who didn't, arrest caused twice as much repeat violence in three of the four tests. Last week, as the TIMES reported, we released a 23-year followup of one of those studies. The evidence was clear, statistically significant, and shocking. Victims in Milwaukee had been more likely to die of *natural* causes if the offenders had been *arrested* than if they had been *warned*. Black victims suffered the most from arrest, with almost a 100% increase in death rates from the random assignment of arrest to their partners. "Collateral" harm to victims from the processes of justice have never had such clear empirical evidence.

So far, this new finding has been largely ignored, in both the US and UK. It is hardly a match falling in a dry forest, especially without a demonstrably "positive" alternative to arrest. But many such alternatives are already being tested. The prospects for the restorative justice conferences, for example, working well in domestic violence cases look even more favourable today than they did a decade ago. Professor Linda Mills at NYU, who is herself a victim of domestic violence, has tested such conferences in the US on a court-ordered basis, with encouraging results. But with 100,000 arrests per year in the UK for these crimes, most of which receive immediate release and no further action, there is good reason to focus new testing on police decisions at the scene.

More than one English police agency is now interested in *targeting* domestic assault for further *testing* of arrest versus alternatives, while *tracking* the results in different kinds of cases and communities. We should do no less to develop new knowledge about other kinds of crime, such as cyber-crime and financial crimes, that cause massive harm with low visibility. I am especially pleased that the Cambridge Institute of Criminology is close to creating a new Centre on Economic Crime that would help target patterns and networks of offending and test new ways to prevent and punish the offences.

For these and other reasons, there is much evidence for hope about transforming justice with knowledge. Whether welcome or not on first arrival, new knowledge can help to improve justice in the long run. At the very least, such systematic analysis reminds us of the problems causing the greatest true harm, as opposed to the greatest outrage in each day's list of scandals. Systematic knowledge can keep institutions of justice focused on the greatest opportunities for reducing harm, whatever the daily headlines. In the long run, such may create much clearer picture of how we can reduce harm without creating collateral harm, or at least creating less of it. That would a transformation devoutly to be wished.

For further reading see links below.

Parole decisions in Israel: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3084045/

Parole decisions in Maryland: Forthcoming University of Maryland study by Darren G. Fisher et al

Targeting, testing, tracking and the Crime Harm Index: http://cebcp.org/wp-content/evidence-based-policing/Sherman-TripleT.pdf

James Lind and the story of scurvy treatment http://en.wikipedia.org/wiki/James_Lind

Hot spots policing http://www.campbellcollaboration.org/lib/project/24/

Restorative Justice conferences http://www.campbellcollaboration.org/lib/project/63/

Forecasting Homicide
http://www.crim.cam.ac.uk/people/academic_research/lawrence_sherman/mjrss2008.pdf

Victim deaths from domestic violence arrests
http://www.crim.cam.ac.uk/news/documents/MilDVE%20Victim%20Mortality%20JEC%20FINAL%20ALL.pdf

http://time.com/12682/when-not-to-arrest-an-abuser-in-a-domestic-violence-case/