Statistical Standard for Ethnicity 2005

Definition

Rationale for Standard

This standard was developed because Ethnicity is a key social factor used with other topics in describing the New Zealand population. Information collected on ethnicity is used to inform, plan, and evaluate services and policies by a wide range of organisations, local authorities and government agencies.

Major uses of ethnicity data are:
- to monitor and report changes and disparities in outcomes among ethnic groups over time
- to monitor the changing ethnic diversity of New Zealand's population at national, regional and local levels, so that appropriate services may be delivered
- to estimate future trends through population estimates and projections for Māori, European, Pacific and Asian populations
- to monitor the demographic, social and economic progress of, and outcomes for, ethnic groups
- to evaluate the impact of central and local government policies on the economic and social well-being of ethnic groups
- to model the impacts and costs of policy changes, and to forecast expenditure on services for particular groups
- to assist in the delivery of services in a culturally appropriate way and to plan social services which meet the special needs of ethnic groups, and
- to identify significant communities of interest for liaison and development purposes.

The Statistical Standard for Ethnicity was developed to ensure that ethnicity is collected consistently for all surveys and administrative collections. Data from a large number of collections is combined with other sources, such as the population census, to produce official measures in a range of areas such as education, health, employment and unemployment, income, housing and crime. Unless consistent ethnicity data is available, valid and reliable measures cannot be produced. Lack of consistency across different collections means data may not be comparable.

For all of these considerations the Statistical Standard for Ethnicity 2005 encourages the use of a standardised concept, definition, collection, coding method, and output, as given in this standard, to promote data consistency and comparability in all official statistics.

This Statistical Standard has been developed from the recommendations in the Report of the Review of the Measurement of Ethnicity released in June 2004. This report established the continued need for collection of

detailed ethnicity data and the need to measure ethnicity in a consistent way across all official statistics.

Key recommendations from the Report of the Review of the Measurement of Ethnicity:

- The 2006 Census will use the same ethnicity question that was used in the 2001 Census.
- Statistics New Zealand will continue to educate respondents, users and producers of ethnicity data about the concept of ethnicity.
- A comprehensive programme of research into the measurement of ethnicity in official statistics will be completed by the end of 2009.
- All collections of official statistics measuring ethnicity should have the capacity to record and report six ethnicity responses for each individual, or the minimum of three responses when six cannot be implemented immediately.
- The method of reporting ethnicity in all collections of official statistics should be self-identification.
- Responses of 'New Zealander' and 'Kiwi' and similar responses will be classified separately from New Zealand European responses.
- The practice of prioritising ethnic group responses to one per individual will be discontinued.

Definition

Ethnicity is the ethnic group or groups that people identify with or feel they belong to. Ethnicity is a measure of cultural affiliation, as opposed to race, ancestry, nationality or citizenship. Ethnicity is self perceived and people can belong to more than one ethnic group.

An ethnic group is made up of people who have some or all of the following characteristics:

- a common proper name
- one or more elements of common culture which need not be specified, but may include religion, customs, or language
- unique community of interests, feelings and actions
- a shared sense of common origins or ancestry, and
- a common geographic origin.

This definition is based on the work of Smith (1986).

Operational Issues

Collection

Collection of ethnicity presents some difficulties. People report a range of aspects of their identities such as cultural affiliation, ancestry, nationality and race when asked for ethnic group identification. Evidence suggests that people may answer the question easily but not understand the ethnicity concept being asked for. Another difficulty is that some may

report one ethnic group but identify with more than one, or report more but in fact identify with fewer groups. Finally, a number of people object to answering an ethnicity question and may refuse to answer or may answer facetiously.

When collecting ethnicity information, people need to be able to state their specific ethnic groups without being forced to identify themselves in a more general category. Detailed ethnic group information is to be collected in order to allow categorisation at the most detailed level of the Standard Classification of Ethnicity, level four. Data can be aggregated into a smaller number of categories as users require.

Where it is not possible to collect data at level four of the classification, for instance in administrative data collections where written responses are not able to be coded, ethnic group information should be collected at level two of the classification which is less detailed.

All collections of official statistics measuring ethnicity should have the capacity to capture six ethnicity responses per person. Where this cannot be implemented immediately, it is recommended that a minimum of three ethnicity responses be collected.

Ethnic group changing over time (ethnic mobility)
The ethnic group or groups that someone identifies with may change over time. It is necessary to allow for ethnic mobility in longitudinal surveys and administrative databases. Ethnic mobility also affects the integration of different data sets as the same person may have given different ethnic group answers in different collections. Rather than using both data sets' responses, the decision on what is appropriate to use needs to be decided on a case by case basis.

Ethnic group changing with context
A difficulty that is not easily overcome when collecting ethnic group information is the possibility that a person may give a different response depending on the context. For example, when filling in a self-administered form a person may respond differently from when asked his/her ethnic group by an interviewer. Also, the social or cultural setting may affect the ethnicity response reported. A decision on what is appropriate to use for integrated data sets needs to be decided on a case by case basis.

Legal age
Ethnicity is self-defined. While no legal or recommended age has been set at which a child can respond on their own behalf, the expectation is that teenagers will self-identify their ethnicity.

Ethnicity collection by proxy
Statistics New Zealand collects a self-identified concept of ethnicity. In some circumstances a person may be unable to answer this question and the next-of-kin, parent, spouse or partner needs to respond on their behalf, for example, in the case of birth or death, or incapacity because of disability, injury or sickness. Also parents, caregivers or guardians of a child may complete an ethnicity question on behalf of their child.

Multiple ethnicity

People may identify with more than one ethnic group so when collecting ethnicity data there needs to be provision to collect multiple ethnic groups for each individual. It is recommended that six ethnic group responses per person be collected where possible. The ability to collect three responses is the minimum requirement to meet the standard. See Coding Process for more detail.

Explanatory Notes

Race/Ancestry/Citizenship/Ethnic origin

Ethnicity should not be confused with other related terms. Race is a biological indicator and an ascribed attribute. Ancestry is a biological and historical concept and refers to a person's blood descent. Citizenship is a legal status. These terms contrast with ethnicity which is self-perceived and a cultural concept. Ethnic origin is a person's historical relationship to an ethnic group, or a person's ancestors' affiliation to an ethnic group, whereas ethnicity is a person's present-day affiliation.

Business, family and household

Ethnicity is a personal attribute and therefore it is not valid to attribute an ethnicity to a business, family or household based on the ethnicity of an individual within that business, family or household.

New Zealander responses

Prior to the introduction of this Statistical Standard, a New Zealander response was included in the New Zealand European category. In the Standard Classification of Ethnicity, New Zealander and similar responses like Kiwi are classified to a separate ethnic group category at the most detailed level, level four. This category is called 'New Zealander'. For time series purposes, the counts of the New Zealander category can be added to the counts of the New Zealand European category to recreate a count for the New Zealand European category which will be comparable to those from previous data collections.

Classification and Coding Process

Classification Criteria

The criterion for classification of ethnic groups is self-identification with one or more ethnicities. Detailed ethnic group information is collected so that responses can be classified to specific ethnic group categories at the most detailed level of the classification, level four. Where this is not possible, information may be classified to less detailed levels, level two or level three. Level one is used solely for output.

Individual ethnic groups are aggregated into progressively broader ethnic groups from level three up to level one, according to geographical location or origin, or cultural similarities. The classification reflects responses received and is made up of geographic, nationality and ethnic group terms.

The Standard Classification of Ethnicity has been developed from the recommendations in the Report of the Review of the Measurement of Ethnicity released in June 2004.

Classification

The Standard Classification of Ethnicity is a hierarchical classification of four levels.

Level one of the classification has seven categories, including one residual category, and is used solely for output.

Level two of the classification has 27 categories, including six residual categories.

Level three of the classification has 42 categories, including six residual categories.

Level four of the classification has 239 categories, including six residual categories.

Levels two, three and four are used for collection without the 'not further defined' (nfd) categories and for output with the nfd categories.

Coding pattern

Level	Digits	Code	Descriptor
One	one	3	Pacific Peoples
Two	two	32	Cook Islands Maori
Three	three	321	Cook Islands Maori
Four	five	32121	Rarotongan

Level one

1 European
2 Māori
3 Pacific Peoples

4 Asian
5 Middle Eastern/Latin American/African
6 Other Ethnicity
9 Not Elsewhere Included

Level two
10 European nfd
11 New Zealand European
12 Other European
21 Māori
30 Pacific Peoples nfd
31 Samoan
32 Cook Islands Maori
33 Tongan
34 Niuean
35 Tokelauan
36 Fijian
37 Other Pacific Peoples
40 Asian nfd
41 Southeast Asian
42 Chinese
43 Indian
44 Other Asian
51 Middle Eastern
52 Latin American
53 African
61 Other Ethnicity
94 Don't Know
95 Refused to Answer
96 Repeated Value
97 Response Unidentifiable
98 Response Outside Scope
99 Not Stated

Note: Nfd categories and residual categories are not used for collection. They are used for coding and output purposes. Residual categories may be aggregated to one category, Not Elsewhere Included, for output.

Level three and level four
Levels three and four of the classification can be found in Appendix 1.

The residual categories 'Don't Know', 'Refused to Answer', 'Repeated Value', 'Response Unidentifiable', 'Response Outside Scope' and 'Not Stated' are defined under the heading 'Glossary and References'.

The Statistical Standard for Ethnicity and the Standard Classification of Ethnicity can be downloaded from the Statistics New Zealand website: www.stats.govt.nz.

Within Statistics New Zealand, the classification is stored in the Classifications and Related Standards System (CARS) under the subject 'Ethnic' and the topic 'Ethnicity'.

Coding Process

The ethnicity question allows for multiple responses from each individual as described by the definition in this Standard. Ethnicity is self-perceived and some people identify with more than one ethnicity.

In the Census of Population and Dwellings and large surveys up to six responses per person are recorded for the ethnicity question. For administrative collections and other surveys, three may be the maximum number of responses recorded per person. When more responses are given than can be recorded, a random method for reducing the number of responses selects the ethnicities to be retained (see Appendix 2 for the description of a manual methodology for reducing multiple ethnicity responses; see Appendix 3 for a link to a technical paper for software developers).

Codefile

Statistics NZ maintains a codefile which is updated with responses that occur in the Census of Population and Dwellings. A codefile is used to classify ethnic group responses. It is a list of probable survey responses and the classification categories to which they are coded. For example, the codefile for ethnicity contains the names of countries, similar terms used by ethnic groups to describe themselves, such as Slav and Slavic, abbreviations, and some common misspellings. Collectors of ethnicity information who wish to use the codefile may obtain it from Classifications and Standards, Statistics New Zealand.

Coding multiple-worded responses

These coding guidelines are given to ensure consistency between collections. There are a number of ethnic groups that are multiple-worded responses but are one ethnic group. Some common examples are given here:
Fijian Indian
Turkish Cypriot
Cook Islands Maori
French Canadian
Malaysian Chinese
American European
South African European.

There are responses that may be hyphenated or linked in some way, or written without linkage, that need to be classified as two responses. For example:
Polish-Hungarian
Tongan-Māori
French/Austrian
Australian/Lebanese
Chinese New Zealander
SerboCroatian.

Coding responses

Iwi
An iwi response to an ethnicity question is coded to Māori.

Country
A country response is coded to an appropriate ethnic group term, for example, Korea is coded to Korean.

Religion
Religious responses to the ethnicity question indicating an ethnic group are coded to the specific category in the classification. For example, Jewish and Sikh have separate categories at the most detailed level. Religious responses which are not an ethnic group, for example, Muslim, are not coded to an ethnic group, but to 'response outside scope'.

New Zealander
A New Zealander response and like responses such as 'Kiwi' or 'NZer' are coded to a separate category, 'New Zealander', at level four in the Other Ethnicity group.

Questionnaire Module

Requirements

The questionnaire module needs to collect ethnicity information in keeping with the ethnic group concept contained in this Statistical Standard. The ethnic concept used is self-identification. In recognition that some people belong to more than one ethnic group, provision for multiple responses per individual is required. Also, people need to be able to state their specific ethnic groups without being forced to identify themselves in a more general category. To meet these requirements the question needs to provide for collection of detailed information to allow classification at level four. This will mean that a write-in area for responses is mandatory.

Data collected at the most detailed level of the classification can be aggregated into a smaller number of broader categories for output as users require.

It is recognised that some collections of ethnicity data, and in particular administrative data collections, may be unable to collect information at this level of detail. In these cases collection should use level two of the classification. Multiple responses for each individual need to be collected.

Note: Level two collection. Analysis of data may be compromised by those people who choose to write a response when a valid tick box is available. For example, people who answered Rarotongan rather than ticking Cook Islands Maori cannot be accurately coded unless write-in responses are processed. Another issue with level two collection is that it may not be comparable with data collected at level four if those who identify as New Zealander do not mark the 'other' tick box. These issues need to be taken into consideration when deciding to collect at level two of the classification.

Example
Below are examples of questionnaire modules by mode. Other questionnaire modules may vary in format but should conform to the requirements contained in this standard.

1. Self-administered forms

This is the ethnicity question for the 2001 and 2006 Censuses:

11 Which ethnic group do you belong to?
Mark the space or spaces which apply to you.

- New Zealand European
- Māori
- Samoan
- Cook Island Maori
- Tongan
- Niuean
- Chinese
- Indian
- other (such as DUTCH, JAPANESE, TOKELAUAN). Please state:

2. Internet

There is a button to click on for each option below.

Which ethnic group do you belong to?
Mark the space or spaces which apply to you.

* New Zealand European
* Māori
* Samoan
* Cook Islands Maori
* Tongan
* Niuean
* Chinese
* Indian
* Other

Please state: eg Dutch, Japanese, Tokelauan
(Another box opens to write in when 'Other' is clicked)

3. Interviewer administered

The interviewer states: please use this card to tell me which ethnic group or groups you belong to.

Showcard

O   New Zealand European
O   Māori
O   Samoan
O   Cook Islands Maori
O   Tongan
O   Niuean
O   Chinese
O   Indian
O   Other (such as Dutch, Japanese, Tokelauan). Please state.

The interviewer ticks all that apply. Asking the question in this way allows for more than one ethnicity to be selected. It also allows reporting of all other ethnic groups chosen by the person in the open ended 'other' category. It facilitates self-identification and allows the person to pick one or a number of categories that they identify with. This method reduces interviewer bias.

4. Telephone interviewing

The interviewer states: I am going to read out a list of ethnic groups. Can you tell me which ethnic group or groups you belong to:

New Zealand European?
Māori?
Samoan?
Cook Islands Maori?
Tongan?
Niuean?
Chinese?
Indian?
Another ethnic group such as Dutch, Japanese or Tokelauan? Please say what it is.
The interviewer should read out each of the categories and wait for a yes/no answer to each. When a yes answer is given, the interviewer continues asking the rest of the list until it is completed.

Asking the question in this way allows for more than one ethnicity to be selected. It also allows reporting of all other ethnic groups chosen by the person in the open ended 'Another ethnic group' category. It facilitates self-identification and allows the person to pick one or a number of categories that they identify with. This method reduces interviewer bias.

5. Administrative and other collections without write-in coding facilities for 'Other'

Which ethnic group do you belong to?
Mark the space or spaces which apply to you.

New Zealand European
English
Australian
Dutch
Other European
Māori
Samoan
Cook Islands Maori
Tongan
Niuean
Tokelauan
Fijian
Other Pacific Peoples
Filipino
Chinese
Indian
Japanese
Korean
Cambodian
Other Asian
Middle Eastern
Latin American
African
Other

Note: Responses to Other are output as Other Ethnicity.

A coding tool is available from Classifications and Standards, Statistics New Zealand, to enable coding of responses to the correct category in level two of the classification, if people are unsure where to mark their response(s).

Output

Standard Output

Ethnicity Total Response Standard Output
Total response output shows the counts of all responses given for each ethnic group (see glossary). Output tables need to state how many ethnicities are output per person as the number of ethnic groups recorded may differ between collections.

Standard footnote for total response ethnicity output
People who reported more than one ethnic group are counted once in each group reported. This means that the total number of responses for all ethnic groups can be greater than the total number of people who stated their ethnicities.

Note: Where output is for explicit ethnic groups then the footnote should state the names of those groups. For example, "People who reported more than one ethnic group within the Pacific Peoples ethnic group are counted once in each group reported".

Total response standard output: Level one
European
Māori
Pacific Peoples
Asian
Middle Eastern/Latin American/African*
Other Ethnicity
Not Elsewhere Included

* Middle Eastern/Latin American/African may be abbreviated to MELAA with a note explaining its composition

Total response standard output: Level two
European nfd*
New Zealand European
Other European
Māori
Pacific Peoples nfd*
Samoan
Cook Islands Maori
Tongan
Niuean
Tokelauan
Fijian
Other Pacific Peoples
Asian nfd*
Southeast Asian
Chinese
Indian
Other Asian

Middle Eastern
Latin American
African
Other Ethnicity
Not Elsewhere Included

* not further defined

Total response standard output: Level three
European nfd*
New Zealand European
Other European nfd*
British and Irish
Dutch
Greek
Polish
South Slav
Italian
German
Australian
Other European
Māori
Pacific Peoples nfd*
Samoan
Cook Islands Maori
Tongan
Niuean
Tokelauan
Fijian
Other Pacific Peoples
Asian nfd*
Southeast Asian nfd*
Filipino
Cambodian
Vietnamese
Other Southeast Asian
Chinese
Indian
Sri Lankan
Japanese
Korean
Other Asian
Middle Eastern
Latin American
African
Other Ethnicity
Not Elsewhere Included

* not further defined

Total response standard output: Level four
The classification categories with counts of 1,000 or more may be output separately.

Single/Combination Output
Single/combination output places each person into only one category (see glossary).

Standard footnote for single and combination output
People are counted just once according to the ethnic group or combination of ethnic groups they have reported. This means that the total number of responses equals the total number of people who stated an ethnicity.

Total response output shows the counts of all responses given for each ethnic group (see glossary). Output tables need to state how many ethnicities are output per person as the number of ethnic groups recorded may differ between collections.

Examples of single/combination outputs are given below. The amount of detail supplied in single/combination output will depend on individual surveys and collections.

Level one single/combination output can be used for large collections where there is sufficient data. This comprises 14 groups and a residual category. For sample surveys and small collections seven groups and a residual category may be used. The detailed single/combination output is for the use of population surveys, such as the census.

It is important when using single/combination output that level one total response output is also shown. This enables the counts of a level one ethnic grouping to be calculated (for example, Māori) and is necessary because some of the combination categories do not name all of the ethnic groups contained within them. For example, 'Four to Six Groups' may contain people from any of the level one categories.

Single/combination output (15 Groups)
European Only
Māori Only
Pacific Peoples Only
Asian Only
MELAA* Only
Other Ethnicity Only
Māori/European
Māori/Pacific Peoples
Pacific Peoples/European
Asian/European
Two Groups Not Elsewhere Included
Māori/Pacific Peoples/European
Three Groups Not Elsewhere Included
Four to Six Groups
Not Elsewhere Included

* MELAA refers to Middle Eastern/Latin American/African

Note: It is essential that total response counts (Grouped Total Responses) are reported with single/combination data so that the counts of each level one group are available.

Single/combination output (8 Groups)
European Only
Māori Only
Pacific Peoples Only
Asian Only
MELAA* Only
Other Ethnicity Only
Two or More Groups
Not Elsewhere Included

Detailed single/combination output

Single Ethnic Group
European Only
Māori Only
Pacific Peoples Only
Asian Only
MELAA* Only
Other Ethnicity Only

Two Ethnic Groups
European/Māori
European/Pacific Peoples
European/Asian
European/MELAA
European/Other Ethnicity
Māori/Pacific Peoples
Māori/Asian
Māori/MELAA*
Māori/Other Ethnicity

Pacific Peoples/Asian
Pacific Peoples/MELAA
Pacific Peoples/Other Ethnicity

Asian/MELAA
Asian/Other Ethnicity

MELAA/Other Ethnicity

Three Ethnic Groups
European/Māori/Pacific Peoples
European/Māori/Asian
European/Māori/MELAA
European/Māori/Other Ethnicity
European/MELAA/Other Ethnicity
Pacific Peoples/European/Asian
Pacific Peoples/European/MELAA
Pacific Peoples/European/Other Ethnicity
Asian/MELAA/European
Asian/European/Other Ethnicity

Māori/Pacific Peoples/Asian
Māori/Pacific Peoples/MELAA
Māori/Asian/MELAA
Māori/Asian/Other Ethnicity
Māori/MELAA/Other Ethnicity
Pacific Peoples/Māori/Other Ethnicity
Pacific Peoples/MELAA/Other Ethnicity
Asian/Pacific Peoples/MELAA
Asian/Pacific Peoples/Other Ethnicity
Asian/MELAA/Other Ethnicity

Four Ethnic Groups
Five Ethnic Groups
Six Ethnic Groups
Not Elsewhere Included

* MELAA refers to Middle Eastern/Latin American/African

Standard footnote for single and combination output
People are counted just once according to the ethnic group or combination of ethnic groups they have reported. This means that the total number of responses equals the total number of people who stated an ethnicity.

Residual Categories for Output
94 Don't Know
95 Refused to Answer
97 Response Unidentifiable
98 Response Outside Scope
99 Not Stated

The residual categories may be output separately or combined. Where a combination item of residuals is to be used in output, this item should be labelled 'Not Elsewhere Included' and should have a footnote indicating its composition. 'Repeated Value' should not be used in any official output.

The output category sets are stored in the Classifications and Related Standards System (CARS), Statistics New Zealand, under the subject 'Ethnicity' and the topic 'Ethnicity'.

Data Comparability

If three ethnicities per person are collected, any data comparisons with other collections should be made with a similar method and the same number of ethnicities per person. For example, comparing a survey and census data: if three ethnicities are collected by the survey then data collected by census must be reduced to three ethnicities per person using the same methodology before comparison takes place.

Output at level one and two may be affected if collection at level two fails to collect New Zealander responses with the 'Other' tick box. Comparability may also be affected if people do not choose a valid tick box from the list and mark the 'Other' tick box instead. See the paper *Understanding and Working with Ethnicity Data* (details in Appendix 4) for more details.

Time Series

A concordance that links the categories from the previous classification ETHNIC to ETHNIC05 will be available from Classifications and Standards, Statistics New Zealand.

Level one data is not comparable to previous level one data with the introduction of the level one groups, MELAA and Other Ethnicity. For time series purposes, the counts of the New Zealander category can be added to the counts of the New Zealand European category (except where people have given both New Zealander and NZ European responses). This will recreate a count for the New Zealand European category which will be comparable to those from previous data collections. See the paper *Understanding and Working with Ethnicity Data* (details in Appendix 4) for more details.

Time series of level two data may be affected if collection at level two does not produce similar data (see data comparability and questionnaire module requirements).

Related Classifications and Standards

New Zealand

Ethnicity
The ETHNIC classification used prior to this standard will be concorded to the revised ETHNIC05 classification.

Māori descent
This classification records whether a person is descended from a Māori. It is a biological indicator and not an ethnic group indicator. A person may have Māori descent but not indicate that they belong to the Māori ethnic group. Also, some people indicate that they belong to the Māori ethnic group but do not have Māori ancestry.

Iwi
Ths is a classification of iwi and is used to record iwi affiliation for those descended from Māori.

International

There are currently no other countries measuring ethnicity in the same way as New Zealand. Other countries collect statistics on one or more of these related variables: nationality, ethnic origin, race, citizenship, immigrant status, country of birth, language, religion and ancestry. Definitions of these variables may differ from New Zealand use and will make data incomparable. For example, the United Kingdom uses the term ethnic group in its 2001 Census but asks respondents to give an answer based on colour. Fiji collects what appears to be ethnic group but it is described as ethnic origin and also as a person's racial origin.

Glossary and References

Glossary

Classifications and Related Standards System (CARS)
The Classifications and Related Standards System (CARS) is a computer system for the storage of all economic, social and geographic classifications data used in Statistics New Zealand. The data stored includes standard classifications, historical classifications required for the analysis of historical data, and survey specific classifications which are not standard.

Cultural affiliation
Cultural affiliation is the social, historical, geographical, linguistic, behavioural, religious, and self-perceived affinity between a person and an ethnic group.

Detailed output for counts of 1,000 or more
The detailed ethnic group categories at level four are output when the total number of responses to an ethnic group is 1,000 or more. All the remaining ethnic group responses are counted in the category 'Other Ethnic Groups'.

People who do not report their ethnicity are counted in the 'Not Stated' category or 'Not Elsewhere Included' group.

Single and combination ethnicity output
People are counted just once according to the ethnic group or combination of ethnic groups they have reported. A person reporting just one ethnic group is counted once in the relevant 'only' group (that is, single ethnic group). People reporting two or more ethnic groups are counted once in the relevant 'combination' group. This means that the total number of responses equals the total number of people who stated their ethnicity.

When someone reports two or more ethnic groups within the same level one group the person would be counted once in the single group. For example, a person who reported 'English' and 'Scottish' ethnic groups would be counted once in the 'European Only' output group.

People who do not report their ethnicity are counted in the 'Not Stated' category or 'Not Elsewhere Included' group.

Level one output groups are European, Māori, Pacific Peoples, Asian, MELAA (Middle Eastern, Latin American, African) and Other Ethnicity. Level four is a list of individual ethnic groups which are grouped together to make up the level one output groups.

Residual Categories

Not elsewhere classified (nec)
A 'not elsewhere classified' (nec) ethnic group category contains ethnicity responses that are infrequent or unanticipated. For example, Asian nec contains ethnic groups with small counts such as Bhutanese.

Not further defined (nfd)
A 'not further defined' (nfd) ethnic group category contains responses that are not specific ethnic group responses but are able to be placed in a broader category in the ethnicity classification. For example, Continental European, African.

Don't Know
The use of this category is necessary when the person is unsure of their ethnic group in an interviewer administered survey or writes this in as a response.

Refused to Answer
This category is only used when it is known that the respondent has purposefully chosen not to respond to the question. Use of this residual category in processing is optional. Its use is most applicable in face-to-face or telephone interviews, but may be used in self-completed questionnaires if the respondent has clearly indicated they refuse or object to answering the question.

Repeated Value
It is used when a respondent has given two responses that have the same code. This may be two written responses, or one tick box response and one written response. For example, someone may tick the NZ European tick box and write in NZ European.

Response Unidentifiable
This category is used when there is a response given, but the response is illegible, or it is unclear what the meaning or intent of the response is – this most commonly occurs when the response being classified contains insufficient detail, is ambiguous or is vague.

Response Outside Scope
This category is used for responses that are positively identified, that is, the meaning and the intent are clear but clearly fall outside the scope of the classification/topic as defined in the standard. For example, a response of vegetarian falls outside the scope of the ethnicity classification.

Not Stated
This category is only used where a respondent has not given any response to the question asked in a self-administered questionnaire, that is, it is solely for non-response.

References

Lang, K (2001). *Review of the Measurement of Ethnicity – Policy Perspectives Paper*, Statistics New Zealand, Wellington.

Smith, A (1986). *The Ethnic Origins of Nations*, Blackwell, Oxford.

Statistics New Zealand (2004). *Report of the Review of the Measurement of Ethnicity*, Statistics New Zealand, Wellington.

Further Information

For internal users, further information on current and historical practice in surveys and output is available through the Survey Information Manager (SIM). This database stores information about the surveys and outputs administered by Statistics New Zealand.

For external users, this information can be accessed through Statistics New Zealand's information service by emailing info@stats.govt.nz.

Appendix 1

Standard Classification of Ethnicity (ETHNIC05)

Level one
1 European
2 Māori
3 Pacific Peoples
4 Asian
5 Middle Eastern/Latin American/African
6 Other Ethnicity
9 Residual Categories

Level two
10 European nfd
11 New Zealand European
12 Other European
21 Māori
30 Pacific Peoples nfd
31 Samoan
32 Cook Islands Maori
33 Tongan
34 Niuean
35 Tokelauan
36 Fijian
37 Other Pacific Peoples
40 Asian nfd
41 Southeast Asian
42 Chinese
43 Indian
44 Other Asian
51 Middle Eastern
52 Latin American
53 African
61 Other Ethnicity
94 Don't Know
95 Refused to Answer
96 Repeated Value
97 Response Unidentifiable
98 Response Outside Scope
99 Not Stated

Level three
100 European nfd
111 New Zealand European
121 British and Irish
122 Dutch
123 Greek
124 Polish
125 South Slav
126 Italian

127 German
128 Australian
129 Other European
211 Māori
300 Pacific Peoples nfd
311 Samoan
321 Cook Islands Maori
331 Tongan
341 Niuean
351 Tokelauan
361 Fijian
371 Other Pacific Peoples
400 Asian nfd
410 Southeast Asian nfd
411 Filipino
412 Cambodian
413 Vietnamese
414 Other Southeast Asian
421 Chinese
431 Indian
441 Sri Lankan
442 Japanese
443 Korean
444 Other Asian
511 Middle Eastern
521 Latin American
531 African
611 Other Ethnicity
944 Don't Know
955 Refused to Answer
966 Repeated Value
977 Response Unidentifiable
988 Response Outside Scope
999 Not Stated

Level four
10000 European nfd
11111 New Zealand European
12100 British nfd
12111 Celtic nfd
12112 Channel Islander
12113 Cornish
12114 English
12115 Gaelic
12116 Irish
12117 Manx
12118 Orkney Islander
12119 Scottish
12120 Shetland Islander
12121 Welsh
12199 British nec
12211 Dutch

12311 Greek
12411 Polish
12500 South Slav nfd
12511 Croatian
12512 Dalmatian
12513 Macedonian
12514 Serbian
12515 Slovenian
12516 Bosnian
12599 South Slav nec
12611 Italian
12711 German
12811 Australian
12911 Albanian
12912 Armenian
12913 Austrian
12914 Belgian
12915 Bulgarian
12916 Belorussian
12917 Corsican
12918 Cypriot nfd
12919 Czech
12920 Danish
12921 Estonian
12922 Finnish
12923 Flemish
12924 French
12925 Greenlander
12926 Hungarian
12927 Icelandic
12928 Latvian
12929 Lithuanian
12930 Maltese
12931 Norwegian
12932 Portuguese
12933 Romanian
12934 Gypsy
12935 Russian
12936 Sardinian
12937 Slavic
12938 Slovak
12939 Spanish
12940 Swedish
12941 Swiss
12942 Ukrainian
12943 American
12944 Burgher
12945 Canadian
12946 Falkland Islander
12947 New Caledonian
12948 South African nec
12949 Afrikaner

12950 Zimbabwean
12999 European nec
21111 Māori
30000 Pacific Peoples nfd
31111 Samoan
32100 Cook Islands Maori nfd
32111 Aitutaki Islander
32112 Atiu Islander
32113 Mangaia Islander
32114 Manihiki Islander
32115 Mauke Islander
32116 Mitiaro Islander
32117 Palmerston Islander
32118 Penrhyn Islander
32119 Pukapuka Islander
32120 Rakahanga Islander
32121 Rarotongan
33111 Tongan
34111 Niuean
35111 Tokelauan
36111 Fijian
37111 Admiralty Islander
37112 Australian Aboriginal
37113 Austral Islander
37114 Palau Islander
37115 Bismark Archipelagoan
37116 Bougainvillean
37117 Caroline Islander
37118 Easter Islander
37119 Gambier Islander
37120 Guadalcanalian
37121 Chamorro
37122 Hawaiian
37123 Kanak
37124 Kiribati
37125 Malaitian
37126 Manus Islander
37127 Marianas Islander
37128 Marquesas Islander
37129 Marshall Islander
37130 Nauruan
37131 New Britain Islander
37132 New Georgian
37133 New Irelander
37134 Banaban
37135 Papua New Guinean
37136 Phoenix Islander
37137 Pitcairn Islander
37138 Rotuman
37139 Santa Cruz Islander
37140 Tahitian
37141 Solomon Islander

37142 Torres Strait Islander
37143 Tuamotu Islander
37144 Tuvaluan
37145 Ni Vanuatu
37146 Wake Islander
37147 Wallis Islander
37148 Yap Islander
37199 Pacific Peoples nec
40000 Asian nfd
41000 Southeast Asian nfd
41111 Filipino
41211 Cambodian
41311 Vietnamese
41411 Burmese
41412 Indonesian
41413 Laotian
41414 Malay
41415 Thai
41499 Southeast Asian nec
42100 Chinese nfd
42111 Hong Kong Chinese
42112 Cambodian Chinese
42113 Malaysian Chinese
42114 Singaporean Chinese
42115 Vietnamese Chinese
42116 Taiwanese
42199 Chinese nec
43100 Indian nfd
43111 Bengali
43112 Fijian Indian
43113 Gujarati
43114 Tamil
43115 Punjabi
43116 Sikh
43117 Anglo Indian
43199 Indian nec
44100 Sri Lankan nfd
44111 Sinhalese
44112 Sri Lankan Tamil
44199 Sri Lankan nec
44211 Japanese
44311 Korean
44411 Afghani
44412 Bangladeshi
44413 Nepalese
44414 Pakistani
44415 Tibetan
44416 Eurasian
44499 Asian nec
51100 Middle Eastern nfd
51111 Algerian
51112 Arab

51113 Assyrian
51114 Egyptian
51115 Iranian/Persian
51116 Iraqi
51117 Israeli/Jewish
51118 Jordanian
51119 Kurd
51120 Lebanese
51121 Libyan
51122 Moroccan
51123 Omani
51124 Palestinian
51125 Syrian
51126 Tunisian
51127 Turkish
51128 Yemeni
51199 Middle Eastern nec
52100 Latin American nfd
52111 Argentinian
52112 Bolivian
52113 Brazilian
52114 Chilean
52115 Colombian
52116 Costa Rican
52117 Latin American Creole
52118 Ecuadorian
52119 Guatemalan
52120 Guyanese
52121 Honduran
52122 Malvinian
52123 Mexican
52124 Nicaraguan
52125 Panamanian
52126 Paraguayan
52127 Peruvian
52128 Puerto Rican
52129 Uruguayan
52130 Venezuelan
52199 Latin American nec
53100 African nfd
53112 United States Creole
53113 Jamaican
53114 Kenyan
53115 Nigerian
53116 African American
53117 Ugandan
53118 West Indian
53119 Somali
53120 Eritrean
53121 Ethiopian
53122 Ghanaian
53199 African nec

61111 Central American Indian
61112 Inuit
61113 North American Indian
61114 South American Indian
61115 Mauritian
61116 Seychellois
61117 South African Coloured
61118 New Zealander
61199 Other Ethnicity nec
94444 Don't Know
95555 Refused to Answer
96666 Repeated Value
97777 Response Unidentifiable
98888 Response Outside Scope
99999 Not Stated

Appendix 2

Reducing multiple ethnic responses - manual methodology

Method for recording six ethnicities responses

If there are more than six responses per individual, then a random method for reducing the number of responses selects the six ethnicities to be retained. This manual method mirrors the software application method in Appendix 3.

To make your selection random, use a random number chart and methodically assign a number by either choosing a column or row to follow. After assigning the random numbers choose the lowest number as the response to be removed. This is explained fully in the method below.

Every level one ethnic group category that is represented by an individual's ethnicities must be represented in the final selection of responses. All level one ethnicity categories will be retained when the number of responses is reduced to six, as there are six categories at level one.

Responses which would be coded to a residual category are removed first. For example, a response of vegetarian would be coded to the residual category 98888 Response Outside Scope and would be the first response removed. If there are still more than six ethnicities then identify the level one categories they belong to.

Retain the responses from the level one categories that have just one ethnicity response belonging to them. All responses with the same first digit belong in the same level one category.

The level one categories that are represented by more than one ethnicity response are selected to reduce the number of responses through a random method. At least one ethnicity response representing each level one category is retained. Randomly select a level one category with more than one response. Do this by assigning each level one category a random number from a chart and select the lowest random number. This is the category from which an ethnicity will be selected to be removed. Next, randomly remove one response from the level one category selected. Do this by assigning each ethnicity in this category a random number and then remove the ethnicity with the lowest random number.

If this reduces the number of ethnicities to six then the random selection procedure can stop as the maximum number of responses has been retained. Otherwise, repeat this process until six responses remain.

The final selection must meet the requirement of no more than six ethnicities for processing and also retain information at level one of the classification for ethnicity.

Example of reducing to six responses

An individual's responses are French, Niuean, Cambodian, Vietnamese, English, Algerian and New Zealander.

There are no residual categories to remove.

Classifying each ethnicity to their respective level one category has the following result:

French and English are classified within the level one European category.
Niuean is classified within the level one Pacific Peoples category.
Cambodian and Vietnamese are classified within the level one Asian category.
Algerian is classified within the level one Middle Eastern/Latin American/African (MELAA) category.
New Zealander is classified within the level one Other Ethnicity category.

For three of the level one categories there is only one response given and these responses must be retained. They are Niuean, Algerian and New Zealander. That leaves two level one categories with more than one response from which to select the remaining responses to be retained.

Assign each level one category a random number. In this example, European is assigned 393 and Asian is assigned 214. The Asian category has the lowest random number and is the category from which an ethnicity will be selected to be removed.

Assign each ethnicity in the Asian category a random number: Cambodian is assigned 149 and Vietnamese is assigned 613. Cambodian has the lowest random number and so this response is removed.

The six ethnicities to be retained from the given responses are:
French and English within the level one European category.
Niuean within the level one Pacific category.
Vietnamese within the level one Asian category.
Algerian within the level one MELAA category.
New Zealander within the level one Other Ethnicity category.

This selection meets the requirement of six ethnicities for processing and retains all the level one ethnicity information given by the individual.

Method for recording three ethnicity responses

If a maximum of three responses is retained for a collection and more than three ethnicity responses are given, a random method is used to select the three responses. Where possible each level one category must be represented from the responses retained. This manual method mirrors the software method. (see Appendix 3 for the software application methodology and a more detailed explanation)

To make your selection random, use a random number chart and methodically assign a number by either choosing a column or row to follow. After assigning the random numbers choose the lowest number as the response to be removed. This is explained fully in the method below.

Responses which would be coded to a residual category are excluded first. For example, a response of vegetarian would be coded to the residual category 98888 Response Outside Scope and would be removed first. If there are still more than three ethnicities then identify which level one categories the ethnicities belong to. All responses with the same first digit belong in the same Level one category.

If a person has given more than three responses, but these are all classified within one, two or three level one categories, assign each ethnicity response a random number and remove the response with the lowest random number. If this reduces the responses to three then the random selection procedure can stop as the maximum number of responses has been retained. Otherwise, repeat the process for the remaining level one categories with more than one response. Avoid removing responses from a level one category when it is represented by only one ethnicity.

Repeat this process until three responses have been selected. In this situation, all level one ethnicity information has been retained.

If a person's responses are classified to more than three level one categories, a random selection is made to remove responses, one at a time, from the level one categories with more than one response. This will result in one response only in each level one category but at this point there will still be more than three responses. Start by assigning each category a random number. Select the category with the lowest random number and then assign each ethnicity in that category a random number. Remove the ethnicity with the lowest random number. Repeat this process until each level one category is represented by one ethnicity.

Although each level one category is represented up to this point, there will still be too many responses to retain. To reduce the number of level one categories to three, assign a random number to the remaining ethnicities and select the lowest random number to remove. Repeat this process until three ethnicity responses remain.

The final selection must meet the requirement of no more than three ethnicities for processing and also retain information at level one of the classification for ethnicity as far as possible.

Example of more than three responses and less than three level one categories

An individual's responses are French, German, Swiss, Dutch, Amhara and Tigrean.

There are no residual categories to remove.

Classifying each ethnicity to their respective level one category has the following result:

French, German, Swiss, and Dutch are classified within the level one European category.
Amhara and Tigrean are classified within the level one Middle Eastern/Latin American/African (MELAA) category.

Assigning each ethnicity response a random number results in: French 641, German 239, Swiss 873, Dutch 458, Amhara 210 and Tigrean 632. The lowest random number is Amhara and that response is removed. The MELAA level one category is represented by the Tigrean response only and so this ethnicity is retained. This leaves four responses from the European level one category. The lowest random number of the European responses is German 239 and this is removed. One more response needs to be removed so selecting the next lowest random number removes Dutch 458.

The three ethnic groups to be retained from the given responses are:
French and Swiss within the level one European category, and
Tigrean within the level one MELAA category.

This selection meets the requirement of three ethnicities for processing and retains all the level one ethnicity information given by the individual.

Example of more than three responses and more than three level one categories

An individual's responses are French, Niuean, Cambodian, Vietnamese, New Zealander and English.

There are no residual categories to remove.

Classifying each ethnicity to their respective level one category has the following result:

French and English are classified within the level one European category.
Niuean is classified within the level one Pacific Peoples category.
Cambodian and Vietnamese are classified within the level one Asian category.
New Zealander is classified within the level one Other Ethnicity category.

As there are more than three level one categories represented, a random selection is made to remove responses, one at a time, from the level one categories with more than one response. This will result in one response only in each level one category.

The European and Asian categories are represented by more than one ethnicity. Assign each ethnicity in the European and Asian categories a random number and select the lowest random number to remove from each category. French is assigned 987 and English is assigned 109 so

English is removed from the European category. Cambodian is assigned 291 and Vietnamese is assigned 312 so Cambodian is removed from the Asian category.

This reduces our total number of responses to four: French, Niuean, Vietnamese and New Zealander and represents each level one category up to this point.

To reduce the number of level one categories to three, assign a random number to the remaining ethnicities and select the lowest random number to remove. French is assigned 367, Niuean 183, Vietnamese 671 and New Zealander 948. Niuean with the lowest random number is removed.

The three ethnic groups to be retained from the given responses are:
French within the level one European category.
Vietnamese within the level one Asian category.
New Zealander within the level one Other Ethnicity category.

This selection meets the requirement of three ethnicities for processing and retains as far as possible, all the level one ethnicity information given by the individual.

Appendix 3

Technical paper for software developers:

When Individual Responses Exceed Input Storage - A Procedure For Unbiased Reduction

This methodology paper is a technical paper for the development of software systems that support the inputting of large numbers of responses in surveys and administrative data sets. It outlines the treatment of responses where the number of ethnic groups given by an individual exceeds the number being output. Scenarios are described for reducing the number of multiple responses to six per individual, and to three responses, and examples are given of each.

As a supporting paper to the Statistical Standard for Ethnicity 2005 it needs to be read in conjunction with the Standard and the paper 'Understanding and Working with Ethnicity Data' (see Appendix 4).

This paper is available on the Statistics New Zealand website:

http://www.stats.govt.nz/analytical-reports/review-measurement-ethnicity/papers.htm

Appendix 4

Understanding and Working with Ethnicity Data

This technical paper provides support to users of ethnicity data. It addresses issues arising from the Review of the Measurement of Ethnicity and advises on best practices. Among the topics covered are: ethnic mobility, contextual effects and the dynamics of ethnicity; sources of data; collection issues including consistency, and reduction of multiple responses; comparing collections; output choices, time series analysis, and data integration.

This paper is available on the Statistics New Zealand website:

http://www.stats.govt.nz/analytical-reports/review-measurement-ethnicity/papers.htm