Do Toronto Police Engage in Racial Profiling?

Ron Melchers
Department of Criminology
University of Ottawa

On 19 October 2002 the *Toronto Star* began a series of articles "Race and Crime,"[1] making claims that "justice is different for blacks and whites" ("Singled out" 2002), "Blacks arrested by Toronto police are treated more harshly than whites" ("Singled out" 2002), and that "Police target black drivers" ("Police target" 2002: A1). Subsequent stories suggested that Toronto police were engaging in *racial profiling*, defined by the *Star* as "the practice of stopping people for little reason other than their skin colour" ("Police target" 2002: A8). Published interviews with black community leaders and advocates added the weight of anecdote and presumption to charges of racial profiling, and University of Toronto criminologist Scott Wortley, whose research has focused on race issues in criminal justice, deemed the *Star* analysis "clear evidence of what, until now, has been based largely on assumption" ("Singled out" 2002: A13).

Representatives of the police responded angrily, denying accusations of singling out blacks. The Toronto Police Service commissioned an independent review of the *Star*'s analysis by a prominent criminal lawyer (Alan Gold) and a University of Toronto sociology professor (Edward Harvey) (see Harvey and Liu 2003; Harvey 2003; Gold and Harvey 2003). Their review concluded that the *Star* analysis was "junk science" and the conclusions of the articles "completely unjustified, irresponsible and bogus slurs" to be "put down at once" (Gold and Harvey 2003). The police union went further and on 17 January 2003 launched a $2.7 billion class action libel suit on behalf of its 7,200 members ("Police union" 2003). The furor has since spread throughout the Ontario criminal justice system, with judges, attorneys, crown prosecutors, and police officials making additional controversial statements supporting or refuting the allegations of racial profiling in the criminal justice system. The media have given considerable attention to the issue, as befits its renewed prominence in public debate, but perhaps also in defence of one of their own. The debate is likely to be given further play in municipal, provincial, and federal election campaigns.

Whatever else may come out of these events, it is already clear that the Toronto Police Service has now joined the ranks of North American police organizations that must now devote considerable effort to addressing accusations that they engage in racial profiling. The development of data collection and analysis systems to respond to accusations of racial bias in policing has become something of a growth industry in the U.S. (e.g., Ramirez, McDevitt, and Farrell 2000; McMahon, Garner, Davis, and Kraus 2002; Fridell, Lunney, Diamond, and Kubu 2001: 118; Engel, Calnon, Bernard 2002: 250, 262–263). Widespread public belief that police engage in racial profiling undermines public confidence in the police, as well as the credibility of the testimony and evidence submitted by police officers in criminal proceedings. It may now have become difficult to prosecute criminal cases involving accused persons from groups claiming police bias, without extensive expert testimony on the issue of racial profiling. In Toronto (as elsewhere across North America), the issue of racial profiling has become a significant threat to the ability of police to maintain order, ensure public safety, and prosecute those accused of criminal offences. This issue also affects the ability of the courts to weigh evidence appropriately, assess criminal responsibility, and determine appropriate sanctions.

The evidence for claims of racial profiling

What evidence is there that Toronto Police engage in what is called *racial profiling*? The claims made by the *Toronto Star* were based on the newspaper's own analysis of arrest data from the Toronto police's Criminal Information Processing System (CIPS), obtained under a freedom of information request. The data were recorded between late 1996 (when CIPS was first implemented on a trial basis) and early 2002.

The *Toronto Star*'s investigative team, under the supervision of Dr. Michael Friendly, professor of psychology and director of consulting services for York University's Institute for Social Research, worked with a database consisting of 483,614 incidents in which someone had been arrested, charged, or ticketed. These incidents resulted in more than 800,000 charges being laid under criminal and other statutes or by-laws. Of these charges, 301,551 were for *Criminal Code* or drug offences.

No individual identifying information was included in the information obtained by the *Star* under the freedom of information request and offences were aggregated into broad categories to preserve individ-

ual's privacy. As well as incident, charge, and police disposition detail, information was also provided about the age, gender, skin colour (white, black, brown, other), immigration and residency status in Canada, employment information, and country of birth of those arrested, charged, or ticketed. The data also provided some information on the criminal histories of individuals arrested: previous convictions, bail status, probation orders, or conditional release status. Complete information, the *Star* reported, was not available in every instance, and the analysis excluded those cases where specific information was missing. As with all police-recorded data, the more serious offences would tend to be over-represented among completed records and the least serious offences under-represented (Melchers 2001).

CIPS records only incidents or police actions in which a person is arrested, charged, or ticketed. The purpose of CIPS is to manage information that is likely to be required at subsequent stages in criminal justice processing. Therefore, no information is reported for an incident where there is no subsequent action taken and information is often incomplete, especially in cases where there is little expectation of follow-up. Depending on the type of offence, incidents reported may be only a small portion of the total number of instances in which the police intervened. CIPS coverage is, furthermore, not consistent throughout this period, as the system was being progressively implemented, initially on a voluntary basis.

The *Star* analysis focused its attention, although not exclusively,[2] on two categories of offences: arrests, charges, or ticketing for *out-of-sight* driver offences (such as driving without a valid licence or without insurance) arising from vehicle stops; and police dispositions of persons charged with single counts of simple drug possession offences. These were chosen because they offered two of the more commonly suspected, if not also more likely, opportunities to observe racial bias in policing. The belief that police stop black drivers for reasons of skin colour alone (known as DWB "Driving While Black" in some circles) is widespread.[3] Drug offending is a stereotypical "black crime" in the minds of some, an idea reinforced by local media portrayals of "Jamaican posses." Both situations further rely on considerable officer discretion in enforcement.

Traffic stops

Traffic stops were chosen, following the lead of U.S. studies according to the *Star*, because of the high degree of discretion (therefore potential

for discrimination) in most police decisions to stop vehicles and because police would have no prima facie indication of these infractions at the time of their decision to stop a vehicle. The researchers assumed that the racial distribution of these charges would be representative of the racial distribution of all vehicle stops and also that any differences in this distribution from that of the population would be evidence of bias or racial profiling.

The *Star* reported that of 4,696 out-of-sight offences reported over the five-year period for which skin colour was recorded for purposes of identification, 33.6% involved drivers described as "black." The *Star* wrote, "It's assumed random checks would generate a pattern of charges that mimics the racial distribution of drivers in society as a whole," citing U.S. studies that consider these figures a "bellwether for racial profiling" ("Police target" 2002: A1). Skin colour is not identified on driver's licences in Ontario, nor is any record of kilometres driven available for registered vehicle owners by skin colour. So the *Star* used as a proxy the proportion (8.1%) of the Toronto population who reported themselves as black on the 1996 Census forms. The difference between these two proportions was considered by the *Star* to be evidence of racial profiling.

The use of population data in studies of racial profiling

Contrary to the assurances of the *Star*, this method is not well supported in the literature – quite the contrary in fact. For example, McMahon et al. concluded from their review of racial profiling research that studies of racial bias in traffic stops "too often base their conclusions on comparing preliminary data on traffic stops to aggregate city demographics without establishing credible benchmarks for comparison purposes. These superficial evaluations are dangerous, in that they may foster incorrect conclusions and generate inappropriate corrective measures" (McMahon et al. 2002: 1). There are two problems with the assumption that proportions of drivers stopped by police should be identical to proportions within the population. The first relates to the use of population data and the second to the assumption of randomness in police vehicle stops.

Incidence versus prevalence

The reference to population data results in two errors. First, data on traffic stops and population data belong to two very different categories of statistics and cannot be combined or compared without intro-

ducing considerable distortions, which compromise the ability to draw accurate conclusions. Counts of traffic stops measure *incidence*. They are counts of events, not individuals. Indeed, a single individual may be involved in more than one event over any defined period. This is more likely, the longer the period covered. Thus, the proportion of stops that involve drivers with any given characteristic (incidence) cannot be assumed to be the same thing as the percentage of drivers stopped who have that characteristic (known as *prevalence*). Population statistics are only an appropriate base for statistics that measure prevalence. By comparing incidence to population, one inevitably creates the false impression that any group with some number of members who are stopped frequently is over-represented as a whole. When the nominator and the base in a rate do not have the same units of count, or when the units of counts are insufficiently interrelated, this is called *base error*.

The impact of repeat offenders (and also repeat victims) on aggregate crime counts receives a great deal of attention from specialists. Crime and victimization are highly concentrated phenomena. Indeed, much research suggests that a much smaller portion of the population experiences crime than the reported numbers of offences or victimizations divided by the total population would suggest.[4] A recent U.S. study found 10% of offenders involved in more than 50% of crimes (Eck, Gersh, and Taylor 2000). Fully one half of incidents reported by the Canadian General Social Survey in 1988 were repeat victimizations (Sacco and Johnson 1990). A recent Statistics Canada study of recidivism among convicted youth and young adults between the ages of 18 and 25 found that 60% of offenders convicted in 1999/2000 had prior convictions, 43% multiple prior convictions (Thomas, Hurley, and Grimes 2002). This can often result in very large errors in interpretation. These sorts of errors are common when incidence statistics are used to infer prevalence. It can result in a small but very active group having an inordinate impact on how a more diverse larger group encompassing them is perceived. For example, a single address in a street block to which police are frequently called may result in an otherwise peaceful and law-abiding neighbourhood's being branded as "crime ridden." This is called *aggregation error*.

To illustrate how these errors together may skew interpretations, consider an arbitrary group of 10 people. One of them commits a crime and is charged 10 times for a criminal offence, over a set period of time. Ten percent of the group has therefore been charged with a criminal offence (*prevalence*). Yet simply placing the number 10 *(incidence)* next

to the total number of people in the group as a base, without any information about repeats, one could just as easily conclude, as for any number from 1 to 10, that all 10 members of the group (100%) had been charged with an offence. One might then be tempted to conclude spuriously that whatever common trait the group shared was in some way related to the "fact" that such a seemingly large proportion of the group appeared to have engaged in crime. The nominator in this illustration is not logically connected to the base and produces a false and highly misleading aggregate rate. Comparing the number of incidents in which a charged person's skin colour is black with the total population sharing this characteristic, even if such a statistic were available and reliable, is an example of aggregation and base errors.

The measurement of racial identity

The second error that arises in the use of population data is one of measurement. *Black* identity as measured in the Census of the population is not the same thing as *black skin colour* as reported by a police officer for identification purposes.[5] The measurement of collective identity is not as simple a matter as one might think and it has long been one of the most challenging measurement issues for demographers and census designers. Some countries have historically used *blood-quantum* definitions of racially defined groups, a measurement subject to a host of reliability problems. In Canada, skin colour was only included in the Census of the population for the first time in 1996, in response to the needs created by equity legislation. Before asking about membership in "equity groups," the Census long form, completed by one-fifth of respondents, first of all asks respondents to identify their ancestry, defined by the question as ancestral "ethnic or cultural group(s)." Among the options on the list provided for illustration, we find examples of groups defined by territorial origin, residence, national citizenship, language, bloodline, faith, ethnicity, and culture. "African," for instance, is not included in the illustrative list compiled from the most frequently reported answers. Respondents who do not report Aboriginal ancestry are then asked to complete a further question. Only two of the responses refer to skin colour: "white," which is the default for elimination from the count of visible minorities, and "black," the only colour among a list of what are otherwise visible-minority groups defined by geographical origin. In 1996 8.1% of respondents for the City of Toronto were estimated as black. In 2001, respondents to the 20% Census sample identifying themselves as black were by inference estimated to represent 204,075 individuals, making up 8.3% of the total population. Judging by definitions for the

17 To which ethnic or cultural group(s) did this person's **ancestors** belong?

For example, Canadian, French, English, Chinese, Italian, German, Scottish, Irish, Cree, Micmac, Métis, Inuit (Eskimo), East Indian, Ukrainian, Dutch, Polish, Portuguese, Filipino, Jewish, Greek, Jamaican, Vietnamese, Lebanese, Chilean, Somali, etc.

▶

Specify as many groups as applicable

19 [   ]

20 [   ]

21 [   ]

22 [   ]

Figure 1

19 Is this person:

Mark "⊗" more than one or specify, if applicable.

This information is collected to support programs that promote equal opportunity for everyone to share in the social, cultural and economic life of Canada.

05 ◯ White
06 ◯ Chinese
07 ◯ South Asian (e.g., East Indian, Pakistani, Sri Lankan, etc.)
08 ◯ Black
09 ◯ Filipino
10 ◯ Latin American
11 ◯ Southeast Asian (e.g., Cambodian, Indonesian, Laotian, Vietnamese, etc.)
12 ◯ Arab
13 ◯ West Asian (e.g., Afghan, Iranian, etc.)
14 ◯ Japanese
15 ◯ Korean
     Other — Specify
16 [   ]

Figure 2

Census sample, people subjectively reporting black identity could clearly be either a much smaller or a much larger number than the number of people who might be objectively described as having black skin colour. The two definitions are essentially unrelated. McMahon (2002) also expressed concern about the accuracy of Census data, especially as regards under-coverage of specific populations, such as non-documented immigrants, as well as minority populations in urban areas in general.

Is traffic policing conducted randomly?

Another assumption in the *Star* analysis that bears examination is that of the randomness of reported vehicle stops. Much research into discrimination in policing practices comes from the analysis of searches following highway traffic stops. Unlike urban traffic policing, highway traffic stops of speeding vehicles are essentially random. The distribution of drivers according to visible traits, particularly for a limited-access toll highway, can be observed, and most studies have found

that stops themselves follow that distribution closely. However, searches of stopped vehicles have been observed in several U.S. jurisdictions to be systematically more frequent when the driver is perceived to be black.[6]

The likelihood of any driver being stopped in an open urban traffic situation is never random. First of all, it is unreasonable to assume that either vehicle ownership or driving patterns are identical among all groups of drivers. If police stops were random, then the number of kilometres driven would be considered the most appropriate measure of the potential of being stopped by police. But urban vehicle stops are not random. Entirely random motorized patrol deployment and stopping of vehicles at random would be wasteful of police resources. The purpose of traffic policing is to ensure the safe and orderly flow of traffic, and police patrol is primarily preventative. Patrols are most effectively deployed when they focus on when and where problems are most expected. For example, Smith and Petrocelli (2001: 4–27), in their study of traffic stops, found that the level of crime and hence deployment of police resources in a neighbourhood was a good predictor of police decisions to stop.

Obviously, increased vigilance at certain times and places will result in an increased likelihood of those present being stopped and subsequently over-represented in the data. But these are broad systemic biases unrelated in any direct way to skin colour, not evidence of unlawful discrimination in discretionary decision making. Therefore, rather than being distributed randomly, the likelihood of any individual driver being stopped by police, even if we exclude for just a moment non-random factors in police decision making, is a factor of both the number of kilometres driven in the city by that driver and the pattern in space and time of police deployment. Together these define the universe of opportunities for police stops of vehicles.

Even within this universe, police do not stop vehicles randomly. Rather, they stop specific vehicles for a variety of legally relevant reasons, on the basis of observation, information, and reasoned judgment. These may be the consequence of formally prescribed practices or policies structuring the use of discretion. In less well-ordered police organizations they may more often be the product of ad hoc experience. Some stops are low discretion, for example of vehicles running a red light, excessively speeding, or identified by police as stolen or having been involved in other offences. Other stops are high discretion stops. Even in these cases, however, the age and condition of the vehicle (for

example malfunctioning lights or under-inflated tires), the day and time, the location, in some circumstances the age of the driver, not to mention various driving behaviours (for example, indication of improper seatbelt use, unusual driving patterns, or minor traffic violations) may all legitimately result in vehicles receiving greater or lesser attention from police officers. This would in turn influence the detection, after the stop has been made, of other out-of-sight infractions.

Nor is the police decision to lay or not lay charges random, and it is unreasonable to assume so in the analysis of data arising from charges. In the case of some of the infractions retained for examination as out-of-sight offences, charging is mandatory. The decision to lay charges, where it is not mandatory to do so, depends on a number of factors, such as, for example, previous charges on record. It can also depend upon driver behaviours judged by the officer to be aggressive or otherwise suspicious. In other cases, prompt pleading or an uncontested finding of guilt may result in incomplete records. But the information available from the CIPS data obtained by the *Star* makes it impossible to parse such factors out one from the other. Without information on the most significant legally relevant factors involved, no reliable conclusion can be made about the presence or relative importance of non-legally relevant factors, such as skin colour. While the existence of a variety of non-legally relevant factors in police decision making is certainly possible – even plausible, in the minds of many careful observers – the information presented by the *Toronto Star* cannot be taken as either evidence of or a measure of the existence of such factors.

Finally, in addition to all of these problems in reasoning about the data, there are further limitations arising from potential *sample bias* in the data used by the *Star* in its analysis. The 4,696 complete records comprise 63% of the total 7,511 incidents reported for the specific infractions retained over a five-year period. For 2,815 of the total number recorded, infractions of skin colour was not recorded. In order to infer the skin colour distribution of incidents for total number of charges for out-of-sight infractions from that of those incidents for which skin colour was recorded, it was assumed that the smaller group was a random subset of the large one; that is, that no systematic factors intervened in the omission of skin colour from some records.

This assumption cannot be reasonably made. We should expect that vigilance in reporting all details might vary according to the expectation of the officer that the information will subsequently be required for identification purposes. Many legally relevant factors in evidence

at the time of the stop might influence this expectation and the consequent thoroughness of the record, or even whether or not a record is made. We would expect to find systematic differences that would limit the ability to infer from the smaller number to the larger. In a presentation to the Toronto Police Services Board, the *Star* analysts and Professor Friendly pointed out that even if all retained cases where skin colour went unreported were assumed to be white, the 34% figure would only be reduced to 21%, which Friendly still considers evidence of racial profiling.

Notwithstanding the less than scientific nature of this assurance, not even the larger number of 7,511 infractions can be assumed to be a representative or actual total of the number of incidents of all these specific types of violations investigated by police. This is still a very small number, a daily average of just four vehicle stops, for a city of nearly 2.5 million residents, when compared with the total number, certainly well into the many hundreds of thousands, of vehicle stops made by police in a region the size of Toronto over a five-year period.[7] This larger number is unknown, as police are not required to report each and every vehicle stop. One could only infer from this smaller, nested number to the total of all vehicle stops if it were a random subset of the larger whole. This is not a reasonable assumption. At the very least, statements such as "the observed differences suggest police use racial profiling in deciding whom to pull over" or that police are "stopping people for little reason other than their skin colour" ("Singled out" 2002: A1) are overly ambitious in the extreme as interpretations of the evidence examined.

Simple drug possession dispositions

The central argument made with respect to the choice of drug possession as an offence category is not so much (as it is with regard to vehicle stops) that the difference between this number and the proportion of blacks in the population demonstrates differences in police vigilance in enforcing drug laws. Rather, it is the use of police discretion both at the scene and at the station that is seen to reflect discriminatory treatment. Officers may decide to release a suspect at the scene, to take a suspect to a station for booking before being released, or to hold a suspect for a bail hearing. It was assumed that any differences between the skin colour distribution of arrests and that of subsequent police dispositions would be evidence of bias or racial profiling.

The *Toronto Star* investigative team examined 10,729 police disposi-

tions of persons charged. Because of the high likelihood of follow-up and thus of identifying information's being subsequently required, 93.8% of records were complete. According to published reports from the newspaper, 23.6% of incidents involved persons whose skin colour was identified as black and 63.8% as white. White-skinned suspects were released at the scene in 76.5% of cases, whereas those identified as black, only 61.8% of the time. The *Star* also reported that 15.5% of blacks were held for a bail hearing, as compared with 7.3% of whites. This, the *Star* concludes, is evidence that police treat blacks more harshly.

The *Star* article claimed to have taken "into account a number of factors that might influence police decision making, including a suspect's age, criminal history, employment, immigration status, and whether or not the person had a home address." The authors stated that only "police history" had some impact, but that "the difference in treatment between blacks and whites remained" ("Singled out" 2002: A1). But neither the manner in which this analysis was conducted nor the detailed results are reported in the newspaper articles so that they might be examined. The series of articles reported only descriptive statistics, for the entire city and also by patrol divisions, and some breakdowns by age, type of drug, and police history, though only single-variable breakdowns. Descriptive information is provided only for the total number of persons arrested: gender; the proportion with previous criminal convictions (50.3%), out on bail (23.5%), on probation (17.6%), or *unescorted temporary absence* (parole) from a correctional institution (1.6%) at time of arrest; and the percentage subsequently charged with a major violent offence (14.6%). None of these factors is further broken down by race. Only the number of charges laid is broken out by race. It appears from what is reported that only single-variable descriptive statistics were examined.

It is impossible to assess this part of the work conducted by the *Star* analysts on the basis of the published accounts. To determine whether and how much of the difference in treatment between black and white suspects might or might not be explained by any single factor, multivariate analysis would be required. In multivariate analysis, each factor is weighed to determine its specific influence, controlling at the same time for the influence of all other factors in a given model. In this case, where the information available is binomial, multinomial, or expressed as constrained proportions, the most appropriate procedure would be log linear models or logistic regression. The results of these procedures, expressed as relations among different statistical probabil-

ities, are notoriously difficult to interpret and to translate into plain language explanations. There is no evidence in the newspaper accounts of the analysis that this was done.

However, in his presentation on behalf of the *Star* before the Toronto Police Services Board, Michael Friendly did present such an analysis. Friendly presented two models, one examining the likelihood of release at the scene and another examining the decision to hold for a bail hearing by five independent factors: gender (M or F), employment status (employed or not), citizenship (Canadian or not), age (young or not), and across years. Friendly concluded that employment status, bail status, a variable listed as "CPIC" (a national database of criminal records) but not precisely defined in the notes of the presentation, and citizenship were among the variables, with skin colour, that had the most effect on the decision to release at the scene. Controlling for these factors, Friendly reported that the likelihood of blacks being taken to the station was still between 1.3 and 1.7 times greater than that of whites. Friendly also reported that blacks were between 1.3 and 1.9 times more likely than whites to be held for a bail hearing. However, it is difficult from the documents of that presentation alone to fully grasp the exact definitions of the variables used and the details of the analyses that were conducted. Were this work to be submitted for scholarly publication, it could be better assessed.

However, even using appropriate statistical procedures it would be impossible to eliminate all possibility of having committed the error of attributing the observed differences to statistically significant yet wrong factors, known as *spuriousness.* To illustrate the idea of spuriousness, consider someone with the firm belief, based on careful observation, that the carrying of umbrellas by transit passengers is the cause of rain. Even a casual observer would be compelled to admit that rain occurs more frequently on days that umbrellas are most in evidence. The statistics would be irrefutable as long as one's knowledge was limited only to that gained from the observation of rain and umbrellas.

The problem here is of course that there are other unobserved influences at work that cause both rain and the carrying of umbrellas. Once the influences of these factors are recognized, understood, and observed, the original hypothesis may then be re-examined, tested, and perhaps understood to have been incorrect or *spurious*. Thus, even a statistically significant and strong relationship or model may ultimately turn out to be spurious, as the researcher gains more knowledge of the underlying forces behind a phenomenon and becomes

capable of more complex thinking about how these all interrelate. There are many more factors than those used, principally as a result of the convenience of being more or less available, by the *Star* analysis of police arrest decisions. Mere availability is always the poorest of reasons for including a factor within an explanatory model.

This limitation is what prompts scientists to make only modest claims for explanations their observations enable them to exclude. Occasionally stilted and wooden, the language of scientific reasoning requires a modesty that may frustrate many eager to leap to conclusions but that nonetheless best serves careful public consideration of complex and compelling issues.

Does racial profiling actually exist?

Polls and some studies consistently show that a majority of the public, and yet a larger majority of some visible minority groups (notably "Blacks"), believe that police are racially biased. The term *profiling* has grown in popular usage throughout the last decade to describe more or less formal practices of police and other public officials for singling out individuals by specific traits for investigation. Popular media and the entertainment industry were initially responsible for the success of the expression and for its entry onto the political stage.

The use of the expression *racial profiling* is for rhetorical value alone. It attempts to redefine racial discrimination as more than individual bias, or even as improperly tolerated individual wrongdoing in an organization, but rather as the official policy and sanctioned practice of organizations. Drawing upon myths of the news and entertainment media, the term *racial profiling* dramatizes public discourse on the issue of racial bias and discrimination. In June 1999 U.S. President Clinton officially consecrated the expression, calling it a "morally indefensible, deeply corrosive practice" and further stating that "racial profiling is in fact the opposite of good police work, where actions are based on hard facts, not stereotypes. It is wrong, it is destructive, and it must stop" (U.S. Department of Justice 1999: 22–23).

Today many police and other public service organizations routinely monitor the skin colour and ethnic backgrounds of persons stopped, questioned, or investigated to determine whether their personnel are acting in a discriminatory manner. Racial Profiling Data Collection systems have become a common feature of public administration, to respond to widespread public belief that such practices exist and to

attempt to maintain or restore public confidence. Do police and other enforcement agencies in fact engage in racial or other sorts of offender profiling? As we have seen in the preceding discussion, the assessment of available evidence to answer this question is rarely as simple a matter as many might initially think. Another question that might be asked is whether various profiling practices could ever be justified by their expected and observed results. Given the attention the issue has received, the lack of any scientific assessment of any form of profiling is surprising.

Canter and Alison (as cited in Ainsworth 2001: 109–115) note, from an examination of various offender psychological profiles, that "... a careful examination of the content of their profiles reveals a severe lack in the accounts of any systematic procedures or any substantive, theoretical models of behaviour. There is no reference to any commonly accepted psychological principles – pathological or social." Muller (as cited in Ainsworth 2001: 109–115), in an examination of profiles of serial killers, the most widely popularized application of profiling, deplores that "there have never been any published empirical studies on the difference between various subtypes of serial offenders." Muller goes on, "one of the biggest hurdles standing in the way of acceptance of criminal profiling is that there is very little authoritative material on it, and almost nothing in the way of scientific studies to support the claims of the profilers." Peter Ainsworth concludes from his review of scientific assessment of offender profiling "that the early attempts at profiling may have captured the public's imagination but that they may not be based upon a scientific bedrock of data-gathering or empirical research" (Ainsworth 2001).

A few illustrations suffice to demonstrate that most of what we call "profiling" is ill-founded. Consider a hit-and-run collision involving a taxicab and an eyewitness account. The witness describes the cab as green in colour. Only 25% of the 4,000 licensed taxicabs in the city are green, so this narrows the investigation considerably. Any eyewitness account is prone to error. In this case let us say we can set the likelihood of witness error at 20%, a small error rate, given what research on the topic has indicated. What is the likelihood that any single green cab pulled over by police in the course of the investigation was the one seen fleeing the scene of the accident? Should police devote their investigative efforts to checking the logs of every green cab in the city?

The answer to the first question is one in 1,250 (4,000*0.25/0.8). There are 1,000 green cabs to be pulled over, but the probability that the cab

was actually green is only 0.8, because of the possibility of witness error. The answer to the second question should be obvious: it would not be a good use of police effort to investigate all green cabs. If the error factor were higher than 50%, as witness reliability often is, such profiling would more often steer police in wrong than in right directions. The point of this illustration is to point out that error rates multiply when additional factors are brought into any prediction. Statistical likelihood would be a very poor alternative policing strategy to traditional evidence gathering, as many police investigators have discovered at their expense.

Another problem creeps in when prevalence rates are low, as is the case with most criminal offending. Take another example: assume a very hypothetical (and clearly ludicrous) birth test for delinquency, with an accuracy of 95%. However, delinquency is not a frequently occurring phenomenon in a population. Let us say that 3% of a population at birth will become delinquent. Out of 1,000 times, the test will correctly identify future delinquents only 29 times ($0.03 \times 0.95 = 0.029$). At the same time, it will fail to identify future delinquents 49 times ($0.97 \times 0.05 = 0.049$) and will furthermore incorrectly identify two infants as future delinquents ($0.03 \times 0.05 = 0.002$). The total number of incorrect predictions (51:1,000) exceeds that of correct predictions of delinquency (29:1,000) by nearly two to one. Balanced against its ineffectiveness and injustice, the contribution of such a test to early identification of delinquency and to public safety would be questionable at the very least. Yet the sorts of predictions made by offender profilers have, albeit unknown, predictably far higher error rates and may often involve prevalence rates as low as one to a base of total population in the case of the identification of a single suspect. It is simply implausible that actual profiling policies or practices on any basis, be it psychological or racial, would ever be officially adopted by any rationally behaving organization as an alternative to traditional, evidence-gathering, investigative practices.

How is it, then, that so many observers are confident that police and other investigative organizations have adopted official policies and practices of racial profiling? First, this may be because so many organizations have indeed done so, despite overwhelming indications of the folly of such approaches. One need only think of current U.S. border controls that target all travellers of Middle East and Central Asian origin. There are also carefully conducted, scientifically credible studies that offer clear evidence that at least some members of some police organizations have also done so (Lamberth 1999).[8] Scientific ground-

ing has never been a sine qua non requirement in public policy nor individual discretion.

Furthermore, whether or not racial profiling is officially sanctioned in an organization, the statistical results of police operations often look the same. That is because so much of what explains the over-representation of some groups in the criminal justice system is so systematically and seemingly intractably integrated into the functioning of all social institutions and behaviours: poverty, deprivation, isolation from broader social values, lack of community social cohesion, and so on. At least some groups of Canadians – for example those of Aboriginal or African ancestry – have experienced and continue to experience discrimination in so many ways that it would be unreasonable for the consequences not to show up in virtually every aspect of Canadian society.

Finally, the same sort of feedback loop of self-fulfilling prophecy described in the *Star* articles by Scott Wortley to explain how police may themselves come to adopt discriminatory beliefs may explain how beliefs about racial profiling emerge and appear increasingly founded in a cycle of ever-heightening apprehension. Let me paraphrase his illustration so as to demonstrate an equally plausible yet opposite conclusion. Consider a situation in which drivers are indeed stopped randomly. Drivers stopped who are convinced, incorrectly in this hypothetical example, that they will be stopped because of their skin colour will conclude their apprehension to have been borne out, further reinforcing their belief that skin colour is the cause. Those not sharing this belief, whether correctly or incorrectly, will attribute the event to randomness or other causes. The latter may not view their individual experience as a significant aspect of a shared skin-colour experience or identity and may consider it unworthy of recalling or repeating in conversation. On the other hand, the greater simplicity and appeal of the belief that skin colour is the singular cause of police stops guarantees its iterative value, so that, when placed in competition with a belief in randomness, it will become the dominant shared collective view, a sociological rendering of Gresham's Law, "bad currency drives out good." We seem culturally – perhaps because of widespread innumeracy – indisposed to accept randomness. The popularity of various forms of gambling is evidence of this.

In conclusion, while it is highly plausible that, once all legally relevant factors have been accounted for, differences in the treatment of groups

according to race will remain, even this in and of itself may not be evidence of actual discriminatory practices, as opposed to any number of equally valid explanations of these differences. The best research can conclude in such cases is the modest statement that the possibility of discrimination cannot be excluded. In the absence of compelling evidence, to make any more ambitious statement goes against the scientific ethic.

Notes

1. The complete series of articles, accompanied by subsequent additional material, commentary, and news items can be found on the Web site of the *Toronto Star*: http://thestar.ca/

2. The *Toronto Star* also looked at cocaine possession and violent offences, as well as a number of other issues. However, in these cases, it is less clear from published reports what precise definitions and procedures were used. It is therefore difficult to review these in any detail.

3. For example, McMahon et al. report, "Based on a Washington Post survey, 52% of African-American males believe they have been victims of racial profiling, while a Gallup poll indicates that about 60% of Americans believe racial profiling exists" (2002: 17).

4. See, *inter alia*, Robert 2001, who reports that 9% of French victims surveyed in 1996 reported two-thirds of assaults and almost the totality of personal thefts.

5. See Roberts (1994) for further discussion of this point.

6. One of the earliest and best known of these studies is that of John Lamberth (1994) for the ACLU.

7. Information reported by McMahon et al. from urban traffic stops in Baltimore, a less densely policed city than Toronto, suggested that the number of stops in Toronto could be expected to be close to 500,000 annually (2002: 33ff). The same study suggested that only 1% of vehicles stopped are searched and highlights the difficulties in drawing conclusions from such nested, multilevel, non-random samples.

8. For a discussion of the Lamberth study, see Harris 2002.

References

Ainsworth, P.
  2001   Offender Profiling and Crime Analysis. Devon: Willan Publishing.

Eck, J.E., J.S. Gersh, and C. Taylor
  2000   Finding crime hot spots through repeat address mapping. In V. Woodsworth, P.G. McGuire, J.H. Mollenkopf, and T.A. Ross (eds.), Analyzing Crime Patterns: Frontiers of Practice. Thousand Oaks, CA: Sage.

Engel, R.S., J.M. Calnon, and T.J. Bernard
  2002   Theory and racial profiling: Shortcomings and future decisions in research. Justice Quarterly 19(2): 249–273.

Fridell, L., R. Lunney, D. Diamond, and B. Kubu
  2001   Racially Biased Policing: A Principled Response. Washington, DC: Police Executive Forum. http://www.policeforum.org/racial.html

Gold, Alan D. and Dr. Edward B. Harvey
  2003 February   Executive Summary of Presentation on behalf of the Toronto Police Service. Toronto: Toronto Police Service. http://www.torontopolice.on.ca/

Harris, David A.
  2002   Profiles in Injustice: Why Racial Profiling Cannot Work. New York: New Press.

Harvey, Edward B.
  2003 February   An Independent Review of the Toronto Star Analysis of Criminal Information Processing System (CIPS) Data Provided by the Toronto Police Services (TPS): A Summary Report. Toronto, Toronto Police Service. http://www.torontopolice.on.ca/

Harvey, Edward B. and Richard Liu
  2003 March   An Independent Review of the Toronto Star Analysis of Criminal Information Processing system (CIPS) Data Provided by the Toronto Police Services (TPS). Toronto: Toronto Police Service. http://www.torontopolice.on.ca/

Lamberth, John
  1994 November   A Statistical Analysis of the Incidence of Police Stops and Arrests of Black Drivers/Travelers on the New Jersey Turnpike

between Exits or Interchanges 1 and 3 from the Years 1988 through 1991. Washington, DC: ACLU.

Lamberth, John
  1999   Driving while black: A statistician proves that prejudice still rules the road. Washington Post, August 16: C1.

McMahon, Joyce, Joel Garner, Ronald Davis, and Amanda Kraus
  2002   How to correctly collect and analyze racial profiling data: Your reputation depends on it – Final report for racial profiling: Data collection and analysis. Washington, DC: Government.

Melchers, R.
  2001   What Can (and Can't) Statistics Tell Us about Crime and Criminal Justice. Ottawa: Office of the Auditor General for Canada.

Police target black drivers
  2002   Toronto Star, 20 October: A1.

Police union sues Star over race-crime series
  2003   Toronto Star, 18 January: A6.

Ramirez, D., J. McDevitt, and A. Farrell
  2000 November   A Resource Guide on Racial Profiling Data Collection Systems: Promising Practices and Lessons Learned. Washington, DC: U.S. Department of Justice Monograph. (NCJ 184768)

Robert, Ph.
  septembre 2001   Enquêtes de victimisation et statistiques policières: 1'expérience française. Séminaire La rationalité pénale moderne, Université d'Ottawa.

Roberts, J.V.
  1994   Crime and race statistics: Toward a Canadian solution. Canadian Journal of Criminology 36(April): 2.

Sacco, Vincent and Holly Johnson
  1990 March   Patterns of Criminal Victimization in Canada. Ottawa: Statistics Canada, Canadian Centre for Justice Statistics. (11–612E, no 2)

Singled out
  2002   Toronto Star, 19 October: A1.

Smith, Michael, and Matthew Petrocelli
  2001   Racial profiling? A multivariate analysis of police traffic stop data.
         Police Quarterly (March): 4–27.

Thomas, M., H. Hurley, and C. Grimes
  2002   Pilot analysis of recidivism among convicted youth and young adults
         – 1999/00. Juristat 22(9).

U.S. Department of Justice
  1999 June   Attorney General's Conference on Strengthening Police-Com-
              munity Relationships – Report on the Proceedings. Washington, DC:
              U.S. Department of Justice.