Guidelines for Community Odour Assessment

Report No 2706/1

March 1997

Lincoln Environmental
PO Box 84
Lincoln University
Canterbury
New Zealand

TABLE OF CONTENTS

Page

Page

LIST OF APPENDICES:

LIST OF TABLES:

LIST OF FIGURES:

acknowledgements

Lincoln Environmental acknowledges the financial support received from the following organisations in the production of these guidelines:

- Ashburton District Council
- Christchurch City Council
- Environment Waikato
- Hawkes Bay Regional Council
- Manukau City Council
- Ministry for the Environment
- NZ Fertiliser Manufacturers' Research Association
- The NZ Refining Co Ltd
- Wanganui District Council

Technical Terms

Assessment area An area defined for measuring and assessing odorant emissions. The area is defined according to the objectives of the investigation. For measurement purposes, the area is covered by a grid of equidistant measurement points.

Assessment square The grid measurement points are taken as assessment squares to show how the measurement results vary from place to place within the assessment area.

Bias The difference, averaged over many samples, between the sample estimate and the true population value.

Cluster sampling Items in the sampling frame are put into groups, so that the groups are as similar as possible. Groups often correspond to geographic areas. Some of the groups are then sampled. Typically this is not a very efficient method of sampling, but because all those in the sample live in just a few geographic areas, the costs of interviewing are lower than other methods.

Confidence interval The range of values within which the true population value lies for a certain (high) percentage of all possible samples.

Control zone A zone that is selected outside the environmental load area having comparable type and density of buildings but without any air quality load due to other odour emitters.

Correlation A statistic that measures the association between two variable. A correlation has the following properties: (1) a positive correlation implies that as one variable increases, the other variable also increases; (2) a negative correlation implies that as one variable increases, the other decreases; (3) a correlation can take values from -1 to 1, and a value close to zero implies that there is little association between the two variables.

Definitive survey A council or odour producer seeks information for legal purposes - as part of the resource consent process, or when enforcement orders or abatement notices are called into question.

Detection threshold The concentration of odour substance in which 50% of cases exposed to the stimulus can smell something.

Efficiency An unbiased estimate is more efficient than another if, for a given sample size, it has the smaller variance.

Estimation Calculating the properties of the population from the properties of a sample.

Grid point (measurement point) The grid points are the corners of the assessment squares and are the points at which emission measurements are performed.

Immission or environmental load area The reach of impact from a source of emissions being considered.

Indicative survey A council or odour producer seeks information for its own internal use - do a few complaints represent a wider problem; how to better manage odour? There is no intention to use this information for legal purposes (see 'Definitive survey').

Inference The process of drawing conclusions about a population of interest from a sample taken from that population.

Investigation area The area (within the immission area) covered by the investigation.

Investigation zone That part of the investigation area in which a sufficiently homogeneous emission load from one or more sources can be assumed.

Measure of association A statistic that measures in some way the tendency of particular values of two variables to occur together in sample data. A correlation is an example of a measure of association.

Population of interest The group of people whose opinions you want to record. A formal definition will most likely include statements about the area in which these people live or work, and a time period during which these people may have found odour offensive.

Precision Precision is the variability of an estimate about an average value if you were to repeatedly sampled the population - if the estimate isn't biased, then this average value will be the true value for the population.

Recognition threshold The concentration of odour substance in which 50% of cases exposed to the stimulus results in identification ('I can smell *x*').

Sample design The statistical aspects of running a survey - developing a sampling frame, selecting the sample, estimation, adjusting estimates to compensate for non-response.

Sampling frame A list of all those items in the population that are, for practical reasons, available for selection.

Simple random sampling A method of selecting a sample. Each item in the sampling frame has the same probability of ending up in the sample. Often abbreviated to SRS.

Stratified sampling Items in the sampling frame are put in groups, so that items in each group are as similar as possible, and the groups differ as much as possible. A simple random sample is then taken from each group. This has the potential to be a very efficient method of sampling.

Variance The variance of an estimate is a measure of how precise the estimate is.

With replacement sampling An item is selected for the sample from the sampling frame, and then returned to the sampling frame before the next item is selected.

Summary

1 Introduction

This project has been partially funded under the Ministry for the Environment's Sustainable Management Fund and contributes towards maintaining air quality in New Zealand. The management of effects of odour-producing activities has been limited to some degree by the fact that firm guidelines for gathering complaints had not been developed and that the international air quality regulatory community appeared to have been slow to develop standardised procedures for carrying out odour surveys and for determining overall community response to actual or perceived odour problems. In other words, there has been a need for procedures with which to collect subjective information using recognised objective approaches.

The purpose of this document is to provide guidance on a range of standardised techniques that can be used to establish when community odour problems exist and the magnitude of the problem. The guidelines focus on sociological and associated methods of community odour assessment and provide means by which local authority officers or the operators of odour-producing facilities can investigate whether facilities are causing adverse effects (in terms of

the Resource Management Act 1991 (RMA)) or nuisance or offensiveness (in terms of the Health Act 1956).

The odour assessment techniques proposed were identified through consultation with the tangata whenua and the National Air Quality Working Group, a review of the relevant international literature, a survey of local authority staff and interviews with people in communities where odour is a problem.

Guidelines are provided for:

- community surveys (questioning on one occasion);
- odour diaries;
- community odour panels (questioning on several occasions);
- public meetings;
- working parties;
- consultation with the tangata whenua; and
- community consultation.

We have assumed two broad situations in which the techniques might be used. One is for internal monitoring purposes where the organisation does not intend to use the data collected in legal proceedings. In this instance the guidelines can serve as a guide to good practice. The second situation where information could be or is expected to be presented as evidence in legal proceedings. In this latter instance it is important that a recognised expert in question and sample design and data analysis be engaged to ensure that the evidence meets the requirements for admissibility.

2 Procedures for Gathering Odour Complaints

One question facing odour regulators is whether or not odour complaints provide a good indication of annoyance in the community. Complaints may be considered as good indicators of sudden instances of odour pollution such as industrial accidents, but they are poor indicators of the general level of satisfaction of the population. In other words they represent one of a number of means of providing an indication if there is a community problem with odour. Suggestions are given on information to be gathered that could also be used to complement data derived using any of the other techniques discussed below.

The guidelines on gathering odour complaints have been drawn in large part from the complaints procedures developed by the State of New Jersey's Department of Environmental Protection to investigate whether their regulation prohibiting air pollution has been or is being violated. These procedures are similar in many ways to those used by local authorities in New Zealand and the regulation is not dissimilar to requirements controlling discharges to air under the RMA and the Health Act .

3 Community Odour Assessment Techniques

For each technique we have suggested instances in which they could be used as well as examples of possible costs of using each technique.

We have included a brief section on Council staff and the range of roles they may have in odour management. The following techniques can assist them in carrying out those roles.

The word 'survey' can have different meanings for different people in different situations. In these guidelines we use the expression 'community survey' to mean the systematic gathering of quantitative data by way of 'questioning on one occasion'. A representative sample of people is chosen, they are all questioned on one occasion over a period of say one week or more, depending on the size of the survey, and their responses are collated and analysed at the end of that period. This kind of survey could be repeated on a subsequent occasion, say in a year or two, to see whether people's perceptions of odour have changed, for example.

The results of a community survey can give a gradation of annoyance/offensiveness measurement, provide information on the cumulative experience of the community over a period of time, as well as collecting data on the effects of annoyance/offence if that is required.

The method does not enable coverage of individual odour events in the way that the use of odour panels (or odour diaries) do and is therefore not suitable for describing real variations in annoyance/offence over time. On the other hand it is suitable for measuring the cumulated sensation or experience of offensiveness over time.

Guidelines for personal interviews, telephone surveys and postal questionnaires appear in the same section because basically they are different methods of systematically collecting data.

Issues relating to survey management, questionnaire development, sample design, and data analysis are provided.

The RMA permits the Environment Court to receive or call for anything in evidence to assist it in its deliberations. Market surveys have been admissible as evidence for some time and existing case law provides guidance as to whether (a) a survey may be admissible and (b) if it is, what weight could or should be put on that evidence. The guidelines provide a discussion of this case law.

'Odour diaries' (sometimes referred to as 'odour surveys'), and 'panel surveys' or 'odour panels' differ from what we refer to as 'community surveys' in that data is gathered from the same sample or group of people through 'questioning on several occasions'. They may be instructed to record information every day, every time they detect an odour, or every time they experience an 'offensive or objectionable' odour over a specified period although their responses are likely to be submitted to the person organising the exercise on a frequent basis.

This technique broadly involves people living in the vicinity of a suspected, known, or potential odour source keeping a record of odour occurrences (and sometimes non-occurrences). In addition to the FIDO factors (frequency, intensity, duration and offensiveness), diarists may be asked to record the time, date, prevailing weather conditions and suspected source.

We found no written information formally or scientifically describing diaries as odour assessment techniques during the course of developing these guidelines. The odour diary technique as it is used in New Zealand appears to embody components of other odour assessment techniques such as community odour panels, trained independent observers/field

inspections, and odour community surveys. The guidelines are based on the authors' interpretation of similarities and differences of odour diaries with these other techniques, and with information gathered during interviews with selected local authority staff who had used odour diaries.

There are several different types of odour panel with each type being used to achieve particular objectives. These include:

- small panels of volunteers who are invited to have their noses calibrated and to assess odour concentrations through the use of an olfactometer;

- field inspection panels where carefully selected observers are trained to assess odours in ambient air 'in the field' at regular intervals to determine whether odours are discernible (and if so, the type and intensity);

- a sample of neighbours (and possibly out-of district panellists) on the down-wind side of an activity assess whether the size of a proposed buffer zone is sufficient to accommodate adverse effects of odour;

- relatively large panels of randomly-selected volunteers who indicate at regular intervals (generally once weekly) from their own homes whether they detect and are annoyed or offended by odour.

These guidelines focus on the last type mentioned. This approach involves the panellists standing outside their home at a specified time and on a specified day and checking for odour. They then complete a specially designed postcard which they return immediately to the organisation carrying out the exercise.

The term 'public meeting' is self-explanatory and implies meetings that are called for a particular purpose that are open to any interested person.

A Working Party is a group that is set up generally as a problem-solving entity with a specific task or tasks in mind. It may act as the driving force behind various techniques that are used to investigate issues.

The observations on consultation with the tangata whenua came from representatives of groups that have been involved in odour issues.

A range of approaches can be used when consulting the community. One approach (Meek, 1993) termed the '**LUCAT**' approach emphasises *'listening', 'understanding', 'credibility' and 'trust'*. It is based on seven principles of good consultation. These principles are also relevant to the other techniques that are presented in this document.

1 INTRODUCTION

1.1 Scope and Purpose of the Guidelines

This project has been partially funded under the Ministry for the Environment's Sustainable Management Fund and contributes towards maintaining air quality in New Zealand. The management of effects of odour-producing activities has been limited to some degree by the fact that firm guidelines for gathering complaints had not been developed and that the international air quality regulatory community appeared to have been slow to develop

standardised procedures for carrying out odour surveys and for determining overall community response to actual or perceived odour problems. In other words, there has been a need for procedures with which to collect subjective information using recognised objective approaches.

The overall purpose of this study (Lincoln Environmental, 1997) has been:

- to develop guidelines for gathering odour complaints;
- to assess various techniques that can be used to establish whether a community odour problem exists; and
- to provide detailed guidelines on how to use a range of appropriate techniques (including odour surveys).

The purpose of this document is to provide guidance on a range of standardised techniques that can be used to establish when community odour problems exist and the magnitude of the problem. These guidelines focus on sociological and associated methods of community odour assessment. They provide means by which local authority officers or companies can investigate whether facilities are causing adverse effects (in terms of section 17 of the Resource Management Act) and .... nuisance or offensiveness ..... (in terms of section 23 of the Health Act). Use of the techniques described in this report provides the opportunity for 'ordinary people' to express a view. They provide ways by which the perception of the 'ordinary person' can be captured through systematic and rigorous approaches.

The odour assessment techniques proposed were identified through consultation with the tangata whenua and the National Air Quality Working Group, a review of the international literature, a survey of local authority staff and interviews with people in communities where odour is a problem. Guidelines are provided for:

- community surveys (questioning on one occasion);
- odour diaries;
- working parties;
- community odour panels (questioning on a number of occasions);
- public meetings;
- consultation with the tangata whenua; and
- community consultation advice that can be utilised in a range of situations.

The techniques described can be for internal use such as monitoring by an odour-producing facility operator or for gathering information that could eventually be presented in legal proceedings. The guidelines are intended to provide a framework for good practice when investigating the community's perceptions and experiences of odour.

1.2 Legislative Context

The Ministry for the Environment's paper (1995) *Odour Management under the Resource Management Act* (1995, pp15-20) provides a broad overview of legislation that relates to the control of air pollution. Prior to the Resource Management Act (RMA) the Clean Air Act of 1972 was the core air pollution control statue. Transitional provisions for scheduled activities under the Clean Air Act 1972 (repealed by the RMA in 1991) were carried over until

October 1996. The RMA 1991 and the Health Act 1956 are now the major statues of air pollution control in New Zealand.

The RMA 1991 has as a major focus the controlling of the 'effects' of activities rather than activities per se. Section 5(2) of the RMA states that "...'sustainable management' means managing the use, development, and protection of natural and physical resources in a way, or at a rate, which enables people and communities to provide for their social, economic, and cultural well being and for their health and safety while ..... Avoiding, remedying, or mitigating any **adverse effects** of activities on the environment".

The 'environment' is defined (in Section 2) as including:

"(a) Ecosystems and their constituent parts, including people and communities; and
(b) All natural and physical resources; and
(c) Amenity values; and
(d) The social, economic, aesthetic, and cultural conditions which affect the matters stated in paragraphs (c) to (c) of this definition or which are affected by those matters:"

'Amenity values' are defined (Section 2) as:

"those natural or physical qualities and characteristics of an area that contribute to people's appreciation of its pleasantness, aesthetic coherence, and cultural and recreational attributes".

The meaning of 'effect' has been specifically set out in section 3 of the RMA. It can be argued that this section does not prescribe or define what the word effect actually means, but it does provide an indication of the broad range of effects that the term includes. These are:

"(a) Any positive or adverse effect; and
(b) Any temporary or permanent effect; and
(c) Any past, present, or future effect: and
(d) Any cumulative effect which arises over time or in combination with other effects; regardless of the scale, intensity, duration or frequency of the effect, and also includes -
(e) Any potential effect of high probability; and
(f) Any potential effect of low probability which has a high potential impact."

Sections 7(c), (d), (f) and 8 of the RMA are also of particular relevance to odour management.

*"**7 Other matters ....***
(c) The maintenance and enhancement of amenity values:
(d) Intrinsic values of ecosystems: ....
(f) Maintenance and enhancement of the quality of the environment: ....

8 Treaty of Waitangi ....
....shall take into account the principles of the Treaty of Waitangi (Te Tiriti o Waitangi)."

The management of odour under the RMA is facilitated through the development of national policy statements, regional policy statements and air quality plans, and in the resource consent process (where a company or local authority is proposing to establish a plant or plans to expand or modify an existing plant and conditions need to be placed on a consent, when a consent is to be renewed, or conditions reviewed).

The Health Act 1956 (section 23) also empowers local authorities to investigate conditions likely to be injurious to health of offensive. Section 29 of this Act contains 'nuisance' and 'offensive trade' provisions which, although making no specific references to odour, are "wide enough to cover odour where it is offensive or likely to be injurious to health" (Ministry for the Environment, 1995, p15). The Health Nuisance notice requires immediate abatement and has provision for an authority to take action and to recover costs from the discharger.

It should be noted that an offence under the Health Act is independent of a Regional or District Plan or Proposed Plan. Even though an activity might be permitted under a Plan, it may still be an offence under the Health Act (McLaren, 1996, pers. comm.).

Section 276(1) of the RMA permits the Environment Court to:

> *"(a) Receive anything in evidence that it considers appropriate to receive; and*
> *(b) Call for anything to be provided in evidence which it considers will assist it to make a decision or recommendation; and*
> *(c) Call before it a person to give evidence who, in its opinion, will assist it in making a decision or recommendation".*
>
> *"Pursuant to the RMA guidelines have no legal status in that they are not part of the framework allowed by the Act ..... Guidelines are merely indicators as to a way of conducting procedures and the Court will consider guidelines as just one factor in addressing whether evidence should be admitted and whether weight should be placed upon that evidence"* (Somerville, 1997, pp5-6).

Although the Environment Court (RMA, section 276(2)) "*is not bound by the rules of law about evidence that apply to judicial proceedings"* experience has shown that:

> *"... the Environment Court does generally test evidence by considering the principles of evidentiary law to determine whether or not reliance can be placed on any material*." (Somerville, 1997, p6).

If there is any likelihood that information obtained through the use of these techniques might be used in legal proceedings, then you should employ the services of an expert to design your survey, community odour panel, or odour diary and to carry out the necessary statistical analyses. The guidelines will help you to understand what your expert is doing. If you plan to use the material for internal purposes within your organisation, the design and analysis could be done 'in house', although you may wish to pay for some professional peer review.

1.3 Odour Nuisance

Odour nuisance is a significant environmental issue in New Zealand, one that results in numerous complaints to regulatory authorities.

Those involved in odour management face the multiple challenges of determining the source of an odour, the potential and actual adverse effects of that odour on the community, the extent of those effects, and trying to develop emission standards that aim to avoid, remedy or mitigate the adverse effects of odour emissions. They have the option of issuing enforcement orders or abatement notices if an activity is "*likely to be noxious, dangerous, offensive, or objectionable to such an extent that it has or is likely to have an adverse effect on the environment"* (RMA 1991, s.17(3)(a)). Provisions such as no '*discernible*' or '*objectionable*' odour beyond the boundary, for example, are now appearing in air plans and resource consent conditions.

> *"The evaluation of odour pollution problems can be based on chemical-analytical, psychophysical [including olfactometry] or sociological methods. However, results from each of these methods give information only on some part of the problem."* (van Langenhove *et al.*, 1988, p2509).

An underlying premise of these guidelines is that the people affected by offensive or objectionable odours are the best source of information on the offensiveness and objectionable character of those odours. The use of sociological and associated methods are therefore the most appropriate to use in evaluating this aspect of the odour 'problem'.

> *"Whether an odour is judged to be annoying, depends mainly on its character and on the expectation of the human perceiver about its occurrence in a certain context. Only when unacceptable smells are concerned, odour concentration starts playing a role in the amount of annoyance experienced, but even there the relationship is not always simple. A smell which at low concentrations is judged acceptable and even slightly pleasant in a given context (e.g. manure in the countryside), can all of a sudden become annoying at higher concentrations."* (Punter *et al.*, source unknown, p151).

The perception of odour effects is subjective in nature *"... but are effects to be taken as effects perceived by the general community and people or one or two particularly sensitive people?"* (Hearn, 1994, p6). The Planning Tribunal (now Environment Court) has proposed that:

> *"The correct test, I am satisfied, is whether or not the ordinary person, neither hypersensitive nor insensitive, would find the extraneous thing to be ...."* (*Zdrahal v Wellington City Council* referred to in Ministry for the Environment, 1995, p18)

These guidelines provide means by which the perceptions and experiences of "the ordinary person" can be assessed.

2 GUIDELINES FOR GATHERING ODOUR COMPLAINTS

2.1 Limitations of Odour Complaints Records

"Approximately half of all air pollution complaints in Western society involve odor" (Porteus cited in Tapper and Sudbury, 1991, p433) despite the fact that very few people register formal complaints with regard to odour problems. Generally the number of different complainants is less than 15% of the number of documented residences (USEPA cited in Jones, 1994, p2).

One of the questions facing odour regulators is whether odour complaints provide a good indication of annoyance in the community. Köster *et al.* (1984, pp3-4) present a number of reservations:

> *"1. Complaints are ungraded, all-or-nothing, responses. They are unfit to measure small amounts of annoyance in a sensitive way. They only occur when a certain threshold of dissatisfaction has been surpassed.*
>
> *2. Complaints are human responses with a refractory period. After making a complaint it will take some time before the same person will make a complaint again.*
>
> *3. People differ very strongly in the heights of their thresholds for making complaints and in the length of their refractory period. A large part of the population will never make a complaint. Others will be habitual complainers. Complaining behaviour therefore is not representative for the annoyance experienced by the total population.*
>
> *4. The frequency of complaining behaviour is influenced by factors other than the annoyance itself. Thus, the accessibility of the authorities and the costs, both financial and in terms of human effort, of making complaints will influence it.*
>
> ***Also, people may lose faith in the effectiveness of making complaints. This especially is true in cases of odour pollution, where no immediate remedy is available and where it may take months or years to cure a problem.*** (Authors' highlighting)
>
> **[Members of the public may well be reluctant to state publicly their objections to an odour if, for example, they are employed at the facility or are related to someone who works there.]**
>
> **[Perceptions of the quality of the person/company generating the odour can also influence complaining behaviour.]**
>
> *5. The frequency of complaining behaviour is subject to the influence of factors that, although they change the feeling of annoyance, are extraneous to the odour pollution itself. Press publications may influence the tolerance for annoyance or they may lower the threshold for making complaints temporarily. Pressure groups also can exert a strong influence on the number of complaints registered in a certain period. Even if such a group represents only a small part of the population, it may influence the number of complaints dramatically when all members decide to double the frequency of their complaints.*
>
> *In general, complaints may be considered as good indicators of sudden instances of odour pollution such as industrial accidents, but they are poor indicators of the general level of satisfaction of the population".*

2.2 Potential Uses for Complaints Data360Wind direction

- **Prevailing weather conditions (e.g. overcast, misty, hot and still, humid)**

Air iscuss the diversity of implementation techniques, costs and likely response rates to each apard

Tables 3.2.2 to 3.2.6 illustrate the kinds of findings that could have emerged from an indicative survey (either council recommended as being the best person to participate)

APPENDIX A:

Sample Design

Table of Contents:

APPENDIX A : *SAMPLE DESIGN*

A1 Introduction - How to Use this Appendix

To use this appendix, you will need to know some basic statistics - a good pass in sixth form statistics or mathematics should be enough. You should first read section 3.2.5 of this report, and then sections A2 to A4 of this appendix. These sections are about survey objectives, defining a population of interest, how to represent a community's views, and some general principles of sample design. Once you have read these sections, you will know which of the specific sample designs (sections A5 to A8) is going to be of most use. If you're planning an indicative survey, you will find appropriate sample designs in sections A5 and A7. Sample designs for definitive surveys are discussed in sections A6 and A8, but we strongly recommend that you seek the advice of an expert in sample design if you think the results of your survey could be used in legal proceedings (see sections 3.2.4 and 3.2.7).

Section 3.2.5 of this report covers the different objectives councils and odour producers may have when surveying a community. While a council could also be an odour producer, in general there will be three parties: a council, an odour producer, and the rest of the community. Both councils and odour producers need to select a sample of people to represent the community. A council often has a list of ratepayers from which it can draw its sample. An odour producer will typically not have this sort of information. So, councils and odour producers will tend to use different sorts of surveys because their objectives differ, and because councils already know where people live and work in the community.

Section A2 of this appendix is about how to define a population of interest. The population of interest is the group of people whose opinions you want to record. Obviously, if you define your population as those living within a certain area, and this area is very large, then the proportion of people who perceive an odour problem may be very small. So this definition of your population is crucial to the subsequent use you can make of your survey's results.

Section A3 of this appendix is about how to represent the community's views. While it is relatively easy to sample households and businesses in an area, sampling individuals within each household or business adds another layer of complexity to the sampling process. You need to consider whether this added complexity is necessary.

Section A4 of this appendix is about general principles of sample design - the statistical aspects of running a survey. Four major steps in the sample design process are: developing a sampling frame (a list of those in the population of interest), selecting people for the survey, estimation (calculating results once the survey is completed) and adjusting results for non-response. Sample design takes place in conjunction with questionnaire development and survey management planning. Each of these three operations affects the others. Statistics New Zealand's 'A guide to good survey design' (Statistics New Zealand, 1995) is full of good practical advice on running a survey.

A2 Defining a Population of Interest

The starting point for any survey is a definition of the population of interest. The definition should be clearly stated. The definition will most likely include statements about an area in which the community to be surveyed lives, and about a time period of interest. The aim is to define an area and a time period within which a community is likely to have found odour offensive.

Those planning a definitive survey must take care that the definition used can withstand legal challenge. For example, if odour quickly dissipates with distance from its source, increasing the area of interest will reduce the proportion in the survey who perceive a problem. A survey at the end of the summer is not a sensible way to assess an odour problem that occurs only in winter.

To decide on a population of interest, and to later defend that decision, you could use: addresses and dates of complaints received; results of atmospheric dispersion modelling (Ministry for the Environment, 1995, p54); documented visits by you or others to different areas; populations of interest used elsewhere for similar problems; discussion of how local climate and topography influence odour dispersion. Much of this information can be displayed using maps.

The Ministry for the Environment suggests that 'an offensive odour generally becomes noticeable when its concentration reaches 5-10 OU' (odour units) (Ministry for the Environment, 1995, p11). So atmospheric dispersion modelling, which gives contours in odour units, could be used to set a boundary to an area of interest. However, since models generally assume constant climatic conditions and crudely approximate topographical features, the addresses of those complaining will be the better guide *(ibid.*, p54).

For a definitive survey, it may be necessary to survey outside your population of interest as well. This is to demonstrate that you have done a good job in defining those likely to find odour a nuisance.[1] However, you can use different sampling fractions in and outside the population of interest, sampling fewer outside the population of interest in order to keep the total sample size down.

A2.1 *Given the addresses of complaints you've received, you define a population of interest as those living or working inside a circle with a radius of 1.0 km. At the centre of the circle lies a factory you suspect is causing an odour problem. To show that few people outside this area find odour offensive, you also survey those living between 1-2 km away from the centre of the circle.*

A3 Representing a Community's Views

A community consists of those living and working in an area, those visiting an area and those investing in an area. But individuals in these last two categories will be hard to identify. In

[1] Conditions under which survey results would be judged admissible have included 'That the proper universe was examined' or 'interviewees must be selected so as to represent a relevant cross-section of the public' - *Auckland Regional Authority v Mutual Rental Cars*. (1987) 2 NZLR, p680.

most cases, it should be enough to survey just those living and working in an area, and make some comment about how these other two groups could be affected in the light of your survey's results. In special circumstances (perhaps a large park, tourist attraction, or shopping centre), surveying these other two groups may be necessary. These guidelines do not address how to survey visitors nor investors.

So you will need to sample both businesses and households. It is relatively easy to tell how many businesses and households are in an area. You might use a list of ratepayers, a list of telephone numbers, aerial photographs or electoral role information (the 'Habitation Index'). And you can then sample some of these businesses and households. But you need to talk to individuals, and so you need some way of choosing which person or people to interview within each of the businesses or households in your sample. This means, in sampling terms, a second stage of selection: the first stage is to sample households and businesses; the second stage, to sample individuals from within the previously selected households and businesses. If possible, you want to avoid this second stage of selection for two reasons.

Firstly, you don't know how many people are in each primary sampling unit (the business or household). So the interviewer has to find this out, and then choose at random a person to interview. This means more (and difficult) work for the interviewer, and interviewers will need more training as a result.

Secondly, estimates from a survey are not enough. You need to show how precise your estimates are. To do this, you need to calculate the variance in what you are trying to measure. With two stages of sampling, both stages contribute to the variance (Sarndal *et al.*, 1993, p137). But to estimate the contribution to the variance from the second stage, you need to interview more than one person in each household. This is difficult to do without the answers of the first person interviewed influencing the answers of the second person interviewed (Kish, 1965, p398).

One way around the problem is to always interview 'the household member who is most knowledgeable' (Fowler, 1985, p33). We suggest 'the person most often at home' should be the most knowledgeable member of the household on matters of offensive odours, and a Health and Safety Officer (if on site) should be the most knowledgeable employee in a business. That is, if the business has such a person on site: you would have to approach the CEO first, and ask to speak to the most appropriate person (the person who would receive any complaints about odour from those employed by the business).

However this solution means that the results of your survey will be in terms of households (and businesses), not individuals. That is, '15% of households within 500 metres of the source found the odour offensive', not '15% of the individuals living within 500 metres'. And you will need to explain your procedure for choosing someone to represent the views of each household or business (particularly if it is a definitive survey).

Selecting individuals this way has some real practical advantages. Firstly 'the person most often at home' will be available for an interview more often than someone chosen at random from those in the household. This will reduce the number of households you have to visit (or phone) a second time in order to interview the person you want. Secondly the survey may be repeated some time in the future, if estimates of change are required. It is more accurate to use the same sample (both 'before' and 'after') so that the change is not partly due to a difference between two different samples. If the same sample is used, a sample of differences

('after' minus 'before') is then used to estimate the average difference in the population. Finding 'the person most often at home' in a second survey (and this person could legitimately change between surveys) is much easier than finding some specific individual, chosen at random in the first survey.

On the other hand, a community is really made up of individuals, not households and businesses. 'The person most often at home' is more likely to find an odour offensive than someone chosen at random from the household. So using the 'person most often at home' approach will over-estimate how many individuals perceive a problem. Of course, as an odour problem becomes more and more obvious, this bias disappears because it won't matter who you talk to in the household - everyone will tell you there's a problem!

If you are running an indicative survey, a slight over-estimate won't really matter. With a small sample, the bias (the difference, on average, between what the survey tells you and the true value) will be small compared to the variance in the estimate. When the variance is large, this is saying the estimate isn't known precisely anyway. Selecting individuals this way will actually be an advantage to the odour producer who wishes to gather information in order to better manage odour. In this case, you want to talk to those who think there is a problem.

In a definitive survey, an over-estimate of the individual's views will not be a problem if it does not help your cause. For example, the odour producer who wishes to show there is a low level of community concern will not benefit from using results from 'the person most often at home' as representing the level of individual concern.

However, if a definitive survey of individuals is necessary and an over-estimate is going to be to your advantage, then you may need to randomly sample individuals within households and businesses. At this point, consult a professional. Briefly, the common method of selecting the individual who has the next birthday is not recommended.[2] Kish gives a better method: the interviewer ranks household members by increasing age (the interviewer doesn't actually need to know these ages), then uses a random number table to choose one person from this list (Ministry for the Environment, 1995, p11). Estimates need to be adjusted to take account of the variable selection probabilities. The variance in these estimates can be calculated as if there were no second stage of sampling. This will under-estimate the true variance (Sarndal *et al.*, 1993, pp139-140), but Kish (1965, pp400, 403) suggests the difference will not be large provided most households and businesses are about the same size (say one to six people). An appreciable number of larger businesses in the population of interest means that businesses should be treated differently. A sample of people will need to be selected from each business, and second stage variances calculated for the business component of the community. Alternatively, businesses could be sampled with replacement (see section A4.2) and then there is no need to calculate second stage variances (Sarndal *et al.*, 1993, p 151).

[2] Asking for the household member with the nearest birthday gives too much control to the person first contacted. This person often claims to be the person with the nearest birthday when they are not; this person may not know the birthdays of all the others in the house; or they may decide that visitors are not part of the household. Or the interviewer may decide that it is easier to interview this first contact rather than someone else who isn't perhaps at home. Other methods of selecting a household member force the interviewer to do a proper job.

A4 General Principles of Sample Design

A4.1 Sampling frame

The sampling frame is a list of all households and businesses available for selection. If you are lucky, this list will consist of every household and business in your population of interest. Often there is some mismatch between the sampling frame and the population of interest. For example, a household without a phone or with an unlisted phone number will not be in a sampling frame compiled from the telephone directory. Duplicates in the sampling frame will cause results to be biased towards the opinions of the duplicated group. It is good practice to report any mismatch between sampling frame and the population of interest and discuss how this could affect results. As a result of this mismatch, do you expect estimates to be too high or too low for certain questions?

Some sort of sampling frame is necessary in any survey. A district or city council will be able to use a list of ratepayers. A regional council or odour producer may have to use some other sort of list. In what follows, we will refer to a *'council'* as though all councils can use a list of ratepayers as a sampling frame, and we will refer to an *'odour producer'* as though all odour producers cannot use a list of this sort. This is not true and so a council may find itself using a sample design more commonly used by an odour producer and the other way around.

A survey's objectives (see section 3.2.5) and the sampling frames available strongly influence sample design. So each combination of *'council'* or *'odour producer'* and *'indicative'* or *'definitive'* survey will require a different sort of sample design.

A4.2 Selecting the sample

A sample is then selected from the sampling frame. Random sampling ensures that one can estimate (without bias) the properties of the population of interest from the sample. In a simple random sample, each item in the sampling frame has an equal chance of being selected for the survey. Other sorts of random sampling are more efficient than simple random sampling. That is, for a given sample size, other sorts of random sampling can give more precise estimates (precision is measured by the variance of the estimate - see section A4.3). So before selecting the sample, you need to calculate a suitable sample size but this calculation depends on the sort of random sampling you're going to use. Appropriate methods of random sampling will be introduced as needed, together with how to calculate a suitable sample size.

The simple random sample (<u>without</u> replacement) is a component in many of the more complicated methods of random sampling. Sampling <u>with</u> replacement means that when an household or business is selected from the population for the sample, it is observed and then returned to the population before the next draw. Sampling <u>without</u> replacement ensures that once an household or business is selected for the sample, it cannot be selected a second time. Simple random sampling is usually <u>without</u> replacement because it is more efficient (Sarndal *et al.*, 1993, p73).

The easiest way to take a simple random sample (<u>without</u> replacement) is to use a computer program (such as a spreadsheet or statistical package) to select your sample. You could read your sampling frame into the computer program and then select the sample. Alternatively, you could create a column of numbers inside the computer program, so that each number

represents one item in your sampling frame; then select numbers from this column and use these numbers to identify the items in your sampling frame that are in the sample. Most computer programs will have simple random sampling without replacement as the default, but it pays to check.

If you can't use a computer to draw the sample, you can take a systematic sample from your sampling frame by hand. To draw a systematic sample of roughly size n, choose a random number between one and k, and then select every k^{th} ratepayer from the sampling frame starting with the ratepayer corresponding to the random number. Calculate k as the integer part of N/n, where N is the total number of ratepayers in the sampling frame.

A4.2.1 *You calculate the required sample size to be 60. There are 893 ratepayers in the sampling frame, so k = 893/60 = 14.88 $\Rightarrow$ 14. You now need to find a random number between 1 and 14, as a starting point for sampling from the list. You get a scientific calculator to choose a random number between 0 and 1: the calculator selects 0.271. Multiply this number by k; add one; and then just use the integer part of the answer: (0.271´14) +1 = 4.794 $\Rightarrow$ 4. So select the fourth ratepayer in the list, and then every 14th ratepayer after that, until you get to the end of the list.*

A systematic sample from a list arranged alphabetically can be treated as if it were a simple random sample (Cochran, 1977, pp212-213). If the list is not in alphabetic order, make sure the order is essentially random. Particularly watch out for any order in the list that roughly coincides with the frequency of sampling from the list (*ibid.*, pp217-219).

A4.2.2 *In the example above (section A4.2.1), suppose the sampling frame consisted mainly of households but every 14th ratepayer in the sampling frame was the owner of a business. The systematic sample would then either have no businesses in it at all or would consist entirely of businesses, depending on the random starting point chosen. While an exact periodic effect like this is unlikely, roughly periodic variation in the list is a possibility and this could lead to biased estimates. If in doubt, take a simple random sample instead.*

A4.3 Estimation

With a simple random sample, the sample average is an unbiased estimate of the population average. This is not true for most other random sampling methods. Once results are in, estimates of population averages or totals must be calculated. Mostly, estimate of averages will be required: 'on the last occasion odour was perceived as annoying, the odour lasted on average for 1.8 hours'. Note that a proportion is a special sort of average. Proportions are converted to percentages by multiplying the proportion by 100. So '15% of respondents felt the odour was offensive' is a statement about a proportion. If each person who feels odour is offensive is recorded as a one, and a zero is recorded otherwise, adding up the ones and zeros and dividing by the number of respondents yields a proportion. Yet this process is just finding an average.

The likely variation in estimates must also be calculated. This 'variance' is a measure of how precisely answers are known. In statistics, precision refers to how much estimates would vary about an average value if you repeatedly sampled your population - but if the estimate is

biased, its average value will not be the true value. In general, one can calculate the precision of an estimate but not its bias.

For legal purposes, knowing the precision of the estimate is as important as knowing the estimate itself. If a survey is to be used as evidence, details may be required of 'any tests applied and the results of any tests applied to determine the extent to which the survey or results of the survey can be trusted' (*Auckland Regional Authority v Mutual Rental Cars* (supra)).

One appropriate test is the 95% confidence interval. This interval is approximately the estimate (of an average, total or proportion) plus or minus two times the square root of the variance of the estimate.

$$\text{ie. } 95\%\ \text{CI} = \text{estimate} \pm 2.\sqrt{\text{variance of estimate}} \qquad (1)$$

The interpretation of this interval is that for 95% of all possible samples, the true value for the population of interest will lie within the interval. Hence the true value is to be found within this interval with a high degree of confidence. As a general rule, to be reliable a confidence interval needs to be based on a sample size of at least 30 (Cochran, 1977, p27; Berenson *et al.*, 1988, p227).

A4.3.1 *You take a simple random sample of 30 households from a population of 2000. In five of these 30 households, the person 'most often at home' says that odour is a problem. Therefore the proportion of households where odour is perceived as a problem is 5/30 or 0.17 (17%). Let's say the variance of this proportion is 0.0047 (you'd use equation 11 in section A5.3). The 95% confidence interval is then $0.17 \pm (2 \times 0.07)$, or from 0.03 to 0.31. That is, in 3% to 31% of households in the population of interest, the 'person most often at home' says that odour is a problem.*

Other higher or lower percentage confidence intervals could be useful. Most statistics textbooks give examples of how to do this. The 95% confidence interval has a long history of use in science, and should be acceptable for legal purposes. But there's nothing magic about this 95% confidence interval - for indicative surveys, 90% or 80% confidence intervals may well be more useful. Indicative surveys are likely to use small samples. The result will be 95% confidence intervals that are very wide: perhaps so wide as to be rather uninformative (as in section A4.3.1 above). The 80% interval will give you a narrower interval within which the true value is likely to lie, although you cannot be quite so certain that the true value is within this interval. The 80% confidence interval is approximately the estimate (of an average, total or proportion) plus or minus 1.3 times the square root of the variance of the estimate:

$$\text{ie. } 80\%\ \text{CI} = \text{estimate} \pm 1.3.\sqrt{\text{variance of estimate}} \qquad (2)$$

A4.4 Non response

Finally, those who do not respond to a survey are often different to those who do. Non-response introduces a bias: on average estimates from the survey differ from the true values for the population. Many things can be done, as part of survey management, to reduce non-response. With a definitive survey, you should try to contact those not at home at least three times before giving up. With an indicative survey, try to contact those not at home at least twice (or at least try some them a second time - see section A5.4). Other ways to reduce non-response include contacting people in advance to arrange an interview time, assurances of privacy and confidentiality, providing some small incentive (Department of Statistics, 1992, p33). Good questionnaire design helps too. But there will always be some degree of non-response. Once answers are in, statistical methods can be used to adjust estimates so they are less biased.

You should always report your survey's response rate. As a rule of thumb, the response to your survey should be 70% or better. Otherwise 'most researchers would have more accurate and useful estimates if they reduced the sample size and devoted the saved resources to obtaining responses from a higher percentage of the sample' (Fowler, 1985, p147).

A5 Council: Indicative Survey

A5.1 Sampling frame

A list of ratepayers (both households and businesses) in the area of interest will be the best sampling frame. Those selected from the list will be contacted, and the person most often at home will be asked a short series of simple questions.

A telephone survey will be the quickest and cheapest way to reach those selected (*ibid.*, pp68-9). Postal surveys are not recommended because of their low response rates (*ibid.*, pp66-67). A telephone survey needs a short questionnaire and simple questions. A short questionnaire is a good selling point when you're trying to persuade someone to participate in your survey. Simple questions are needed for telephone interviews; otherwise the respondent may have difficulty understanding what is being asked. [In face to face interviews, it's easier to tell if the respondent doesn't understand the question and needs additional information from the interviewer. The respondent can also see a copy of a long question, or the choice of answers as the question is being asked by the interviewer.]

Treat flats as businesses. In the first instance, interview the owner (as the CEO): he or she may then suggest you speak to a tenant (as a more appropriate person to answer your questions) and may help you contact this person. For an indicative survey, it won't really matter if you have to interview the owner when you'd rather interview a tenant (see comments in section A3 about bias versus variance in estimates from indicative surveys).

A5.2 Selecting the sample

Take a simple random sample (or systematic sample) from your list of ratepayers (section A4.2). To calculate the required sample size, identify the most important question or questions in your survey. For each question, consider the width of the confidence interval

you want to end up with. Calculate an approximate sample size for each question as (Berenson *et al.*, 1988, pp259-261):

$$n_0 = \frac{Z^2 \sigma^2}{\partial^2} \quad (3)$$

n_0 = first approximation to required sample size;
Z = 2.0 for a 95% confidence interval;
= 1.3 for an 80% confidence interval;
∂ = half the width of the desired confidence interval;
σ^2 = population variance (as yet unknown).

Now adjust this approximate sample size given the number in the sampling frame (N) (Berenson *et al.*, 1988, pp266-268; Cochran, 1977, p76):

$$n = \frac{n_0 N}{n_0 + N} \quad (4)$$

n = required sample size;
N = population size (number in the sampling frame).

Do this for each important question and take the largest value of n as the sample size required for your survey. Remember, you need a sample size of at least 30 to be able to calculate reliable confidence intervals (section A4.3).

Of course you don't know the population variance before surveying, and the trick is to make an intelligent guess. If you are going to estimate an average, think of the largest and smallest values you are likely to get in answer to your question. The population variance is then roughly: [3]

$$\sigma^2 = \left[\frac{(\text{largest - smallest})}{6} \right]^2 \quad (5)$$

[3] Equation 5 follows if what you are trying to measure is distributed according to a normal distribution. Then 99.9% of observations (ie. almost all) lie within 3 standard deviations on either side of the mean. So the range of what you observe is about 6 standard deviations - and the variance is just the square of a standard deviation. If you think that what you are trying to measure should follow a rectangular or triangular distribution, use similar methods - see Cochran, Working Group (1977): *Sampling techniques (third edition).* New York: John Wiley, p81.

A5.2.1 *You are going to ask respondents: 'Think back to the last time you were annoyed by odour. How long did the odour last?'. You want to estimate how long offensive odours last. You think the answers you'll receive will range from 0 to 15 hours. Your rough estimate of the population variance is then* $[(15-0)/6]^2=6.25$. *If you decide you want to calculate an 80% confidence interval for this mean with width plus or minus half an hour, your initial sample size is* $1.32\times6.25/0.52 = 42$. *If there are 2000 ratepayers in your sampling frame, the required sample size is adjusted to* $(42\times2000)/(42+2000) = 41$. *Select say 60 people to allow for a 70% response rate. Note that in this example, using equation 4 doesn't really change the sample size because here the sample size is small relative to the population size.*

If you are going to estimate a proportion, note that the population variance for n variable that can only take the values zero or one is approximately:[4]

$$\sigma^2 = P.(1-P) \qquad (6)$$

This equation is at a maximum when the population proportion (P) is 0.5 (that is, 50%). So think of what the population proportion is likely to be; then choose a value closer to 0.5. Without any information at all, you might use P = 0.5; but you risk being too conservative and ending up with a much larger sample size than you really need.

A5.2.2 *You think around 10% of the population will find a particular odour offensive. You want a 95% confidence interval with width plus or minus 10%. [That is, if the estimate of this proportion turns out to be 0.12, you want to end up with a 95% confidence interval from 0.02 to 0.22.] To be slightly conservative, you assume that the population proportion (P) is 0.15. Your initial sample size is* $2.02\times0.15\times0.85/0.12 = 51$. *Notice that if you'd assumed P = 0.5, your initial sample size would be 100. If there are 200 ratepayers in your sampling frame, the required sample size is adjusted to* $51\times200/(51+200) = 41$ *- say 60 people to allow for a 70% response rate.*

[4] For a variable taking only values zero or one (see section A4.3), the population variance is N/(N-1) *x* P *x* (1-P). Since N/(N-1) will be close to one, the population variance is therefore approximately P *x* (1-P) - see Cochran, 1977, p51.

A5.3 Estimation

Once the survey has been completed, estimate an average as (Cochran, 1977, pp20-21):

$$\bar{\hat{Y}} = \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} \quad (7)$$

$\bar{\hat{Y}}$ = estimate of population average;

$\bar{y}$ = sample average;

$\sum_{i=1}^{n} y_i$ = add up each sample observation;

n = sample size.

The variance of this estimate is (*ibid.*, p26):

$$v(\bar{y}) = \frac{(N - n)}{N} . \frac{s^2}{n} \quad (8)$$

$v(\bar{y})$ = variance of sample average;

N = population size;

s^2 = sample variance.

Use the statistical function of a scientific calculator to find the sample variance (s^2), or calculate this by hand as:

$$s^2 = \frac{\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}}{n-1} \quad (9)$$

$\sum_{i=1}^{n} y_i^2$ = add up the square of each sample observation;

$\left(\sum_{i=1}^{n} y_i\right)^2$ = add up each sample observation, then square the result.

A5.3.1 *Each person in a sample of 40 tells you how long (in hours) the odour lasted in their most recent experience of an offensive odour (see section A5.2.1). The sum of each sample observation is 72, and the sum of each sample observation squared is 320. The variance is therefore [320-(722/40)]/39 = 4.88. The mean is*

72/40 = 1.8 hours, the variance of the mean (equation 8) is [(2000-40)/2000]×(4.88/40) = 0.12. From section A4.3, an 80% confidence interval for this mean is 1.8 ± (1.3×0.35) - that is 1.3 to 2.3 hours.

Estimate a proportion (see section A4.3) and its variance as (Cochran, 1977, p51):

$$p = \frac{a}{n} \tag{10}$$

p = sample proportion (see section 3.2.6.3.3);
a = number in sample with value one rather than zero;
n = sample size.

The variance of this proportion is (*ibid.*, p52):

$$v(p) = \frac{(N - n)}{N} \cdot \frac{p(1 - p)}{(n - 1)} \tag{11}$$

v(p) = variance of sample proportion;
N = population size.

If you then calculate a confidence interval (section A4.3) and find that zero or one is included in this interval, you will need to use a more accurate method (*ibid.*, pp57-60).

A5.3.2 *You ask a sample of 40 people (from the 200 in your sampling frame - see section A5.2.2) if they have experienced any offensive odours in the last year. Of the 40, 5 answer 'yes'. Your estimate of the proportion is 5/40 = 0.125 and the variance of this proportion is [(200-40)/200]×(0.125×0.875/39) = 0.0022. From section A4.3, a 95% confidence interval for the proportion experiencing offensive odours is 0.125 ± (2×0.047) = [0.03, 0.22] - that is, 3 to 22%.*

A5.4 Non response

If you have the resources, you should make at least two attempts to contact 'the person most often at home'. One way to save resources in this process is follow up just a simple random sample of those who weren't contacted in the first attempt.

You take a sample of n people: you have contacted n_1 of these people but you haven't been able to reach n_2 ($n_1 + n_2 = n$). Take a simple random sample (or systematic sample - see section A5.2) of size r from the n_2 people you haven't contacted. Make a major effort to get responses from these r people. Instead of using the sample average, estimate the population average as:

(12)

$$\bar{y}' = \frac{(n_1\bar{y}_1 + n_2\bar{y}_r)}{n}$$

$\bar{y}'$ = adjusted sample average;
$\bar{y}_1$ = average for those contacted initially;
n_1 = number contacted initially;
n_2 = $n - n_1$;
$\bar{y}_r$ = average for the r respondents contacted subsequently.

The difference between this estimate and the normal sample average should be reported. This difference is a measure of the bias that results from being unable to contact some people. Note that since a proportion is just a special sort of average (section A4.3), you can replace averages in the above equation with the appropriate proportions.

A5.4.1 *You carry out the survey described in section A5.2.2. The response rate in your survey is lower than the 70% you hoped for. Of the 60 ratepayers in your sample, you manage to contact only 35 on your first attempt. Of these 35, 4 say they have experienced offensive odours in the last year. From the 25 non-respondents you choose 10 at random, and pursue these people until at last you make contact. Of these 10, 3 answer 'yes' to your question. Your proportion adjusted for non-response is [35×(4/35)+ 25×(3/10)]/60 = 0.19 (19%). Note that if you had not contacted some of the non-respondents, you would have estimated the proportion answering 'yes' as 4/35 = 0.11 (11%). Calculate the variance of the proportion as if you hadn't contacted the non-respondents. Using equation 11, the variance is therefore [(200-35)/200] ×(0.11×0.89/34) = 0.0024. From section A4, a 95% confidence interval for the proportion experiencing offensive odours is then 0.19 ±(2×0.049) = [0.09, 0.29] - that is, 9 to 29%. This confidence interval will be a bit wider than it should be - it's possible to calculate a variance that takes account of the contacted non-respondents but it's a little more complicated and with an indicative survey, there's really no point in being that accurate when calculating the variance* (Cochran, 1977, pp370-371).

A6 Council: Definitive Survey

A6.1 Sampling frame

A list of ratepayers (households and businesses) in the area of interest will be the best sampling frame. You could still take a simple random sample (or systematic sample) from this list, and then simply follow the guidelines in section A5 above. You would probably take a larger sample for a definitive survey, to increase the accuracy of estimates for legal purposes. [Indicative surveys typically use small samples and as a result, 95% confidence intervals are likely to be rather uninformative - see section A4.3.] Remember that if a definitive survey of individuals is necessary and an over-estimate of the severity of an odour problem will help your cause, then you should consult a professional. You may need to randomly sample individuals within households and businesses (see section A3).

In a definitive survey you may wish to interview people face-to-face rather than phoning, so you can ask more complex questions (perhaps on the FIDOL factors) (Ministry for the Environment, 1995, p14). In this section, we will assume you have decided to visit households and businesses, interviewing 'the person most often at home' from each selected household and a suitable representative from each selected business (see section A3). If a selected address turns out to be a flat, then you will end up interviewing a tenant rather than the owner of the property. The tenant 'most often at home' will usually be the most appropriate person to answer your questions.

Often you will want to make estimates for different parts of a population - perhaps for both those inside and outside an affected area. The easiest way to do this efficiently is by using stratified random sampling (section A6.2). You divide the population up into parts (strata), so that later you can make estimates for each stratum. Ideally each stratum is as similar within and as different between strata as possible (that is, similar and different in terms of the community's perception of odour).

In practice, the result might be something like this: you form say three strata as concentric rings around a point source (assuming local topography does not affect odour or its influence is unknown). The closest stratum is the area in which most of those who have complained live. The second stratum out has the odd complainant. And the third stratum is an area outside your population of interest, so you can show you have covered everyone likely to find odour a problem (see section A2). You might divide up some or all of these circles into sectors, making more strata, based on your knowledge of the prevailing winds. Do complaints seem more prevalent 'down-wind' of a suspected odour source? Remember, the trick is to form another stratum only if a those living in a particular area are likely to have a very different perception of the problem.

A6.2 Selecting the sample

It does not pay to have too many strata, because you now have to list all the ratepayers in each stratum. You then take a simple random sample (or systematic sample) of ratepayers in each stratum, using the methods given in section A4.2. Use a different random number to start each systematic sample.

The big advantage of stratified sampling is that it can be far more efficient than simple random sampling (that is, a smaller sample size can give estimates of the same precision). To

calculate the overall sample size, first identify the most important question or questions in your survey. For each question, you need to consider the width of the confidence interval you want to end up with. For each question, you need to guess the population variance in each stratum, using the methods in section A5.2. Once you have this information, you can calculate the overall sample size. Having done this for those questions you consider important, you take the largest overall sample size as the required sample size and allocate this sample among the strata.

Calculate the overall sample size for a given question as (Cochran, 1977, p98):

(13)

$$n = \frac{\left(\sum_{h=1}^{L} N_h \sigma_h\right)^2}{(N^2 \partial^2 / Z^2) + \sum_{h=1}^{L} N_h \sigma_h^2}$$

n = overall sample size;
N = overall population size;
N_h = population size in stratum h (h = 1, ..., L);
Z = 2.0 for a 95% confidence interval;
∂ = half the width of the desired confidence interval;
σ_h^2 = population variance in stratum h;
σ_h = square root of σ_h^2.

Now allocate this sample size among the strata (Cochran, 1977, p98):

(14)

$$n_h = n.\frac{N_h \sigma_h}{\sum_{h=1}^{L} N_h \sigma_h}$$

n_h = sample size alloated to stratum h.

Note that sometimes n_h turns out to be larger than N_h. If this happens, you'll need to make a slight adjustment to this method (*ibid.*, p104).

This method of allocation means that if you are taking systematic samples, each stratum will have a different value of k (the integer part of N_h/n_h). There are other methods of allocating the overall sample size which keep k constant, but the method shown here (known as Neyman allocation) ensures that strata where responses are more varied are sampled more intensively. This improves sampling efficiency (Cochran, 1977, pp370-371).

Simple random sampling within each stratum ensures that you can make estimates (of averages or proportions) for any single stratum, as well as overall estimates for all those surveyed. However, if you want to find a confidence interval for a single stratum, as a rule of thumb that stratum needs a minimum sample size of 30 (section A4.3).

A6.2.1 *You are going to ask respondents if they have experienced any offensive odours in the last year. You want a 95% confidence interval (with width plus or minus 5%) for the proportion that answer 'yes' to this question. You form three strata - concentric rings around the factory you believe is releasing odours. The first column in the table below gives your conservative guess of the proportion likely to answer 'yes' to your question in each stratum. You can then estimate the variance in each stratum as $P_h.(1-P_h)$ - see equation 6. The third column shows the population in each stratum - the total population is 600. Using equation 13, the total sample size is (231.72)/(225+94) = 168. The fourth column in the table shows how equation 14 allocates the total sample size among the three strata. Note that the sampling fraction (n_h/N_h) is highest in the stratum where you think more people are finding odour offensive. Note too that the sample size in each stratum is over 30, so you will be able to calculate confidence intervals for each of your strata. You will want to increase the sample size in each stratum to allow for non-response (see section A5.2.2).*

Stratum	Ph	[h2	Nh	nh
Inner ring	0.5	0.25	100	36
Middle ring	0.3	0.21	200	67
Outer ring	0.1	0.09	300	65
Total			600	168

A6.3 Estimation

Once the survey has been completed, estimate an average as (Cochran, 1977, p91):

$$\hat{\bar{Y}} = \bar{y}_{st} = \frac{\sum_{h=1}^{L} N_h \bar{y}_h}{N} \qquad (15)$$

$\hat{\bar{Y}}$ = estimate of population average;

$\bar{y}_{st}$ = average of a stratified random sample;

$\bar{y}_h$ = sample average in stratum h.

The variance of this average is (*ibid.*, p95):

$$v(\bar{y}_{st}) = \frac{1}{N^2}\sum_{h=1}^{L}\frac{N_2}{n_2}.(N_h - n_h).s_h^{\,2} \qquad (16)$$

$v(\bar{y}_{st})$ = variance of $\bar{y}_{st}$;

$s_h^{\,2}$ = sample variance in stratum h.

Use the statistical function of a scientific calculator to find the sample variance in each stratum (s_h^2), or calculate this by hand using equation 9 in section A5.3.

Estimate a proportion as (*ibid.*, p107):

$$p_{st} = \frac{\sum_{h=1}^{L} N_h p_h}{N} \qquad (17)$$

p_{st} = proportion for a stratified random sample;

p_h = sample proportion in stratum h.

The variance of this proportion is (Cochran, 1977, p108):

$$v(p_{st}) = \frac{1}{N^2}\sum_{h=1}^{L}\frac{N_h^2(N_h - n_h)}{(N_h - 1)}.\frac{p_h(1-p_h)}{(n_h - 1)} \qquad (18)$$

$v(p_{st})$ = variance of p_{st}.

When using equations 7, 9 or 10 - to calculate the sample average, variance or proportion in stratum h - just use sample observations from stratum h. That is:

$$\overline{y_h} = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}, \quad p_h = \frac{a_h}{n_h} \qquad (19)$$

A6.3.1 *You carry out the survey described in section A6.2.1. In the inner ring, 12 out of 30 respondents answer 'yes', they have experienced offensive odours in the last year. In the middle ring 15 out of 60 respondents say 'yes', and in the outer ring 5 out of 60 respondents say 'yes'. Your estimate of the overall proportion is*

(40+50+25)/600 = 0.19. The variance of this proportion is (58.5+89.4+93.5)/6002 = 6.706 x 10-4. So a 95% confidence interval for the proportion is 0.19 ± (2.0x0.026) = [0.14, 0.24]. You could also calculate a 95% confidence interval (section A5.3) for say the inner ring as 0.40 ± (2.0x0.076) = [0.25, 0.55].

A6.4 Non-response

To allow for those you cannot contact, use the procedure outlined in section A5.4 to adjust your estimate of the average. Take a random sample of those who were not available in your first attempts to get an interview. You could do this in just one stratum or take a random sample in each of several strata. Make a real effort to interview these non-respondents. Adjust the estimate of the stratum average using the procedure in section A5.4. Use adjusted stratum averages instead of the normal stratum averages when combining estimates from each stratum to give the average of a stratified random sample (in section A6.3). This works for proportions too.

You can use a similar adjustment to account for those who refuse to be interviewed. This method is called the 'basic question' approach (Bethlehem and Kersten, 1985, pp287-300). You should identify one question (maybe two questions at most) that best summarises what your survey is about. Because this question needs to be asked to both respondents and non-respondents in as similar circumstances as possible, ideally this 'basic question' is one of the first you ask in your questionnaire. This approach will not be possible if the most important question is 'buried' in the questionnaire to de-sensitise the issue.

When someone refuses to participate, you ask them to answer just this one 'basic question'. Later adjust each stratum average (or proportion) and use adjusted stratum averages instead of the normal stratum averages in your calculation of the average for a stratified random sample. The adjustment to be made to each stratum average is:

1A6.4.1 For the survey described in section A6.2.1, your sample size calculations suggest you need a sample of 36 from the inner ring. You increase this to a sample of 50 to allow for a 70% response. But when you carry out the survey, you manage to complete 20 full interviews in the inner ring. Of the remaining 30, 15 were not at home the <u>three</u> times you called, 5 refused to talk to you and 10 refused an interview but gave answers to just the 'basic question'. This question was whether they had experienced any offensive odours over the last year. In the full interviews, 15 out of 20 said 'yes', but only 2 of the 10 answering just this 'basic question' said 'yes'. Your adjusted estimate of the proportion answering 'yes' in the inner ring is then [20x(15/20)+15x(2/10)]/(20+15) = 0.51. If you had just used information from full interviews, you would have estimated this proportion as 15/20 = 0.75. You now need to repeat this process for the other two strata, and then you can use these adjusted proportions in equation 17. You can use the data from both sources when calculating the variance (equation 18) - in the inner ring, you have data for your 'basic question' from 20+10 = 30 respondents in total.

For both overall and single stratum estimates, compare the adjusted average for this 'basic question' with an average calculated the usual way. The comparison will indicate whether non-response is likely to cause appreciable bias in estimates for other questions. If

appreciable bias seems likely, you can get a statistician to adjust estimates for other questions. [Estimates for other questions can be adjusted by regression, using the information you have just collected about your 'basic question' (Cochran, 1977, p107).]

A6.4.2 *For the survey described in section A6.2.1, you require estimates for each ring (as well as overall estimates). Section A6.4.1 shows that in the inner ring, there's a large difference between estimates for the 'basic question' with and without the extra information (0.51 versus 0.75). This difference suggests appreciable bias is possible in estimates for other questions. You should get a statistician to adjust the estimates for these other questions. Otherwise you could be seriously over-stating the problem.*

A7 Odour producer: Indicative Survey

A7.1 Sampling frame

As an odour producer, you will not have a list of ratepayers from which you can draw a sample. A telephone survey will be quicker and cheaper than household interviewing. Remember to keep your questionnaire short and simple (section A5.1). If you are lucky, your population of interest will coincide with a local calling area. You can simply take a systematic sample from a phone book. More commonly, the population of interest will be only a small part of a local calling area, or spread over several local calling areas. In this case talk to Telecom Directories, Directory Information Services. They can select all the phone numbers within a certain geographic area. They may not be able to exactly match your area of interest - you should ask for the smallest area that completely includes your area of interest. They will sell you a list of phone numbers, without names and addresses (because of the Privacy Act 1993), and at the time of writing their prices seemed reasonable.[5] The rest of section A7 covers this situation.

One other alternative is to contract a council to run the survey for you, using a questionnaire that is acceptable to both you and the council. [The council is unlikely to be able to give you a sample of their ratepayers, because of the Privacy Act 1993.]

A7.2 Selecting the sample

Your population of interest is contained within a local calling area in a phone book, or within a list provided by Telecom's Directory Information Services. The problem is some of the numbers in your sampling frame are for people living outside the area you are interested in. You need to take a sample of telephone numbers, and just telephone those in your population of interest. But even if you have an address as well as a phone number, you may not know if a particular household is in your population of interest. In this case, when you phone you have to find this out first, before you ask the rest of your questions. You only want to record

[5] On 1/11/96 the cost (GST exclusive) of selecting phone numbers was 3.5 cents per number with a set up cost of $150. Alternatively, Directory Information Services would take a random sample for you, for 18.5 cents per number and with no set up cost. With both these alternatives you had to buy a minimum of 3000 numbers. So if you need only a small sample for an indicative survey, these prices aren't really relevant - phone 0800-501-515 for price and product options. Some time in the future, you may be able to specify your area of interest using Statistic New Zealand's meshblock classification system (see section A8.1).

answers for those who are in your population of interest. This situation is called sampling a subpopulation or domain (Cochran, 1977, p34).

To calculate the required sample size, identify the most important question or questions in your survey. For each question, consider the width of the confidence interval you want to end up with. Calculate an approximate sample size for each question as:

$$n_0 = \frac{Z^2 \sigma^2}{\partial^2} \quad (20)$$

n_0 = first approximation to required sample size;
Z = 2.0 for a 95% confidence interval;
= 1.3 for an 80% confidence interval;
∂ = half the width of the desired confidence interval;
σ^2 = population variance (as yet unknown).

Of course you don't know the population variance before surveying, and the trick is to make an intelligent guess. Some ideas on how to do this are in section A5.2.

Now you need to adjust this approximate sample size given the number in your population of interest (N_j). You probably won't know what this number is: work out the number of phone numbers in your sampling frame (N), and roughly estimate the proportion of these that belong to households in your population of interest (p_j). Multiply these two numbers together as an estimate of N_j. You will be conservative if you make N_j slightly larger than you think it probably is:

$$n_j = \frac{n_0 N_j}{n_0 + N_j} \quad (21)$$

n_j = required sample size;
N_j = size of population of interest,
= $N . p_j$.

Remember, n_j should be at least 30 so that later you can calculate reliable confidence intervals (section A4.3).

You now know how many households you want to sample from the population of interest. But to find these households, you need to take a sample from the sampling frame and then reject those households that turn out not to be in your population of interest. So the sample

you initially select (n) needs to be larger than the sample you want to end up with (nj):

$$n = \frac{n_j}{p_j}. \qquad (22)$$

A7.2.1 *You think around 10% of the population will find a particular odour offensive. You want a 95% confidence interval with width plus or minus 10%. To be slightly conservative (section A5.2.2), you assume that the population proportion (P) is 0.15. Your initial sample size is 2.02x0.15x0.85/0.12 = 51. There are 300 phone numbers in your sampling frame, and you expect 80% of these numbers to belong to those in your area of interest. So the size of the population of interest (Nj) is 300x0.80 = 240. The required sample size is then adjusted to 51x240/(51+240) = 42. But to find these 42 people, you will need to select 42/0.8 = 53 phone numbers from the sampling frame - say 75 phone numbers to allow for a 70% response.*

You need to follow through this process for each important question and take the largest value of n as the sample size required for your survey.

Now take a simple random or systematic sample of size n from the N phone numbers in your sampling frame (section A4.2). When you phone each number in the sample, first check that the household you've called is in your population of interest. You'll never get exactly the sample size you planned (n_j), but you should get close enough. But if your population of interest makes up only a small percentage of the sampling frame, you'll need to phone a lot of numbers to get the (roughly) n_j you need.

A7.3 Estimation

You can treat this sample as a simple random sample. Use the equations in section A5.3 to calculate averages, proportions and their variances. Provided you just collect responses from those in your population of interest, these equations work if you replace n with nj, and N with Nj (Bethlehem and Kersten, 1985, pp287-300). That is, use the sample size and population size of your population of interest, not the sampling frame sample size and population size. Now that you have selected the sample, you are in a position to make a better estimate of N_j:

$$N_j = \frac{n_j}{n}.N. \qquad (23)$$

A7.3.1 *Of the 75 phone numbers in section A7.2.1, you make contact with 60 households, and 45 of these turn out to be in your population of interest. In 10 out of the 45 households, 'the person most often at home' had noticed an offensive odour during the last year. Your estimate of Nj is therefore (45/60)x300 = 225 - your estimate before taking the sample was 240. Your estimate of the proportion is 10/45 = 0.22 (equation 10) with variance [(225-45)/225]x0.22x0.78/44 = 0.0031 (equation 11). An 80% confidence interval (equation 2) is 0.22 ± (1.3x0.06) = [0.15, 0.29].*

A7.4 Non-response

You may wish to sample some of those you are initially unable to contact. By finding some of these people, you can adjust estimates to reduce the bias that results from non-response. The method is a variation on equation 12 and it's not that easy. At this point you either consult your friendly statistician, or try to make some sense out of the following example.

A7.4.1 *Of the 75 phone numbers in section A7.2.1, you manage to make contact with only 40 households, of which 30 turn out to be in your population of interest. Of these 30, 6 have noticed an offensive odour during the last year. From the remaining 35 phone numbers, you take a simple random sample (or systematic sample) of 15 and pursue these at length. Of these 15, 10 households turn out to be in your population of interest, and of these 10, 3 have noticed an offensive odour. Your estimate of the proportion (equation 12) is [30x(6/30) + n2 x (3/10)]/[30+n2]. The trouble is, you don't know n2 - the number in your population of interest among those you haven't contacted. Your best estimate of n2 is 35x(10/15). This makes your estimate of the proportion (6+7)/53.3 = 0.24. If you hadn't sampled non-respondents, you would have estimated this proportion as 6/30 = 0.20. Calculate the variance of the proportion as if you hadn't sampled non-respondents. The resulting confidence interval will be a bit wider than it needs to be - it's possible to calculate a variance that takes account of the non-respondents you finally contacted but things are complicated enough already (Cochran, 1977, p51).*

A8 Odour Producer: Definitive Survey

A8.1 Sampling frame

You could use the indicative sampling frame (section A7.1) for a definitive survey. You would probably take a larger sample for a definitive survey, to increase the precision of estimates for legal purposes. [Indicative surveys typically use small samples and as a result, 95% confidence intervals are likely to be rather uninformative - see section A4.3.]

In a definitive survey you may wish to interview people face-to-face rather than phoning, so you can ask more complex questions (perhaps on the FIDOL factors) (Cochran, 1977, p52). In this section, we will assume you have decided to visit households and businesses, interviewing 'the person most often at home' from each selected household and a suitable representative from each selected business (see section A3).

There are a number of ways you can construct a sampling frame with addresses; no one method will suit all circumstances. The size of the area you want to survey will largely determine which of the following methods is most suitable. In order from small scale to large, possible sampling frames are: (1) a list of houses and businesses from aerial photographs; (2) a list of houses and businesses for each streets in a town or city suburb; (3) the phone book for a local calling area; (4) the 'meshblock' classification system used by Statistics New Zealand.

The first method is very small scale. Using aerial photos, simply number the houses in the area of interest. Take a simple random sample of these numbers using a computer.

The second method requires a list of all the addresses of households and businesses for each street in your area of interest. It will be easiest to assemble this sampling frame in a computer spreadsheet. You can get all the streets (for your population of interest) off a map, and then visit each street to find the last number in the street. This will tell you roughly how many houses there are in the street. Or if this is too much work, visit a Registrar of Electors (New Zealand Post) and look at the 'Habitation Index' for the electorate. The 'Habitation Index' has streets in alphabetical order, and the names and addresses of registered voters who live in each street.6 Find each street of interest in the 'Habitation Index', and record the last address in the street. Obviously, the 'Habitation Index' will not be very useful for a new subdivision. Once you know roughly the last address in the street, you know roughly how many households are in the street. With a spreadsheet, you can then generate a number for every household in every street and this is your sampling frame.

This trick of using the last address to tell you how many households are in a street is not exact - but often you'll be close enough. In a mathematical sense, the number of the last address in the street will under-estimate the number of households in the street. If the 'Habitation Index' does not contain a high percentage of all the addresses in the street, use the following unbiased estimate (Rayner, 1994):

$$\hat{N}_s = \left(\frac{(m+1) Y_{max}}{m} \right) - 1 \qquad (24)$$

$\hat{N}_s$ = an estimate of the number of houses in a particular street;
Y_{max} = the largest number among the m numbers listed in the Habitation Index for that street.

In a practical sense, even this alternative (equation 24) is likely to under-estimate the number of households in the street - because multiple households at the same street number will be more frequent than parks and empty sections. So your sampling frame will under-represent those living in flats and retirement villages and if these households represent a significant part of your population, you will just have to visit those streets where there are a lot of multiple households at the same street number and record all the letter box numbers in these streets. You can get a pretty good idea of whether a street has a high proportion of flats or a retirement village from the names and addresses in the 'Habitation Index'.[7]

A sampling frame of this sort will contain some street addresses that simply do not exist, and will miss out some addresses that do exist. The best you can do is to ensure that these additions and subtractions are as 'random' as possible. That's why you should visit and amend your sampling frame with actual letter box numbers where multiple households are

---

6 The 'Habitation Index' itself is not suitable as a sampling frame - too many eligible voters do not register, and so many households are missing. But from the 'Habitation Index' you can get an idea of how many households are in each street, and using this information you can put together your own sampling frame.

7 You could also use Supermap (section A8.4) to find areas with a high proportion of flats or retirement villages. 'Dwelling type' data include percentages for 'flats or houses joined together' and for 'homes for the elderly'.

prevalent or in newly subdivided areas. Remember to document your procedures so that later you can explain what you did and why. You should also 'post-stratify' estimates (see section A8.4). If your sampling frame under-represents certain groups, estimates may be biased - with sensible 'post-stratification', adjusted estimates will be less biased.

The third method uses the phone book. Since the phone book gives addresses, you could use this for a sampling frame. But it's more likely your population of interest does not lie within a single local calling area; or your population of interest is only a small part of a local calling area; or you want to make estimates for different parts of your population of interest. In each case, the phone book is not going to be a satisfactory sampling frame. [Telecom's Directory Information Services can only supply you with phone numbers, not addresses.]

The fourth method is surveying on a large scale. Statistics New Zealand uses a 'meshblock' classification system to divide the country up into areas roughly the size of a city block (rural meshblocks tend to be larger). In theory, you could take a sample of meshblocks from those containing your population of interest, then list all the households in each of the sampled meshblocks, and sample some of these households and businesses. This is called cluster sampling. Several stages of sampling are involved: first a sample of meshblocks, then a sample of houses and businesses within each sampled meshblock. While this is the best way to survey a large or high density area (such a city), you are going to need a statistician. Calculating the required sample size, calculating averages, proportions and their variances are all more difficult with a cluster sample.

One other alternative is to contract a council to run the survey for you, using a questionnaire that is acceptable to both you and the council. [The council is unlikely to be able to give you a sample of their ratepayers, because of the Privacy Act 1993].

A8.2 Selecting the sample

With numbered houses on an aerial photo, use a computer to draw a simple random sample. Do not use a systematic sample in this situation because you are unlikely to number the houses in a random order (see section A4.2). On the other hand, take a systematic sample if you are using the phone book as the sampling frame. Use section A5.2 to calculate the sample size for a simple random or systematic sample.

If you construct a sampling frame of street names and house numbers in a spreadsheet, you could take a simple random sample, or use a systematic sample provided the streets are in alphabetical order. Or you could arrange the streets into strata, and take a simple random sample (or systematic sample) from each stratum. Each stratum should be as similar within and as different between as possible (that is, similar and different in terms of the community's perception of odour). Read section A6.1 on how to divide a population up into strata; read section A6.2 on how to calculate the sample size for a stratified random sample.

A8.3 Estimation

Use section A5.3 to calculate estimates from a simple random or systematic sample. Use section A6.3 to calculate estimates from a stratified random sample.

A8.4 Non-response

To account for those you cannot contact, use the method described in section A5.4 (for simple random sampling) and in section A6.4 (for stratified random sampling). A random sample of those not available in first attempts at an interview is used to adjust the estimate of an average or proportion.

You can use a similar adjustment to account for those who refuse to be interviewed. This 'basic question' approach is described in section A6.4. You should identify one question (maybe two questions at most) that best summarises what your survey is about. Because this question must be asked of both respondents and non-respondents in as similar circumstances as possible, this 'basic question' has to be one of the first you ask in your questionnaire. When someone refuses to participate, ask if they will just answer this one question. Answers from those who complete a full interview and from those who answer just the 'basic question' are then combined to give an adjusted estimate of an average or proportion.

The next method of adjustment can reduce not only the bias due to non-response (section A4.4), but also the bias in estimates due to inadequacies in a sampling frame. So this method of adjustment will be particularly useful if you've had to construct the sampling frame yourself, using a street map and the 'Habitation Index' (see section A8.1).

This third method is called 'post-stratification'. But you need to think about its use before you survey, because you have to find out which 'post-strata' each respondent belongs to by asking the appropriate questions in your survey. 'Post-strata' typically involve groups based on say age, sex or ethnicity. Like the usual sort of geographically based strata (section A6.1), 'post-strata' should be as similar within the group and as different between groups as possible - similar and different with respect to perceptions of odour. While you could form 'post-strata' within geographical strata, this would involve a large overall sample size. Each 'post-stratum' needs a sample size of at least 20 (Sarndal *et al.*, 1993, p267), so you are most likely to use 'post-stratification' in conjunction with simple random sampling. There is a more efficient way to 'post-stratify' across (rather than within) geographical strata, but the calculations are not for the faint-hearted (*ibid.*, pp268-269). We will just consider 'post-stratification' as it applies to simple random sampling.

You need to know the frequency with which each group occurs in your population of interest. The easiest way to find this out is using Supermap - a Statistics New Zealand database on CD ROM. You will find Supermap at major public libraries, polytechnics and universities. Using Supermap, you can identify the meshblocks (see section A8.1) that make up your population of interest. Supermap will give the number of people in the meshblock at the last Census, by age, sex, ethnicity and many other variables.

So to use this method, perceptions of odour should vary between groups and you have to be able to get data from Supermap for each of these groups. Groups based on age or on work status (full time, part time, or not in the labour force) are likely to fit these two criteria. [Even if perceptions don't vary between groups, 'post-stratification' won't increase the bias in estimates.] If those in flats and retirement villages make up a significant part of your population and you think they are likely to be under-represented in your sampling frame (and this may bias estimates), then form groups based on age. If you are more concerned about bias in estimates because of low response rates, then form groups based on work status. You might form groups based on other variables - it depends on what sort of people you think are

under-represented in your sample (either because of problems with your sampling frame or because of non-response). Since you have to ask questions in your survey to establish group membership, you could look at the way these questions were asked in the last Census. You may also need to check the definitions used in the last Census - for concepts such as 'part time' or 'not in the labour force' (Department of Statistics, 1991).

To adjust estimates, replace the usual sample average ($\bar{y}$ 2 in equation 7) with (Cochran, 1977, p134):

$$\bar{y}_w = \sum_{g=1}^{G} \frac{N_g}{N} . \bar{y}_g \qquad (25)$$

$\bar{y}_w$ = post – stratified estimate of the sample mean;
N_g = number of people in group g;
N = add up number of people in all G groups;
$\bar{y}_g$ = sample average for those in group g.

Since a proportion is just a special sort of average, you can replace the averages in the above equation with the appropriate proportions.

The 'post-stratified' estimate weights each group average by the frequency with which that group occurs (N_g/N). That's why it doesn't matter too much if meshblock boundaries don't coincide exactly with the boundaries of your population of interest. If there's a slight mismatch, it probably won't change the group frequencies much. 'Post-stratification' reduces both the bias in estimates, and the variance. So you can use the usual variance calculation (equation 8 or 11) because the result will be conservative. You could get a statistician to calculate a more accurate variance or to 'post-stratify' using several variables (perhaps using both age and work status) or to 'post-stratify' across geographical strata (Cochran, 1977, p107) .

A8.4.1 *You are going to ask respondents if they have experienced any offensive odours in the last year. You identify all the streets and part streets in your population of interest from a street map. You use the 'Habitation Index' and visits to construct a sampling frame, and you then take a simple random sample of 80 out of 400 households. You think your sampling may under-represent flats in this lower socio-economic area. From Supermap, you find that the area you're interested in includes most of six meshblocks. You add up the people in these six meshblocks, for each of three age groups: 0-29, 30-59, 60 and over. In your questionnaire, you ask respondents which age group they belong to and after the survey, you estimate a sample proportion for each age group (pg). The data are in the table*

below. The 'post-stratified' estimate for the proportion who say 'yes' is (0.50x0.33)+(0.30x0.24)+(0.20x0.18) = 0.27. The usual sample proportion is 20/80 = 0.25.

'Post-stratum'	Ng	Ng/N	'yes'	n	p_g 3
0-29	600	0.50	7	21	0.33
30-59	360	0.30	10	42	0.24
60+	240	0.20	3	17	0.18
Total	1200	1.00	20	80	

APPENDIX B:

Calculating the Chi-Squared Statistic[8]

Table of Contents:

8 Includes contributions from Debbie Singh, Department of Sociology, University of Auckland.

APPENDIX B : CALCULATING THE CHI-SQUARED STATISTIC

In this appendix we show how to calculate the chi-squared statistic and its p-value. The data in Table 3.2.11 (section 3.2.13.3) is used as an example:

Origin of Odour	Offensiveness		
	'Not at all'	**'A little'**	**'Very'**
Source of interest	4 (8.1)	15 (16.7)	15 (9.2)
Other sources	3 (3.1)	8 (6.4)	2 (3.5)
Don't know	8 (3.8)	8 (7.9)	0 (4.3)

There are three steps involved:

1. Calculate the expected counts for each cell;
2. Calculate the chi-squared statistic from observed and expected counts;
3. Use a chi-squared table in a statistical textbook to find the p-value.

B1 Expected Counts

The expected count in each cell is given by:

$$E_{ij} = \frac{R_i \times C_j}{n},$$

E_{ij} = expected count in row i column j,

R_i = total count in row i, 0

C_j = total count in column j,

n = total count in table.

So for the cell in the first row and first column, the expected count is (4+15+15) × (4+3+8) / 63 = 8.1 as shown. Note adding up all the observed counts gives the total count for the table as n = 63.

B2 Chi-Squared Statistic

The chi-squared statistic (symbol χ^2) is calculated as:

$$\chi^2 = \sum_{ij} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}, \quad 0$$

O_{ij} = observed count in row i column j.

The chi-squared statistic adds up contributions from each cell. Each cell's contribution is the squared difference between observed and expected counts, divided by the expected count.

For the first cell, the contribution is $(4\text{-}8.1)^2 / 8.1 = 2.08$. Contributions for all nine cells are 2.08, 0.17, 3.66, 0.00, 0.40, 0.64, 4.64, 0.00, 4.30 and these add up to 15.89.

B3 P-Value

A computer programme uses a complicated formula to calculate the p-value. This is not something that is easily done by hand. But most statistical textbooks have a chi-squared table at the back in an appendix. The table will have rows labelled 'Degrees of freedom' or 'DF'; the columns will be for different values of probability; and the cells of the table give the value of chi-squared for that combination of DF and probability.

The appropriate degree of freedom is given by DF = (I-1) × (J-1), where I is the number of rows in your table and J is the number of columns. So for this example DF = (3-1) × (3-1) = 4. Having calculated the DF, find the appropriate row in the textbook's table, look along the row to find chi-squared values higher and lower than the one you calculated, and then look up to see the probabilities associated with these higher and lower values.

For DF = 4 in Berenson, Levine and Rindskopf (1988), the row starts with chi-squared 0.2 for probability 0.995 and finishes with chi-squared 12.8 for probability 0.005. At 15.9, the calculated chi-squared statistic is larger than the last value in this row, so all we can say is that the probability of the calculated chi-squared is less than 0.005. So with a textbook we would conclude that the p-value was $p < 0.005$. A computer program would calculate a p-value of $p = 0.003$.