Using data to make a difference

Understanding crime through victimisation surveys

Jean Redpath*

redpath@iafrica.com

Victimisation surveys have the potential to deepen our understanding of crime in South Africa. Using the example of a survey conducted in Galeshewe, this article considers the challenges facing analysts in analysing victimisation surveys and suggests ways to increase the information that can be mined from local and national victimisation surveys.

Surveys of a sample of a population can be used to show trends in the whole population. Mathematics has shown that a large enough random sample drawn from a population will show the same distribution of characteristics as appears in the population as a whole. Thus if a quarter of the population has experienced crime, then a quarter of a randomly drawn sample of the population will also have experienced crime. 'Random' in mathematics does not mean choosing respondents by convenience, but implies a rigorous method of choosing so that every person in the population has an equal chance of being selected.

Because national and local victimisation surveys can provide accurate information about crime in the population as a whole, they are important tools for measuring changes in crime trends. An independent national measure of crime rates is necessary because not all crimes are reported to,

or recorded by, the police. For example, the British Crime Survey (BCS) is a national survey conducted on a regular basis, providing an independent measure of crime rates in England and Wales against which to gauge the impact of state policies on crime rates.[1] The figure on the next page shows crimes reported to the police in England and Wales as compared to crimes measured using the BCS, over the period 1981 to date. The number of BCS crimes is calculated by multiplying the distribution of victims found in the survey by the known size of the population of England and Wales (obtained via Census estimates).

What is apparent is the large gap between recorded crime and BCS trends – the BCS counts more than double the crimes counted in reported crime data. Furthermore the number of crimes reported to the police showed little change over the entire period; by contrast the BCS showed an increase in early 1990 followed by a decline (which the Blair government naturally attributed to the impact of its policies).

Similarly, small-scale local victimisation surveys provide a baseline measurement of crime rates

---

*   Jean Redpath is an independent researcher currently analysing the Galeshewe victimisation survey and the first national victimisation survey for the country of Moldova. She recently visited the British Crime Survey offices at the UK Home Office.

Figure 1: BCS and recorded crime trends 1981-2008/9
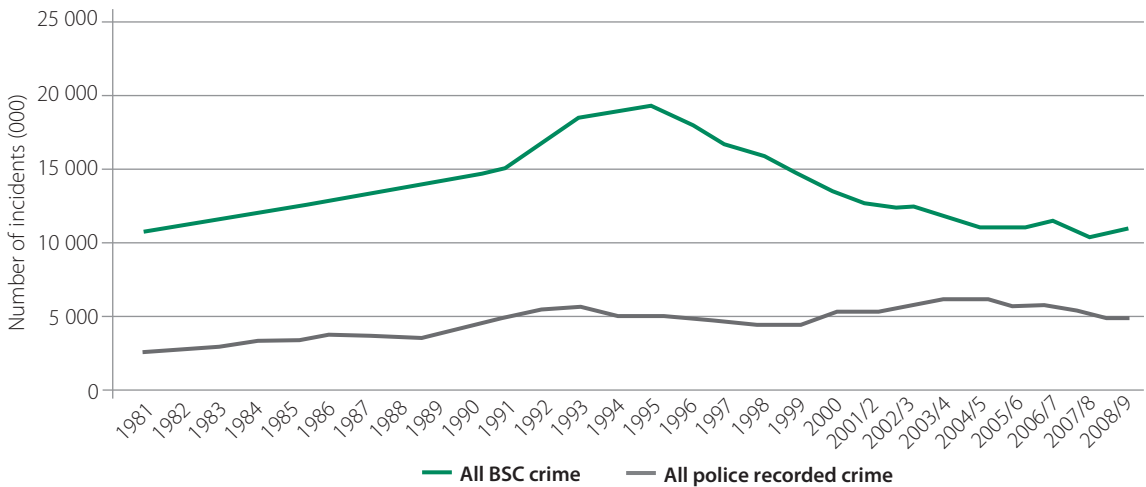
Source: Presentation by Rachel Murphy and Phil Hall 'Introduction to the British Crime Survey' UK Home Office 19 May 2010

against which the success of crime prevention interventions can be measured. For example, the Galeshewe victimisation survey, conducted ahead of the implementation of a crime prevention plan, involved interviewing 800 respondents selected randomly by street address to provide a baseline measure of crime rates in Galeshewe. This baseline measure was considered necessary not only because not all crimes are reported or recorded, but also because crime prevention initiatives may lead to an increase in trust in the police, which in turn may lead to an increase in reporting rates and consequently higher rates of recorded crime. This means that successful interventions may seem unsuccessful if the only benchmark available is reported crime data. Such data would tend to show an increase rather than a decrease in crime rates.

Apart from providing a benchmark against which to measure change, local victimisation surveys can also inform the design of crime prevention initiatives by highlighting at-risk groups, areas and behaviours, and fill the gaps in information available from official sources.

However, the problem with victimisation surveys is that very large sample sizes are necessary to capture sufficient crimes for analysis. In 1981 the BCS began with a sample of 11 000 respondents.[2] This increased to 20 000 in 2000 and to 45 000 in 2004/5. In 2005 the population of England and Wales was around 53 million,[3] only slightly larger than South Africa's current population of around 48 million, suggesting South Africa needs a national victimisation survey nearly as large.

WHY ARE LARGE SAMPLE SIZES NECESSARY?

The size of the whole population is not the only factor determining the sample size for a survey. What is more important is the relative size of the issue being investigated. In survey analysis, the general rule of thumb is that at least 30 occurrences of any category of interest (such as robbery) are necessary for legitimate analysis of that category of interest.[4] This implies that if reported crime data indicate a robbery rate of around 500 per 100 000 people per year, and the population size is around 100 000, a sample size of 6 000 would be necessary for a simple random survey to yield roughly 30 counts of reported robbery in the last year which could then be legitimately analysed (500 robberies per 100 000 is equivalent to 30 robberies per 6 000). A representative random sample of 1 000 people in a

population of 100 000 would yield only five people who would have experienced a robbery in the last year that they reported to the SAPS (500 robberies per 100 000 is equivalent to five robberies per 1000).

Large sample sizes have important cost implications. The size of a survey is frequently determined by the budget available to fund it, rather than by the requirements of analysis. For the analyst, one way to get around the problem of small samples for each category of crime is to confine analysis to larger categories of interest (such as 'all crime in the last year' or 'all violent crime').

However, conflation of crime categories makes interpretation of the data far more difficult, as there may be different trends pertinent to different crime types that influence the analysis in different ways. For example, whether a person is a victim of housebreaking or not may be influenced by the characteristics of the neighbourhood in which she lives, such as the availability of street lighting. Theft, by contrast, may occur outside the home, in the workplace, in recreational areas, and so on, and thus not be largely influenced by neighbourhood characteristics such as street lighting. An analysis of 'all economic crime' against whether respondents have sufficient street lighting around their homes may fail to show any significant association between street lighting and economic crime, because the inclusion of theft in the category 'all economic crime' weakens that association.

However, low *reported* crime rates (in comparison to actual crime rates) mean that the actual number of victims of crime in the population is larger than the number predicted by reported crime rates. Thus if the reported rate of robbery is 500 per 100 000 (or 5 per 1 000) and the reporting rate for robbery is 50%, this means the actual crime rate is 1 000 per 100 000 (or 10 per 1 000) and a sample of 3 000 would yield the required 30 victims of robbery. But because reporting rates vary widely by crime type and are not known prior to the survey being administered, reporting rates are calculated using the survey

itself (see below) – they cannot be used to suggest an appropriate sample size. However, reporting rates uncovered in prior surveys or surveys of similar populations can be used as a guide to likely reporting rates and can thus inform sample size.

Another factor that increases the number of victims of crime in a sample, compared to that predicted by police crime statistics, is the tendency of respondents to say they have been a victim of crime 'in the last 12 months' when in reality the crime occurred before the period of interest. This is known as 'telescoping' – because the crime looms large in respondents' lives it seems to them to have occurred more recently than it in fact did.

While telescoping does have the effect of bumping up category sizes, an analyst has no way of knowing which respondents have telescoped and which have not. This can affect the analysis. For example, the analyst may be investigating whether a victimisation prevention course conducted before the period under investigation had any effect on the likelihood of course participants being victims of crime, compared to the rest of the population. If respondents include crimes from previous years (i.e. before the intervention), evidence of any actual effects of the intervention will be weakened, because these crimes could not have been affected by whether or not they attended the intervention.

It is possible to estimate the overall extent of telescoping, using reporting rates derived from the survey and reported crime figures. Respondents are asked whether they reported the crimes against them or not, and the percentage claiming to have reported the crime committed against them is the reporting rate. Multiplying the reporting rate by the victimisation rate would result, in the absence of telescoping, in a figure which approximates the reported crime rate.

To illustrate the application of these concepts it is useful to consider a practical example. At the time of writing the Centre for Justice and Crime Prevention (CJCP) was developing a safety plan

for Galeshewe in the Sol Plaatje (Kimberley) Municipality. As part of the research to inform the safety plan, CJCP conducted a victimisation survey in Galeshewe. The victimisation survey will provide baseline data against which the success of the plan can be measured and will inform its design. Budget constraints dictated that the Galeshewe survey could involve interviewing 800 respondents selected randomly from the population of 100 000.

The Galeshewe survey asked respondents whether they had ever been a victim of six broad categories of crime (theft, robbery, assault, burglary, hijacking and sexual assault). If respondents had ever been victims of any of these crimes, they were asked whether the crime occurred during the last 12 months, and if so, whether they reported it. Further various details of the crime were then recorded, as well as demographic and other characteristics of the respondent.

Table 1: Galeshewe SAPS DATA and Galeshewe victimisation survey data by crime type

	2008/9 SAPS	2008/9 SAPS Total	Counts	Weighted counts	Counts reported	Weighted reported counts	Reporting %	Reported crime divided by reporting rate	Extent of tele-scoping	Tele-scoping %
Total sexual crimes	256									
Total sexual assault	256	256	9	1 125	5	625	56%	461	205	80%
Attempted murder	102									
Assault with the intent to inflict grievous bodily harm	1 617									
Common assault	870									
Total assault	2 589	2 589	58	7 250	46	5 750	79%	3 264	675	26%
Burglary at business premises	86									
Burglary at residential premises	593									
Total burglary	679	679	64	8 000	51	6 375	80%	852	173	25%
Theft of motor vehicle and motorcycle	46									
Theft out of or from motor vehicle	201									
Stock theft	18									
All theft not mentioned elsewhere	685									
Total theft	950	950	37	4 625	14	1 750	38%	2 511	1 561	164%
Robbery with aggravating circumstances	571									
Common robbery	285									
Robbery at business premises	16									
Robbery at residential premises	4									
Total robbery	876	876	67	8375	49	1750	73%	1198	322	37%
TOTAL	5 350	5 350	235	29 375	165	16 250	70%	7 620	2 270	42%

In Galeshewe an estimated 'telescoping rate' for reported crime of 42% was calculated (see Table 1). In other words, 42% more respondents referred to a crime in the last 12 months which they said they had subsequently reported to the SAPS, than was indicated by the SAPS crime data. Analyses by crime type, however, suggest telescoping rates in Galeshewe varied widely from 25% (housebreaking) to theft (164%). Despite telescoping, some categories of crime did not yield sufficient respondents for analysis, necessitating analysis by larger categories of interest, such as lifetime experience of crime.

It was expected that lifetime experience of crime among the sample population would be far greater than experience of crime in the last 12 months. However, the Galeshewe results showed a small difference between experience in the last year and lifetime experience. While only just over a third of people indicated that they had ever been a victim of these crimes, as much as a quarter had been a victim of these crimes in the last 12 months.

There is no reason to believe that the last 12 months were any different from preceding years, i.e. there was no crime wave in Galeshewe in the last 12 months according to SAPS statistics. In fact, reported violent crime has dropped from about 5 000 to 4 000 crimes over the last five years).[5] So, how do we account for the fact that as much as a quarter of the population was victimised in the last year, but only a third had ever been victimised in their lifetimes (including the last year)? Obviously telescoping may also be at work here: crimes which occurred more than 12 months previously are remembered as having occurred in the last year, thus boosting the 'victim in the last year' figure.

Telescoping does not however account for the apparently low *lifetime* victimisation rate. As much as two-thirds of the sample population say they have *never* in their lifetimes been a victim of any of these crime types. A possible reason for this may be the tendency of respondents to forget crimes that occurred a long time in the past (thus reducing the 'victim ever' figure).
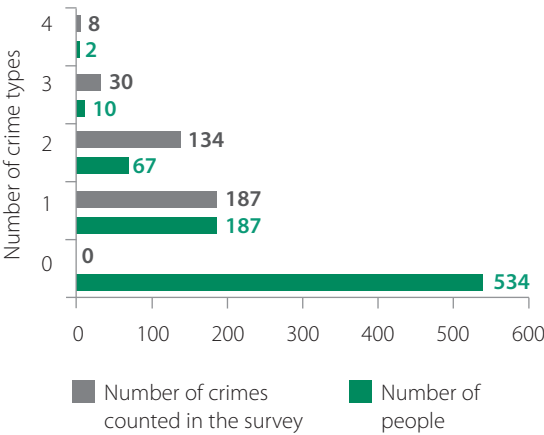
But such collective amnesia cannot on its own account for the phenomenon. Other factors must be at work to account for the lower than expected lifetime victimisation rate. Just over 5 000 relevant crimes (i.e crimes that the victimisation survey sought to find out about) are reported to the SAPS each year in Galeshewe. If 5 000 crimes are reported each year, in six years around 30 000 crimes will have been reported. Thus in six years enough crimes are reported in Galeshewe to match the third of the Galeshewe population who have 'ever' been victims of crime. From year seven onward, the number of crimes begins to exceed the number of people who have ever been victims of crime. This suggests many people in Galeshewe must have been victims of more than one crime.

Statistical analysis of the Galeshewe survey reveals that only 29% of those who were victimised in the last year, had never been previously victimised. This suggests that having been a victim of crime at some stage in life in Galeshewe is a strong predictor for more recent victimisation ('victimisation in the last year'). Furthermore, the data suggest that victims of more than one crime type account for a disproportionate share of crimes counted in the survey; but a diminishing share of crimes reported. That is, a person victimised in respect of more than one crime type is less likely to report crime.

This conclusion was reached by analysing the extent to which respondents had been victim to more than one crime type. The survey, like many other victimisation surveys, did not record whether respondents were victims of more than one of the same crime type in the last year. However the survey did record different crime types separately. Thus it is possible to deduce from the data whether respondents were victims of *more than one crime type* ever or in the last year (see Figure 2 on page 14). Each respondent would thus record being a victim of between no and six crime types.

Almost half (48%) of *all crimes* counted by the survey (only the most recent of each crime type was counted by the survey) were committed against people who were victims of more than one
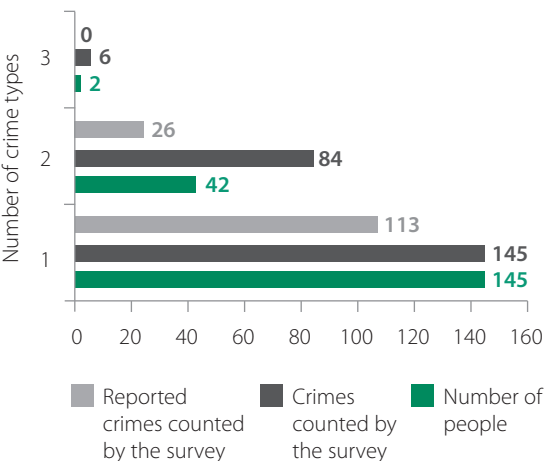
Figure 2: Lifetime experience of crime by number of different crime types experienced

When considering *crimes against victims in the last year*, the data show how SAPS reported crime data may tend to mask the degree to which there are multiple victimisations of the same person, due to poorer reporting rates after a greater degree of victimisation. Some 78% of those who were victim to only one crime type in the last year said they reported the most recent of these crimes to the police. This compares to only half (50%) of those victim to more than one crime type in the last year, who reported all of the relevant most recent crimes.

crime type. Ten per cent of the sample population experienced more than one crime type, 20% experienced only one crime type, and 70% experienced no crime.

In other words, half of the crimes counted were committed against less than a third (30%) of the victims identified – or perhaps more starkly, half of the crimes counted were committed against only 10% of the sample population. Unfortunately, because the survey only counted and considered in detail the most recent of each crime type, the true extent to which victims may have been victimised (repeated instances of the same crime type) cannot be established.

These findings suggest that crime prevention initiatives in Galeshewe should target past victims of crime. This is because past victims are highly at risk of crime. Targeting past victims successfully is likely to yield greater reductions in real overall crime rates than more general interventions, because past victims are likely to comprise the majority of new victims. Note however that the impact of a successful intervention aimed at preventing re-victimisation is unlikely to be properly reflected in *reported* crime data. Without a survey to benchmark such change, the likelihood is that official data will fail to highlight the successful prevention of multiple victimisations.

Figure 3: Last 12 months experience of crime by number of different crime types experienced

Furthermore, these findings suggest that victimisation surveys – whether at local or national level – need to be designed adequately to capture the extent to which re-victimisation may be occurring. Understanding repeat victims would require surveys to ask only an additional two questions about *the number of times* ever and in a specific year (e.g. 'during 2009') the respondent had been victim of the crime type under consideration. Further, more detailed, questions could then relate only to the most recent of these crimes, thus keeping the survey short.

Victimisation surveys can also collect a range of additional information about both the victim and the offender, and information regarding the attitudes and perceptions of the general public in relation to the criminal justice system. This information can then be mined for informative

correlations to inform crime prevention as well as trust in the criminal justice system. The BCS has four sub-components on such issues that are randomly applied to respondents. Thus each respondent will only answer one of the sub-components, keeping the survey interview to manageable length.[6]

THE PROBLEM OF SMALL CATEGORY SIZES AND SERIOUS CRIME

The prevention of serious crime such as aggravated robbery and rape is the most fervent goal of crime prevention initiatives. Yet it is precisely these serious (and relatively rare) categories that suffer from small numbers in victimisation surveys, thus preventing detailed analysis which may inform crime prevention initiatives.

Sexual crimes in particular are understood to suffer the most from under-reporting, both in surveys and in official data.[7] The reasons for the under-reporting by victims of sexual crimes to both victimisation surveys and crime data are well known and need not be repeated in detail here. The Galeshewe survey recorded a count of nine victims of sexual assault in the last year, of whom five said they had reported the crime to the police. Because of the small number, it is not possible for more detailed analysis by demographic, social and behavioural characteristics to be carried on the category of sexual assault victims, which would assist in understanding sexual assault in Galeshewe.

However the figure of only nine victims is actually larger than expected. Published SAPS figures indicate reported sexual assaults of 265 in 2008/9 (per 100 000). Given that the population of Galeshewe is around 100 000 and the sample size is 800, the likely category size for reported sexual assault is 265/100 000 x 800 = 2. This is less than half the number of reported rapes recorded in the survey (5).

This gives the magnitude of the combined impact of the opposing effects of under-reporting to interviewers (which would tend to reduce the incidence by excluding some sexual assaults) and telescoping (which would tend to increase the incidence by including sexual assaults from previous years).

SAMPLE SIZE AND THE VALUE OF VICTIMISATION SURVEYS

The Galeshewe results provide some guidance on sample size for a national survey by estimating a reporting rate for sexual offences.[8] The average national yearly reported incidence of sexual assault of around 45 per 100 000 suggests a sample size of 67 000 (30 per 67 000 is the same as 45 per 100 000) is necessary if serious relatively 'rare' crime types such as sexual assault are to be adequately interrogated by a simple random survey. If it is assumed the Galeshewe reporting rate (5/9) can be applied to the rest of the country, the sample may be roughly halved (to around 33 000).

If clustering is adopted, this will necessitate an increase in the sample size. Clustering means that instead of conducting 30 000 interviews in 30 000 disparate households, interviewers conduct for example 10 interviews in 3000 areas – this reduces costs per interview. Because the survey is no longer random, the impact of clustering on survey reliability must be taken into account – but this can only be determined after the survey is completed.

Apart from providing a more accurate picture of crime trends, is there any other reason why it is worth going to the expense of ensuring a large sample-size victimisation survey?

SAPS data are collected at a rich level of detail (down to street address level) but this level of data is not made publicly available.[9] SAPS data are usually only available in the form of crime counts by policing area per year. Analysis of SAPS crime data by policing area can only provide correlates of crime by area characteristics, not correlates by person characteristics. This is one of the key advantages of victimisation surveys. To give an idea of the kind of insights which can be obtained, it is worth returning to Galeshewe.

In Galeshewe, five statistically significant demographic risk factors emerged in relation to ever being a victim of crime. The independent effect of these risk factors is listed below in order of magnitude or effect.

- Age 27-31 (43% ever-victim versus 31% among persons of other ages) (39% increase in risk)
- Having ever taken drugs (41% ever-victim versus 30% among the drug-free) (37% increase in risk)
- Being male (39% ever victim versus 29% of women) (34% increase in risk)
- Not having children (40% ever victim versus 31% among those with children) (29% increase in risk)
- Having completed high school (38% versus 30%) (27% increase in risk)

These risk factors were considered together in a multivariate analysis. Multivariate analysis looks at the impact of one variable, assuming the others remain the same, e.g. is the effect of being male still positive (in the sense of increasing risk), taking past drug use into account? The multivariate analysis found that all the factors listed above remained positive (increased risk) even when considered together. This suggests that the risk factors are robust. Interestingly, the likely victim in Galeshewe is not that different from the likely perpetrator, suggesting that risk-taking may play a role.

Further analysis of the Galeshewe survey indicates that characteristics frequently presumed to be 'good' or 'bad' for crime in an area have different effects at the individual level. Thus while it may be presumed that a high degree of social capital activities (participating in sports, attending church and the like) work to prevent crime at the area level, at the individual level, participation in these activities may operate to increase personal risk in areas such as Galeshewe.

This in turn suggests that there is a 'threshold level' of participation required in an area before the benefit of such activities is realised at the individual level. Similarly service delivery

variables such as the following were generally found not to be significant in relation to 'ever' victimisation at the individual level:

- Living in a shack/house/RDP house or not
- Living in a household of four or more or not
- Going hungry or not
- Having a toilet in the home or not
- Being unemployed or not
- Having to collect water outside the home or not
- Living in a household which receives a social grant or not

These findings that some service delivery indicators are generally *not* a risk factor for victimisation in Galeshewe may seem counter-intuitive, particularly to those who espouse social crime prevention views. This is because these results speak to the *risk experienced by individuals in the Galeshewe context*, and not to the impact of *changes to the context*, with which social crime prevention is concerned.

In other words, these results should not be interpreted as providing support for the view that, for example, rolling out a universal social grant would have no impact on crime trends in Galeshewe. What the result *does* mean is that for any individual in the current Galeshewe context, given the current levels of social grant provision, whether or not their household receives a social grant has no impact on whether or not that individual is likely ever to have been a victim of crime in Galeshewe.

It is notable that one of the exceptions (to the general irrelevance of service delivery indicators to *individual* risk in Galeshewe) was that increasing distance from a police station was correlated with greater risk of *victimisation of more than one crime type*:

- Long travel time to police station (>30 minutes) (12% v 8%) (50% increase in risk)

This suggests that the proximity of the police, at least in Galeshewe, does have an impact at the individual level via a reduced risk of victimisation by more than one crime.

One marker of context was also found to be a significant correlate in Galeshewe of being a multiple victim:

• High perceived prevalence of graffiti (17% v 9%) (89% increase in risk)

This may please those who hold to 'crime and grime' theories.

The fact that there are any correlates of crime at all for individuals who live in Galeshewe suggests that crime in Galeshewe is not entirely random. People who have certain characteristics, live in certain contexts, and engage in certain behaviours are more at risk of crime than those who do not share these characteristics. This is comforting, because it suggests there are things that can be done about crime. The more that is understood about crime the more specific, targeted and successful crime prevention interventions can be.

Furthermore, to understand crime trends in South Africa properly requires an expanded, reliable, national victimisation survey. Police performance is currently measured by changes in recorded crime rates. Measuring the performance of the police in terms of data they themselves collect provides an incentive to police officials to either not record crimes reported (the evidence suggests this happens frequently in relation to sexual offences) or for down-classifying crimes. For example, aggravated assault may be recorded as assault, robbery may be recorded as theft, and murder may be recorded as culpable homicide (see the articles by Bruce and Faull in SACQ 31).

Consequently if it is desired to measure police performance in terms of actual changes in crime, these changes in crime should be measured independently. A national victimisation survey is one way to do this.

An expanded large-scale national victimisation survey in conjunction with better access to detailed official SAPS data may begin to provide us with a very rich understanding of risk, and thus inform crime prevention for South Africa as a whole, by providing a deep understanding of victimisation in context. For a country as deeply affected by crime as South Africa is, it is of serious concern how little data are really available to inform our approaches to crime. Without data, designing crime prevention interventions amounts to no more than a stab in the dark.

To comment on this article visit
http://www.issafrica.org/sacq.php

NOTES

1 Presentation by Rachel Murphy and Phil Hall, Introduction to the British Crime Survey, UK Home Office, 19 May 2010.
2 Ibid.
3 Ibid.
4 Central limit theorem: David Nelson, *Penguin Dictionary of Mathematics*, 2nd edition, 1998, London. If n is more than 30 then a normal distribution is (often) approximated. If n is small (less than about 30) the normal distribution is no longer approximated and one has to use the Student t distribution for any analysis. As standard analytical software does not use one distribution for cells above 30 and another for cells below 30, it is advisable that cells with a small n are treated with great caution (or ignored). See http://www.andrews.edu/~calkins/math/ webtexts/prod12.htm (accessed 5 July 2010).
5 See www.saps.org.za for crime data for Galeshewe and Kagiso, which together cover the area of interest.
6 Presentation by Rachel Murphy and Phil Hall 'Introduction to the British Crime Survey' UK Home Office 19 May 2010.
7 L Artz & D Smythe, Case Attrition in Rape Cases: A Comparative Analysis, *South African Journal of Criminal Justice*, August 2007.
8 Note that this is the rate of reporting to the police of incidents reported to survey interviewers. Surveys cannot indicate the true reporting rate (the rate of reporting to police of all incidents).
9 For some offences it may be desirable for such data to be widely available – for the purposes of preserving the anonymity of victims. However SAPS does not release such information even to other government departments responsible for crime prevention, such as provincial departments of safety and security.