Part 2

Methodology report

TABLE OF CONTENTS

1 BACKGROUND AND OBJECTIVES

Food Standards Australia New Zealand (FSANZ) is an independent bi-national organisation that has the role, in collaboration with other organisations, to protect the health and safety of the people in Australia and New Zealand through the development of food standards for both countries and through related functions.

Hence one of the principle objectives behind the development of new food standards is to ensure that labels are easier to interpret and that they deliver information that is easy to understand and use, thereby enabling consumers to make informed choices about the foods they purchase.

In response to suggestions by the Australian National Audit Office in 1999, FSANZ has developed a means to provide evaluation about the impact of the new joint *Australia New Zealand Food Standards Code* (the Code), how well the regulatory arrangements are working, and the level of monitoring and enforcement activity.

In late 2001 and early 2002, NFO Donovan Research conducted qualitative studies with consumers and stakeholders in Australia and New Zealand about food labels. The information gained from these studies has formed the basis of the current quantitative study with consumers. The results from the overall research program will form the basis of recommendations to the FSANZ Board regarding the labelling provisions in the Code.

This Report summarises the methodology for the study. A separate report documents the survey pilot in more detail than is covered in this report.

1.1. Research objectives

The objectives for this study are to quantify the extent of consumer awareness, knowledge and behaviour in relation to food labelling by providing baseline data on:

- what information consumers currently look for in the labels of packaged food;
- whether consumers are familiar with the various major labelling components and new labelling information;
- whether and how consumers use, or would intend to use such information, and their reasons;
- concern about the clarity and trust worthiness of label elements;
- whether changes to labelling have changed consumers' purchasing decisions,
- which label elements consumers find difficult to interpret.

2 METHODOLOGY

2.1. Questionnaire Design

The development and testing of the questionnaire included three key stages:

1. Consultation and extensive feedback from key FSANZ staff in Australia and New Zealand, on a **questionnaire framework** developed by NFO Donovan Research (Appendix A);
2. Question **pre-testing** using cognitive interviewing and group discussions; and
3. Formal **pilot** 'dress rehearsal' via a two stage process, with 30 interviews in Australia that resulted in some adjustments to the questionnaire that were then tested by 20 interviews in New Zealand, resulting in n=50 interviews. Pilot results were not included in the final analysis.

In addition to the first stage of feedback against the questionnaire framework, key FSANZ staff in both countries and the FSANZ project team provided ongoing and valuable feedback and comment on the questionnaire throughout its development.

Stimulus materials were also pre and pilot tested. They included a set of *picture cards* of labelling elements, *label mock ups* and *show cards* with response choices typed on them.

The questionnaire and stimulus materials are attached at Appendix B and C. A more detailed discussion of the three stages of questionnaire development, and FSANZ's involvement is reported in the Pilot Methodology Report.

2.2. Sample Definition and Quotas

The sample included males and females aged over 18. Due to the substantial number of interviews being conducted, demographic quotas such as age and gender were not set, however people were excluded from participating in the research if they were aged under 18 or worked in the areas of advertising, market research, nutrition or the food industry.

Proportional quotas by location were set, to ensure that the results correctly reflected each country's proportion of the population in metropolitan areas.

2.3. Data Collection

Interviewing in Australia was sub-contracted to an Interviewer Quality Control Australia (IQCA) accredited field team, SurveyTalk. In New Zealand, interviewers were conducted by the NFO New Zealand field team. Prior to the survey pilot, interviewers attended a 3-4hr face to face briefing. This briefing included the provision of comprehensive briefing notes covering project background and the purpose of the survey and the pilot, an overview of food labelling changes and the 14 labelling elements included in the questionnaire, and the use of the stimulus

materials. Particular attention was paid to probing techniques and the importance of consistency in probing amongst interviewers. The briefing included role play and interviewers practiced interviewing each other to help them become more familiar with the questionnaire, and to prompt them to ask questions and clarify potential problems as they progressed through the interview. This training was repeated by the fieldwork supervisors for all interviewers involved in the main survey.

Sample households were selected using the random drawing of start points and "Clover Leaf Method" for the door-to–door interviews. Start points were allocated randomly using the North, South, East, West principle, based on either the number of dwellings in an area or the population of an area. This selection ensures that every dwelling within an area had an equal chance of being randomly selected, thus maximising the representativeness of the sample. A Map area was passed on to the interviewer and from their start point they followed a specified route (observing 'left shoulder to the wall' or 'right hand rule'), until they had interviewed the desired number of respondents.

Interview locations in Australia included metropolitan cities only in all States and Territories (Sydney, Melbourne, Brisbane, Perth, Adelaide, Hobart, Canberra and Darwin). In New Zealand interviewing was conducted in the North and south Islands in Auckland, Christchurch, Whangerei, Hamilton, Tauranga, Rotorua, Gisborne, Invercargill, Napier, Hastings, Palmerston North, Dunedin, Wanganui, Nelson, and Timaru.

The interviews were conducted on weekdays, weekday evenings and weekend day times between 17th August and 28th September 2003 in both Australia and New Zealand. The average interview length was 30 minutes, with the shortest interview being 6 minutes, and the longest 96 minutes.

The refusal rate for the interview was 16%, which is extremely low for a door to door study (the refusal rate is usually around 40%). Therefore the survey samples are more likely to be representative of the populations surveyed.

In order to maintain an acceptable interview duration and minimise respondent fatigue, several of the label interpretation questions were evenly rotated within the sample. Approximately half the sample was asked question 11 and half was asked question 12. Similarly questions 15-20 and asked in two batches, where approximately half the sample were asked questions 15-18, and the remainder were asked questions 19-20.

2.4. Sample Characteristics

	Australia (%)	New Zealand (%)	Total (%)
Grocery Shopper			
Yes	92	93	92
No	8	7	7
Any special dietary needs?[1]			
Yes	56	69	63
No	44	31	37
Age			
18-24	12	12	12
25-34	20	21	20
35-44	20	20	20
45-54	19	15	18
55-64	13	15	14
65+	15	16	16
Gender			
Male	29	36	33
Female	71	64	67
Language			
English	89	91	89
Other language	11	10	11
Highest Level of Education			
Primary school or less	5	2	4
Secondary	48	44	46
Trade/Diploma	25	25	24
Tertiary or higher	22	29	23
Income (p.a)			
under $40,000	41	46	43
40,000-74,999	23	23	23
75,000+	18	19	17
Household			
live alone	18	17	17
live with children	47	47	47
live with adults only	8	7	8
Age of Children			
under 3 years	11	9	11
3-11	33	39	35
12-17	22	26	23
18+	33	25	30
Health conscious			
Not at All Concerned	3	2	3
Moderately	38	49	42
Highly	58	49	55

[1] *Participants who specified a special dietary need e.g general health, food allergies etc.*

2.5. Comparisons to Census Data

The demographic characteristics of participants in both countries were compared to census data collected by the Australian Bureau of Statistics in 2001 and Statistics New Zealand in 2001 for age and education.

In both Australia and New Zealand, the proportion of people in each age group matched very closely to that of census data. This result is notable for the younger age groups, as the 18-24 year old age group typically has low participation rates in surveys. The New Zealand sample was statistically higher (3% points) in the proportion of 55-64 year olds than the Statistics NZ data. The Australian sample was statistically lower in the proportion of people aged 65+, compared to ABS data (5%).

Comparisons to population data for education were more difficult because of the different way that the data are categorised in each country. In Australia the sample contained a significantly greater proportion of people who have a degree than in the national population (sample 17%; population 10%). A similar trend was present in the New Zealand sample (sample 19%; population 7%).

Comparisons for other demographic characteristics such as income were not possible due to the lack of consistency in the way demographic data are reported by these organisations.

The gender distribution of the sample reflected that for household shoppers, although non-household shoppers were not excluded from the survey.

Due to the close proximity of the sample to all but one of the key demographic variables (level of education) and the lack of available comparable national statistics between each country, the data were not weighted.

2.6. Sampling Error and Statistical Significance

A result is statistically significant if the difference between it and another result is sufficiently large enough to make the possibility of sampling error or chance sample fluctuation low. In this survey, a test with a 95% confidence level was used, which means that in 95 cases out of 100, the difference in results reflect a 'real' difference rather than being a function of sampling error[2].

Sampling Error

Sampling error reflects the difference between an estimate derived from a survey and the 'true value' that would be obtained if the whole target population were surveyed. Several factors can affect the size of the sampling error, the main one being sample size. In general, larger samples give rise to smaller sampling error.

The larger the sample size, the greater confidence one can have that even small percentage changes in results reflect 'real' changes.

When interpreting data gathered from survey samples, the existence of sampling error must be taken into account, both when :

i) Making 'population estimates' from a single percentage, or

ii) When comparing between percentages.

i) Population Estimates

The table below presents the margin of error surrounding estimates from various sample sizes.

Table 1a : Sampling Error When Estimating From a Single Percentage (95% Confidence Level)

Sample Size	Survey Result		
	10% or 90%	20% or 80%	40% or 60%
50	8%	11%	14%
100	6%	8%	10%
200	4%	6%	7%
500	3%	4%	4%
1000	2%	2%	3%
2000	1%	1%	2%

Example :- A reported percentage of 20%, based on a random sample of n=1000, has an error rate of plus or minus 2%. That is, there is a 95% probability (ie 95% confidence) that the actual population percentage is between 18% and 22%.

[2] *The test method used is a t-test on mean scores at a 0.95 confidence level. The overlap formula was used.*

ii) Difference Between Sub-Groups

Whenever sub-group comparisons are made, it is important to distinguish between differences that are reliable (ie statistically significant) and those that are not (ie could be due to chance sample fluctuations, or sampling error). To assist in distinguishing between statistically reliable and unreliable differences the table below lists the size of the differences required to reach statistical significance (again at the 95% confidence level), for various sample sizes.

Table 1b : Difference Required to be Significant When Comparing Two Percentages (95% Confidence Level)

Average Sample Size of Groups Being Compared	Average of Two Percentages is . . .		
	10% or 90%	**20% or 80%**	**40% or 60%**
50	12%	16%	19%
100	8%	11%	14%
200	6%	8%	10%
500	4%	5%	6%
1000	4%	4%	5%

Example :- Say you are comparing the results from two sub-samples as follows :

Result A : 23% *agree* : n=750
Result B : 17% *agree* : n=1250

The average of the two sample sizes is 1000. The average of the two 'satisfaction' results is 20%. From the table above, the difference needed to be statistically significant is at least 4%. Therefore, you would be 95% confident that the difference between the two results is statistically reliable.

Reporting of Significant Differences

All analyses were conducted in the SPSS statistical analysis program, and significance tests were conducted using a Microsoft Excel file adapted by NFO Donovan Research for this purpose.

The results for the total sample population are reported, along with significant differences between the subgroups outlined in the previous table (see Section 2.4).

2.7. Presentation of Results

The presentation slides present sub group comparisons only where significant differences exist, or where they existed for a related question and thus are reported for the purposes of comparison. Statistical differences between two sub groups are marked with a star (☆).

It should be noted that where percentage totals do not add up to 100% within the tables or text, this is because 'don't know' and 'refused' responses have not been excluded from the total sample, and are not commented on. Also, overlap will occur for the some multiple response questions.

2.8. Segmentation analysis

Advanced analyses, such as logistic regressions were conducted to investigate the strength of predictive variables in determining label use, for example the special needs variable. These results were inconclusive and are therefore not reported in the Summary Report.

Segmentation analysis was conducted using GEMSegment, a sophisticated segmentation analysis that derives first the number of segments that exist, based on the greatest probability of that solution, versus alternative solutions with fewer or more segments. The GEMSegment analysis then allocates each respondent to a segment based on the greatest probability of an individual occurring in that segment, versus all other segments.

Segmentation analysis is conducted in order to group individuals into segments with like qualities. It is a way of examining whether the sample population is homogeneous (ie basically one single groups with similar attitudes and behaviour) or heterogeneous (ie comprised of several distinct segments, each with different priorities, preference, attitudes or behaviours).

Traditionally, evaluation analysis involves the disaggregation of respondents by demographic characteristics such as age, gender, income, education etc. In some cases, differences in label use can be explained by these types of characteristics however, if two 'like' consumers (with the same demographic characteristics) use different numbers of label elements, there are other factors at play which cannot be explained by traditional analysis.

This segmentation analysis and model attempts to understand what drives high or low label element use, in order to assist FSANZ in developing food standards in the future. The results could also be useful for developing education strategies.

The 'dependent' variable is therefore the number of label elements a consumer uses. The segment produces 'higher' and 'lower' users, compared to the 'average' number of label elements used by the sample population. The segmentation is an

attempt to simplify a very complex and partly irrational behaviour (both impulsive and habitual), and should therefore be used bearing this in mind.

The premise underlying the segmentation analysis emerged from the preceding qualitative research, which indicated that consumers' use of food labels varied enormously depending on their motivation, (and the reasons behind their motivation), and their capacity, including their success in previous attempts to interpret labels. The Model is therefore built around two dimensions, 'motivation' and 'capacity' to use labels:

- "Motivation" = health consciousness, special health needs, interest in food label information, and importance/usefulness of food label information.

- "Capacity" = past success in finding food label information; sufficient time to read labels while shopping; perceptions of label element clarity and trust of label elements.

Levels or strength of "motivation" and "capacity" are broadly expressed as 'high', 'moderate' and 'low' - these are qualitative terms applied to give relative meaning to motivation/capacity between the segments. Intuitively, it appears that the "motivation" dimension is stronger than the "capacity" dimension, but this proposition has not been tested statistically.