USER SATISFACTION:

FURTHER INFORMATION

(TO ACCOMPANY 2011/12 GUIDANCE AND SURVEYS)

March 2011

CONTENTS

A.	Checklist for the survey process	3 - 7
B.	Explanation of probability sampling methods	8 - 18
C.	Calculating sample sizes	19
D.	Guide to survey analysis	20 - 37
E.	Example of an Introduction Letter	38
F.	Service Recovery Template	39
G.	Data Processing Agreement	40 – 44
H.	Confidentiality Agreement	45
I.	Indemnity Agreement	46
J.	Data Protection Explanation Letter	47
K.	Data Protection FAQ	48 - 49

ANNEX A

CHECKLIST FOR SURVEY PROCESS

Section	Stage	Checklist:	Implemented - Y/N?
	Identify victim population	Does the sample include domestic burglary victims involved in crimes classified 28 and 29?	
		Does the sample include violent crime victims involved in crimes classified 8F, 8G, 8K, 34B and 105A?	
		Does the sample include vehicle crime victims involved in crimes classified 37.2, 45 and 48?	
		Does the sample include victims of racist incidents?	
		Is the sample used for reporting satisfaction measures restricted to these four groups?	
	Calculate required number of responses	Is there a required number of responses identified for each of the four user groups?	
		Do these required numbers provide sufficient data to report the results at 95% confidence +/- 4% for each user group at Force level when considered over a 12 month period?	
		Have the required sample numbers taken account of the estimated number of exclusions within each user population?	
		Is there a required number of responses for each BCU?	
		Is this in line with the BCU sample size calculator recommendations?	
		Does it generate BCU results accurate to +/-4% at 95% confidence over a 12 month period?	

		Are details of how these required numbers were calculated retained for audit purposes?	
		Are details of how the sample has been drawn retained for audit purposes?	
		Has the sample been drawn using probability sampling methods?	
	Exclude sensitive cases	Is the sample structured to include only one respondent from each crime?	
		Does the sample exclude those under 16?	
		Does the sample exclude victims of domestic violence, sexual offences, police officers assaulted in the course of their duty?	
		Does the sample attempt to exclude cases involving family members as offenders and unsuitable/unwilling cases?	
		What other criteria are being applied to exclude cases?	
		Is there a record of those included and excluded from the sample (reference number (or name) and reason)?	
Field Work	**Design survey**	Do all surveys include the core satisfaction questions, following the standard wording and in the required order?	
		Do all surveys include the standard response options, in the required order, for all core satisfaction questions?	
		Do all surveys explore the service behind the expressed levels of satisfaction, by including diagnostic questions?	

		Do the surveys include the required diagnostic questions on actions and treatment?	
		Are respondents invited to give free-text responses to issues not covered by the core and diagnostic questions?	
		Do all surveys contain the standard demographic questions, and the standard response options?	
		Do all surveys contain a question asking the respondent if they are willing for identifiable responses to be passed back to the force/authority?	
		Do the surveys seek contact details for further contact and consultation?	
		Are the surveys tailored to different service users?	
		Do the surveys follow the recommended examples	
		Are copies of the survey formats retained for audit purposes?	
	Survey users	Are surveys conducted within 6-12 weeks of the incident being reported?	
		Is there surveying of each group of users each quarter?	
		Has a representative sample of racist incident victims been obtained?	
		Have steps been taken to increase the numbers of responses from racist incident victims, if necessary	
		Is there a Data Protection protocol in place?	
		Does the survey process comply with the protocol and include gaining consent to pass back identifiable responses to the force/authority?	

		Are details of samples and completed surveys retained for audit purposes?	
		Are any steps taken to forewarn victims about the survey process?	
	Telephone survey	Are all interviews being carried out using telephone survey methodology?	
		Do the interviewers make every effort to contact those identified in the sample? Do they make repeated attempts and make use of all available numbers? Is there a record of the calls (successful and unsuccessful)?	
		Does the survey script introduction cover issues of identification, confidentiality and establish the purpose of the call?	
		Do the interviewers work to clear standards? Are they well-trained to carry out their role? Are these standards documented?	
		Where surveys have been contracted out, have data protection issues been covered in the contractual arrangements?	
Analysis and Use	**Ensure sample is representative**	Is the profile of respondents representative of the profile of victims?	
		Has the required number of responses been obtained?	
		Are the survey results for the satisfaction questions accurate to +/-4% at 95% confidence when considering 12 months data?	

	Analyse results	Are the results of the user satisfaction surveys being reported to senior managers and the police authority at Force and BCU level?	
		Have the survey results been analysed to provide knowledge and understanding of service quality and how service could be improved?	
	Identify ways to improve service	How are the results influencing action to improve service delivery?	
		How do the results inform training?	
	Submit quarterly return	Are the quarterly results reporting the findings of the surveys carried out that quarter?	
		Are the returns submitted on time?	

EXPLANATION OF PROBABILITY SAMPLING METHODS

Malcolm Hibberd

We sample for reasons of economy. When we conduct a questionnaire survey, we are likely to be interested in some feature of a population of people. This feature is called a 'parameter', and in police User Satisfaction Surveys this is the satisfaction rate of a group of 'customers'. (Other examples of parameters would be fear of crime, or confidence in the police, especially in surveys of the general public.)

To know the value of the population parameter for sure, we would need to measure every member of the population, by asking them a survey question. However, that is likely to be too expensive, so instead we select a subset of the population – called a sample – and measure each of those individuals. The results from the sample are then used as an estimate of the population parameter.

But because the sample is not the complete population, the estimate of the population parameter may not be accurate. How accurate it is will depend on two factors – the size of the sample, and how the sample is selected.

Bias in sampling

If the results from a sample are to be valid and reliable, the sample should be free from bias. In practice it is rarely possible to be entirely free from bias, so in designing a sample the task is more likely to be one of minimising, rather than eliminating bias.

To understand bias, how it threatens our results, and what can be done about it, we need to recognise that there are two different types – random bias and systematic bias.

Random bias threatens the reliability of a sample. It results from statistical sampling error – the fact that the results we get depend on the particular combination of entities included in the sample. Random bias can be reduced by increasing the sample size, provided probability sampling is used (see below).

Systematic bias threatens the validity of the conclusion drawn from the exercise. There are three main sources of systematic bias:

- distortions in the sampling frame, including incomplete records;
- the sampling method used (if we use probability sampling, for example, we avoid systematic bias dependent on the method);
- non-response – the problem that those who do not respond are in some (possibly unknown) way consistently different from those who do.

Dealing with Systematic Bias – Sampling Methods

Systematic bias in sampling is dealt with by the method used to draw the sample from the population, or sampling frame. Police forces are required to use probability sampling in carrying out User Satisfaction Surveys.

Probability sampling

We use probability sampling to eliminate bias in the way sample members are selected from the sampling frame. If all else were perfect, probability sampling would eliminate systematic bias altogether. However, in the real world, there is often a mismatch between the population and sampling frame, and we can never achieve a perfect response rate, so elimination of systematic bias is practically impossible.

Probability sampling involves each member of the population (although in practice, this will usually be the sampling frame) having an equal chance (probability) of selection for the sample.

There are two types of probability sampling – random sampling and systematic (or interval) sampling; either of these may be modified to give stratified sampling.

We shall discuss sampling methods in relation to User Satisfaction Surveys, covering

- simple random sampling,
- systematic/interval sampling,
- stratified sampling,
- booster sampling,
- quota sampling,
- not sampling.

The first two are the basic methods for probability sampling. **Quota sampling is not a probability sampling method, and is therefore not acceptable for User Satisfaction Surveys.**

Simple random sampling

Simple random sampling is usually held to be the purest form of probability sampling, avoiding any possibility of systematic bias. There are four steps in simple random sampling.

1. Determine the sample size (n), with over sampling to take account of non-response.
2. Sequentially number the sampling frame from 1 to N (N being the population size).
3. Generate n different random numbers from the range 1 to N.
4. Use the random numbers to identify the members of the sampling frame who are to be approached to take part in the survey.

There are two practical problems with this approach. The first relates to the nature of the population for the User Satisfaction Surveys. When an annual survey sample is designed, the population cannot be identified, as they have not yet become (for example) victims of burglary. Because of this we need to keep the sampling flexible, and monitor through the course of the year (as discussed in the worked example below).

The second problem concerns the extent to which we should over sample, as we cannot know in advance what the response rate is going to be. We would usually estimate it, probably on the basis of the previous year's results. But here again we

need to be flexible enough to change our sampling if this year's response rate turns out to be higher or lower than expected.

Systematic sampling

This method is also called interval sampling. Again, there are four steps involved.

1. Determine the sample size (n), with over sampling to take account of non-response.
2. Sequentially number the sampling frame from 1 to N (N being the population size).
3. Select every x^{th} member of the sampling frame, where x is equal to N divided by n (the reciprocal of the sampling fraction). This should start from a randomly determined point among the first n members of the sampling frame, otherwise it is not probability sampling.
4. Approach the selected members of the sampling frame, inviting them to take part in the survey.

This method suffers from the same two practical problems that beset simple random sampling.

Stratified sampling

Both random and systematic sampling may be modified by stratification, without losing any of the advantages of probability sampling. Stratification is carried out to ensure that the sample gives us a close match to the population with respect to key variables, usually demographic.

The reason we may decide to stratify our sample is related to sampling error. Even the ideal of simple random sampling cannot be relied on to give a perfect match between the characteristics of a sample, and those of the population (or sampling frame) from which it is drawn. Chance will usually give rise to slight over- or under-representations of certain groups.

This may cause problems if a demographic variable is particularly important in the interpretation of the results of the survey, an obvious example being BCU. These problems are exacerbated if the population split on the variable is uneven, for example if one BCU contributes disproportionately to the number of victims: this will make the distortion proportionately greater for BCUs with fewer victims.

Stratified sampling is rather more complicated, and is described under the following six stages.

1. Determine the sample size (n), with over sampling to take account of non-response.
2. Identify the relevant demographic characteristics of the population by which you want to stratify. These must be characteristics which can be identified from information we have on members of the sampling frame – e.g. the BCU in which a domestic burglary was recorded. Each subgroup identified will be a stratum of the sampling frame.
3. Determine the sample size (n_A) for each stratum, using the same sampling fraction that has been used for the whole sample size.
4. Sequentially number each stratum, 1 to N_A (N_A being the population size for a given stratum).

5. Use simple random or systematic methods (as described above) to select sub samples separately from each stratum.

6. Approach the selected members of the sampling frame, inviting them to take part in the survey.

The complications of simple random and systematic sampling are multiplied as soon as we attempt to stratify the sample. Stratification requires that we have complete (or nearly complete) information on the key variables for all members of the sampling frame. For this reason, and because of the obvious complications involved, we would be unlikely to attempt to stratify by more than two or at most three variables.

It is also possible that response rates vary consistently between strata. In that case we would have to make separate estimates of the likely response rates for the strata, and vary the over sampling rate accordingly. This approach is illustrated in the worked example at the end of the document.

Booster sampling

The tendency to respond to surveys is not evenly distributed across a population: certain groups tend to be under-represented, such as certain ethnic groups, young males and people of lower educational attainment. This, of course, may threaten the validity of our findings, as the under-represented groups may differ consistently in their attitudes: young males, for example, tend to express consistently more negative attitudes to the police.

There are two possible solutions to this. The first is to weight the results, to give a set of hypothetical findings of what the results would show if the sample were to match the population break down.

The second solution is to supplement our sample with a booster sample, which involves over-sampling more from under-represented groups. This requires that we know the population breakdown according to the relevant characteristics (in practice this usually means knowing the sampling frame breakdown), and also that we know the response rates of the relevant groups in the sample.

There are two ways of doing this – prospectively and retrospectively.

To do it prospectively, we would establish differential response rates from last year's survey, and over-sample significantly under–represented groups in this year's survey. (It is important that we set some threshold for identifying under-represented groups, by using a statistical significance test such as chi squared). This method assumes that last year's response rates will be repeated this year, so it may not work.

To do it retrospectively, we would identify under-represented groups at the end of each quarter, and boost the sampling in the light of this under-representation. This will be more sensitive to variations in the current response rate, but is more cumbersome administratively.

Whichever way you do it, you should only boost groups which are significantly under-represented; to determine this you should use a technique such as the chi squared goodness of fit test. **Members of the booster sample should be selected using probability methods.** Finally, you should do everything you can to encourage accurate and complete recording of victim details for the sampling frame.

Quota sampling

Quota sampling is an example of non-probability sampling, and is widely used as an alternative to random sampling, especially when the absence of a reliable sampling frame means that random sampling is not feasible.

Quota sampling is not appropriate for User Satisfaction Surveys, where sampling frames are available, and is not permitted under the Home Office Guidance. It has been included in this discussion for clarification.

The method can be described in three stages.

1. Describe the population statistically, using frequency distributions by two or three key demographic variables.
2. Design a sample which comprises proportionate quota for each combination. For example, the percentage of the sample who are 'white males aged 16-24' would be the same as the percentage in the population.
3. Fill the quotas by identifying respondents fitting the demographic profiles.

If it is done carefully, quota sampling gives similar results to random sampling. However, 'doing it carefully' relies on having trustworthy and conscientious interviewers, and the method is difficult to audit.

Because quota sampling is not a probability sampling method, confidence intervals should not be used. **Furthermore, probability sampling should be used throughout the survey process, and carried through to completion, and not stopped when numerical quotas have been achieved.**

Not sampling

Sometimes the required sample size (worked out according to standard formulae) will exceed the total population size. In this case, you may approach every member of the population (or sampling frame).

Technically, this is not sampling at all, and any shortfall in the 'sample' will be due to non-response, which is a source of systematic rather than random error.

Dealing with Random Bias – Confidence Intervals

The confidence interval is a margin of error which gives us the likely range of a population parameter on the basis of a random sample. It allows us to express the uncertainty of a sample result arising from sampling error.

The theoretical basis for the confidence interval is the Central Limit Theorem, which tells us that the 'sampling distribution'[1] of a parameter is normally distributed, with a standard deviation (called the 'standard error') inversely proportionate to the sample size, provided the sample is random, and larger than 30.

The standard error for the percentage result, P, from a random sample of size n is given by the following formula.

[1] The sampling distribution is the theoretical frequency distribution of the results from repeated random samples of the same size.

$$\sqrt{\frac{P(100-P)}{n}}$$

The confidence interval is worked out by multiplying the standard error by 1.96[2], and adding the result to and subtracting it from the survey result. This is summarised by the following formula.

$$P \pm 1.96\sqrt{\frac{P(100-P)}{n}}$$

The two resulting figures give us an upper and a lower limit within which we are 95% confident that the 'true' (i.e. population) result lies. This is known as the 95% confidence interval, which is the most widely used, and stipulated in the Home Office Guidance for User Satisfaction Surveys.

Worked example

If a satisfaction rate of P = 68% were achieved from a sample of n = 418, the confidence interval would be

$$68 \pm 1.96\sqrt{\frac{68(100-68)}{418}}$$

$$= 68 \pm 1.96\sqrt{\frac{68(32)}{418}}$$

$$= 68 \pm 1.96\sqrt{5.2057}$$

$$= 68 \pm 1.96(2.2816)$$

$$= 68 \pm 4.4720$$

Thus we would be 95% confident that the satisfaction rate for the population would lie somewhere between 63.5% and 72.5% (i.e. 4.5% either side of 68%).

Finite Population Correction, FPC

The preceding discussion of confidence intervals makes no reference to the size of the population. But, as you would expect, population size does make a difference to the confidence interval. However, if the sample is less than about 2% of the population, the effect is lost if the survey result is rounded to one decimal place. This can be seen in the following table, which shows the confidence interval arising from a sample of 200, with a satisfaction rate of 65%, taken from various population sizes.

[2] This figure is a constant for the 95% confidence interval.

sample	population	sample %	result	CI to 3 d.p.s	C.I. to 1 d.p.
200	20,000,000	0.001%	65%	±6.610	±6.6
200	2,000,000	0.01%	65%	±6.610	±6.6
200	200,000	0.1%	65%	±6.607	±6.6
200	20,000	1%	65%	±6.577	±6.6
200	2,000	10%	65%	±6.271	±6.3

To take account of the population size, we use the Finite Population Correction, or FPC, which modifies the standard formula for the confidence interval given above. For a percentage result P, from a sample of size n selected randomly from a population of size N, the 95% confidence interval is given by the following formula.

$$P \pm 1.96\sqrt{\frac{P(100-P)}{n}\left(1-\frac{n}{N}\right)}$$

As explained above, the FPC only really needs to be used when the sample is more than two per cent of the population. However, with User Satisfaction Surveys this will usually be the case. Furthermore, because of the ease with which the formula can be written into a spreadsheet, it can be included as a matter of course.

The formulae used in the Home Office Guidance incorporates the FPC, as does the Sample Size Calculator.

Response rates and confidence intervals

Suppose we draw a random sample (approached sample, that is) of 700 from a population of burglary victims, and invite them to take part in a telephone survey. Further suppose that 420 of them agree to take part, giving us a response rate of 60%, and that the results show a satisfaction rate of 70%.

To remind ourselves, the purpose of this exercise is to use the sample result as an estimate of the population parameter. Technically, this is called inductive inference, moving from what we know to be true for the sample to what we think might be true for the population. And, of course, nothing is certain, so we hedge the inference with a margin of error, the confidence interval, which in this example would be ±4.4% (leaving aside the Finite Population Correction for simplicity's sake).

But are we entitled to draw this inference about the population? Remember, only 60% of those we approached agreed to take part. What if this statistic reflects a second population parameter – the tendency to respond to the survey? If this second parameter is related to the one in which we are primarily interested (in this case, satisfaction with the police) then there will be a systematic bias, undermining the validity of the conclusion.

Strictly speaking, therefore, we are only entitled to comment on that segment of the population who might be referred to as 'survey responders'. (We are assuming that the tendency to respond is distributed in the sample in a similar way to the

population. Responding then becomes a theoretical parameter in its own right, and should properly have its own confidence interval, but let's not get into that!)

The views of the 280 non-responders are, of course, unknown, by definition. But let us look at the implications of the two extreme positions.

First, suppose they had all been 'satisfied'. In that case the satisfaction rate would not have been 70%, but 82%.

Secondly, if all the non-responders had been not satisfied, the overall satisfaction rate would be 42%. (It's never that clear cut, of course, but it shows the possible range of variation.)

The important conclusion to arise from this is that confidence intervals cannot adjust for systematic error, only for random error. In practice there are various reasons for non-response, not all of which will produce a serious bias, and some of which may cancel each other out. But these are unknowns. To take account of this, we should collect information on reasons for non-response, and in the light of this speculate on the net effect of response bias, and report on it accordingly. This can include recording demographic characteristics of non-respondents and weighting the results to give a picture of what the results would have been from a more representative sample.

Sampling Plan – Worked example

The following hypothetical example is for a systematic sample, stratified by ethnicity, for the year 2006/07. The same principle can be used to stratify samples by BCU or crime type.

Information

Suppose in 2005/06, we had 6,394 victims of burglary, of whom 1,023 (16%) were BME (call this stratum 1), and 5,371 (84%) were white (stratum 2).

Conclusion

In 2006/07, we need a sample of 400 burglary victims, and we want our sample to reflect the ethnicity of the population. We therefore decide to stratify the sample.

If we assume that burglary victims in 2006/07 will show the same ethnic breakdown, this means we need 64 BME respondents (16% of 400), and 336 W respondents (84% of 400).

Of course, these are achieved samples, and we will need to over sample. The calculations for over sampling are complicated by the fact that (let us suppose) the response rates in 2005/06 were different for the two strata.

Information

In 2005/06, the response rate for BME was 52%, while for W it was 64%.

Conclusion

Assuming we have the same response rates this year for the two groups, we will need to over sample accordingly.

For BME, we need to approach (64 × 100 ÷ 52) = 123.

For W we need to approach (336 × 100 ÷ 64) = 525.

A further complication arises from the need to sample across the four quarters of 2006/07.

Conclusion

Our approached sample size for BME burglary victims is 123 (see above). Therefore we need to approach 31 per quarter. (This actually gives us a total of 124, but a variance of one or two can be treated as negligible.)

Similarly, our approached sample for W burglary victims is 525, so we need to approach 131 per quarter (giving us a total of 524).

Unfortunately, we do not know how many burglaries are going to happen. The best guide is likely to be the number that happened last year. We can then work out the appropriate sampling fractions.

Information

To remind ourselves, last year there were 1,023 BME and 5,371 W burglary victims.

Conclusion

We need to approach 124 BME burglary victims. Therefore we should approach every 8th burglary victim (1,023 ÷ 124 = 8.25). Starting with a random number between one and eight[3], we would approach every eighth BME burglary victim, and invite them to take part in the survey.

Similarly, we need to approach 524 W burglary victims. Therefore, we need to approach every 10th W victim (5,371 ÷ 524 = 10.25). We would start with a random number between one and ten, and approach every tenth W victim.

This gives us the following sampling plan for the year:

In theory…		Quarter 1	Quarter 2	Quarter 3	Quarter 4	Year
BME	*approach*	31	31	31	31	124
	achieve	16	16	16	16	64
W	*approach*	131	131	131	131	524
	achieve	84	84	84	84	336

[3] It is particularly important that we start with a random number, otherwise it will not be probability sampling.

This plan is a starting position, and is based on the following assumptions:

- that burglary victims in 2006/07 will show the same breakdown by ethnicity as in 2005/06;
- that this year's response rates by ethnicity will be the same as last year's;
- that there will be the same number of burglaries this year as last year;
- that there is no significant (i.e. predictable) variation in the numbers of burglaries across the four quarters.

It is very unlikely that all these assumptions will turn out to be valid. Therefore, we should follow the plan, but monitor at the end of each quarter, and make adjustments in the light of what actually happens.

There are two dangers – that we end up not achieving our sample size, or that we end up with too big a sample size, which we do not want to do for reasons of cost. Similarly those two dangers could affect the either BME or W sub-sample separately, giving rise to distortions in the findings.

Let's see what happens in our hypothetical example at the end of the first quarter.

Review at end of first quarter

Let us suppose – as is likely – that things do not turn out as planned!

Information

In the first quarter, there were more burglaries than expected committed against BME victims, so our sampling plan yields more than the planned 31; furthermore, a higher response rate of 66% in the first quarter for BME victims means we end up with 25 interviews, rather than the planned 16.

For W victims, there were fewer crimes recorded, so our sampling plan yields fewer than the planned, so that with a similar response rate (62%) we have under sampled.

This is summarised in the following table.

In practice…

		Quarter 1	Quarter 2	Quarter 3	Quarter 4	Year
BME	*approach*	38				124
	achieve	25				64
W	*approach*	106				524
	achieve	66				336

We now must adjust the sampling plan.

Conclusion

We still wish to achieve a sample of 64 BME burglary victims, but we already have 25. Therefore we need another 39, which neatly divides into 13 per quarter.

The over sampling formula should be changed in the light of the higher response rate achieved in the first quarter of 2006/07. Remember, last year the BME response rate was 52% across the year; the latest quarter's response rate was 66%. We would combine these rates as an average, but giving last year's response rate four times the weight of the response rate for the latest quarter.

This gives us a composite response rate of 55% – [((4 × 52) + 66) ÷ 5]. Therefore we will need to over sample by (13 × 100 ÷ 55) = 24.

Similarly for W victims, we need 336 interviews, we already have 66, therefore we need another 270, or 90 per quarter. The composite response rate for the latest quarter plus last year is 64% – [((4 × 64) + 62) ÷ 5]. Therefore we need to approach (90 × 100 ÷ 64) = 141 per quarter.

This gives us the following revised sampling plan for the remaining three quarters of 2006/07.

Revised plan after Quarter 1 …

		Quarter 1	**Quarter 2**	**Quarter 3**	**Quarter 4**	**Year**
BME	*approach*	38	24	24	24	110
	achieve	25	13	13	13	64
W	*approach*	106	141	141	141	529
	achieve	66	90	90	90	336

Similar adjustments will be made at the end of Quarters 2 and 3, using the same principles and methods.

This is far from perfect, of course. For example, it still gives a BME/W breakdown for your sample based on last year's burglary figures, which may be inaccurate in the light of what happens in 2006/07. However, further refinements could be made to the method to take this into account as well. The requirement to do this needs to be clearly communicated to the contractor; furthermore, a mechanism will need to be established for monitoring compliance with the requirement.

Ultimately, any method that depends on using the past to predict the future will be imperfect, giving only an approximation. However, it should help to achieve a reasonable compromise between the dangers of not achieving your sample and of costly over sampling, and in a way that gives you a reasonable chance of achieving a representative demographic split.

ANNEX C

CALCULATING BCU SAMPLES SIZES

A sample size calculator has been developed to simplify calculating sample sizes for BCUs.

The spreadsheet calculates the sample sizes that need to be achieved (i.e. the number of valid responses needed, not the number of attempts to contact) in 12 months at BCU and Force level.

These calculations are based on the number of crimes and incidents which took place in each BCU over a 12-month period and estimates of the likely percentage of exclusions.

The spreadsheet calculates the achieved-composite-sample size needed to give an 8% width (i.e. +/- 4%) with 95% confidence presuming a 50:50 result, and then applies the finite population correction. The spreadsheet allows you to enter other response ratios, and will calculate required numbers of responses accordingly.

The achieved-composite-sample is then divided amongst the 3 service-user types (burglary, vehicle crime and violent crime) in proportion to the mix within the composite population. For example, if 25% of the adjusted composite population is burglaries, then 25% of the achieved-composite-sample will be burglaries.

Where the sum of the BCU samples will not meet the Force requirements for representative samples for the user groups, the BCU samples are "topped up" automatically by the spreadsheet calculator. These top ups are distributed across the BCUs in proportion to the amount of that crime/incident occurring in the BCU. For example if a top-up of 100 burglaries needs to be added, and one BCU has 25% of the volume of burglaries, then that BCU sample of burglaries will be topped-up by 25.

The sample size calculator also allows you to generate a force total for numbers of racist incident responses.

Notes about the spreadsheet:
The Sample Size calculator is built in Excel. It contains macros, and these need to be enabled. The latest version is v3.1 (March 2011). The spreadsheet has been developed by Lancashire Constabulary. Many thanks to Alan Tattersall and Peter Langmead-Jones for their work to produce this.

The spreadsheet requires forces to input the number of crimes/incidents which occurred in each of the user groups over 12 months. It also asks for an estimate of the proportions of crimes/incidents that will need to be excluded on the groups outlined in Section 4. As a guide, Lancashire estimate that they will need to exclude 25% of domestic burglaries, 30% of vehicle crimes and 45% of violent crimes.

~~The sample size calculator – and all the rest of the guidance documents – can now be downloaded at:~~

~~*www.police.homeoffice.gov.uk/performance-and-measurement/assessment-methods*~~

A GUIDE TO SURVEY ANALYSIS

by Malcolm Hibberd

This guide sets out guidance for a methodology for analysing quality of service surveys. It assumes the analyst has a basic knowledge of statistical methods. It also provides guidance on how to carry out the recommended analysis using SPSS, the package that is most widely used in survey analysis in police forces.

The purpose of analysis

Quality of service surveys provide a way of assessing the service delivered to certain specified groups of 'customer'. Satisfaction rates are the familiar performance indicators derived from these surveys. They provide a useful proxy for what we are interested in. But they are no more than a proxy. They do not stand alone, but need to be interpreted.

Why the need to interpret? To answer this question, we need to make a basic and very important distinction between performance *measurement*, and performance *management*.

The enterprise of measuring performance should never be seen as an end in itself. Whether looked at from the external viewpoint of accountability, or the internal viewpoint of management, the ultimate purpose of performance measurement should be to assist efforts to improve the services provided by an organisation, or to maintain those services if they are judged to be at an acceptable level.

Strategic objectives – the outcomes an organisation strives to achieve – are measured by performance indicators. This is performance measurement. To achieve objectives, staff in an organisation must engage in the activities that are likely to lead to those outcomes. This depends on knowing what those activities are. This is performance management.

The analysis of quality of service surveys should serve both performance measurement and performance management. It should serve performance measurement by processing questionnaire data in order to provide performance indicators which are comparable over time and between police forces or basic command units. It serves performance management by providing the knowledge and understanding to inform decisions about action aimed at improving service delivery (as distinct from customer satisfaction). It does this by attempting to identify and put in context the outputs that are likely to have the greatest influence on the desired outcome. Hence the need to interpret.

This suggests two objectives for analysis: first, to provide performance indicators, thereby serving performance measurement; secondly to provide knowledge and recommendations for improving service, thereby serving performance management.

This guide set out a methodology for achieving these two objectives.

Types of question

The quality of service questionnaire has been designed partly to reflect the strategy for analysis outlined in following sections. This means that different questions serve different purposes. This should determine how they are used in analysis.

There are four basic types of question – outcomes, outputs, context and demographics. Each type is described below, with reference to how they should be used in the analysis of a survey.

Bear in mind that a question can serve more than one purpose. For example this question

> Firstly, did you contact the police about this incident yourself?
> YES/NO

acts as a filter to identify respondents who do not have to answer certain questions; it also serves as a contextual question (see below). This section only discusses the analytic functions of questions.

Outcomes

These are the questions that attempt to capture the respondent's evaluation of the service they have received. They provide the performance indicators, which are the starting point for attempts to assess quality of service, whether it has risen or fallen over time, and whether it varies between areas.

The outcomes have been set up as seven point scales, to maximise the scope for the expression of satisfaction, without threatening the reliability of the measure.

An example of an outcome question is shown below.

> Are you satisfied, dissatisfied or neither with the actions taken by the police?
>
> Completely satisfied
> Very satisfied
> Fairly satisfied
> Neither satisfied nor dissatisfied
> Fairly dissatisfied
> Very dissatisfied
> Completely dissatisfied
> Don't know

These questions will provide the headline figures for performance reports in this area. They also present the biggest challenge to the analyst, who should aim not only to assess satisfaction, but to explain it in terms of other questions, particularly in terms of outputs.

Outputs

The idea of using measures of satisfaction as a performance indicator is premised – albeit implicitly – on the assumption that it is an outcome which the police can influence. (If the police have no influence on satisfaction, it cannot reflect police

performance – we may just as well use a measure of rainfall as a performance indicator.)

It is reasonable to suppose (and there is already substantial evidence to support this) that the police influence satisfaction through the actions they take (or fail to take) when dealing with a victim of crime, visitor to the front office, or any other type of 'customer'. As a shorthand, we can call these actions outputs.

The questionnaire contains several output questions, which together amount to a summary of the qualitative nature of the service received by any customer. Consider the following, taken from the questionnaire.

Thinking about the actions taken by the police officers and staff who dealt with your incident once they had the initial details.

Did they…

Provide you with a reference number? YES/NO/N.A./D.K

Provide you with a contact name and number for someone dealing with your case? YES/NO/N.A./D.K

Offer contact details for Victim Support? YES/NO/N.A./D.K

These are three output questions, capturing for each respondent whether the police performed a particular action. Note that they are not in themselves an assessment of service, but factual questions[4] about what the police did.

Note also that for the purposes of hypothesis testing (see below), we would normally exclude the *not applicable* or *don't know* responses[5]. This means that each question is effectively a dichotomy, defining two groups of respondents for whom the action was either taken or not taken. Most of the output questions in the questionnaire give dichotomies in this way, although there are exceptions, such as the following, which has six options, excluding *don't know/can't remember*.

How long did it take for them to arrive?

Within 10 minutes
11-30 minutes
31 minutes – 1 hour
1 – 4 hours
More than four hours
At the agreed or appointed time
Don't know / Can't remember

[4] Strictly speaking, they are not factual questions, as they represent facts *as perceived or remembered* by the respondent – sometimes called pseudo-factual questions. It may be that the officer did, for example, offer contact details for Victim Support, but a respondent who did not remember this happening may answer 'no'. This should be borne in mind in the interpretation and reporting of the findings.

[5] That is not to say that the *don't know* responses are unimportant. There is evidence to suggest that a high level of *don't knows* is indicative of poor quality of service. Furthermore, for certain outputs – e.g. whether an offender is arrested – the *don't know* response is associated with lower levels of satisfaction.

In analysis and reporting, output questions should be treated in two ways. First, the percentage of respondents who received each output should be reported. This allows service delivery to be described, and monitored over time. Secondly, they should be analysed against satisfaction (outcome) in order to identify reasons for variations in satisfactions.

Context

Understanding satisfaction (as opposed to merely describing it) is not simply a matter of crosstabulating output questions with outcome questions. The relationship between service and satisfaction may depend on the context in which the service is delivered. This context must be taken into account in the analysis.

There are several contextual questions in the questionnaire which describe certain basic features of the situation in which the service was delivered. An example is shown below.

After you had reported the details was your incident dealt with

Entirely over the 'phone

By an officer or other member of staff visiting you at home
or attending the crime scene

At the police station

Other

Suppose we find about half of respondents said that the person who dealt with them *explained what was going to be done and why* (output), and that there was a strong link between this and satisfaction (outcome). Such a finding would have clear practical implications. But before making any recommendations for action, we should find out if the delivery of the output varies by context. If we were to find, for example, that *explanations of what was going to be done and why* tended not to be given to respondents dealt with by telephone or at the police station, but were always given by officers visiting, then this would sharpen up our recommendations, and arguably make them more useful. In this example, the question shown in the box above provides the context for understanding the relationship between the output and the outcome.

(We may also be interested in finding out if there are differences in satisfaction between respondents who were dealt with in different ways, which means this question can also be seen as an output. As explained above, a question may serve more than one purpose.)

Demographics

Demographic questions allow us to classify people into groups, independent of the subject matter of the survey. Typical examples are so familiar, that we need do no more than list sex, age and ethnicity as the three demographic questions asked in the questionnaire.

We should, however, give some thought to their functions in analysis.

Their first main function is to provide a simple, high level description of the sample. This is important in its own right, and also has two derivative purposes. The first

applies only if we have corresponding data on the population from which a sample is drawn. If we do, then we can check whether the sample is representative of the population. The second is to compare samples over time, or between areas. Where outputs or outcomes are associated with demographic variables, changes in the constitution of the sample can help explain variations in patterns of service or satisfaction.

The second main function served by demographic questions is formally to establish whether there are significant variations in service or satisfaction by demographic groups. This is important for equality of service delivery.

The third main function is contextual. A link between an output and an outcome may vary according to demographic group. For example, whether or not a victim was given a referral to Victim Support may have a greater influence on satisfaction of older respondents.

Approaching analysis – a three stage model

Analysis is best approached in three stages. The first stage is superficial; the second and third stages take you deeper into the data. And the deeper you go, the more likely the results are to have practical implications.

Stage One – Description

This is the most basic level of analysis, and is in truth so superficial that it doesn't really merit such a grand word as analysis at all. It consists of the production of descriptive statistics, summarising the answers given by members of a sample to each question in turn. These descriptive statistics will be in the form of frequency distributions – simple head-counts, reporting on how many respondents gave each answer.

This will yield the performance indicator questions, which will be a simplified frequency distribution, derived by collapsing together the first three categories of the seven-point scale (*completely*, *very* and *fairly satisfied*), and expressing this as a percentage of the valid sample size. This may also be supplemented with the result of collapsing the first two categories (*completely* and *very satisfied*).

There may also be some breakdowns of key indicators by the main demographic variables, and also by BCU within forces. However, since the purpose of such breakdowns is to establish whether there are differences (e.g. between male and female respondents, or between BCUs), this should really be seen as hypothesis testing, which is dealt with under the second stage of analysis.

Key indicators should be accompanied by a confidence interval, giving a margin of error for the result. Although the sample sizes have been designed to give a confidence interval of 4%, do not assume that this is what the confidence interval will turn out to be. This is because the sample size has been set on the assumption of an underlying indicator level of 50%; where the result is different from this, so too will the confidence interval be different. (Also, if the sample size turns out to be larger than that recommended by the guidance, this too will change the confidence interval.) The confidence interval should be worked out using the Finite Population Correction.

The end product of description will be a report (or section of a report) which presents frequency and percentage frequency distributions on

- sample characteristics
- delivery of service (outputs)
- performance indicators (outcomes)

This will be presented at force level, and will also be likely to be broken down by BCU.

Stage Two – Hypothesis testing

In the second stage of analysis we are concerned with the possible relationships between patterns of answers to two or more questions. To do this, we identify hypotheses about possible relationships.

Broadly, these will be of two types. First, to establish whether or not there are consistent differences in service delivery or satisfaction according to basic sample characteristics – including demographic and geographical (such as BCU) variables. The sort of hypotheses here would be in the form of

- does overall satisfaction vary by BCU?
- are female victims more likely to be given reassurance than male victims?

The second type of hypothesis is to establish whether there is a link between what the police do, and how satisfied the respondent is. It is here that we link policing outputs (activities, elements of service delivered) with the outcome of satisfaction (i.e. the performance indicator). The sort of hypotheses here would be in the form of

- are respondents more satisfied (OUTCOME) if they are told what will happen next (OUTPUT)?
- does satisfaction (OUTCOME) vary with response time (OUTPUT)?

The purpose of the analysis here is to identify the predictors of satisfaction, and thereby bring satisfaction more under the influence of police managers.

There are different techniques for testing the hypotheses, some of which are discussed below. However, as a general principle, the statistics used should tell us two things – first, whether the pattern is statistically significant (which means whether the finding is reliable, or if it is merely a random pattern, or coincidence), and secondly how strong the relationship is, which gives us an idea of the relative importance of different outputs.

Always remember, however, that the outcome of a hypothesis test is merely a statistical pattern. If we are to work through to its practical implications, we need to understand why there is a statistical link, which takes us on to the third stage.

Stage Three – Explanation

Explanation is the culmination of analysis. It should put us in a position of saying what a finding means. This is the most important part, as it is likely to lead to more informed decision-making, leading in turn to improvements in police performance,

thereby providing a bridge from performance measurement to performance management.

The results of a statistical test of a hypothesis will be a pattern – for example, a statistical relationship between two variables, which we may be confident is not random, and whose strength we will have assessed.

We are now concerned with why there should be such a link. This is the part that involves the hardest thinking: everything up to this point has either been simple (Stage One) or more sophisticated (Stage Two) number crunching. Stage Three is real analysis.

So how do we do it? I find it best to approach it in two stages – opening up, and closing down.

Opening up Starting with the results of a hypothesis test, we consider possible explanations for what it should mean.

Suppose we look at this question

> **Have you had any contact with the police about this crime/collision since your report and any immediate police actions?**

Crosstabulating it against satisfaction, suppose we find that those who had had contact were significantly less likely to be satisfied than those who hadn't. This is a surprising, and possibly disturbing finding, and it has potential practical implications for policing. (*It is also a real finding, from a recent survey of victims of violent crime.*) What could it mean? I can think of at least three different possible explanations for this finding:

- that respondents found it difficult to make subsequent contact;
- that the staff who dealt with the respondents were unhelpful, or otherwise wanting in manner;
- that contact was made to chase up information that should have been provided.

These are all plausible explanations for the finding. You may be able to think of others, but that is not the point. That point is that each of these explanations has different implications for police action – the first explanation suggests we should change the management of communications, the second suggests a change in the social skills of the staff who deal with subsequent enquiries, the third suggests that we should pre-empt the chasing up by providing the necessary information promptly. What action we take therefore depends not on the finding, but on what the finding means. Good analysis allows us to identify deficiencies in the service we provide, and at the same time suggests what can be done to put things right.

Closing down Each of the three options above is only a suggestion – each represents what *might* be the explanation of the finding. The task now is to close down on the options – to reach a (more or less confident) conclusion about what the most likely explanation is.
To do this, we need to bring more information to bear. Some of the information we need may be available within the questionnaire. For example, another question asks

> **[Would you say] You were kept informed of progress only after asking**

This can then be combined to begin to resolve the uncertainty. Suppose we look at those who had needed to contact the police again, and compare those who were satisfied overall and those who were not. If we find that those who were dissatisfied were more likely to say they were kept informed only after asking, then this suggests that the third explanation, namely

- that contact was made to chase up information that should have been provided.

may be the correct one. On the other hand, if we find that those who were dissatisfied were not more likely to say they were kept informed of progress only after asking, then this would tend to suggest that this explanation is unlikely to be true.

If the necessary additional questions have not been asked, it might be worth considering including extra questions to allow more detailed analysis. Other sources of additional information to help resolve competing interpretations are responses to open-ended questions in the survey, and the separate exercise of running focus groups.

You should also be interested in explaining the absence of any pattern of difference. For example, in a recent survey of burglary victims it was found that there was no link between response time and satisfaction with response time. This is a markedly counter-intuitive finding. It was resolved by establishing that longer response times tended to be accompanied by a satisfactory explanation of why the police would not be there quickly.

Not all results of hypothesis tests will need such involved analysis. For example, if we find a link between satisfaction and the answer to this question

> **[Thinking about how you were treated by the police officers and other police staff who dealt with you, did they:] Treat you politely?**

we would probably take this finding at face value – being treated impolitely leads to dissatisfaction. It is a matter for the analyst's judgement to decide where the more detailed approach is needed; and it is a matter for the analyst's skill to carry out the analysis with thoroughness and rigour.

Statistical Principles

Any analysis that goes beyond simple description will usually be seeking to establish whether there is a relationship between specified variables. This involves hypothesis testing, as described in the preceding section.

One of the commonest instances of this – and certainly the most important practically – is the test of a link between a policing output, and the outcome of satisfaction. If we are to carry out such a test, we need to know two things:

- statistical significance, and
- strength of the relationship (or size of effect).

The following section explains these two ideas.

Statistical significance

We need to establish statistical significance so that we can be confident that the finding is reliable, as opposed to a chance (or random), occurrence. If a finding is statistically significant, this means that it is unlikely to be a random occurrence. Such findings are usually quoted with significance levels, which tell you the probability that the finding is random. So, for example, if a finding is significant to the 0.01 level, this means that there is less than a 0.01, or 1%, chance that it is random; the other way of looking at this is to say that we are 99% confident that the finding is non-random.

There are two important reasons for being cautious when using statistical significance.

First, statistical significance is not the goal of analysis. In fact, it can be as interesting – and certainly worth reporting – to establish that there is no significant relationship between two variables.

Secondly, the more significance tests we carry out, the more likely we are to identify random patterns as statistically significant. To guard against this, we should

- only test hypotheses that have been specified and thought through in advance, and
- in situations where we are carrying out large numbers of significance tests, we should use a more exacting criterion for statistical significance – for example, requiring significance at the 0.01 or even 0.001 level for a link to be established.[6]

Even when approached cautiously, however, statistical significance never tells the whole story – we also need to know the strength of the relationship.

Strength of relationship

The strength of the relationship is a separate issue from the statistical significance of the relationship. The simplest way to explain this is to consider the influence different police outputs have on satisfaction. Suppose we wish to examine the effects of the following two outputs on satisfaction:

- keeping the victim informed of the progress of the enquiry, and
- offering the victim a referral to Victim Support.

Both of these are likely to be positively related to satisfaction – in other words, our hypotheses would be likely to state that respondents who received each of the two outputs would be more likely to be satisfied. However, while both may show a statistically significant association with satisfaction, it may be that one of those factors has a stronger influence than the other. Let us suppose that

- those who are referred to Victim Support are *a little more likely* to be satisfied, while
- those who are kept informed of progress are *much more likely* to be satisfied.

[6] A more formal approach to this is to apply the Bonferroni correction, adjusting the effective significance level according to the number of tests to be performed.

In this scenario, being kept informed has the stronger effect on satisfaction. This is clearly important for performance management: if a police manager knows the police activities that are most likely to influence satisfaction, he or she is in a better position to plan activity aimed at improving satisfaction.

Unfortunately, statistical significance alone does not tell us this. (The reasons for this are technical. In brief it is to do with a complex interaction between statistical significance, size of effect, and size of sample. The smaller the effect, the larger the sample that is needed in order to register it as statistically significant. With large samples, quite small effects will register as statistically significant.) Therefore we also need some statistical measure of how strong the effect is.

Summary

When testing the link between two variables we need

- to test for statistical significance, so that we can be confident that the finding is not merely a random pattern, and
- a measure of the strength of relationship to identify the most important influences, which in turn should inform performance management.

Statistical Techniques

As any statistics text book will demonstrate, there is a bewildering range of statistical methods which can be used to test hypotheses. In this section we discuss three basic methods. However, they are by no means comprehensive – merely a selection of techniques that are likely to be useful.

1. Chi-squared

The chi-squared test is one of the most familiar statistical significance tests. It is usually applied to a bivariate frequency distribution (also called a cross-tabulation, or contingency table) to establish whether there is an association between the two variables.

SPSS note

*Chi-squared is available through **ANALYZE – DESCRIPTIVE STATISTICS – CROSSTABS – STATISTICS**.*

The output table for chi-squared gives several values. In most circumstances, you should use the first one – called Pearson Chi-Square – and evaluate the probability in the right hand column: if it is less than .050, the association is significant to the 0.05 level; if it is less than .010, it is significant to the 0.01 level; and if it less than .001 (effectively, shown as equal to .000), it is significant to the 0.001 level.

If the crosstabulation is a two-by-two table (i.e. two rows and two columns), then SPSS will automatically produce an additional chi-squared value, labelled 'Continuity Correction'; this is the one to use for a two-by-two table.

The interpretation of the probability is quite straightforward. If the value given is less than 0.05, then the association between the two variables is significant to the 0.05 level, which means that we can be 95% confident that it is not random. If – as will often be the case – you see the value 0.000, then we can say it is significant to the 0.001 level, and we are 99.9% confident that it is not a random pattern.

You should also establish that the table satisfies what is sometimes called the twenty per cent rule – that is, that no more than twenty per cent of the expected cell frequencies are less than five, and that none of them is less than one. A message at the foot of the chi squared table tells you this.

If this rule is satisfied, then you may safely use the chi squared value. If it is violated, then the chi-squared value is invalid, and may not be used. You can get round this by excluding or collapsing categories, according to the circumstances.

In the analysis of customer satisfaction questionnaires, violations of the twenty per cent rule tend to arise from one of two circumstances, or a combination of the two.

The first is with a satisfaction scale, which typically shows relatively few people endorsing the individual dissatisfied options, which will tend to result in large number of low expected values. The solution is simply to collapse two or three categories together into a general dissatisfied category. While this loses the fine distinctions between *fairly*, *very* and *completely dissatisfied*, it allows us to produce a reliable analysis. (In any case, if there aren't many in each of the dissatisfied options, the differences between them would not be reliable.) Do bear in mind that collapsing the dissatisfied options does NOT mean that you also have to collapse the satisfied options. The other thing to remember (and this is commonsense, really) is the collapsing must produce an intelligible aggregate category, which makes sense in the context of the research.

The second common reason why the twenty per cent rule is violated is where one (or both) of the variables has a catchall option – such as *don't know*, or *not applicable* – which has been ticked by a small number of the sample. In that case, there may be a justification for excluding all the "*don't knows"* from the analysis. This is best done on SPSS by defining it as a missing value, under variable definition.

SPSS note

If the twenty per cent rule is violated, SPSS will still perform a chi-squared test if requested. However, there will be a message at the foot of the chi-squared table telling you what percentage of the expected values are less than five, and what the minimum expected value is.

If it is violated, you will need to establish which cells in the table are causing a problem. For this reason, it is always a good idea to ask SPSS to print the expected values on the table. This is done through the ***CROSSTABS*** *menu, via the* ***CELLS*** *option button: Observed Count is the default option; you should select Expected Count as well.*

Categories can be collapsed by recoding, which is available (in the data window only) through ***TRANSFORM – RECODE****; it is usually better to recode* ***INTO DIFFERENT VARIABLE****, as this will make the original, uncollapsed variable available for other forms of analysis, and the new collapsed variable may then be used repeatedly.*

*Categories can be excluded either by defining the appropriate value as **MISSING** in variable definition, or through (again in the data window only) **DATA – SELECT CASES – IF CONDITION IS SATISFIED**, and then writing a logical statement to select out the unwanted values.*

If the twenty per cent rule is violated on a two-by-two table, it is not possible to combine or exclude without losing the bivariate table. In that case, the Fisher Exact Probability test can be used. This is produced automatically by SPSS for a two-by-two table.

SPSS Note

When a chi-squared test is performed on a two-by-two cross-tabulation, SPSS will automatically provide a value for Fisher's Exact Test. This will only be used if the message at the foot of the table shows that the twenty per cent rule has been violated.

Two separate probability values are provided – two-sided and one-sided. If you have made a prior prediction of the direction of the association (which will normally be the case) you may use the one-sided value; otherwise, you should use the two-sided value.

The probability value itself is interpreted as for chi-squared.

Size of effect

Chi squared is merely a test of statistical significance; it does not tell you how strong the effect is. We therefore need to supplement chi squared with a measure of effect size. The best way is to use the coefficient Cramér's V or phi. These two are closely related; the difference is that phi is for two-by-two tables, while Cramér's V is for tables with three or more rows or columns. Both give you a number on a scale of zero to one – the higher the value, the stronger the relationship. This enables you to compare the likely influence of two or more outputs on the outcome of satisfaction – the one with the higher Cramér's V/phi value is the stronger relationship.

SPSS Note

*Phi and Cramér's V are available through **ANALYZE – DESCRIPTIVE STATISTICS – CROSSTABS – STATISTICS** (as with chi-squared).*

Phi and Cramér's V should be used only if the relationship is statistically significant.

Chi-squared – advantages and disadvantages

Chi-squared is familiar and can be used in a wide variety of situations. However, the cost of this versatility is that it is something of a blunt instrument.

The first problem is that while chi-squared tells you that there is an association *somewhere* in the table, it doesn't tell you exactly *where* it lies. There are two ways of resolving this:

- first, to establish this we have to examine the residuals (the magnitude of the difference between the observed and expected valued);

- second, for more complex tables, to partition the table to find the source of the difference (this involves splitting a complex table up into a series of two-by-two tables and testing them separately).

For large tables – i.e. with many rows and columns – the results can be difficult to interpret, and even harder to report or explain.

The second problem with chi-squared is that it is susceptible to the twenty per cent rule. While this can be overcome, it is at a cost – such as the loss of the distinctions between the levels of dissatisfaction.

Thirdly, chi-squared is not a particularly powerful technique. This means that it is not the most efficient method for testing statistical significance – other tests are better at detecting relationships.

Finally, when it is applied to a satisfaction scale, it fails to exploit the ordering of the categories. Technically, a satisfaction scale is ordinal (once any don't knows are excluded), but chi-squared treats the variable as a nominal scale, which is therefore wasteful of data.

2. Mann Whitney

The Mann Whitney test allows you to compare two groups on a measurement scale that is at least ordinal. It is ideal for testing for a difference in the level of satisfaction of a sample according to whether or not they have received a particular element of service.

SPSS Note

*The Mann Whitney U test is available under **ANALYZE – NONPARAMETRIC TESTS – 2 INDEPENDENT SAMPLES**, where it is the default option.*

*The test variable will be the satisfaction scale, and the grouping variable will define the two groups you wish to compare; the codes identifying the two groups must be entered under **DEFINE GROUPS**.*

The output comprises two tables. The first gives the mean rank for each group, and allows you to determine which group scores higher. A lower mean rank indicates lower scores (i.e. on a 7-point sale, code 1 is the lowest and 7 the highest).

The second table gives the result of the significance test. For large samples this is a normal approximation, with a Z score (number of standard deviations on a normal distribution). The probability is the last figure given, labelled 'asymp. sig. (2-tailed)', and needs some explanation:

- *if your hypothesis does not predict which group will be higher (e.g. if you were comparing satisfaction of men against women), this is a two-tailed (non-directional) test, and you would interpret the probability in the usual way – i.e. concluding that there is a significant difference if the probability is less than .050*
- *if your hypothesis does predict which group will be higher (e.g. you would normally predict that those who receive an element of service would be more satisfied than those who do not) then this is a one-tailed (directional) test, you should **halve** the probability before evaluating it against the usual criterion.*

Size of effect

There is no formal test for the size of the effect between the two groups, but it is possible to derive one from the mean ranks given in the SPSS output. Take the following example from a survey of victims of violent crime.

Worked example

The Mann Whitney U test established that initial satisfaction (the old BVPI23) was higher where the first police officer gave some practical help to the victim, and that this was statistically significant to the 0.001 level.

The mean ranks (taken from the SPSS output) were as follows:

practical help given	75.61
practical help not given	108.94

(The mean rank is lower for practical help given, because the coding of the scale ran from *1 = totally satisfied* to *7 = totally dissatisfied*, thereby indicating that those who were given practical help were more likely to be satisfied.)

The mean rank for *practical help given* is higher by 33.33 (108.94 – 75.61); this means it is 31% higher than the mean rank for *practical help not given*.

In the same survey, initial satisfaction was compared according to whether the first officer gave reassurance. Again, satisfaction was higher where reassurance was given, and this was statistically significant to the 0.001 level.

The mean ranks were:

reassurance given	79.43
reassurance not given	130.24

The mean rank for *reassurance given* is higher by 50.81 (130.24 – 79.43); this means it is 39% higher than the mean rank for *reassurance not given*.

The conclusion we would draw here is that *reassurance* appears to have a stronger effect (39%) on satisfaction than has *practical help* (31%).

Mann-Whitney – advantages and disadvantages

Compared to chi-squared, the Mann-Whitney U test has some distinct advantages, but also a limitation.

The first advantage is that it is a more powerful technique, which exploits the ordinal nature of the satisfaction scale (which, as we have seen, chi-squared does not do), and will pick up differences that chi-squared will fail to detect.

The second advantage is that it allows you to draw a more precisely focused conclusion – for example, that when reassurance is given, satisfaction is significantly higher. By contrast, chi-squared only allows the conclusion that there is an association between the two variables.

A third advantage is that there is no equivalent of the twenty per cent rule for Mann-Whitney, and hence no need to 'waste' data by combining categories.

The main limitation of Mann-Whitney is the flip-side of one of its advantages. Because it allows you to conclude that one group is higher than another, it only allows you to draw that conclusion. A significant relationship of a different nature may not be picked up by Mann-Whitney. For example, if the effect of an output is to polarise satisfaction – i.e. leading to people making more extreme judgements in either direction – Mann-Whitney would fail to detect this. The blunter – and consequently more versatile – chi-squared test will pick this up.

Because of this, it is a good idea to use both Mann-Whitney and chi-squared: Mann-Whitney to test the hypothesis that satisfaction is higher or lower; chi-squared to pick up any different patterns that may emerge.

Mann-Whitney can only be applied in a situation in which two groups are being compared on an ordinal scale. This means that if there are any questionnaire options that do not fall on a scale – a *don't know* option, for example – these must be excluded from the analysis.

Furthermore, Mann Whitney only allows you to compare the satisfaction ratings of two groups of respondents. If you wish to compare three or more groups, there is an equivalent test called the Kruskal-Wallis H.

SPSS note

The Kruskall-Wallis H test is available under ***ANALYZE – NONPARAMETRIC TESTS – K INDEPENDENT SAMPLES****, where it is the default option.*

The procedure is similar to that for the Mann-Whitney test, but instead of entering the codes to identify the two groups, the range of values (maximum and minimum). This means that if there are values within the range that you do not want to include in the comparison, they must be excluded, either by defining them as missing values, or by using ***DATA – SELECT CASES****.*

The output for Kruskall-Wallis gives you a chi squared value, and a probability for assessing statistical significance. Look for a value at least below .050, which would tell you that there is a significant difference between the groups; the pattern of difference must be established by looking at the table of mean ranks – low mean ranks corresponding to low scale values (i.e. on a satisfaction scale, 1 is low, 7 is high). Unlike Mann-Whitney, there are no circumstances under which this probability would be halved

3. Kendall's tau

Kendall's tau is a correlation coefficient which can be applied to questionnaire scale data. It allows you to establish whether there is a consistent link between two satisfaction indicators.

Unlike chi-squared and Mann-Whitney (which are significance tests which have to be supplemented with a measure of size of effect), Kendall's tau (like other correlation coefficients) is a measure of the strength of a relationship – the closer the coefficient is to one, the stronger the relationship. The correlation coefficient is supplemented by a test of statistical significance (using the t-distribution for a small sample, and z for a large sample).

SPSS note

*Kendall's tau is available under **ANALYZE – CORRELATE– BIVARIATE**; the Pearson coefficient, which is the default option, should be switched off, the **KENDALL'S TAU-B** should be selected.*

*If your hypothesis predicts the direction of correlation – which it is likely to – then you should select **ONE-TAILED** under test of significance; otherwise leave it as **TWO-TAILED**.*

Select the variables (at least two) that you wish to correlate. SPSS will carry out the correlation on every possible pairing of the specified variables.

The output gives a correlation matrix, with each selected variable appearing in both the rows and columns. For a correlation between a given pair of variables, locate one in the columns and the other in the rows (it doesn't matter which way round you have them), and the results will be at the intersection of the row and column.

The results show the correlation coefficient itself, followed by the probability (significant if less than .050), and the sample size. A significant correlation is shown by an asterisk.

*Kendall's tau is also available through **ANALYZE – DESCRIPTIVE STATISTICS – CROSSTABS – STATISTICS**, as with chi-squared. This offers you the choice of **KENDALL'S TAU B** and **KENDALL'S TAU-C**: if your table has the same number of rows and columns, you should use **TAU-C**; otherwise, use **TAU-B**.*

Kendall's tau is the ideal method to use for looking at the relationship between two questions each of which form an ordinal scale.

4. Assessing the significance of change

As explained above, the ultimate goal of measuring and analysing customer satisfaction is to improve service delivery. Changes in the levels of satisfaction may be used to establish whether there has been any underlying change in service delivered.

One way of testing the statistical significance of the change is to carry out a simple chi-squared test. However, a more powerful method is to apply the Z test (derived from the central limit theorem) to the difference between two percentage results.

The calculation compares single percentage results (effectively a dichotomy – such as the percentage at least very satisfied) between two samples. These may be

samples at different points in time, such as successive years, or from different geographical locations, such as one BCU against the force, or one force against its most similar forces. For each sample we need to know the percentage result on the indicator, the sample size, and (to adjust for the finite population correction) the population size. It also requires access to and familiarity with the table of the standard normal deviate, Z; these are available in the appendix of most statistics textbooks.

If you are attempting to assess whether satisfaction has changed, or if your force's satisfaction rate is different from that for the most similar forces, you would use a two-tailed test. If you are specifically predicting an improvement – for instance, if you are evaluating an initiative to improve service delivery, then you should use a one-tailed test. This determines the value of Z you use in the calculation.

The values of Z for the usual significance levels are set out for reference below.

significance level	*Z (one-tailed)*	*Z (two-tailed)*
0.05	1.65	1.96
0.01	2.33	2.58
0.001	3.08	3.30

SPSS note

This technique is not available on SPSS, but the calculation is simple, and it is easy to set up an Excel spreadsheet to perform the calculation.

Multivariate techniques

There are more advanced methods available for addressing some of the analytic problems of surveys, including multivariate techniques such as logistic regression. SPSS also offers CHAID (chi-squared automatic interaction detector), which carries out large numbers of chi-squared tests on survey data.

While there is scope for using these, I believe it is better to use more basic, bivariate methods, introducing additional (contextual) variables having thought through the possible interpretation of links between outputs and outcomes. In my experience, this gives the analyst a closer understanding of the statistical relationships, and puts her/him in a better position to make informed recommendations.

A note on reporting

This guide has dealt with some of the statistical technicalities of survey analysis. Very little of this is suitable for managerial consumption. The technicalities belong 'behind the scenes'.

The end product of analysis should consist of

- a summary of delivery of outputs and achievement of outcomes,
- an account of the factors that influence satisfaction, and
- recommendations for action.

This may best be presented as an executive report, which should make no reference to Mann-Whitneys, one-tailed tests, the twenty per cent rule, or other statistical

technicalities. However, the basic principles of statistical significance, confidence intervals and strength of relationship should be understood by readers, so that they may be confident in general terms about how conclusions have been reached.

A full report of the analysis should also be produced, with statistical technicalities where appropriate (but never for their own sake). Such a report would be made available on request or for reference.

EXAMPLE OF AN INTRODUCTORY LETTER

Date

Dear sir/madam

I am writing following your recent report of an incident to …….. We recognise the impact that an incident of this nature can have on those who suffer from it.

Your views on how you have been treated by ……. in relation to this incident are very important to us.

To enable us to monitor and improve our service, I have asked an independent research company, ……, to carry out a telephone survey with you. One of the researchers will try to contact you within the …… I would be grateful if you could take the time to answer their questions when they call you. You will not be asked any questions about the incident itself; only about how the police responded and how you felt about the way you were dealt with.

Your details have been passed to ……. for the purposes of this research only. Your details and the answers that you give to the questions will be treated with the strictest confidence and in accordance with Market Research Society Codes of Conduct. Under no circumstances will your details be passed to, or used for any other purpose. Your details will not be passed back to ……. unless you explicitly say so.

Thank you in advance for taking the time to participate in this important research.

Yours sincerely,

Name
Role
Department

ANNEX F

SERVICE RECOVERY TEMPLATE

Respondent Details
Name:
Telephone Number:
Gender:
Ethnicity:
Age:
Date of Incident:
Date of Interview:
BCU:
District:
SNT:
Crime Number:
Incident Type:

Results of satisfaction indicators

Ease of Contact	Fairly satisfied
Actions	Very dissatisfied
Follow-up	Completely dissatisfied
Treatment	Fairly satisfied
Whole Experience	Completely dissatisfied

REASONS FOR DISSATISFACTION:
Ease of contact

Knew how to get in touch with the police:	Yes
Easy to contact the police when first needed to:	Yes
Dealt with in reasonable time:	Yes
First person able to take all details:	No

Actions

Gave practical help:	No
Appeared to know what they were doing:	Yes
Explained what was going to happen and why:	Yes
Carried out investigation of scene:	No
Provided reference number:	Yes
Provided contact details:	No
Offered Victim Support details:	Yes
Offered advice:	No
Made further visits	No

FOLLOW-UP
Free Text Comments
TREATMENT
Free Text Comments
OTHER COMMENTS
Free Text Comments

ANNEX G

DATA PROCESSING AGREEMENT

Forces/authorities must ensure that the Data Protection Principles are adhered to when conducting surveys. The Data Protection Act states that all data must be processed fairly. This is set out in the first data protection principle. The situation of forces/authorities which contract their survey work to a third party has been clarified with the office of the Information Commissioner. Market research companies carrying out surveys on behalf of forces/authorities are acting as their processors of information.

The following needs to be taken into account when out-sourcing telephone surveys:

- While the Data Protection Notification makes provision for information to be passed back to a police force, the Market Research Society Code of Conduct (details at: *www.mrs.org.uk)* states that interviewers need the respondent's explicit consent before they can pass back any identifiable survey responses to the force/authority.

- The interviewer should make it clear to the respondent that their responses are extremely important to forces/authorities in maintaining levels of public service and must, therefore, seek consent to pass back data to them. If the respondent states that they are not happy with identifiable responses being passed back to the force/authority, they must be reassured that their responses will remain anonymous and the survey company must ensure their data is anonymised before being passed back to the force/authority.

- Responses can identify the respondent in a number of ways:
  - Name
  - Telephone number
  - Address or post code
  - Crime or incident number

- The Market Research Society states that partial postcodes (the first section) may be passed back but in some areas one postcode could refer to a single household, therefore full postcodes are not always available to be passed back.

- Some questions have been added to the survey script to cover the requirement to seek consent.

6.2 Forces/authorities should consider notifying users at the point when they first report the crime, or soon afterwards, that the information may subsequently be used for the purposes of research by the service. This notification could be done through the non-emergency call handling, by a reference in any victims' information leaflet, or through a follow-up letter or call.

6.3 The *data controller – data processor* relationship should be defined in the contractual arrangements. This meets the requirement of the Data Protection

Act. Survey companies should take care when introducing themselves to users to be clear that they are working on behalf of xxx police, and that the information gathered will be provided directly to the police, and used for no other purpose.

6.4. Some Forces have adopted the use of a Frequently Asked Questions (FAQ) sheet in order to explain some of the most frequent queries surrounding Data Protection. These sheets can be sent with the Victim of Crime letters and distributed to Controls Rooms & Switchboard. They can also be sent out to respondents who call in for clarification around Data Protection and provided to the Market Research Companies. An example of Gloucestershire's FAQ sheet is in Annex K.

6.5. Other Forces have a letter that is signed off by the Data Protection Officer explaining adherence to the Data Protection Act, this has been useful in appeasing dissatisfied respondents. A copy of this can be found in Annex J.

Any individual queries concerning data protection issues should be referred to the individual force Data Protection Officer.

This is an agreement that sets out the terms and conditions under which personal data held by the specified 'data controller' will be processed by the specified 'data processor'. This agreement is entered into with the purpose of ensuring compliance with the Data Protection Act 1998. Any processing of data must comply with the provisions of this Act.

The data controller for the purposes of this agreement is

1. **XXX Constabulary/Police**

The Force will be known as the **Data Controller** in this document.

The data processor for the purposes of this agreement is ……..?

Terms of Processing.

1. The data controller agrees to provide the data processor with the relevant data required for conducting telephone surveys under **Home Office** guidelines.

   The data controller agrees to disclose personal data to the data processor or instruct the data processor to process personal data on the understanding that the data processor will be processing that data on behalf of the data controller and in accordance with the statutory duty under which the data controller are empowered to collect, hold and store that data. The data processor will process the personal data only to achieve the following objectives:

   Achieve the number of interviews required under **Home Office** guidelines

   The data processor will not, by act or omission, place the data controller in breach of the Data Protection Act 1998.

2. **The data controller agrees to provide the data processor access to personal data under the following conditions:**

(a) The data processor will process personal data on behalf of the data controller purely for the purposes outlined in Paragraph 1 and will not retain or process personal data for any other purposes.

(b) Personal data will not be processed by the data processor to support measures or decisions with respect to particular individuals.

(c) Personal data will not be used by the data processor to approach any identifiable individual unless under instruction from the data controller

(d) Personal data will not be processed in such a way that substantial damage or substantial distress is, or is likely to be, caused to any data subject.

(e) Personal data will not be disclosed to any third party without the written authorisation of the data controller.

(f) Personal data will not be disclosed to any employee of the data processing organisation who has not signed the Confidentiality Agreement at Annex H.

(g) Any results and/or publications produced from the processing activity that contains personal data or personal identifiers will be made available ONLY to the data controller and not to any third party. Those results and/or publications will be retained by the data processor only for the period that the data processor is under contract to the data controller. At the end of the contract all results and/or publications will be passed to the data controller and no copies will be retained by the data processor.

(h) All personal data held by the data processor will be returned to the data controller or will be destroyed by a date to be agreed by the relevant parties. After this date the data processor must provide a written declaration confirming that all the data has been returned/destroyed.

(i) The data processor will not transfer the personal data outside the European Economic Area.

3. The data controller and data processor must ensure that arrangements for the transfer and storage of and access to personal data are accompanied by appropriate security measures to ensure compliance with the 'seventh data principle' (see page 44). In addition to the security measures outlined in section 2 above, the data controller must satisfy themselves that the data processor is taking reasonable steps to ensure compliance with the seventh principle. The completed questionnaire will be reviewed by the data controller who will undertake an inspection of premises if required.

4. To ensure the security of confidential data that is collected from the data controller, the following guidelines will be followed by the data processor.

Data Collection – Personal details of those to be interviewed will be sent to …… on disk which is password protected via recorded delivery. The password will be disclosed via email to ……who will be advised of the date of sending. ……. will then confirm that they have received the disk.

Data Storage – The disk will be kept in a safe at …… premises until all information has been used, the disk will then be returned to XXX Constabulary for disposal.

5. The data processor agrees to any staff that will be involved in the processing of the data controller's personal data being subject to vetting checks by the data controller. The data processor will be informed only of the employees' suitability to have access to data owned by the data controller.

6. The data processor agrees to the data controller monitoring compliance with this agreement.

7. The data processor undertakes and agrees to indemnify the data controller in the terms set out in Annex I.

Declaration

8. I agree to abide by the terms and conditions of this agreement. In doing so, I am aware of and understand the relevant provisions of the 1998 Data Protection Act, and I agree to abide by these provisions as specified by this Agreement.

Signature of XXX Constabulary/Police Date

Signature of the data processor Date

Terms and Conditions

The principal terms used in this agreement are based upon the definitions laid out in section 1(1) of the 1998 Data Protection Act.

'data controller' means, XXX Constabulary/Police who determines the purposes for which and the manner in which any personal data are, or are to be processed.

'data processor' in relation to personal data, means any person (other than an employee of the data controller) who processes the data on behalf of the data controller.

'personal data' means data which relate to a living individual who can be identified:

a) from those data, or
b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller,

and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual.

The Seventh Data Protection Principle – 'Appropriate technical and organisation measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data'.

ANNEX H

CONFIDENTIALITY AGREEMENT

For the purpose of the User Satisfaction Survey all directors and employees of ……..that have access to the data controller's data will abide by this confidentiality agreement.

1. You will only access/make use of the Consortium's data in connection with the work undertaken under the terms of the contract and in accordance with the Data Processing Agreement between …….and the data controller.

2. You will not disclose police information obtained by you during the term of the Contract to anyone (including friends and family) without the authority of the data controller. Where such authority is given, disclosure will only be made for a policing purpose and to persons entitled to receive it; you will ensure that enquiries are made with any person requesting disclosure to ensure that they are entitled to have the information.

3. You will keep all information to which you have access (computer and manual) secure. You will comply with the appropriate physical and system security measures for any information you have access to.

4. You may not copy any data or software.

5. You are required to comply with the terms of:

    a. Data Protection Act 1998
    b. Computer Misuse Act 1990

    These Acts outline your **personal liability** for any wilful or reckless act regarding misuse of information.

I have read, understood and accept the above. I accept that a breach of this agreement may result in the termination of this contract and potentially may render me personally liable for offences under the aforesaid legislation.

Name……

Signed……..

Date……

INDEMNITY AGREEMENT

1. In consideration of a party to this agreement (the Provider) providing information in accordance with the terms of this agreement and protocol to another party. The recipient hereby agrees to indemnify the provider against any liability, which may be incurred, by the provider as a result of the provision of the information.

Provided that this indemnity shall not apply:

(a) where the liability arises from information supplied which is shown to have been incomplete or incorrect, unless the provider establishes that the error did not result from any wilful wrongdoing or negligence on its part;

(b) unless the provider notifies the recipient as soon as possible of any action, claim or demand to which this indemnity applies, and permits the recipient to deal with the action, claim or demand by settlement or otherwise and renders the recipient all reasonable assistance in so dealing;

(c) to the extent that the provider makes any admission, which may be prejudicial to the defence of the action, claim or demand.

2. The recipient of any information provided under this protocol hereby indemnifies the provider against any liability which may be incurred by the provider arising from:

a. the recipient's disclosure of the information to any third party unless the provider has given permission to disclose,
b. any breach by the recipient of the terms of this Agreement.

DATA PROTECTION EXPLANATION LETTER

Dear

Following our telephone conversation yesterday I said I would write to confirm my response to your concerns about the use of your personal details.

I understand your concerns, particularly as I work daily with the legislation aimed at ensuring that organisations make fair and lawful use of peoples' personal information. This legislation is the Data Protection Act 1998 and more information about the Act and your rights can be found on the Office of the Information Commissioner website *www.informationcommissioner.gov.uk* or by contacting them on 01625 545745.

Your information was passed to [Market Research Company] under a contract requiring the company to conduct customer satisfaction surveys on behalf of the Constabulary. These surveys are an essential part of the Constabulary's responsibility and desire to ensure that we are providing a high standard of service to the public. The use of people's personal information for this purpose is covered on our notification under the Data Protection Act 1998 as 'Administration and Ancillary Support for Policing Purposes'.

The Data Protection Act enables an organisation to pass personal information to another company to use on their behalf but requires that there is a written contract specifying the limitations of use of the personal information. In addition to the contract securing the services of [Market Research Company] we also have a data processing contract, which deals specifically with issues relating to the use of personal information. This contract states that [Market Research Company] may use the information only for the purpose of conducting interviews on behalf of the Constabulary. It also states that the personal information is not to be used or retained by [Market Research Company] for any other purpose, that approved security measures are put in place to protect the data and that all [Market Research Company] staff having access to the information must sign a confidentiality agreement pointing out their personal liability for criminal prosecution should they misuse the information.

Responding to the survey is voluntary and the identity of who is conducting the survey is an important element in deciding whether or not to take part. I can confirm that the agreed script for [Market Research Company] staff has as its second sentence the words "I am calling from [Market Research Company] conducting a survey on behalf of XXX Constabulary".

Our **Citizen Focus Section in the Corporate Support Department** is the contact point for ensuring that your details are included on our register of individuals who do not wish to be contacted for the purpose of surveys.

I apologise for any concern that this has caused you.

Yours sincerely
Data Protection and Freedom of Information Manager.

DATA PROTECTION FREQUENTLY ASKED QUESTIONS

USER SATISFACTION SURVEY

Telephone surveys are conducted with victims of domestic burglary, vehicle crime, violent crime and racist incidents).

"Who is [Market Research Company**]?"**
XXX Constabulary/Police has commissioned [Market Research Company], an independent market research company, to undertake a telephone survey on our behalf.

"How did [Market Research Company] **get my details? and/or "My telephone number is ex-directory or Telephone Preference Service (TPS) registered"**

Members of the public who contact XXX Constabulary/Police are asked for their telephone number should we need to subsequently contact them about their crime or incident. This includes seeking feedback on the way we dealt with you. The telephone number continues to be used only for a purpose connected with the incident and XXX Constabulary/Police retains full ownership and responsibility for the information.

The requirement on [Market Research Company] for Data Protection Act compliance in respect of their processing of our information, which includes your telephone number, prevents the use of it for any purpose other than that within the terms of the survey contract. We can assure you that all confidential information is treated securely.

As the reason for contact is with regards to the crime or incident and not to sell a product then we are not in breach of any TPS legislation.

"Is this process legal?"
Yes. [Market Research Company] meet the requirements of the national framework contract developed by the Home Office and this contract complies with the Data Protection Act, Official Secrets Act and confidentiality rules.

"Why have I been asked to complete a questionnaire?"
XXX Constabulary/Police has a legal obligation to monitor and report on the service they provide to the public. One way to find out how satisfied people are with the service they received is to ask them. As someone who has reported an incident or crime to XXX Constabulary/Police your name has been randomly selected. We hope that you can provide us with first-hand knowledge of your experience.

"Can I be identified from my responses?"
I would like to reassure you that all telephone surveys will be conducted in the strictest of confidence. Your name will be **not** be recorded against any response you provide, unless you provide consent that you wish for your comments to be passed to XXX Constabulary/Police. Findings from the surveys are reported at a general level and **not** reported on individual experiences.

"Will my response make any difference?"
Yes it will. We report the findings and provide a summary of the information to senior officers who can and do make changes whenever possible in order to improve the service provided.

“Some of my details are incorrect”
Owing to the circumstances in which the initial details are recorded at the time of the incident, we apologise that *some* information may be inaccurate.

“Is this exercise a waste of money?”
Costs are kept to a bare minimum and all the information gained is fully used, so it is a cost-effective way of seeking to improve the service. The more responses we receive, the more worthwhile the exercise is.

For more information contact: *name and direct telephone number*