Australian Bureau of Statistics

New Issue

Standards for Statistics on Cultural and Language Diversity

1999

W. McLennan
Australian Statistician

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11:30 AM (CANBERRA TIME) MON 22 NOV 1999

ABS Catalogue No. 1289.0
ISBN 0 642 25657 8

Produced by the Australian Bureau of Statistics

INQUIRIES

- For more information on the implementation of the standard variables, or any aspect of the collection and use of cultural and language diversity data, please contact the ABS on Canberra 02 6252 5736 or by email: social.classifications@abs.gov.au. Feedback or comments on this publication are welcome.

CONTENTS

PREFACE

The Australian Bureau of Statistics (ABS) has developed *Standards for Statistics on Cultural and Language Diversity* in response to a widely recognised need for a nationally consistent framework for the collection and dissemination of data on cultural and language diversity.

The publication presents a set of statistical standards which are designed to collect all the cultural and language information considered necessary for consistent and accurate measurement of cultural diversity in Australia. The standards are intended as a replacement for non-English speaking background (NESB), which has previously been used as a broad measure of culturally related need or disadvantage.

Standards for Statistics on Cultural and Language Diversity is a reference document which defines the standards and outlines methods for their use in statistical, administrative and service provision settings. The standards, which were endorsed by the Council of Ministers of Immigration and Multicultural Affairs (COMIMA) in April 1999, include recommended questions, classifications, coding structures and output categories for use in interview-based and self-enumerated data collections.

It is intended that *Standards for Statistics on Cultural and Language Diversity* be used by government, academic and private sector organisations in all relevant data collection activities, as this will improve the compatibility and comparability of data derived from different sources.

The assistance of the many organisations and individuals who provided information and advice during the development of this set of standards is gratefully acknowledged. The ABS would welcome feedback on any aspect of these standards.

W. McLennan
Australian Statistician

LIST OF ABBREVIATIONS

ABS	Australian Bureau of Statistics
APS	Australian Public Service
ARA	any responsible adult
ASCCEG	Australian Standard Classification of Cultural and Ethnic Groups
ASCL	Australian Standard Classification of Languages
ASCRG	Australian Standard Classification of Religious Groups
ASEAN	Association of South East Asian Nations
ATSIC	Aboriginal and Torres Strait Islander Commission
CLIP	Cultural and Language Indicators Pilot Study
COMIMA	Council of Ministers of Immigration and Multicultural Affairs
DIMA	Department of Immigration and Multicultural Affairs
nec	not elsewhere classified
NESB	non-English speaking background
PI	personal interview
SACC	Standard Australian Classification of Countries
SCIMA	Standing Committee of Immigration and Multicultural Affairs
SSCLD	Standards for Statistics on Cultural and Language Diversity

CHAPTER 1 INTRODUCTION

Standards for Statistics on Cultural and Language Diversity (SSCLD) has been developed by the Australian Bureau of Statistics (ABS) to provide a means of standardising the way the ABS and other agencies collect and disseminate information relating to the origins of individuals and cultural diversity in Australia. This is done by providing statistical standards to measure those attributes of persons that relate to their cultural and language background, and a means of using these standard measures in a consistent, harmonious and useful manner. (Discussion of what constitutes a statistical standard and the advantages of using standards is provided in Chapter 3.)

The statistical variables to which these standards relate are those that are considered relevant as cultural and language indicators. This list of variables has been developed by the ABS in consultation with other organisations in response to growing user needs and a request from government. (A definition of the term 'variable' is provided in the Glossary.)

These standards are presented in this publication in summary form, focussing on the essential components of the standards required for use outside the ABS and on their use as a set to provide information on particular aspects of cultural diversity in Australian society. A full specification of these standards can be found in the *Statistical Concepts Library* (Cat. no. 1361.0.30.001) on the ABS Website, and to be released on CD-ROM in early 2000.

A number of the standards make reference to major Australian standard classifications related to language and cultural diversity. The use of these classifications is fundamental to the proper application of the standard variables. The standard variables provide summary information about the classifications concerned. These classifications are fully specified in the following ABS publications.

- *Australian Standard Classification of Languages (ASCL)* (Cat. no. 1267.0).
- *Australian Standard Classification of Cultural and Ethnic Groups (ASCCEG)* (Cat. no. 1249.0) (forthcoming in mid 2000).
- *Australian Standard Classification of Religious Groups (ASCRG)* (Cat. no. 1266.0).
- *Standard Australian Classification of Countries (SACC)* (Cat. no. 1269.0).

Full details of the published classifications are also available on the ABS Website.

PURPOSE OF THE PUBLICATION

This publication has four main purposes:

- to provide standards to identify, define, classify and disseminate particular attributes of a person or group of people that relate to their origins and cultural and language background. Variables relating to the cultural and language attributes which are considered important and relevant, and the reasons for choosing them, are described in Chapter 2;
- to provide a way to identify, measure and monitor service needs associated with advantage or disadvantage related to cultural and language background. For instance, to provide cultural and language data which facilitate access and equity initiatives;
- to provide information that supports a measure of cultural and language diversity in its broader sense. That is, a measure of the cultural and language communities and groups that make up Australian society; and
- to provide information to replace non-English speaking background (NESB) which is widely considered to be no longer appropriate as a general purpose indicator.

There are two ways in which the cultural and language variables included in SSCLD can be used:

- an individual variable can be used to collect a particular item of information in a statistical or administrative collection to meet a particular need. This method, while satisfying a particular information need, may provide only a superficial measure of cultural and language diversity; and
- a set of variables can be used to collect a range of cultural and language information. This approach can provide a broad and relatively balanced method of measuring the cultural diversity of a population or the cultural and ethnic attributes of an individual.

USE OF INDIVIDUAL CULTURAL AND LANGUAGE INDICATORS

Most ABS statistical collections and the collections of many other agencies do not have the measurement of cultural and language diversity as a primary focus. Many of these collections do, however, need to collect information on one or two of the variables included as cultural and language indicators in this publication. This usage is entirely acceptable and if the statistical standards presented in this publication are used, will provide good quality information which is comparable with many other data collections.

USE OF A STANDARD SET OF CULTURAL AND LANGUAGE INDICATORS

It is widely agreed that there are many elements to cultural and language diversity which must be considered to provide an accurate measure of cultural and language background or diversity. To use a single standard variable, such as country of birth, or a non-standard composite concept, such as NESB, is inadequate.

USE OF A STANDARD SET OF CULTURAL AND LANGUAGE INDICATORS *continued*

When there is a need to identify the cultural and language background of a person or group of people, or to measure the cultural diversity of a population, it is appropriate to use a standard set of cultural and language indicators as described in this publication. Since the Standard Set of Cultural and Language Indicators has been developed by the ABS in consultation with many other organisations, and in response to growing user needs and a clearly articulated request from government, it has wide-ranging acceptance in the user community and is being actively promoted by the ABS and other organisations. The history of the development process is described in the following section.

DEVELOPMENT OF THE STANDARD SET OF INDICATORS

There has been an official shift away from using NESB as a measure of cultural diversity and related need or disadvantage. NESB is no longer considered to be an appropriate measure of culturally related disadvantage, in terms of access to government services, for a variety of reasons:

- the term has many conflicting definitions;
- it groups people who are relatively disadvantaged with those who are not disadvantaged;
- it is unable to separately identify the many cultural and linguistic groups in Australian society; and
- it has developed negative connotations.

In summary, NESB is seen as an oversimplified indicator of disadvantage which may result in inappropriate service provision and neglect the positive aspects of cultural and linguistic diversity. Consequently, government agencies at all levels have increasingly sought to develop a more accurate, effective and consistent measure of cultural and language diversity in order to improve strategic planning and evaluation of service programs.

At a meeting of the Council of Ministers of Immigration and Multicultural Affairs (COMIMA) in May 1996, Commonwealth and State Ministers noted the problems associated with the use of NESB and agreed that the term and its acronym be dropped, where possible, from official communications. The ability of all government agencies to capture a common core set of cultural indicator data which will allow a more precise and meaningful assessment of service uptake by different cultural groups across a number of different portfolio services, as well as a comparative assessment across agencies, was seen as an area requiring urgent attention. To progress the need for the development and implementation of standardised cultural indicator data across all levels of government, the ABS was engaged to cost, develop and pilot a data collection instrument and trial it in a number of government agencies.

A working group comprising the Department of Immigration and Multicultural Affairs (DIMA), the Multicultural Affairs Unit of the Department of Premier and Cabinet (Victoria) and the ABS was established and a draft proposal was developed. In November 1997 the

DEVELOPMENT OF THE STANDARD SET OF INDICATORS *continued*

draft proposal outlining the objectives, methodology and costs associated with the pilot study was endorsed by the Standing Committee of Immigration and Multicultural Affairs (SCIMA), with the requirement that the project team report back in November 1998.

The overriding aims of the study, known as the Cultural and Language Indicators Pilot Study (CLIP), were to determine the type of data necessary to replace NESB, and to provide a way for cultural and language diversity to be more effectively built into the strategic planning and reporting processes associated with program or service delivery. The objectives were to propose a standard set of variables to measure cultural and language diversity which could be used in all administrative and service provision settings, and to determine a Minimum Core Set of variables from the full standard set that would effectively replace NESB.

It should be stressed that it was not the intention of the exercise to recommend a single key measure of cultural and language diversity to replace NESB, nor to propose an alternative acronym. It was apparent that precise measurement of cultural and language diversity, and related advantage or disadvantage, required a combination of variables which produced a range of data about a person's background.

The study consisted of two main streams of activity. The first was to design and pilot a data collection instrument, which would test a number of indicator variables relating to cultural and language background on surveys and administrative forms in various settings (e.g. hospital admission forms, DIMA Offshore and Onshore processing offices). The second was to undertake analysis of ABS census and survey data, and to conduct supplementary research to assess the performance and suitability of the proposed indicator variables.

OUTCOMES OF THE CULTURAL AND LANGUAGE INDICATORS PILOT STUDY

This data collection test involved 5,016 clients who used the services provided by a range of local government, State and Territory government and Commonwealth government agencies during the period March to September 1998. This sample size allowed for a broad representation of the community and for a detailed analysis of the results in terms of response rates to individual questions, quality of data collected and the degree of respondent burden. Feedback from the participating agencies on respondent reaction to questions, and respondent concern about sensitivity or privacy of the information sought, was also analysed. Evaluation of the performance and suitability of the variables tested found that only one, Visa Category, would be difficult to implement in particular administrative settings. The remaining variables were found to be suitable indicators of cultural and language diversity, and could be successfully implemented in a range of statistical and administrative collections.

Existing ABS census and survey data were analysed to determine whether there was a relationship between individual cultural and language indicator variables used in ABS collections (and trialed in the CLIP data collection test), and other recognised measures of socioeconomic

OUTCOMES OF THE CULTURAL AND LANGUAGE INDICATORS PILOT STUDY *continued*

disadvantage, such as unemployment rates. The analysis indicated that population groups from certain countries of birth, people who spoke particular languages and those with certain religions had a relatively high correlation with indicators of socioeconomic disadvantage. However, not all people, or even a majority of people, within these groups necessarily exhibited the characteristics of disadvantage. Therefore, attempts to use a single cultural or language variable as a generalised indicator of advantage or disadvantage suffered from the same limitations associated with the use of NESB. This analysis confirmed the need for a range of cultural and language diversity indicators.

The CLIP outcomes, supplementary research and subsequent SCIMA discussions were all used to refine and finalise the Minimum Core Set of Cultural and Language Indicators selected out of the full Standard Set of Cultural and Language Indicators. These indicators are designed to replace NESB and to collect a wide range of cultural and language data.

Because the project was directed primarily towards developing cultural and language indicators to replace NESB, Indigenous Status was not included in the CLIP testing program. However, it was acknowledged that Indigenous Status is a fundamental element of cultural diversity in Australian society and that the existing ABS standard for Indigenous Status, which was at the time the subject of a number of other initiatives, should be included in all data collections where the focus is not restricted to migrants and their descendants.

The Minimum Core Set consists of four variables: Country of Birth of Person, Main Language Other Than English Spoken at Home, Proficiency in Spoken English, and Indigenous Status. Indigenous Status forms part of the core set for those collections which are not specifically focussed on migrants to Australia.

The full Standard Set also includes Ancestry, Country of Birth of Father, Country of Birth of Mother, First Language Spoken, Languages Spoken at Home, Main Language Spoken at Home, Religious Affiliation, and Year of Arrival in Australia. Languages Spoken at Home was not included in the CLIP testing program and, therefore, was not endorsed as one of the Standard Set of indicators by COMIMA. It has been included as one of the full Standard Set in this publication because of its value in providing data on the stock of languages actively used in Australian homes. Any of these variables can be added to the Minimum Core Set variables to collect other relevant cultural and language diversity data to meet particular information needs.

The reasons for these variables being chosen are discussed in Chapter 2. Details of the definition, standard questions, standard classifications, and coding and output procedures for each variable can be found in Chapter 3. Examples of question modules, with correct question ordering and sequencing, for the variables can be found in Chapter 4.

ENDORSEMENT OF THE STANDARD SET OF INDICATORS

The Minimum Core Set and the Standard Set of Cultural and Language Indicators were endorsed by COMIMA in April 1999. COMIMA felt that this range of variables was able to comprehensively measure different aspects of a person's origins and the extent to which persons from certain cultural and language backgrounds are associated with advantage or disadvantage. The variables were also found to support effective planning, evaluation and monitoring of service programs in all administrative and service provision settings.

COMIMA recommended that the Minimum Core Set of variables be implemented in all national and state and territory statistical and administrative collections which require information on cultural and language diversity. COMIMA further recommended that additional variables from the full Standard Set be added to the collection where a wider range of information is required.

NEED FOR A STANDARD SET OF INDICATORS

Government and private organisations currently collect a wide range of cultural diversity data, with a focus on NESB, using many different data collection methodologies (e.g. self-enumerated administrative forms, personal interviews, etc.) and different measures of a person's cultural background and language use. Because standard or consistent measures of cultural and language data have not been frequently used, it is difficult to compare data from different sources in any meaningful way.

The introduction of the Standard Set of Cultural and Language Indicators, with standard questions and data collection procedures, will provide a number of significant benefits, including:

- providing a consistent method for measuring cultural and language diversity in all statistical and administrative collections;
- allowing data from different sources, and different time periods, to be compared and integrated in a meaningful way;
- improving the quality, relevance and accuracy of data produced; and
- reducing development and operational costs for agencies collecting data on cultural and language diversity by providing a ready-made and reliable method for use in all service provision settings.

CHAPTER 2 USING THE STANDARD SET OF CULTURAL AND LANGUAGE INDICATORS

As noted in Chapter 1, the Council of Ministers of Immigration and Multicultural Affairs (COMIMA) has recommended that the measurement of cultural and language diversity be based on the use of the Standard Set of Cultural and Language Indicators in statistical and administrative collections across all states and territories. It has further recommended that any measures currently based on the notion of non-English speaking background (NESB) be replaced by the new method. A Minimum Core Set of the recommended indicators is considered necessary to collect the minimum amount of information needed to replace the key measure of NESB.

Organisations will need to review and adjust their measurement tools and data processing procedures to fully implement the proposed cultural and language indicators in their collections. This could involve:

- the deletion or addition of questions;
- changes to question wording and sequencing;
- changes to definitions;
- changes to the classification of responses; and
- changes to manual and computer coding systems used to capture and manipulate data.

The cultural and language indicators consist of a suite of statistical variables for which the Australian Bureau of Statistics (ABS) has developed statistical standards. The use of statistical standards is considered essential to provide the basis of comparability of information collected within and between agencies, especially comparability with data produced by the ABS. Therefore, to implement the recommendations of COMIMA, organisations will need to accept and use ABS statistical standards for the Minimum Core Set and Standard Set of Cultural and Language Indicators.

While implementation of the ABS standards may involve some initial costs and inconvenience, consistent use of the standards will ultimately result in developmental and operational savings as well as improving the quality of information collected.

THE MINIMUM CORE SET OF CULTURAL AND LANGUAGE INDICATORS

The Minimum Core Set of Cultural and Language Indicators consists of the following four variables:

Country of Birth of Person
Main Language Other Than English Spoken at Home
Proficiency in Spoken English
Indigenous Status

THE MINIMUM CORE SET OF CULTURAL AND LANGUAGE INDICATORS *continued*

It is recommended that the Minimum Core Set of variables (see Chapter 3 for details of the statistical standard for each variable) be collected in all administrative and service provision settings where information on cultural and language diversity is required. However, where the focus is on migrant issues only, it may not be appropriate or useful to include Indigenous Status. In instances where data collection activities are focussed specifically on Aboriginal and Torres Strait Islander peoples, Country of Birth of Person could be omitted as the population of interest would almost all be born in Australia. Chapter 4 provides an example of the Minimum Core Set question module which illustrates appropriate question ordering and sequencing.

Where particular collections currently do not include one of the four Minimum Core Set variables, but include other cultural diversity variables which are better suited to the collection, it is not intended that a core variable necessarily replace a variable currently collected. Instead, it is recommended that the other Minimum Core Set variable be added to the collection. For example, if an organisation collects Country of Birth of Person and Year of Arrival in Australia data because they are useful when used together, it is not proposed that the organisation replace Year of Arrival in Australia with Main Language Other Than English Spoken at Home and Proficiency in Spoken English. Rather, the organisation is encouraged to collect Main Language Other Than English Spoken at Home and Proficiency in Spoken English *as well*.

THE STANDARD SET OF CULTURAL AND LANGUAGE INDICATORS

The Standard Set of Cultural and Language Indicators is as follows:

Country of Birth of Person
Main Language Other Than English Spoken at Home
Proficiency in Spoken English
Indigenous Status
Ancestry
Country of Birth of Father
Country of Birth of Mother
First Language Spoken
Languages Spoken at Home
Main Language Spoken at Home
Religious Affiliation
Year of Arrival in Australia

Any of the non-core variables in the Standard Set (see Chapter 3 for details of the statistical standard for each variable) can be added to the Minimum Core Set to create question modules which enable other relevant data to be collected to meet particular information requirements. The additional indicators can be added, either individually or in combination, to the core set. A general principle might be to ask as wide a range of questions as possible to provide a comprehensive picture of an individual's origins and characteristics.

THE STANDARD SET OF CULTURAL AND LANGUAGE INDICATORS *continued*

The questions outside the core set have not been assigned a hierarchy of priority. Individual requirements will determine which additional questions are needed and the order in which non-core questions are asked. However, some questions should only be asked of certain populations. For example, Year of Arrival in Australia should only be asked of those people born overseas. Chapter 4 contains examples of question modules showing the correct sequencing of questions when variables are added to the Minimum Core Set.

The Standard Set contains several different language variables which each measure a different concept associated with language usage. The language variable, Main Language Other Than English Spoken at Home is included as part of the Minimum Core Set because it was identified, following extensive consultation with users of language data, as the most useful general purpose language variable. It is the language variable used in the ABS Census of Population and Housing and its use therefore enables administrative data to be directly compared or integrated with census data. This would be more difficult to achieve if different measures of language were used in the two contexts.

Organisations should use other language variables in addition to Main Language Other Than English Spoken at Home if additional language data are required. The choice of additional language variables, the order in which they are asked and which language variable is used as a filter for Proficiency in Spoken English will depend on particular information requirements. The relative strengths and weaknesses of the four language variables included in the Standard Set are discussed in the next section.

It should be noted that in statistical collections without a cultural focus, individual variables from the Standard Set can be used without having to use the full Standard Set or Minimum Core Set of variables. In such circumstances it is still necessary to use the standards developed for these variables as detailed in Chapter 3.

CHOICE OF VARIABLES FOR THE STANDARD SET OF INDICATORS

The variables included in the Minimum Core Set and the Standard Set of Cultural and Language Indicators were chosen by the ABS and other agencies involved in the project because they provide a range of information that is pertinent to the measurement of cultural and language diversity, and of related advantage or disadvantage in terms of access to government and other services. They are variables that are being increasingly used in the statistical and administrative collections of the ABS and other organisations. The fact that ABS standards, with standard definitions, question wording and data collection procedures, already existed for these variables, or were in the process of being developed, supported the choice of variables.

The following outlines the main reasons why each variable was chosen for the Standard Set. It also outlines the relative strengths and weaknesses of each language variable, the quality of data collected and the suitability of the language variables as a filter for the Proficiency in Spoken English variable.

Country of Birth of Person

This variable provides fundamental and objective information about a person's origins. It is widely regarded by many organisations as a priority measure of cultural background, and forms a key element of their current data collection practice. The variable readily enables comparison with existing ABS census and survey data, and with overseas data. When used in conjunction with other cultural and language variables, Country of Birth of Person allows for the identification of subgroups within a migrant population. It has limitations in identifying ethnic and cultural groups which form minorities in their country or countries of origin and groups which have significant populations in countries outside their country of origin. See page 20 of Chapter 3 for more details.

Main Language Other Than English Spoken at Home

This language variable provides information on the number of people who speak English only and, if one or more other languages are spoken, the main non-English language used in the home. The variable has the merit of capturing a language other than English, where the main language spoken may be English but a language other than English is still used in the home. This maximises numbers for the more established migrant communities. In some cases, however, this measure may not reflect complete language use, for example, when English is the only language spoken in the home but a language other than English is spoken outside the home, within a person's ethnic or community group. This measure may also record the language usage of those people whose main and preferred language is English but who have learnt another language, which is occasionally but not normally spoken at home, in the 'Other Than English' category.

A weakness of Main Language Other Than English Spoken at Home is that it can only capture language use in the home, and may indicate the use of languages other than English for those people whose main language is usually English and whose use of another language is marginal. As such, this variable filters some people whose main language spoken is English to the Proficiency in Spoken English variable, which ideally should only be asked of those for whom English is a second language. It may also exclude some people who speak a language other than English outside the home but who speak only English at home, although not proficiently, from being asked the Proficiency in Spoken English question. However, the extent of these 'filtering' problems is not known and many users regard the combination of this language variable and Proficiency in Spoken English as the best measure for identifying service needs and the potentially disadvantaged. This variable is also used as a filter for Proficiency in Spoken English by the ABS in the Census of Population and Housing.

Main Language Other Than English Spoken at Home was chosen for the Minimum Core Set not only because of its strengths as a measure of language usage, but also because it is the language variable used in the Census. A major advantage of including this language variable in the Minimum Core Set is that it allows for comparability of language data collected by statistical and administrative collections, with ABS census

Main Language Other Than English Spoken at Home *continued*

data. This allows, for example, rates of usage of services by particular language groups in particular regions to be calculated on the basis of the size of the group in the catchment area as revealed by census data. See page 26 of Chapter 3 for more details about this variable.

Proficiency in Spoken English

This variable provides a broad measure of the English proficiency of non-native English speakers. Proficiency in Spoken English is seen as a key variable by many organisations because it identifies those people most likely to suffer disadvantage, in terms of their ability to access services, due to a lack of competence in spoken English. Analysis of 1996 Census data has also identified a clear association between this variable and socioeconomic disadvantage. As it is designed to assess the ability of people who speak languages other than English, Proficiency in Spoken English must be asked after a language variable which filters out people whose only language is English. See page 32 of Chapter 3 for more details.

Indigenous Status

Indigenous Status provides data on the number of people who identify as being of Aboriginal or Torres Strait Islander origin. Indigenous Status is a fundamental element of cultural diversity in Australian society and should be included in all relevant data collections except for those specifically focussed on migrants and their descendants. See page 35 of Chapter 3 for more details.

Ancestry

This variable provides a self-assessed measure of ethnicity and cultural background by identifying a person's origins and heritage. It can also be used in combination with other variables as a measure of the extent to which people retain the ethnicity and culture of their forebears (e.g. parents and grandparents). However, there are many Australians with origins and heritage which do not, in practice, relate to their current ethnic identity. As such, Ancestry is not considered to be a particularly good measure of service needs and should be used in conjunction with the Country of Birth variables and language variables to provide additional information about a person's cultural identity. It may be of most value for some analytical purposes when the population of interest is restricted to persons born overseas, or who have one or more parents born overseas.

One advantage of including the Ancestry variable in the Standard Set is that it will allow comparisons of Ancestry data from administrative sources with ABS census data, as the ABS plans to include this variable in the 2001 Census. See page 39 of Chapter 3 for more details.

Country of Birth of Father

Country of Birth of Father identifies the country in which a person's father was born. It is regarded as an important variable as it can be used, in association with other cultural and language variables, to determine the extent to which second generation Australians retain their parents' culture, ethnicity or language. See page 44 of Chapter 3 for more details.

Country of Birth of Mother

Country of Birth of Mother identifies the country in which a person's mother was born. It is regarded as an important variable for the same reasons as Country of Birth of Father, and the two variables are generally used together. See page 49 of Chapter 3 for more details.

First Language Spoken

This variable provides accurate information about a person's cultural and linguistic background, as First Language Spoken does not change over a person's lifetime, and is regarded as a good surrogate measure of ethnicity because of its connection with a person's origins and the origins of his or her parents. This variable also provides a good measure of current language use in the community. ABS data show that 95% of Australians whose first language is a language other than English, are still able to use their first language.

In some instances however, depending on age, year of arrival in Australia, and living arrangements, a person's first language spoken will not necessarily be the person's language of greatest competence or the main language he or she currently speaks at home or in the community. As such, like Main Language Other Than English Spoken at Home, this variable may overstate the real level of usage of languages other than English. See page 54 in Chapter 3 for more details about this variable.

Languages Spoken at Home

This language variable provides data on the stock of languages actively used in Australian homes. In some cases, however, this measure may not reflect complete language use when, for example, only one language is spoken in the home but other languages are spoken outside the home, within a person's ethnic community group. This variable puts no restrictions on the number of languages a person can report as being spoken in the home. However, multiple language responses may include languages which play a minor role in a person's communication because they are not the person's first language, the language mainly used or the language of greatest competence. The variable does not determine the frequency with which each language, reported as being spoken in the home, is used.

As it is possible to have multiple responses to this variable which include a mix of English and non-English languages, it does not provide a reliable filter for Proficiency in Spoken English and should not be used as such. See page 60 of Chapter 3 for more details about this variable.

Main Language Spoken at Home

This variable provides information about the language most frequently used by a person at home. It is a good indicator of the language in which an individual is likely to be most at ease. However, Main Language Spoken at Home tends to understate current community language usage, of languages other than English, amongst the longer standing migrant groups who now mainly use English at home. In some instances, it does not provide information about a person's cultural and language background but rather information about an aspect of their living arrangements (i.e. the single language most frequently used in the household in which they live).

Main Language Spoken at Home *continued*

Main Language Spoken at Home can be used as a filter to Proficiency in Spoken English. However, it will sequence people who mainly speak a non-English language at home but who are proficient in English, which is their main language outside the home, to the Proficiency in Spoken English question. It may also sequence past the Proficiency in Spoken English question some individuals who mainly use English at home but are not fully proficient in English. First Language Spoken and Main Language Other Than English Spoken at Home are better filters to Proficiency in Spoken English. See page 64 of Chapter 3 for more details about this variable.

Religious Affiliation

As well as providing data on the number of people who identify with particular religious groups in the Australian community, this variable provides additional data for identifying specific ethnic or cultural groups, when used in conjunction with other cultural and language variables. Some organisations have found data on religious affiliation helpful in delivering culturally relevant services to clients. See page 68 of Chapter 3 for more details.

Year of Arrival in Australia

This variable is used to derive the length of time a person born in another country has spent living in Australia. It is an important variable for many purposes as it gives an indication of how familiar migrants are likely to be with Australian society and practices, how long it took them to overcome settlement difficulties, and how their social characteristics have changed with the length of time they have been here. Year of Arrival in Australia is also related to familiarity with the domestic labour market and may be a major determinant of the economic situation of migrants. See page 73 of Chapter 3 for more details.

COLLECTING CULTURAL AND LANGUAGE DATA

When collecting data on cultural and language diversity, the standards provided in Chapter 3 should be used and the ordering of questions and sequence guides provided in Chapter 4 should be followed. This will ensure that compatible and comparable data are collected across statistical and administrative collections, and over time, for all government and private sector collections. Responses to the cultural and language questions should be classified using standard ABS classifications and associated coding procedures. These classifications are well researched and soundly developed and their use enables the ready comparison of data from different sources. The use of ABS coding indexes designed to complement the classifications will simplify the coding process and improve data accuracy.

It is recommended that data be captured and stored at the most detailed level of the classification wherever possible. This allows the greatest flexibility for the output of statistics, enables more detailed and complex analysis, facilitates comparisons with previous data using different classifications and preserves information so as to provide maximum flexibility for future use of the data.

COLLECTING CULTURAL AND LANGUAGE DATA *continued*

The Minimum Core Set and Standard Set of questions are designed to provide information about the origins and cultural characteristics of an individual or group. They therefore do not include questions about a person's labour force characteristics, educational qualifications, income, etc. If socioeconomic data are needed a number of such questions should be asked, and a range of ABS standards is available to assist where such questions are required. It should be noted that making assumptions about the socioeconomic status of a person or a population on the basis of their origins is not recommended.

USING CULTURAL AND LANGUAGE DATA

It is not the role of the ABS to anticipate how a particular organisation should use data relating to each variable or use the data taken together. This will depend on the data requirements necessary to support the policy and operational aims and objectives of each organisation. However, the following observations are offered to illustrate how data on cultural diversity can be used.

There are two basic perspectives of data which affect the range and uses of data collected about cultural and language diversity.

- Data about individuals.
- Aggregated data about community groups.

Data about Individuals

Data about individuals are used to determine the facilities and services required by an individual (or the facilities and services an individual does not need). One of the aims of the use of the Standard Set of Cultural and Language Indicators is to enable agencies to make decisions about a person's needs on the basis of direct and accurate information about their background, language and English skills without making unfounded assumptions about individuals on the basis of the general characteristics of the community group to which they belong. This was a drawback of NESB which the methods detailed in this publication hope to overcome.

Assumptions about a person's, say, language skills made on the basis of their country of birth are not recommended. For instance, when analysing the characteristics of a population, relatively high levels of correlation between a level of Proficiency in Spoken English and country of birth can be achieved when Year of Arrival in Australia is also considered. When applied to an individual, however, this method can produce inaccurate conclusions.

Generally, it can be said that when replacing NESB as a measure of disadvantage, Proficiency in Spoken English provides a similar, but more precise and meaningful, measure. The Cultural and Language Indicators Pilot Study (CLIP) (see Chapter 1) supported the notion that this variable provides information which is fundamental to target the provision of services to people whose lack of ability in spoken English is potentially a barrier to gaining access to government programs and participating in Australian society on an equal footing with those who are proficient in English.

Data About Individuals *continued*

Use of a language variable not only provides a filter to Proficiency in Spoken English, but also accurately indicates the language characteristics and origins of an individual. Using Country of Birth of Person with a language variable provides a cultural dimension which is likely to indicate a person's familiarity with Australian institutions, labour market, etc. Adding Country of Birth of Mother and/or Country of Birth of Father to this mix of variables will further refine this measure of cultural and language background and provide additional information on a person's potential for accessing services.

Country of Birth of Person is best used in conjunction with Year of Arrival in Australia as this will provide an indication of the extent to which migrants are likely to have adapted to Australian society. It can also be used to determine which community groups have the most difficulty, or take the longest time, to adapt to Australian society.

Aggregated Data

It is necessary to aggregate data collected from individuals to create community group profiles for the purposes of policy setting, service monitoring, analysis and thematic reporting. In such circumstances, the range of data collected should, where possible, be suitable for the aggregation and analysis undertaken. The use of Country of Birth of Person (and perhaps Country of Birth of Mother and/or Country of Birth of Father), a language variable and Religious Affiliation will accurately identify most cultural and ethnic groups.

Because the data from many administrative sources (such as income and occupation data) is limited, extrapolation of group characteristics using correlations developed from, for example, cross-classified census data is often necessary when studying the social and economic functioning of community groups. Although more meaningful data will be obtained if direct questions are asked of the target population, data derived for community groups can be used because such data do not draw direct conclusions about a particular person's characteristics.

Using relationships between language communities and aspects of functioning such as Proficiency in Spoken English, labour force status, educational qualifications, etc. (established using census data), it is possible to draw conclusions about the service needs of particular language communities. This provides useful planning information when these communities are encountered in administrative and service provision settings. Similar relationships can be established between Country of Birth of Person and aspects of functioning for the targeting of service provision to communities originating in particular countries.

Cultural and language indicator data also provide a useful tool for determining if service providers are successfully targeting the client groups in their catchment area. For instance, if census data indicate a certain number of speakers of a language in a particular region, and the service take-up rate for speakers of that language is proportionally low, it

Aggregated Data *continued*

may be deduced that targeting strategies are not working. The usefulness of comparing census data with data from other sources is a major reason that the census language variable Main Language Other Than English Spoken At Home is included in the Minimum Core Set.

A SINGLE MEASURE

As stated in Chapter 1, the intention of the proposed new method of measuring cultural and language diversity, and its associated advantage or disadvantage, is not to replace NESB with another single measure that attempts to synthesise a number of attributes of the background or functioning of a person or group of persons. Such blanket measures tend to inaccurately assign the characteristics of a population to individuals, and may tend to create an inappropriately negative (or positive) view of all people falling within them. Rather, it is intended that the attributes of each person are determined by direct and accurate measurement.

TERMINOLOGY

It is recommended that use of the term 'Non-English Speaking Background' or acronym NESB be discontinued as a label for data however it is measured. If people are classified into English speakers and non-English speakers on the basis of, for example, the language variable First Language Spoken, they should be described as 'First Language Spoken English' and 'First Language Spoken Other Than English'. Generally, the standard names of the variables and the standard names of classification categories should be used when disseminating data. For instance, the Proficiency in Spoken English of people should be described as 'Speaks English Very Well', 'Speaks English Well', etc.

CHAPTER 3 STATISTICAL STANDARDS FOR THE CULTURAL AND LANGUAGE INDICATORS

This chapter presents statistical standards for each of the variables that are recommended for use as cultural and language indicators. The Standard Set of Cultural and Language Indicators is as follows:

Country of Birth of Person
Main Language Other Than English Spoken at Home
Proficiency in Spoken English
Indigenous Status
Ancestry
Country of Birth of Father
Country of Birth of Mother
First Language Spoken
Languages Spoken at Home
Main Language Spoken at Home
Religious Affiliation
Year of Arrival in Australia

As noted in Chapter 2, the variables included in the Standard Set of Cultural and Language Indicators were chosen by the Australian Bureau of Statistics (ABS) and other agencies involved in the project because they provide a range of information that is pertinent to the measurement of cultural and language diversity, and of related advantage or disadvantage in terms of access to government services. They are variables that are being increasingly used in the statistical and administrative collections of the ABS and other organisations. The fact that ABS standards, with standard definitions, question wording and data collection procedures, already existed for these variables, or were in the process of being developed, supported the choice of variables. Detailed reasons for choosing these variables are provided in Chapter 2.

DEFINITION OF A STATISTICAL STANDARD

Many different standards are used in statistics, such as computing, editing, publication, form design, and data quality standards. This publication is concerned with standards which facilitate the collection, aggregation and output of cultural and language data. A statistical standard in this context can be defined as a set of components which, when used together, produce consistent and high quality statistical output (about the concepts which underpin the cultural and language variables) across collections and over time. Statistical standards are the approved rules we apply in the development, collection, processing and dissemination of official statistics.

DEFINITION OF A STATISTICAL STANDARD *continued*

A statistical standard includes many components which specify standard practice at any point in the cycle of data collection, processing and dissemination. The essential set of components that comprise and specify a statistical standard for a variable are as follows:

- Standard Name of the Variable;
- Standard Definition of the Variable;
- Standard Question(s);
- Standard Classification;
- Standard Coding Procedures; and
- Standard Output Categories.

This publication provides summary information about each of these essential components. A full specification of each standard can be found in the *Statistical Concepts Library* (Cat. no. 1361.0.30.001) on the ABS website.

ADVANTAGES OF USING STATISTICAL STANDARDS

Statistical standards are designed to add quality to the data produced by improving accuracy, reliability, relevance and timeliness. ABS standards are finalised following thorough and rigorous development of concepts, definitions, questions, classifications, and processing and dissemination procedures. This rigorous development process is combined with consideration of practical issues such as statistical feasibility and extensive consultation with users of the standard and of the resulting data. This thorough development work leads to long term cost savings by providing an 'off the shelf solution' for most data collections.

Standards improve data comparability and comprehensibility. Because ABS standards are designed to harmonise, as far as is possible, with established Australian and international practices, this comparability may apply internationally, across collections, across time, across agencies, and within a given subject area. Comprehensibility means the ability to be understood, by users of the standard and by respondents to administrative and statistical collections. It involves clarity of definitions, realism in the sense of modelling the real world, and providing a logical and coherent structure for collecting information.

The widespread use of standards will also provide an integrated statistical picture of Australian society. They facilitate the process of drawing together all the data about a particular topic, variable or population, from the full range of statistical sources, in a meaningful and useful way.

MAINTENANCE OF STATISTICAL STANDARDS

Users of the statistical standards should be aware that the standards are subject to an ongoing testing and maintenance program and may change as a result. In particular, the questions are reviewed regularly and the question response categories are updated on the basis of the latest available data. For instance, the Country of Birth of Person response categories will be updated following the 2001 Census if necessary.

The ABS will provide advice on the use and appropriate implementation of the standard variables. Variations from the standards are not encouraged and should only be formulated and used in consultation with the ABS. See contact details on page 100.

COUNTRY OF BIRTH OF PERSON

INTRODUCTION

Country of Birth of Person identifies the country in which a person was born. This variable is primarily used to determine whether or not someone is a migrant to Australia, the country from which they originated, and the community group to which they are likely to be attached. Country of Birth of Person is a key element of the current cultural data collection practices of many agencies, and is regarded as a primary measure of cultural background.

Migrants and their descendants have been identified by program developers and service providers as forming population groups that are likely to experience disadvantage when seeking access to government and community programs and services in Australia. It should be noted, however, that Country of Birth of Person may also indicate the extent to which persons from certain backgrounds are associated with relative advantage, as well as measuring disadvantage. Country of Birth of Person also helps to measure cultural diversity in its broader sense, particularly when used in conjunction with other cultural and language variables.

DEFINITION

Country of Birth of Person is defined as the country the respondent identifies as being the one in which they were born.

ISSUES FOR CONSIDERATION

The ABS recently changed the standard name of the variable from 'Birthplace' to 'Country of Birth of Person' to better describe the geographic units (countries) for which responses are obtained. The term 'Birthplace' could equally describe subsets of countries such as states, regions or towns. It should be noted that the name change of the variable does not affect the definition, scope or meaning of the concept being measured and that 'Country of Birth of Person' statistics and 'Birthplace' statistics are compatible.

STANDARD QUESTIONS

There are two standard questions for Country of Birth of Person. The Question for Detailed Data elicits detailed data on the specific country in which the person was born. There are two alternatives for this question which may be used depending on space and cost considerations. The first provides a set of tick boxes and an 'Other—please specify' response category for writing a country not specified in the tick box list. The second has a tick box for Australia and an 'Other—please specify' write-in category. The second alternative is more compact but will incur additional coding costs. The Question for Minimum Data only contains the tick box options 'Australia' and 'Other country' and therefore does not collect detail on countries other than Australia.

Question for Detailed Data

The standard question for obtaining detailed data for Country of Birth of Person is as follows:

In which country [were you] [was the person] [was (name)] born?

Australia.....
England .....
New Zealand.....
Italy.....
Viet Nam .....
Scotland .....
Greece.....
Germany .....
Philippines .....
Netherlands.....
Other—please specify:

The countries provided with this question have been included primarily on the basis of their statistical significance in the Australian context. These countries accounted for approximately 85% of all Country of Birth of Person responses recorded in the 1996 Census of Population and Housing. They will be reviewed periodically as different countries of birth assume a greater or lesser significance in the Australian context as a result of changing migration patterns.

Although China should be included in the list based on current numbers in Australia, it is excluded because it is not the only country in which Chinese people may be born and, if included, may bias responses. For example, if China were included in the list, Taiwan and Hong Kong would also have to be included or Chinese people born in these countries may be induced to mark the China tick box option. The 'Other' category is included for those people who were born in a country not listed. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in their country of birth.

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list. However, if there is a need for detailed information but space constraints are an overriding consideration, the full tick box list can be replaced by a tick box for Australia and the 'Other—please specify' write-in category, as follows:

In which country [were you] [was the person] [was (name)] born?

Australia.....
Other—please specify:

COUNTRY OF BIRTH OF PERSON

Question for Detailed Data *continued*

All overseas-born respondents would be required to write in their country of birth response. This would result in significant additional coding costs compared to the use of the full tick box list, as there is likely to be a wide range of responses to the 'Other—please specify' write-in category which would have to be coded manually using a coding index (see *Coding Procedures*). These costs could be reduced by tailoring the choice and number of tick boxes included with the 'Other—please specify' write-in category based on the composition of the population to be covered by the collection, as determined by ABS census data. A tailored tick box list would capture (and self-code) the majority of responses, reducing the amount of manual coding required.

Question for Minimum Data

The standard question for obtaining minimum data for Country of Birth of Person is as follows:

In which country [were you] [was the person] [was (name)] born?

Australia..... ☐
Other country..... ☐

This question is suitable for statistical and administrative collections where there is no requirement for detailed data on Country of Birth of Person. It is primarily recommended for collections that simply seek to determine whether or not a person is a migrant.

CLASSIFICATION

The *Standard Australian Classification of Countries* (SACC) is used when collecting, aggregating and disseminating data relating to Country of Birth of Person. The term 'country' is used in the SACC to describe fully independent countries (sovereign nation states), administrative subdivisions of the United Kingdom (i.e. England, Scotland, Wales and Northern Ireland), external territories and dependencies of independent countries, and recognised geographic areas of which the ownership or control is in dispute.

This classification has a three-level hierarchical structure. The third and most detailed level of the classification consists of 244 base or third-level units which are discrete countries (as defined above). Included in the 244 third-level units are five 'not elsewhere classified' (nec) categories, which contain entities that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics relating to country of birth, data should be captured, classified and stored at the base level of the SACC wherever possible.

The second level of the classification comprises 27 minor groups which are groups of neighbouring countries similar in terms of social, cultural, economic and political characteristics. Each minor group lies wholly within the boundaries of a geographic continent. On average, minor groups contain nine countries, with individual minor groups containing between one and 27 countries.

CLASSIFICATION *continued*

The first and most general level of the classification comprises nine major groups which are formed by aggregating geographically proximate minor groups and, therefore, comprise countries which are broadly similar in terms of social, cultural, economic and political characteristics.

For further details on the classification, see *Standard Australian Classification of Countries (SACC)* (Cat. no. 1269.0) or contact the ABS (see contact details on page 100).

CODING PROCEDURES

A coding index has been developed to assist in the implementation and use of the SACC and should be used when coding responses to the Country of Birth of Person questions. The index contains a comprehensive list of the most probable responses to questions relating to countries and their correct classification codes. Use of the coding index enables responses to be coded accurately and quickly to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the SACC publication, or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

OUTPUT CATEGORIES

The hierarchical structure of the SACC allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the major group level, minor group level, or the base or country level. If necessary, significant countries within a minor group can be presented separately while the remaining countries within the minor group are aggregated. For example, the third-level category 'Nauru' can be separately identified as an output category within the 'Micronesia' Minor Group, with the remaining third-level categories in this minor group aggregated in an 'Other Micronesia' category. The same principle can be adopted to highlight significant minor groups within a major group.

COUNTRY OF BIRTH OF PERSON

OUTPUT CATEGORIES
continued

The following is an example of standard output categories for the nine major groups of the SACC, which have either significant minor groups (second-level categories) or countries (third-level categories) highlighted.

Oceania and Antarctica
- Australia
- New Zealand
- Other Oceania and Antarctica

North-West Europe
- United Kingdom
  - England
  - Scotland
  - Other United Kingdom
- Western Europe
  - Germany
  - Netherlands
  - Other Western Europe
- Other North-West Europe

Southern and Eastern Europe
- Italy
- Greece
- Other Southern and Eastern Europe

North Africa and the Middle East
- Lebanon
- Other North Africa and the Middle East

South-East Asia
- Malaysia
- Philippines
- Viet Nam
- Other South-East Asia

North-East Asia
- Chinese Asia
  - China
  - Hong Kong
  - Other Chinese Asia
- Japan and the Koreas

Southern and Central Asia
- India
- Other Southern and Central Asia

Americas

Sub-Saharan Africa

OUTPUT CATEGORIES
continued

The SACC is intended to provide a single framework for classifying all statistical and administrative data by country for both population and economic statistics. Because the geographically based country groupings of the main structure are not always ideal for the presentation of data, alternative groupings have been included in the SACC publication to meet specific output needs.

These country groupings are of two types:

- a set of standard country groupings based on economic and political requirements and comprising associations or organisations of formally recognised member countries (for example, the Association of South East Asian Nations (ASEAN)); and
- a set of authorised standard output options for time series purposes or other data output needs (for example, Former USSR).

Further information on the alternative standard country groupings can be found in the SACC publication.

The standard output categories for the Country of Birth of Person Question for Minimum Data are:

Born in Australia
Born overseas

MAIN LANGUAGE OTHER THAN ENGLISH SPOKEN AT HOME

INTRODUCTION

Main Language Other Than English Spoken at Home identifies languages, other than English, which are spoken within the home. This variable provides information on the number of persons who speak English only and, if one or more other languages are spoken, the main non-English language used. It has been collected in the ABS Census of Population and Housing since 1986. In the Census, this question is used as a filter for Proficiency in Spoken English (see page 32). 'English only' speakers are sequenced past the Proficiency in Spoken English question.

This variable is useful because it picks up a second language apart from English and it maximises numbers for the more established migrant communities where the main language spoken is now English. It can also be used to measure language shift, that is, the extent to which community languages are being retained by their community or replaced by English.

Data on languages spoken in the home are regarded as an indicator of active ethnicity and language use, and are useful for the study of inter-generational language retention. Information on languages is also required for the investigation and development of interpreter or translation services, and for the implementation of national and state policies on language.

DEFINITION

Main Language Other Than English Spoken at Home is the main language, other than English, spoken by a person in his or her home, on a regular basis, to communicate with other residents of the home and regular visitors to the home. If more than one language is spoken, the respondent is asked to report the language other than English which the person speaks at home most often.

ISSUES FOR CONSIDERATION

Main Language Other Than English Spoken at Home maximises numbers for the more established migrant communities where the main language spoken outside the home is English but a language other than English is spoken at home. In some cases, however, this measure may not reflect complete language use, for example, when English is the only language spoken in the home but a language other than English is spoken outside the home, within a person's ethnic community group. This measure may also record the language usage of those people whose main and preferred language is English but who have learnt another language, which is occasionally but not normally spoken at home, in the 'Other Than English' category.

A problem with this variable, particularly for administrative purposes, is that it includes people whose first, main and most proficient language is English in a category with individuals whose use of a language other than English may only be marginal (the extent of this is not known). For this reason it should be collected in conjunction with Proficiency in Spoken English if the goal is to identify the potentially disadvantaged.

ISSUES FOR CONSIDERATION *continued*

The way in which sign languages are treated with regard to Main Language Other Than English Spoken at Home should also be noted. Sign languages fall into two groups, signed English/finger spelling and other sign languages such as Auslan. Signed English/finger spelling is considered to be another form of English and is treated as such. Auslan is regarded as a distinct language and individuals who report their main language as Auslan (or another recognised sign language) are considered to speak a language other than English.

STANDARD QUESTIONS

There are two standard questions for Main Language Other Than English Spoken at Home. The Question for Detailed Data elicits detailed data on the specific main language other than English that a person speaks in the home. There are two alternatives for this question which may be used depending on space and cost considerations. The first provides a set of tick boxes and a 'Yes, other—please specify' response category for writing a language not specified in the tick box list. The second has a tick box for 'No, English only' and a 'Yes, other—please specify' write-in category. The second alternative is more compact but will incur additional coding costs. The Question for Minimum Data only contains the tick box options 'No, English only' and 'Yes, other' and therefore does not collect details as to which language other than English is spoken.

Question for Detailed Data

The standard question for obtaining detailed data for Main Language Other Than English Spoken at Home is as follows:

[Do you] [Does the person] [Does (name)] speak a language other than English *at home*?
(If more than one language, indicate the one that is spoken most often.)

- No, English only ☐
- Yes, Italian ☐
- Yes, Greek ☐
- Yes, Cantonese ☐
- Yes, Mandarin ☐
- Yes, Arabic ☐
- Yes, Vietnamese ☐
- Yes, German ☐
- Yes, Spanish ☐
- Yes, Tagalog (Filipino) ☐
- Yes, other—please specify: ☐

Instructions are included for those respondents who speak more than one language, other than English, in the home. For self-enumerated collections, respondents are instructed to indicate the language other than English spoken most often and to mark *one* box in the checklist or to write in *one* response where tick boxes do not apply. For interview-based collections, an interviewer could use a prompt like 'Which of those languages [do you] [does the person] [does (name)] use most often?' if a multiple response is given after asking the question.

MAIN LANGUAGE OTHER THAN ENGLISH SPOKEN AT HOME

Question for Detailed Data *continued*

The languages provided with this question have been included primarily on the basis of their statistical significance in the Australian context. These languages accounted for approximately 90% of all languages spoken in Australia in the 1996 Census of Population and Housing. The list will be reviewed periodically as different languages assume a greater or lesser significance in the Australian context as a result of changing migration patterns.

Cantonese and Mandarin are both included in the list to ensure that the response 'Chinese', for example, is not erroneously coded to one or the other. The 'Other' category is included for those people who speak a language in the home that is not offered as a tick box response to the question. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in their language response.

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list. However, if there is a need for detailed information but space constraints are an overriding consideration, the full tick box list can be replaced by a tick box for English and the 'Other—please specify' write-in category, as follows:

[Do you] [Does the person] [Does (name)] speak a language other than English *at home*?
(If more than one language, indicate the one that is spoken most often.)

No, English only..... ☐
Yes, other—please specify: ☐

All respondents who mainly speak a language other than English at home would be required to write in their response. This would result in significant additional coding costs compared to the use of the full tick box list, as there is likely to be a wide range of responses to the 'Other—please specify' write-in category which would have to be coded manually using a coding index (see *Coding Procedures*). These costs could be reduced by tailoring the choice and number of tick boxes included with the 'Other—please specify' write-in category based on the composition of the population to be covered by the collection, as determined by ABS census data. A tailored tick box list would capture (and self-code) the majority of responses reducing the amount of manual coding required.

Question for Minimum Data

The standard question for obtaining minimum data for Main Language Other Than English Spoken at Home is as follows:

[Do you] [Does the person] [Does (name)] speak a language other than English *at home?*
(If more than one language, indicate the one that is spoken most often.)

No, English only ☐
Yes, other ☐

This question is suitable for statistical and administrative collections where there is no requirement for detailed language data. It is primarily recommended for collections that simply seek to determine whether or not a person speaks a language other than English at home.

CLASSIFICATION

The *Australian Standard Classification of Languages* (ASCL) is used when collecting, aggregating and disseminating data on language usage in Australia. The term 'language' is used in the ASCL to describe languages, dialects, pidgins, creoles and invented and sign languages.

The classification has a three-level hierarchical structure. The third and most detailed level consists of 193 base or third-level units which are separately identified languages (as defined above). Included in the 193 third-level units are 26 'not elsewhere classified' (nec) and three 'other' categories, which contain entities that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics, data should be captured, classified and stored at the base level of the ASCL wherever possible.

The second level of the classification comprises 48 narrow groups of languages which are similar in terms of their evolution from a common ancestral language (genetic affinity) and geographic proximity in terms of the areas in which they originated. Included in the 48 narrow groups are five 'other' categories which consist of languages which do not fit into a particular narrow group.

The first and most general level of the classification comprises nine broad groups of languages, including one 'other' category. Broad groups are formed by aggregating geographically proximate narrow groups.

For further details on the classification, see *Australian Standard Classification of Languages (ASCL)* (Cat. no. 1267.0) or contact the ABS (see contact details on page 100).

MAIN LANGUAGE OTHER THAN ENGLISH SPOKEN AT HOME

CODING PROCEDURES

Language data are coded to the classification using the ASCL coding index. This contains a comprehensive list of the most probable responses to questions relating to language and their correct classification codes. Use of the coding index enables responses to be coded accurately and quickly to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the ASCL publication or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

OUTPUT CATEGORIES

The hierarchical structure of the ASCL allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the broad group level, narrow group level, or the base or language level. If necessary, significant languages within a narrow group can be presented separately while the remaining languages within the narrow group are aggregated. For example, the third-level category 'Tiwi' can be separately identified as an output category within the 'Northern Aboriginal' Narrow Group, with the remaining third-level categories in this narrow group aggregated to form an 'Other Northern Aboriginal' category. The same principle can be adopted to highlight significant narrow groups or languages within a broad group.

The following is an example of standard output categories for the nine broad groups of the ASCL, which have either significant narrow groups (second-level categories) or languages (third-level categories) highlighted. This example allows separate identification of all languages with more than 35,000 speakers (aged five years and over) according to data on Main Language Other Than English Spoken at Home in the 1996 Census.

Speaks English only

Main other Languages:

Northern European Languages
- German
- Netherlandic
- Other Northern European Languages

Southern European Languages
- French
- Greek
- Italian
- Maltese
- Spanish
- Other Southern European Languages

OUTPUT CATEGORIES
continued

- Eastern European Languages
  - South Slavic
    - Croatian
    - Macedonian
    - Serbian
    - Other South Slavic
  - West Slavic
    - Polish
    - Other West Slavic
  - Other Eastern European Languages
- Southwest Asian and North African Languages
  - Arabic (including Lebanese)
  - Turkish
  - Other Southwest Asian and North African Languages
- Southern Asian Languages
- Southeast Asian Languages
  - Tagalog (Filipino)
  - Vietnamese
  - Other Southeast Asian Languages
- Eastern Asian Languages
  - Chinese
    - Cantonese
    - Mandarin
    - Other Chinese
  - Other Eastern Asian Languages
- Australian Indigenous Languages
- Other Languages

The standard output categories for the Main Language Other Than English Spoken at Home Question for Minimum Data are:

- English only
- Other language

PROFICIENCY IN SPOKEN ENGLISH

INTRODUCTION

Proficiency in Spoken English is used to assess the English speaking ability of people who speak a language other than English.

Data relating to this variable are used primarily to identify people who may suffer disadvantage as a result of a lack of competence in spoken English. This information can be used to target the provision of services to people whose lack of ability in spoken English is potentially a barrier to gaining access to government programs and services, and participating equitably in Australian society.

Proficiency in Spoken English requires a filter language question so that it is only asked of people who speak a language other than English. The language variables Main Language Other Than English Spoken at Home (see page 26), First Language Spoken (see page 54) and Main Language Spoken at Home (see page 64) are used for this purpose.

DEFINITION

Proficiency in Spoken English is defined as the ability to speak English in every day situations. In practice, the variable measures the self-assessed level of ability to speak English, asked of people who speak a language other than English.

ISSUES FOR CONSIDERATION

It is important that the Proficiency in Spoken English question be asked only of those people who speak a language other than English. It is inappropriate to ask people whose only language is English how well they speak it and the results could be misleading. In order to target those people who speak a language other than English, Proficiency in Spoken English requires a filter language question. The variables Main Language Other Than English Spoken at Home, First Language Spoken and Main Language Spoken at Home can all be used to ask respondents what languages they speak and sequence only those people with languages other than English to the Proficiency in Spoken English question. The choice of language filter question will depend on the requirements of the collection.

Other issues to consider with Proficiency in Spoken English are that it is a measure of proficiency in *spoken* English rather than a measure of proficiency in other aspects of communication in English (e.g. listening, writing and reading) and, secondly, that the assessment of how well a person speaks English is subjective. This variable cannot measure proficiency in spoken English with the same degree of precision as a formal test. An answer to the Proficiency in Spoken English question will depend on how the respondent interprets the question or how an interviewer or person who answers for the respondent assesses the person's ability to speak English.

ISSUES FOR CONSIDERATION *continued*

For example, respondents whose spoken English ability enables them to do the shopping and manage many everyday transactions may consider that they speak English very well even though they may not be able to hold a social conversation in English. Conversely, more exacting or self-critical respondents may consider that they do not speak English well in spite of the fact they can communicate with near-native proficiency. ABS testing prior to the first use of this variable in the 1986 Census showed that it is a good identifier of people who are likely to need assistance in the form of interpreter services or English language classes. Those people likely to need assistance will generally respond 'Not well' or 'Not at all'.

It should also be noted that, like people whose only language is English, it is inappropriate to ask people who use sign languages how well they speak English if these people are filtered to the Proficiency in Spoken English question.

STANDARD QUESTIONS

There are two standard questions for Proficiency in Spoken English. The first question is used in self-enumerated collections and the second in collections conducted by interview. The Proficiency in Spoken English question (either the self-enumerated or interview-based version) should be asked after the chosen language question which filters out most people whose only language is English.

Question for Self-Enumerated Collections

How well [do you] [does the person] speak English?

- Very well ☐
- Well ☐
- Not well ☐
- Not at all ☐

Question for Interview-Based Collections

Do you consider [you speak] [(name) speaks] English very well, well, or not well?

- Very well ☐
- Well ☐
- Not well ☐
- Not at all ☐

A different question is required for interview-based collections as, unlike a self-enumerated question where respondents can see the response choices and use them to interpret the question, respondents may be unaware of how to rate their proficiency when the interviewer asks the question 'How well do you speak English?'. Including the options 'very well', 'well' and 'not well' when asking the question gives the respondents some idea of how to rate their proficiency. It is important that this question records the respondent's own perception of how well they rate their English speaking skills.

PROFICIENCY IN SPOKEN ENGLISH

CLASSIFICATION

The standard classification used to code, store and disseminate data collected by Proficiency in Spoken English is a flat classification, having only one level with four categories. These categories are the same as the question response categories ('Very well', 'Well', 'Not well', 'Not at all'). Two supplementary codes are also provided in the classification for 'Not stated' and 'Not applicable' responses (see *Coding Procedures*), which makes the classification applicable to all persons. However, it should be noted that this is a coding convention and that, setting aside these special codes, the scope of the classification, along with the variable, is people who speak a language other than English.

CODING PROCEDURES

The standard code structure for the Proficiency in Spoken English classification is as follows:

1 Very well
2 Well
3 Not well
4 Not at all
0 Not stated
9 Not applicable

Responses are assigned the appropriate codes.

OUTPUT CATEGORIES

The standard output categories for Proficiency in Spoken English data are as follows:

Very well
Well
Not well
Not at all
Not stated
Not applicable

As the distinction between the categories 'Very well' and 'Well' is somewhat subjective, they can be amalgamated into the output category 'Well or Very well'. Similarly, 'Not well' and 'Not at all' can be amalgamated into the output category 'Not well or Not at all'. In this case, the output categories would be as follows:

Well or Very well
Not well or Not at all
Not stated
Not applicable

INDIGENOUS STATUS

INTRODUCTION

Indigenous Status provides data on the number of people who identify as being of Aboriginal or Torres Strait Islander descent. Accurate and consistent statistics about Aboriginal and Torres Strait Islander peoples are needed in order to plan, promote and deliver essential services, to monitor changes in well-being and to administer government policy.

The purpose of this standard is to provide a consistent conceptual framework for the collection of information about Aboriginal and Torres Strait Islander peoples in ABS and non-ABS statistical collections and in administrative and service provision settings. This standard can be used when collecting population estimates and data on births, deaths, hospital admissions, labour force, the economy, education and law and justice.

DEFINITION

An Aboriginal or Torres Strait Islander is a person of Aboriginal or Torres Strait Islander descent who identifies as an Aboriginal or Torres Strait Islander and is accepted as such by the community in which he or she lives. That is, there are three components to the definition: descent, self-identification and community acceptance. In practice, it is not feasible to collect information on the community acceptance part of this definition and therefore questions on Indigenous Status relate to descent and self-identification only. Ideally, descent could be determined by asking if a person has either an Aboriginal or Torres Strait Islander ancestor. Self-identification could be determined by asking if a person identifies culturally as an Aboriginal or Torres Strait Islander. In practice, people are asked if they are of Aboriginal or Torres Strait Islander origin. This question is considered to measure descent and for some, but not all, cultural identity.

ISSUES FOR CONSIDERATION

The production of accurate population estimates for Indigenous people is a difficult undertaking as data relies on self-identification and the willingness of people to identify as Indigenous. However, recent ABS Census of Population and Housing figures have shown that this 'propensity to identify' has changed over time. 1996 Census figures showed that the Indigenous population had increased by 55% since 1986, compared to a 12% increase in the non-Indigenous population. Much of this increase can be explained by a greater willingness of people to report their Indigenous origin on the census form.

The use of correct terminology is especially important when referring to Indigenous people. The ABS, along with other APS agencies, received a formal request from the Chief Executive Officer of the Aboriginal and Torres Strait Islander Commission (ATSIC) in November 1992 to adopt appropriate terminology when referring to Aboriginal and Torres Strait Islander peoples. Acceptable terms for use when referring to this variable are 'Indigenous Status' and 'Aboriginal and Torres Strait Islander origin'. When referring to the population group, 'Aboriginal and Torres Strait Islander peoples' or 'Indigenous peoples' should be used. 'Indigenous Australians' is also an acceptable term for referring to Aboriginal and Torres Strait Islander peoples. 'Indigenous' covers persons who are

INDIGENOUS STATUS

ISSUES FOR CONSIDERATION *continued*

Aboriginal, Torres Strait Islander or both Aboriginal and Torres Strait Islander. When referring to events or circumstances affecting Indigenous people, phrases such as 'Deaths of Indigenous people' are preferred to 'Indigenous deaths'.

References which should be avoided are the use of the word 'Aboriginality', 'ATSI', 'ATSIs', 'ATSI people', 'Aboriginal and TSI', 'A&TSI', 'Islanders', 'TSIs', or 'TIs'. It is also not appropriate to use the word 'Aboriginal' to refer to both Aboriginal people and Torres Strait Islanders. Another commonly misused term when referring to Indigenous people is 'ATSIC' or 'ATSIC people'. ATSIC is the acronym for the Aboriginal and Torres Strait Islander Commission and therefore does not refer to Indigenous people themselves.

STANDARD QUESTION

The standard question for Indigenous Status is as follows:

[Are you] [Is the person] [Is (name)] of Aboriginal or Torres Strait Islander origin?
(For persons of both Aboriginal and Torres Strait Islander origin, mark both 'Yes' boxes.)

No..... ☐
Yes, Aboriginal ..... ☐
Yes, Torres Strait Islander..... ☐

This question is recommended for self-enumerated or interview-based collections. It can also be used in circumstances where a close relative, friend, or another member of the household is answering on behalf of the subject. When someone is not present, the person answering for them should be in a position to do so, i.e. this person must know the person about whom the question is being asked well and feel confident to provide accurate information about them. However, it is strongly recommended that this question be asked directly wherever possible. In circumstances where it is impossible to ask the person directly, such as in the case of death, the question should be asked of a close relative or friend, and only if a relative or friend is not available should the undertaker or other such person answer.

This question should always be asked even if the person does not 'look' Aboriginal or Torres Strait Islander.

CLASSIFICATION

The classification for Indigenous Status has a hierarchical structure comprising two levels. There are four categories at the detailed level of the classification which are grouped into two categories at the broad level. There is one supplementary category for 'not stated' responses. The classification is as follows:

Indigenous
- Aboriginal but not Torres Strait Islander Origin
- Torres Strait Islander but not Aboriginal Origin
- Both Aboriginal and Torres Strait Islander Origin

Non-Indigenous
- Neither Aboriginal nor Torres Strait Islander Origin

Not Stated

CODING PROCEDURES

Responses to the Indigenous Status question are coded to the appropriate category of the classification. Only the second digit of the two digit code need be used for data input and storage purposes. The code structure for the classification is as follows:

1 Indigenous
- 11 Aboriginal but not Torres Strait Islander Origin
- 12 Torres Strait Islander but not Aboriginal Origin
- 13 Both Aboriginal and Torres Strait Islander Origin

2 Non-Indigenous
- 24 Neither Aboriginal nor Torres Strait Islander Origin

09 Not Stated

The Indigenous Status question allows for more than one response. The procedure for coding multiple responses is as follows:

- If the respondent marks 'No' and either 'Aboriginal' or 'Torres Strait Islander', then the response should be coded to either Aboriginal or Torres Strait Islander as indicated (i.e. disregard the 'No' response).
- If the respondent marks both the 'Aboriginal' and 'Torres Strait Islander' boxes, then their response should be coded to 'Both Aboriginal and Torres Strait Islander Origin'.
- If the respondent marks all three boxes ('No', 'Aboriginal' and 'Torres Strait Islander'), then the response should be coded to 'Both Aboriginal and Torres Strait Islander Origin' (i.e. disregard the 'No' response).

INDIGENOUS STATUS

OUTPUT CATEGORIES

The output categories for Indigenous Status are the same as the classification structure.

The full classification is most commonly used in the census, surveys specifically targeted to Indigenous people and some administrative collections. If 'Both Aboriginal and Torres Strait Islander Origin' has too few people to be released separately, then data should be released for the total 'Indigenous' with no further breakdown.

The terms 'Aboriginal or Torres Strait Islander Origin' and 'Neither Aboriginal nor Torres Strait Islander Origin' may be used as alternative labels for the categories 'Indigenous' and 'Non-Indigenous' respectively.

ANCESTRY

INTRODUCTION

Ancestry describes the ethnic or cultural groups to which a person's forebears are or were attached. There is significant community interest in measuring the ethnic and cultural composition of the Australian population and in identifying the characteristics of particular migrant community groups. An important element of cultural and language diversity in Australia is the extent to which Australians retain the culture, ethnicity or language of their parents. Retention of cultural and language diversity is determined by the Country of Birth variables, in association with variables such as Indigenous Status (see page 35), First Language Spoken (see page 54) and Religious Affiliation (see page 68).

Indigenous Status is a separate variable measuring a specific element of ancestry, namely whether a person is of Aboriginal or Torres Strait Islander origin. However, it does not necessarily provide information on all elements of a person's ancestry.

The Ancestry variable can be used in conjunction with these variables in order to identify particular ethnic or cultural groups, as it provides a self-assessed measure of ethnicity and cultural background. However, Ancestry in the Australian context is problematic as there are many Australians with origins and heritage which do not, in practice, relate to their current ethnic identity. Ancestry data alone, therefore, are not considered to be a particularly good measure of service needs or the extent to which persons from certain backgrounds are associated with advantage or disadvantage, and should only be used as a broad measure of cultural diversity.

A major advantage of the Ancestry variable is that it is able to measure an association with ethnic and cultural groups which do not equate directly to countries or languages and thus cannot be readily identified using country of birth or language variables. For instance, the Ancestry variable assists in the identification and measurement of ethnic and cultural minorities which exist or originate within particular countries, ethnic and cultural groups which form a distinct unbroken geographic block across neighbouring country borders, and ethnic and cultural groups which are located in many disparate countries across the world.

The Ancestry variable was included in the 1986 Census of Population and Housing but was not included in the 1991 or 1996 Censuses. However, due to demand for this data by many groups in the community, the ABS plans to include an Ancestry question in the 2001 Census.

DEFINITION

Ancestry describes the ethnic or cultural heritage of a person, that is, the ethnic or cultural groups to which a person's forebears are or were attached. In practice, Ancestry is the ethnic or cultural groups which the person *identifies* as being his or her ancestry. For example, a respondent may indicate four ancestries because each grandparent is from a different ethnic or cultural background (say Italian, Greek, German, English). However, another person with the same ancestry may choose to identify as 'Australian' because one or both parents were born in Australia, or

ANCESTRY

DEFINITION *continued*

because of a cultural or national attachment to this country. Ancestry therefore involves measures of self-identification of ethnic or cultural group affiliation or nationality as well as of descent from one or more particular groups.

ISSUES FOR CONSIDERATION

It should be noted that many people in Australia have a variety of cultural backgrounds and do not relate to a single ethnic or cultural group. These people will give multiple responses to a question on ancestry, ethnicity or cultural identity. Often the responses will indicate an identification with Australia in a national or cultural sense, but will also acknowledge continuing ties with other ethnic or cultural groups (e.g. Irish Australian, Italian Australian). This does not mean that people who identify primarily with other ethnicities, ancestries or cultures do not also regard themselves as Australian in most senses. The problem is with the terminology. 'Australian' is used as the adjective to describe the culture that has developed in this country over the last two hundred years, and also to describe all members of the citizenry regardless of whether or not they regard their ethnicity, culture or ancestry as Australian.

STANDARD QUESTION

The standard question for obtaining detailed data for Ancestry is as follows:

What is [your] [the person's] [(name)'s] Ancestry?
(For example: Hmong, Kurdish, Australian South Sea Islander, Maori.)
(Provide more than one ancestry if necessary.)

English ☐
Irish ☐
Italian ☐
German ☐
Greek ☐
Chinese ☐
Vietnamese ☐
Dutch ☐
Filipino ☐
Australian ☐
Other—please specify: _____

The list of ancestries provided with this question has been included primarily on the basis of an assessment of their statistical significance in the Australian context. It will be reviewed periodically, principally after the 2001 Census, as different ancestries assume a greater or lesser statistical significance in the Australian context.

STANDARD QUESTION
continued

The 'Other' category is included for those people who have an ancestry or ancestries not offered in the list as a response to the question. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in their ancestry response(s).

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list. However, if space constraints are an overriding consideration, the tick boxes can be deleted and space provided for all ancestry responses to be written in. This would result in significant additional coding costs compared to the use of the full tick box list, as there are likely to be a wide range of responses which would have to be coded manually using a coding index (see *Coding Procedures*). The question would be as follows:

What is [your] [the person's] [(name)'s] Ancestry?
(For example: English, Irish, Hmong, Kurdish, Italian, Greek, Maori, Vietnamese.)
(Provide more than one ancestry if necessary.)

Ancestry :

Since the above questions are subject to an ongoing program of testing in ABS collections and are subject to change, it is recommended that agencies wishing to collect Ancestry in their own data collection activities contact the ABS before implementing this standard.

CLASSIFICATION

The forthcoming *Australian Standard Classification of Cultural and Ethnic Groups* (ASCCEG) (due to be released in mid 2000) is to be used when collecting, aggregating and disseminating data relating to the Ancestry variable. The ASCCEG is designed to be used in the classification of information relating to a number of topics such as ancestry, ethnicity, and cultural identity. Although these topics have elements of difference, it is considered that the fundamental concept common to them all, and thus underpinning the classification, is *ethnicity*. The term 'ethnicity' is used in the ASCCEG to describe a shared identity or similarity of a group of people on the basis of one or more factors, e.g. shared history, cultural traditions, religion or language. In this sense, the term 'ethnicity' of course can be applied to all members of the Australian population. 'Australian' or 'Aboriginal' are valid descriptors of ethnicity.

The classification has a three-level hierarchical structure. The third and most detailed level of the classification consists of base or third-level units which are ethnic or cultural groups. Included in these third-level units are a number of 'not elsewhere classified' (nec) categories, which

CLASSIFICATION *continued*

contain ethnic or cultural groups that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics, data should be captured, classified and stored at the base level of the ASCCEG wherever possible.

The second level of the classification comprises narrow groups of ethnic and cultural groups which are similar in terms of the classification criteria (geographic proximity in terms of the areas in which they originated, a long shared history, and similarity in terms of social and cultural characteristics).

The first and most general level of the classification comprises nine broad groups of ethnic and cultural groups. Broad groups are formed by aggregating geographically proximate narrow groups.

For further details on the classification, see *Australian Standard Classification of Cultural and Ethnic Groups (ASCCEG)* (Cat. no. 1249.0) (available mid 2000) or contact the ABS (see contact details on page 100).

CODING PROCEDURES

A coding index is being developed to assist in the implementation and use of the ASCCEG. It contains a comprehensive list of the most probable responses to questions relating to ancestry, ethnicity and cultural identity and their correct classification codes. Use of the coding index enables responses to be coded accurately and quickly to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the forthcoming ASCCEG publication or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

The Ancestry question will record all claims of association with ancestries, ethnicities and cultures. As such, multiple responses are encouraged. The purpose of the ASCCEG is to code the extent to which people associate or identify with particular ethnic and cultural groups. It should therefore be noted that the classification is not intended to classify people, but rather all claims of association with an ethnic or cultural group. Often the responses will indicate an identification with Australia in a national or cultural sense, but will also acknowledge continuing ties with other ancestral, ethnic or cultural groups. Such responses include Irish Australian, Italian Australian, etc. It is therefore suggested that a minimum of two ancestral, ethnic and cultural groups nominated by a person on a statistical or administrative form be coded.

The ABS has developed guidelines for the coding, storage and presentation of multiple responses to questions on ancestry, ethnicity or cultural identity. These guidelines will be included in the forthcoming ASCCEG publication.

OUTPUT CATEGORIES

The hierarchical structure of the ASCCEG will allow users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the broad group level, narrow group level, or the base (ethnic and cultural group) level. If necessary, significant ethnic or cultural groups within a narrow group can be presented separately while the remaining ethnic groups within the narrow group are aggregated. The same principle can be adopted to highlight significant narrow groups within a broad group.

COUNTRY OF BIRTH OF FATHER

INTRODUCTION

Country of Birth of Father identifies the country in which a person's father was born. An important element of cultural and language diversity in Australia is the extent to which second generation Australians retain their parents' culture, ethnicity or language. Retention of cultural and language characteristics is determined by the variables Country of Birth of Father and Country of Birth of Mother (see page 49), in association with variables such as First Language Spoken (see page 54) and Religious Affiliation (see page 68).

Migrants and their descendants have been identified by program developers and service providers as forming population groups that are likely to experience disadvantage when seeking access to government and community programs and services in Australia. It should be noted, however, that Country of Birth of Father may also indicate the extent to which persons from certain backgrounds are associated with relative advantage, as well as measuring disadvantage. Country of Birth of Father also helps to measure cultural diversity in its broader sense, particularly when used in conjunction with other cultural and language variables.

DEFINITION

Country of Birth of Father is defined as the country the respondent identifies as being the one in which the person's father was born.

ISSUES FOR CONSIDERATION

The ABS recently changed the standard name of the variable from 'Birthplace of Father' to 'Country of Birth of Father' to better describe the geographic units (countries) for which responses are obtained. The term 'Birthplace' could equally describe subsets of countries such as states, regions or towns. It should be noted that the name change of the variable does not affect the definition, scope or meaning of the concept being measured and that 'Country of Birth of Father' statistics and 'Birthplace of Father' statistics are compatible.

Country of Birth of Father is closely related to the variable Country of Birth of Mother, and the two questions are generally asked together in collections with the Country of Birth of Father question traditionally asked first. If space constraints only allow enough room for one question on parents' country of birth, it is recommended that Country of Birth of Mother be asked as, in practice, people are likely to know their mother's country of birth with a greater degree of certainty than their father's country of birth.

STANDARD QUESTIONS

There are two standard questions for Country of Birth of Father. The Question for Detailed Data elicits detailed data on the specific country of birth of the person's father. There are two alternatives for this question which may be used depending on space and cost considerations. The first provides a set of tick boxes and an 'Other—please specify' response category for writing a country not specified in the tick box list. The second has a tick box for Australia and an 'Other—please specify' write-in category. The second alternative is more compact but will incur

STANDARD QUESTIONS
continued

additional coding costs. The Question for Minimum Data only contains the tick box options 'Australia' and 'Other country' and therefore does not collect detail on countries other than Australia.

Question for Detailed Data

The standard question for obtaining detailed data for Country of Birth of Father is as follows:

In which country was [your] [the person's] [(name)'s] father born?

- Australia ☐
- England ☐
- Italy ☐
- New Zealand ☐
- Scotland ☐
- Greece ☐
- Netherlands ☐
- Germany ☐
- Viet Nam ☐
- Lebanon ☐
- Other—please specify: ☐

The countries provided with this question have been included primarily on the basis of their statistical significance in the Australian context. These countries accounted for over 80% of all Country of Birth of Father responses recorded in the 1996 Census of Population and Housing. They will be reviewed periodically as different countries of birth assume a greater or lesser significance in the Australian context as a result of changing migration patterns.

Although China should be included in the list based on current numbers in Australia, it is excluded because it is not the only country in which Chinese people may be born and, if included, may bias responses. For example, if China were included in the list, Taiwan and Hong Kong would also have to be included or people with a Chinese father born in these countries may be induced to mark the China tick box option. The 'Other' category is included for those people whose father was born in a country not listed. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in the other country of birth.

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list. However, if there is a need for detailed information but space constraints are an overriding consideration, the full tick box list can be replaced by a tick box for Australia and the 'Other—please specify' write-in category, as follows:

COUNTRY OF BIRTH OF FATHER

Question for Detailed Data *continued*

In which country was [your] [the person's] [(name)'s] father born?

Australia..... ☐
Other—please specify: [ ]

Respondents with overseas-born fathers would be required to write in their father's country of birth response. This would result in significant additional coding costs compared to the use of the full tick box list, as there is likely to be a wide range of responses to the 'Other—please specify' write-in category which would have to be coded manually using a coding index (see *Coding Procedures*). These costs could be reduced by tailoring the choice and number of tick boxes included with the 'Other—please specify' write-in category based on the composition of the population to be covered by the collection, as determined by ABS census data. A tailored tick box list would capture (and self-code) the majority of responses reducing the amount of manual coding required.

Question for Minimum Data

The standard question for obtaining minimum data for Country of Birth of Father is as follows:

In which country was [your] [the person's] [(name)'s] father born?

Australia..... ☐
Other country..... ☐

This question is suitable for statistical and administrative collections where there is no requirement for detailed data on Country of Birth of Father. It is primarily recommended for collections that simply seek to determine whether or not a person's father is a migrant.

CLASSIFICATION

The *Standard Australian Classification of Countries* (SACC) is used when collecting, aggregating and disseminating data relating to Country of Birth of Father. The term 'country' is used in the SACC to describe fully independent countries (sovereign nation states), administrative subdivisions of the United Kingdom (i.e. England, Scotland, Wales and Northern Ireland), external territories and dependencies of independent countries, and recognised geographic areas of which the ownership or control is in dispute.

This classification has a three-level hierarchical structure. The third and most detailed level of the classification consists of 244 base or third-level units which are discrete countries (as defined above). Included in the 244 third-level units are five 'not elsewhere classified' (nec) categories, which contain entities that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics relating to country of birth, data should be captured, classified and stored at the base level of the SACC wherever possible.

CLASSIFICATION *continued*

The second level of the classification comprises 27 minor groups, which are groups of neighbouring countries similar in terms of social, cultural, economic and political characteristics. Each minor group lies wholly within the boundaries of a geographic continent. On average, minor groups contain nine countries, with individual minor groups containing between one and 27 countries.

The first and most general level of the classification comprises nine major groups which are formed by aggregating geographically proximate minor groups and, therefore, comprise countries which are broadly similar in terms of social, cultural, economic and political characteristics.

For further details on the classification, see *Standard Australian Classification of Countries (SACC)* (Cat. no. 1269.0) or contact the ABS (see contact details on page 100).

CODING PROCEDURES

A coding index has been developed to assist in the implementation and use of the SACC and should be used when coding responses to the Country of Birth of Father questions. The index contains a comprehensive list of the most probable responses to questions relating to countries and their correct classification codes. Use of the coding index enables responses to be coded accurately and quickly to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the SACC publication, or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

OUTPUT CATEGORIES

The hierarchical structure of the SACC allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the major group level, minor group level, or the base or country level. If necessary, significant countries within a minor group can be presented separately while the remaining countries within the minor group are aggregated. For example, the third-level category 'Nauru' can be separately identified as an output category within the 'Micronesia' Minor Group, with the remaining third-level categories in this minor group aggregated in an 'Other Micronesia' category. The same principle can be adopted to highlight significant minor groups within a major group. The Country of Birth of Person standard gives an example of standard output categories for the nine major groups of the SACC, which have either significant minor groups or countries highlighted (see page 24).

COUNTRY OF BIRTH OF FATHER

OUTPUT CATEGORIES
continued

The SACC is intended to provide a single framework for classifying all statistical and administrative data by country for both population and economic statistics. Because the geographically based country groupings of the main structure are not always ideal for the presentation of data, alternative groupings have been included in the SACC publication to meet specific output needs. These country groupings are of two types:

- a set of standard country groupings based on economic and political requirements and comprising associations or organisations of formally recognised member countries (for example, the Association of South East Asian Nations (ASEAN)); and
- a set of authorised standard output options for time series purposes or other data output needs (for example, Former USSR).

Further information on the alternative standard country groupings can be found in the SACC publication.

The standard output categories for the Country of Birth of Father Question for Minimum Data are:

Father born in Australia
Father born overseas

COUNTRY OF BIRTH OF MOTHER

INTRODUCTION

Country of Birth of Mother identifies the country in which a person's mother was born. An important element of cultural and language diversity in Australia is the extent to which second generation Australians retain their parents' culture, ethnicity or language. Retention of cultural and language characteristics is determined by the variables Country of Birth of Mother and Country of Birth of Father (see page 44), in association with variables such as First Language Spoken (see page 54) and Religious Affiliation (see page 68).

Migrants and their descendants have been identified by program developers and service providers as forming population groups that are likely to experience disadvantage when seeking access to government and community programs and services in Australia. It should be noted, however, that Country of Birth of Mother may also indicate the extent to which persons from certain backgrounds are associated with relative advantage, as well as measuring disadvantage. Country of Birth of Mother also helps to measure cultural diversity in its broader sense, particularly when used in conjunction with other cultural and language variables.

DEFINITION

Country of Birth of Mother is defined as the country the respondent identifies as being the one in which the person's mother was born.

ISSUES FOR CONSIDERATION

The ABS recently changed the standard name of the variable from 'Birthplace of Mother' to 'Country of Birth of Mother' to better describe the geographic units (countries) for which responses are obtained. The term 'Birthplace' could equally describe subsets of countries such as states, regions or towns. It should be noted that the name change of the variable does not affect the definition, scope or meaning of the concept being measured and that 'Country of Birth of Mother' statistics and 'Birthplace of Mother' statistics are compatible.

Country of Birth of Mother is closely related to the variable Country of Birth of Father, and the two questions are generally asked together in collections with the Country of Birth of Father question traditionally asked first. If space constraints only allow enough room for one question on parents' country of birth, it is recommended that Country of Birth of Mother be asked as, in practice, people are likely to know their mother's country of birth with a greater degree of certainty than their father's country of birth.

STANDARD QUESTIONS

There are two standard questions for Country of Birth of Mother. The Question for Detailed Data elicits detailed data on the specific country of birth of the person's mother. There are two alternatives for this question which may be used depending on space and cost considerations. The first provides a set of tick boxes and an 'Other—please specify' response category for writing a country not specified in the tick box list. The second has a tick box for Australia and an 'Other—please specify' write-in category. The second alternative is more compact but will incur additional coding costs.

COUNTRY OF BIRTH OF MOTHER

STANDARD QUESTIONS *continued*

The Question for Minimum Data only contains the tick box options 'Australia' and 'Other country' and therefore does not collect detail on countries other than Australia.

Question for Detailed Data

The standard question for obtaining detailed data for Country of Birth of Mother is as follows:

In which country was [your] [the person's] [(name)'s] mother born?

- Australia ☐
- England ☐
- Italy ☐
- New Zealand ☐
- Scotland ☐
- Greece ☐
- Viet Nam ☐
- Netherlands ☐
- Lebanon ☐
- Philippines ☐
- Other—please specify: ☐

The countries provided with this question have been included primarily on the basis of their statistical significance in the Australian context. These countries accounted for over 80% of all Country of Birth of Mother responses recorded in the 1996 Census of Population and Housing. They will be reviewed periodically as different countries of birth assume a greater or lesser significance in the Australian context as a result of changing migration patterns.

Although China should be included in the list based on current numbers in Australia, it is excluded because it is not the only country in which Chinese people may be born and, if included, may bias responses. For example, if China were included in the list, Taiwan and Hong Kong would also have to be included or people with a Chinese mother born in these countries may be induced to mark the China tick box option. The 'Other' category is included for those people whose mother was born in a country not listed. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in the other country of birth.

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list. However, if there is a need for detailed information but space constraints are an overriding consideration, the full tick box list can be replaced by a tick box for Australia and the 'Other—please specify' write-in category, as follows:

COUNTRY OF BIRTH OF MOTHER

Question for Detailed Data *continued*

In which country was [your] [the person's] [(name)'s] mother born?

Australia..... ☐
Other—please specify: ☐

Respondents with overseas-born mothers would be required to write in their mother's country of birth response. This would result in significant additional coding costs compared to the use of the full tick box list, as there is likely to be a wide range of responses to the 'Other—please specify' write-in category which would have to be coded manually using a coding index (see *Coding Procedures*). These costs could be reduced by tailoring the choice and number of tick boxes included with the 'Other—please specify' write-in category based on the composition of the population to be covered by the collection, as determined by ABS census data. A tailored tick box list would capture (and self-code) the majority of responses reducing the amount of manual coding required.

Question for Minimum Data

The standard question for obtaining minimum data for Country of Birth of Mother is as follows:

In which country was [your] [the person's] [(name)'s] mother born?

Australia..... ☐
Other country..... ☐

This question is suitable for statistical and administrative collections where there is no requirement for detailed data on Country of Birth of Mother. It is primarily recommended for collections that simply seek to determine whether or not a person's mother is a migrant.

CLASSIFICATION

The *Standard Australian Classification of Countries* (SACC) is used when collecting, aggregating and disseminating data relating to Country of Birth of Mother. The term 'country' is used in the SACC to describe fully independent countries (sovereign nation states), administrative subdivisions of the United Kingdom (i.e. England, Scotland, Wales and Northern Ireland), external territories and dependencies of independent countries, and recognised geographic areas of which the ownership or control is in dispute.

This classification has a three-level hierarchical structure. The third and most detailed level of the classification consists of 244 base or third-level units which are discrete countries (as defined above). Included in the 244 third-level units are five 'not elsewhere classified' (nec) categories, which contain entities that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics relating to country of birth, data should be captured, classified and stored at the base level of the SACC wherever possible.

COUNTRY OF BIRTH OF MOTHER

CLASSIFICATION *continued*

The second level of the classification comprises 27 minor groups, which are groups of neighbouring countries similar in terms of social, cultural, economic and political characteristics. Each minor group lies wholly within the boundaries of a geographic continent. On average, minor groups contain nine countries, with individual minor groups containing between one and 27 countries.

The first and most general level of the classification comprises nine major groups which are formed by aggregating geographically proximate minor groups and, therefore, comprise countries which are broadly similar in terms of social, cultural, economic and political characteristics.

For further details on the classification, see *Standard Australian Classification of Countries (SACC)* (Cat. no. 1269.0) or contact the ABS (see contact details on page 100).

CODING PROCEDURES

A coding index has been developed to assist in the implementation and use of the SACC and should be used when coding responses to the Country of Birth of Mother questions. The index contains a comprehensive list of the most probable responses to questions relating to countries and their correct classification codes. Use of the coding index enables responses to be coded accurately and quickly to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the SACC publication, or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

OUTPUT CATEGORIES

The hierarchical structure of the SACC allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the major group level, minor group level, or the base or country level. If necessary, significant countries within a minor group can be presented separately while the remaining countries within the minor group are aggregated. For example, the third-level category 'Nauru' can be separately identified as an output category within the 'Micronesia' Minor Group, with the remaining third-level categories in this minor group aggregated in an 'Other Micronesia' category. The same principle can be adopted to highlight significant minor groups within a major group. The Country of Birth of Person standard gives an example of standard output categories for the nine major groups of the SACC, which have either significant minor groups or countries highlighted (see page 24).

The SACC is intended to provide a single framework for classifying all statistical and administrative data by country for both population and economic statistics. Because the geographically based country groupings of the main structure are not always ideal for the presentation of data,

OUTPUT CATEGORIES
continued

alternative groupings have been included in the SACC publication to meet specific output needs. These country groupings are of two types:

- a set of standard country groupings based on economic and political requirements and comprising associations or organisations of formally recognised member countries (for example, the Association of South East Asian Nations (ASEAN)); and
- a set of authorised standard output options for time series purposes or other data output needs (for example, Former USSR).

Further information on the alternative standard country groupings can be found in the SACC publication.

The standard output categories for the Country of Birth of Mother Question for Minimum Data are:

Mother born in Australia
Mother born overseas

FIRST LANGUAGE SPOKEN

INTRODUCTION

First Language Spoken identifies the language an individual first speaks or masters. Data on First Language Spoken provide primary information about a person's cultural and linguistic background. It is regarded as a good surrogate measure of ethnicity because of its connection with a person's origins and the origins of his or her parents.

Persons whose first language is other than English have been identified by program providers as a population group that is likely to experience disadvantage when seeking to obtain equal access to government and community programs and services in Australia. Data relating to First Language Spoken is thus an indicator of advantage and potential disadvantage associated with language background or with other factors associated with cultural background.

First Language Spoken provides a good filter to Proficiency in Spoken English (see page 32) as it enables native speakers of English to be excluded from that question. It does, however, allow some people with a first language other than English to filter to the Proficiency in Spoken English question even though they may speak impeccable English.

DEFINITION

First Language Spoken is defined as the first language an individual masters during the language acquisition phase of intellectual development. This is generally the language spoken in the home by the people who have raised the individual from infancy. In practice, First Language Spoken is defined as the language the respondent identifies, or remembers, as being the first language which they could understand to the extent of being able to conduct a conversation.

ISSUES FOR CONSIDERATION

There are four important issues for consideration associated with First Language Spoken. The first concerns the way in which sign languages are dealt with. Individuals who use non-verbal forms of communication make up at least one per cent of the Australian population. Sign languages fall into two groups, signed English/finger spelling and other sign languages such as Auslan. Signed English/finger spelling is considered to be another form of English and is treated as such. Auslan is regarded as a distinct language and individuals who report their first language as Auslan (or another recognised sign language) are considered to have a first language other than English.

The second issue concerns the number of languages accepted as a First Language Spoken by an individual. While there is some interest in the identification of respondents who claim to have spoken two languages as first languages for those studying changing language patterns, cultural affiliation or expected language proficiency, research suggests that dual first language speakers in Australia are likely to be rare and data produced from such a question may not be statistically viable. Therefore, the questions used for this variable are not designed to identify dual first language speakers. ABS testing and experience in a number of surveys has shown that this does not present a problem in data collection.

ISSUES FOR CONSIDERATION *continued*

It is important to note that First Language Spoken provides a good measure of current language use (ABS data show that more than 95% of Australians whose first language is other than English, are still able to use their first language). However, in some instances (depending on age, year of arrival in Australia, and living arrangements), a person's first language spoken will not necessarily be the person's language of greatest competence or the main language a person speaks. As such, First Language Spoken (like Main Language Other Than English Spoken at Home, see page 26) could be considered to overcount the number of people in the 'Other Than English' category.

Finally, it should be noted that there are a number of different names and definitions associated with the concept of first language. Terms such as 'Native Language' and 'Mother Tongue' are widely used internationally. However 'First Language Spoken' is more appropriate within the Australian context. 'First Language' is the term used in Australian government policy documents and is widely referred to in linguistic journals. The term is unambiguous, concise and clearly describes the concept. It is preferred to alternative word orderings (e.g. 'Language First Spoken') primarily because its meaning is clearer, and because it has the additional advantage of being more easily abbreviated to 'First Language' in table headings, etc.

STANDARD QUESTIONS

There are two standard questions for First Language Spoken. The Question for Detailed Data elicits detailed data on the specific first language spoken by the person. There are two alternatives for this question which may be used depending on space and cost considerations. The first provides a set of tick boxes and an 'Other—please specify' response category for writing a language not specified in the tick box list. The second has a tick box for 'English' and an 'Other—please specify' write-in category. The second alternative is more compact but will incur additional coding costs. The Question for Minimum Data only contains the tick box options 'English' and 'Other' and therefore does not collect details as to which language other than English was first spoken.

FIRST LANGUAGE SPOKEN

Question for Detailed Data

The standard question for obtaining detailed data for First Language Spoken is as follows:

Which language [did you] [did the person] [did (name)] *first* speak as a child?
(Mark only one box.)

- English ☐
- Italian ☐
- Greek ☐
- Cantonese ☐
- Mandarin ☐
- Arabic ☐
- Vietnamese ☐
- German ☐
- Spanish ☐
- Tagalog (Filipino) ☐
- Other—please specify: ☐

For interview-based collections, interviewers should be instructed to mark 'English' for respondents who identify English and another language as their first language spoken. If two (or more) languages are identified by the respondent, and none of them are English, interviewers should be instructed to prompt again for the predominant first language spoken. If more than one answer is still provided, they should mark the first language mentioned as the respondent's first language spoken. For self-enumerated collections, respondents should be instructed to mark one box only.

The languages provided with this question have been included primarily on the basis of their statistical significance in the Australian context. These languages accounted for approximately 90% of all responses to the Main Language Other Than English Spoken at Home question in the 1996 Census of Population and Housing. The list will be reviewed periodically as different languages assume a greater or lesser significance in the Australian context as a result of changing migration patterns.

Cantonese and Mandarin are both included in the list to ensure that the response 'Chinese', for example, is not erroneously coded to one or the other. The 'Other' category is included for those people whose first language spoken is not offered in the list as a response to the question. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in their language response.

Question for Detailed Data *continued*

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list. However, if there is a need for detailed information but space constraints are an overriding consideration, the full tick box list can be replaced by a tick box for English and the 'Other—please specify' write-in category, as follows:

Which language [did you] [did the person] [did (name)] *first* speak as a child?
(Mark only one box.)

English..... ☐
Other—please specify: ☐

All respondents who first spoke a language other than English would be required to write in their response. This would result in significant additional coding costs compared to the use of the full tick box list, as there is likely to be a wide range of responses to the 'Other—please specify' write-in category which would have to be coded manually using a coding index (see *Coding Procedures*). These costs could be reduced by tailoring the choice and number of tick boxes included with the 'Other—please specify' write-in category based on the composition of the population to be covered by the collection, as determined by ABS census data. A tailored tick box list would capture (and self-code) the majority of responses reducing the amount of manual coding required.

Question for Minimum Data

The standard question for obtaining minimum data for First Language Spoken is as follows:

Which language [did you] [did the person] [did (name)] *first* speak as a child?
(Mark only one box.)

English..... ☐
Other..... ☐

This question is suitable for statistical and administrative collections where there is no requirement for detailed language data. It is primarily recommended for collections that simply seek to determine whether or not a person first spoke a language other than English.

FIRST LANGUAGE SPOKEN

CLASSIFICATION

The *Australian Standard Classification of Languages* (ASCL) is used when collecting, aggregating and disseminating data on language usage in Australia. The term 'language' is used in the ASCL to describe languages, dialects, pidgins, creoles and invented and sign languages.

The classification has a three-level hierarchical structure. The third and most detailed level consists of 193 base or third-level units which are separately identified languages (as defined above). Included in the 193 third-level units are 26 'not elsewhere classified' (nec) and three 'other' categories, which contain entities that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics, data should be captured, classified and stored at the base level of the ASCL wherever possible.

The second level of the classification comprises 48 narrow groups of languages which are similar in terms of their evolution from a common ancestral language (genetic affinity) and geographic proximity in terms of the areas in which they originated. Included in the 48 narrow groups are five 'other' categories which consist of languages which do not fit into a particular narrow group.

The first and most general level of the classification comprises nine broad groups of languages, including one 'other' category. Broad groups are formed by aggregating geographically proximate narrow groups.

For further details on the classification, see *Australian Standard Classification of Languages (ASCL)* (Cat. no. 1267.0) or contact the ABS (see contact details on page 100).

CODING PROCEDURES

Language data are coded to the classification using the ASCL coding index. This contains a comprehensive list of the most probable responses to questions relating to language and their correct classification codes. Use of the coding index enables responses to be coded accurately and quickly to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the ASCL publication or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

OUTPUT CATEGORIES

The hierarchical structure of the ASCL allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the broad group level, narrow group level, or the base or language level. If necessary, significant languages within a narrow group can be presented separately while the remaining languages within the narrow group are aggregated. For example, the third-level category 'Tiwi' can be separately identified as an output category within the 'Northern Aboriginal' Narrow Group, with the remaining third-level categories in this narrow group aggregated to form an 'Other Northern Aboriginal' category. The same principle can be adopted to highlight significant narrow groups or languages within a

OUTPUT CATEGORIES
continued

broad group. The Main Language Other Than English Spoken at Home standard gives an example of standard output categories for the nine broad groups of the ASCL, which have either significant narrow groups or languages highlighted (see page 30).

The standard output categories for the First Language Spoken Question for Minimum Data are:

English
Other language

LANGUAGES SPOKEN AT HOME

INTRODUCTION

Languages Spoken at Home identifies all languages spoken within the home. This variable provides data on the stock of languages used in Australian homes and is required to provide a full statistical picture of the diversity of home language usage in Australia.

Data on languages spoken in the home are regarded as an indicator of active ethnicity and are useful for the study of inter-generational language retention. The availability of such data may also help providers of language, welfare and community services to effectively target the population groups that need those services.

DEFINITION

Languages Spoken at Home is defined as the language or languages spoken by a person in his or her home, on a regular basis, to communicate with other residents of the home and regular visitors to the home. There is no restriction on the number of languages reported by the respondent as being spoken in the home.

ISSUES FOR CONSIDERATION

Related language variables such as Main Language Other Than English Spoken at Home (see page 26), First Language Spoken (see page 54) and Main Language Spoken at Home (see page 64) are most commonly used in collections to measure aspects of Australian language usage. However, these variables only collect one language response which may not reflect complete language use. It is considered that Languages Spoken at Home should be used in preference to other language variables when the aim is to collect data on the stock of languages used in the home. In some cases, however, this measure may not reflect complete language use when, for example, only one language is spoken in the home but other languages are spoken outside the home, within a person's ethnic community group. It also does not indicate *how often* languages are used in a particular home and therefore may capture languages that are used infrequently or spoken with limited proficiency.

The way in which sign languages are treated with regard to Languages Spoken at Home should also be noted. Sign languages fall into two groups, signed English/finger spelling and other sign languages such as Auslan. Signed English/finger spelling is considered to be another form of English and is treated as such. Auslan is regarded as a distinct language and individuals who report Auslan (or another recognised sign language) are considered to speak a language other than English.

STANDARD QUESTION

The standard question for Languages Spoken at Home is as follows:

Which language or languages [do you] [does the person] [does (name)] speak *at home*?
(Please indicate all languages spoken.)

English ☐
Italian ☐
Greek ☐
Cantonese ☐
Mandarin ☐
Arabic ☐
Vietnamese ☐
German ☐
Spanish ☐
Tagalog (Filipino) ☐
Other—please specify: _____

The languages provided with this question have been included primarily on the basis of their statistical significance in the Australian context. These languages accounted for approximately 90% of all responses to the Main Language Other Than English Spoken at Home question in the 1996 Census of Population and Housing. The list will be reviewed periodically as different languages assume a greater or lesser significance in the Australian context as a result of changing migration patterns.

Cantonese and Mandarin are both included in the list to ensure that the response 'Chinese', for example, is not erroneously coded to one or the other. The 'Other' category is included for those people who speak a language in the home that is not offered as a response to the question. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in their language response.

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list. However, if space constraints are an overriding consideration, the tick boxes can be deleted and space provided for respondents to write in their language responses. This would result in significant additional coding costs compared to the use of the full tick box list, as there is likely to be a wide range of written responses which would have to be coded manually using a coding index (see *Coding Procedures*). The question would be as follows:

LANGUAGES SPOKEN AT HOME

STANDARD QUESTION
continued

Which language or languages [do you] [does the person] [does (name)] speak *at home*?
(Please indicate all languages spoken.)

CLASSIFICATION

The *Australian Standard Classification of Languages* (ASCL) is used when collecting, aggregating and disseminating data on language usage in Australia. The term 'language' is used in the ASCL to describe languages, dialects, pidgins, creoles and invented and sign languages.

The classification has a three-level hierarchical structure. The third and most detailed level consists of 193 base or third-level units which are separately identified languages (as defined above). Included in the 193 third-level units are 26 'not elsewhere classified' (nec) and three 'other' categories, which contain entities that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics, data should be captured, classified and stored at the base level of the ASCL wherever possible.

The second level of the classification comprises 48 narrow groups of languages which are similar in terms of their evolution from a common ancestral language (genetic affinity) and geographic proximity in terms of the areas in which they originated. Included in the 48 narrow groups are five 'other' categories which consist of languages which do not fit into a particular narrow group.

The first and most general level of the classification comprises nine broad groups of languages, including one 'other' category. Broad groups are formed by aggregating geographically proximate narrow groups.

For further details on the classification, see *Australian Standard Classification of Languages (ASCL)* (Cat. no. 1267.0) or contact the ABS (see contact details on page 100).

CODING PROCEDURES

Language data are coded to the classification using the ASCL coding index. This contains a comprehensive list of the most probable responses to questions relating to language and their correct classification codes. Use of the coding index enables responses to be coded accurately and quickly to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the ASCL publication or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

OUTPUT CATEGORIES

The hierarchical structure of the ASCL allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the broad group level, narrow group level, or the base or language level. If necessary, significant languages within a narrow group can be presented separately while the remaining languages within the narrow group are aggregated. For example, the third-level category 'Tiwi' can be separately identified as an output category within the 'Northern Aboriginal' Narrow Group, with the remaining third-level categories in this narrow group aggregated to form an 'Other Northern Aboriginal' category. The same principle can be adopted to highlight significant narrow groups or languages within a broad group. The Main Language Other Than English Spoken at Home standard gives an example of standard output categories for the nine broad groups of the ASCL, which have either significant narrow groups or languages highlighted (see page 30).

MAIN LANGUAGE SPOKEN AT HOME

INTRODUCTION

Main Language Spoken at Home identifies the language most frequently used by a person at home. It provides a good indicator of the language in which an individual is likely to be most at ease. Data on Main Language Spoken at Home are regarded as an indicator of active ethnicity and are useful for the study of inter-generational language retention. The availability of such data may help providers of language, welfare and community services to effectively target population groups that need those services.

Main Language Spoken at Home may be used in preference to other language variables when the aim is to measure active ethnicity or the likely disadvantage experienced by persons whose usual language is not English.

DEFINITION

Main Language Spoken at Home is defined as the main language spoken by a person in his or her home, on a regular basis, to communicate with other residents of the home and regular visitors to the home. If a person reports that he or she speaks more than one language at home, they are asked to report the language spoken most often.

ISSUES FOR CONSIDERATION

Main Language Spoken at Home tends to underestimate current community language usage amongst the longer standing migrant groups who now mainly use English at home. In some instances it does not provide information about a person's origins but rather information about an aspect of their living arrangements (i.e. the single language most frequently used in the household in which they live).

The way in which sign languages are treated with regard to Main Language Spoken at Home should also be noted. Sign languages fall into two groups, signed English/finger spelling and other sign languages such as Auslan. Signed English/finger spelling is considered to be another form of English and is treated as such. Auslan is regarded as a distinct language and individuals who report their main language as Auslan (or another recognised sign language) are considered to speak a language other than English.

STANDARD QUESTIONS

There are two standard questions for Main Language Spoken at Home. The Question for Detailed Data elicits detailed data on the specific main language a person speaks in the home. There are two alternatives for this question which may be used depending on space and cost considerations. The first provides a set of tick boxes and an 'Other—please specify' response category for writing a language not specified in the tick box list. The second has a tick box for 'English' and an 'Other—please specify' write-in category. The second alternative is more compact but will incur additional coding costs. The Question for Minimum Data only contains the tick box options 'English' and 'Other' and therefore does not collect details as to which language other than English is spoken.

Question for Detailed Data

The standard question for obtaining detailed data for Main Language Spoken at Home is as follows:

Which language [do you] [does the person] [does (name)] mainly speak *at home?*
(If more than one language, indicate the one that is spoken most often.)

English ☐
Italian ☐
Greek ☐
Cantonese ☐
Mandarin ☐
Arabic ☐
Vietnamese ☐
German ☐
Spanish ☐
Tagalog (Filipino) ☐
Other—please specify: ☐

Instructions are included for those respondents who speak more than one language in the home. For self-enumerated collections, respondents are instructed to indicate the language spoken most often and to mark *one* box in the checklist or to write in *one* response where tick boxes do not apply. For interview-based collections, an interviewer could use a prompt like 'Which of those languages [do you] [does the person] [does (name)] use most often?' if a multiple response is given after asking the question.

The languages provided with this question have been included primarily on the basis of their statistical significance in the Australian context. These languages accounted for approximately 90% of all responses to the Main Language Other Than English Spoken at Home question in the 1996 Census of Population and Housing. The list will be reviewed periodically as different languages assume a greater or lesser significance in the Australian context as a result of changing migration patterns.

Cantonese and Mandarin are both included in the list to ensure that the response 'Chinese', for example, is not erroneously coded to one or the other. The 'Other' category is included for those people who mainly speak a language in the home that is not offered as a response to the question. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in their language response.

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list.

MAIN LANGUAGE SPOKEN AT HOME

Question for Detailed Data *continued*

However, if there is a need for detailed information but space constraints are an overriding consideration, the full tick box list can be replaced by a tick box for English and the 'Other—please specify' write-in category, as follows:

Which language [do you] [does the person] [does (name)] mainly speak *at home*?
(If more than one language, indicate the one that is spoken most often.)

English..... ☐
Other—please specify: ☐

All respondents who mainly speak a language other than English at home would be required to write in their response. This would result in significant additional coding costs compared to the use of the full tick box list, as there is likely to be a wide range of responses to the 'Other—please specify' write-in category which would have to be coded manually using the coding index (see *Coding Procedures*). These costs could be reduced by tailoring the choice and number of tick boxes included with the 'Other—please specify' write-in category based on the composition of the population to be covered by the collection, as determined by ABS census data. A tailored tick box list would capture (and self-code) the majority of responses reducing the amount of manual coding required.

Question for Minimum Data

The standard question for obtaining minimum data for Main Language Spoken at Home is as follows:

Which language [do you] [does the person] [does (name)] mainly speak *at home*?
(If more than one language, indicate the one that is spoken most often.)

English..... ☐
Other..... ☐

This question is suitable for statistical and administrative collections where there is no requirement for detailed language data. It is primarily recommended for collections that simply seek to determine whether or not the main language a person speaks at home is English.

CLASSIFICATION

The *Australian Standard Classification of Languages* (ASCL) is used when collecting, aggregating and disseminating data on language usage in Australia. The term 'language' is used in the ASCL to describe languages, dialects, pidgins, creoles and invented and sign languages.

The classification has a three-level hierarchical structure. The third and most detailed level consists of 193 base or third-level units which are separately identified languages (as defined above). Included in the 193 third-level units are 26 'not elsewhere classified' (nec) and three

CLASSIFICATION *continued*

'other' categories, which contain entities that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics, data should be captured, classified and stored at the base level of the ASCL wherever possible.

The second level of the classification comprises 48 narrow groups of languages which are similar in terms of their evolution from a common ancestral language (genetic affinity) and geographic proximity in terms of the areas in which they originated. Included in the 48 narrow groups are five 'other' categories which consist of languages which do not fit into a particular narrow group.

The first and most general level of the classification comprises nine broad groups of languages, including one 'other' category. Broad groups are formed by aggregating geographically proximate narrow groups.

For further details on the classification, see *Australian Standard Classification of Languages (ASCL)* (Cat. no. 1267.0) or contact the ABS (see contact details on page 100).

CODING PROCEDURES

Language data are coded to the classification using the ASCL coding index. This contains a comprehensive list of the most probable responses to questions relating to language and their correct classification codes. Use of the coding index enables responses to be coded accurately and quickly to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the ASCL publication or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

OUTPUT CATEGORIES

The hierarchical structure of the ASCL allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the broad group level, narrow group level, or the base or language level. If necessary, significant languages within a narrow group can be presented separately while the remaining languages within the narrow group are aggregated. For example, the third-level category 'Tiwi' can be separately identified as an output category within the 'Northern Aboriginal' Narrow Group, with the remaining third-level categories in this narrow group aggregated to form an 'Other Northern Aboriginal' category. The same principle can be adopted to highlight significant narrow groups or languages within a broad group. The Main Language Other Than English Spoken at Home standard gives an example of standard output categories for the nine broad groups of the ASCL, which have either significant narrow groups or languages highlighted (see page 30).

The standard output categories for the Main Language Spoken at Home Question for Minimum Data are:

English
Other language

RELIGIOUS AFFILIATION

INTRODUCTION

Religious Affiliation identifies the religious belief to which a person adheres or the religious group to which they belong. This variable is used to assess the composition of the Australian community in terms of religion.

Data relating to Religious Affiliation is useful to religious organisations in planning their services to members, and in social research as it can aid in indicating a person's cultural background when used in conjunction with other cultural and language diversity variables. Some agencies and organisations have also found these data helpful in delivering more culturally relevant services to clients based on their religion.

DEFINITION

Religious Affiliation is defined as the religious beliefs and practices to which a person adheres or the religious group to which a person belongs. In practice, Religious Affiliation is self-assessed and measures the religion to which a person declares they have an affiliation. The variable also identifies people who consider they do not belong to any religious group.

ISSUES FOR CONSIDERATION

As Religious Affiliation is self-assessed, the data cannot give an objective measure of the strength of a person's affiliation with a religion and is subject to the respondent's interpretation of the question.

Answering a question about religious affiliation is optional in ABS censuses and surveys. Section 14 (2) of the *Census and Statistics Act 1905* which addresses the issue of refusal or failure to answer questions states that this 'does not apply in relation to a refusal or failure by a person to answer a question, or supply particulars, relating to the person's religious beliefs.'

This variable has previously been termed 'Religious Denomination'. The term 'Religious Affiliation' is considered more appropriate since the term 'denomination' is sometimes regarded as applying particularly to Christian groups.

STANDARD QUESTION

The standard question for Religious Affiliation is as follows:

What is [your] [the person's] [(name)'s] religion?
(Answering this question is OPTIONAL.)
(For example: Salvation Army, Hinduism, Judaism or Humanism.)
(If no religion, mark last box.)

- Catholic (not Eastern Churches) ☐
- Anglican (Church of England) ☐
- Uniting Church ☐
- Presbyterian ☐
- Greek Orthodox ☐
- Baptist ☐
- Lutheran ☐
- Islam ☐
- Buddhism ☐
- Other—please specify: ☐
- No religion ☐

The standard question is suitable for both self-enumerated and interview-based collections without requiring any variation of question wording or instructions. It may help the respondent to answer the question if some examples of possible responses are included. It is also important to instruct the respondent that if they do not have a religion, they are still asked to answer the question.

The list of religions included with the question is based on the statistical significance of adherents in these groups as counted in the 1996 Census of Population and Housing. It implies no view by the ABS of the relative importance of any religious group or groups and will be reviewed periodically as different religions assume a greater or lesser statistical significance in the Australian context.

The qualifier 'not Eastern Churches' is added to the 'Catholic' response category in the tick box list to distinguish between Roman (or Western) Catholic churches and the Eastern Catholic churches (e.g. Maronite Catholic, Melkite Catholic and Ukrainian Catholic churches). The 'Other' category is included for those people who adhere or belong to a religion not listed. 'Please specify' is added to the 'Other' category and a space is provided for respondents to write in their religion.

Use of the full tick box list with the question ensures compatibility across collections and minimises coding costs. Tick box responses are 'self-coded' which reduces the amount of manual or office coding to be performed as the majority of responses are captured by the list.

RELIGIOUS AFFILIATION

STANDARD QUESTION
continued

However, if space constraints are an overriding consideration, the full tick box list can be replaced by a space for respondents to write in their religion and a tick box for 'No religion', as follows:

What is [your] [the person's] [(name)'s] religion?
(Answering this question is OPTIONAL.)
(For example: Judaism, Humanism, Islam, Greek Orthodox, Baptist, Church of England.)
(Please write in your religion or mark the box if no religion.)

No religion ..... ☐

This alternative would result in significant additional coding costs compared to the use of the full tick box list, as there is likely to be a wide range of religion responses which would have to be coded manually using a coding index (see *Coding Procedures*). These costs could be reduced by tailoring the choice and number of tick boxes to be included with the question, along with an 'Other—please specify' write-in category, based on the composition of the population to be covered by the collection as determined by ABS census data. A tailored tick box list would capture (and self-code) the majority of responses reducing the amount of manual coding required.

CLASSIFICATION

The *Australian Standard Classification of Religious Groups* (ASCRG) is used when collecting, aggregating and disseminating data relating to Religious Affiliation. The term 'religious group' is used in the ASCRG to describe groups of religions, separately identified religions, and subsets of religions such as religious denominations, administrative and organisational groupings, groups of churches, individual churches, and breakaway groups. They are described in the classification as religious groups in that each is comprised of a group of people who share common religious beliefs and practices, or belong to organisations that are unified by a common religious theme.

The classification has a three-level hierarchical structure. The most detailed level of the classification consists of 107 base or third-level units which are religious groups (as defined above). Included in the 107 third-level units are nine 'not elsewhere classified' (nec) categories which contain entities that are not listed separately in the classification. Each third-level category has a four digit code which is used to code responses (see *Coding Procedures*). Regardless of the level of aggregation envisaged for the dissemination of statistics relating to religious groups, data should be captured, classified and stored at the base level of the ASCRG wherever possible.

The third-level categories of the classification are aggregated on the basis of similarity of religious beliefs and religious practices, and/or cultural heritage, to form the second-level categories of the classification. The second level of the classification comprises 33 narrow groups of religions.

CLASSIFICATION *continued*

Included in the 33 narrow groups are two 'other' categories which consist of religious groups that do not fit into a particular narrow group.

The first and most general level of the classification comprises seven broad groups of religions, consisting of five of the world's main religions, one 'Other Religions' category and one 'No Religion' category, as follows:

1 Buddhism
2 Christianity
3 Hinduism
4 Islam
5 Judaism
6 Other Religions
7 No Religion

For further details on the classification, see *Australian Standard Classification of Religious Groups (ASCRG)* (Cat. no. 1266.0) or contact the ABS (see contact details on page 100).

CODING PROCEDURES

Religious Affiliation data are coded to the classification using the ASCRG coding index. The coding index contains a comprehensive list of religious beliefs, groups, denominations, organisations and churches in Australia. Each response to the question is matched to an index entry, which gives the four digit code used to assign the response to the appropriate category of the classification. Supplementary codes also exist to enable coding of inadequately described responses. Copies of the coding index can be found in the ASCRG publication or can be obtained in electronic form by contacting the ABS (see contact details on page 100).

OUTPUT CATEGORIES

The hierarchical structure of the ASCRG allows users the flexibility to output statistics at the level of the classification which best suits their requirements. Data can be presented at the broad group level, narrow group level or the base religious group level. If necessary, significant religious groups within a narrow group can be presented separately while the remaining religious groups within the narrow group are aggregated. For example, the third-level category 'Albanian Orthodox' can be separately identified as an output category within the 'Orthodox' Narrow Group, with the remaining third-level categories in this narrow group aggregated in an 'Other Orthodox' category. The same principle can be adopted to highlight significant narrow groups within a broad group.

RELIGIOUS AFFILIATION

OUTPUT CATEGORIES
continued

The following is an example of standard output categories for the seven broad groups of the ASCRG, which have either significant narrow groups (second-level categories) or religious groups (third-level categories) highlighted.

- Buddhism
- Christianity
  - Anglican
  - Baptist
  - Catholic
    - Western Catholic
    - Other Catholic
  - Lutheran
  - Orthodox
    - Greek Orthodox
    - Macedonian Orthodox
    - Russian Orthodox
    - Serbian Orthodox
    - Other Orthodox
  - Presbyterian and Reformed
  - Uniting Church
  - Pentecostal
  - Other Christianity
- Hinduism
- Islam
- Judaism
- Other Religions
- No Religion
  - Atheism
  - Other No Religion

YEAR OF ARRIVAL IN AUSTRALIA

INTRODUCTION

Year of Arrival in Australia collects the year in which a person, born in another country, first arrived in Australia with the intention of living here for one year or more. The year in which an overseas-born person arrived in Australia can be used to derive the variable Period of Residence in Australia, which is included as part of this standard.

There is substantial interest in people's Period of Residence in Australia after migrating here from other countries. The length of time migrants have been in Australia can give an indication of how familiar they are with Australian society and practices. When cross-classified with other data, Period of Residence in Australia is also useful for analysing how the characteristics of migrants change with the length of time they have been here. For example, data collected by Year of Arrival in Australia can be used with data on Country of Birth of Person (see page 20) and Proficiency in Spoken English (see page 32) to examine patterns of settlement difficulties, resulting from inadequate English language skills, for various migrant groups. These data can be used to analyse the speed at which different groups overcome these difficulties.

Year of Arrival in Australia is closely associated with the variable Country of Birth of Person, as the Year of Arrival in Australia question should only be asked of those people who were born in a country other than Australia.

DEFINITIONS

Year of Arrival in Australia is defined as the year in which a person, born outside of Australia, first arrived in Australia from another country with the intention of living here for one year or more. In practice, Year of Arrival in Australia collects the calendar year (e.g. '1985') of first arrival in Australia.

Period of Residence in Australia is the length of time a person born outside Australia has spent living in Australia, calculated as the difference between the Year of Arrival in Australia and the year in which the data were collected. It is derived from Year of Arrival in Australia by subtracting the year of first arrival from the year in which the data were collected, and expressing the result in completed years. To improve the accuracy of this result, it is necessary to subtract any significant time periods spent living outside of Australia after the year of first arrival. Significant periods are absences of one year or more, excluding holidays, visits and business trips overseas. In the case of multiple absences from Australia, the total number of significant years (rounded to the nearest whole number) spent outside of Australia is subtracted from the difference between the year of collection and the year of first arrival.

YEAR OF ARRIVAL IN AUSTRALIA

ISSUES FOR CONSIDERATION

The main issue associated with Year of Arrival in Australia is that the question should be asked only of those people born in a country other than Australia. For this, Year of Arrival in Australia requires the supporting variable Country of Birth of Person. The standard procedure when collecting Year of Arrival in Australia is to ask the Country of Birth of Person question first, and sequence only those people born in a country other than Australia to the Year of Arrival in Australia question.

Another issue to consider with Year of Arrival in Australia is that it measures the year of *first* arrival in Australia and not the year of most recent arrival. Significant periods spent outside Australia after first arrival need to be accounted for when deriving Period of Residence in Australia. Significant periods of absence from Australia of one year or more, excluding holidays, visits or business trips, can be collected in a separate question after the Year of Arrival in Australia question (see *Standard Question*). This second question is optional in this standard but is necessary to accurately measure Period of Residence in Australia.

STANDARD QUESTION

The standard question for Year of Arrival in Australia is as follows:

In what year did [you] [the person] [(name)] first arrive in Australia to live here for one year or more?
(Write in the calendar year of arrival or mark the box if here less than one year.)

☐☐☐☐

Will be here less than one year..... ☐

The Year of Arrival in Australia question is suitable for either self-enumerated or interview-based collections. It is anticipated that for the majority of people, their response to the question will be the year of their *only* arrival in Australia. However, some respondents may have multiple arrivals in Australia. To deal with these cases in self-enumerated collections, an instruction such as 'Please indicate the year of *first* arrival only' should be included with the question. For interview-based collections, clear directions should be available to the interviewer so that the respondent can be advised that the year of first arrival is to be reported if there is any doubt.

STANDARD QUESTION
continued

The second (optional) question included to improve the accuracy of Period of Residence in Australia data by collecting all significant absences since first arrival is as follows:

Excluding holidays, visits or business trips overseas, [have you] [has the person] [has (name)] had any periods of living overseas for one year or more since first arriving here?
(If 'Yes', please list each absence from Australia of one year or more and its duration in years and months.)

No..... ☐
Yes, first absence: ___yrs ___mths
Yes, second absence: ___yrs ___mths
Yes, third absence: ___yrs ___mths
etc.

This question is asked directly after the Year of Arrival in Australia question. In practice, it may be necessary for the respondent to approximate the time spent outside Australia.

CLASSIFICATION

Year of Arrival in Australia data are collected and stored as four digit calendar years (e.g. a response of '1985' is coded as '1985' in the classification).

Period of Residence in Australia is calculated as two digit categories '00' to '99', representing the number of years spent living in Australia. Period of residence data are rounded to the appropriate category in terms of completed years. For example, a period of residence in Australia of seven years and ten months would be assigned code '07'.

As special codes are set aside for 'Born in Australia' and 'Not stated/ Inadequately described' (see *Coding Procedures*), the classifications for Year of Arrival in Australia and Period of Residence in Australia are applicable to all persons in Australia. However, it should be noted that this is a coding convention and that, setting aside these special codes, the scope of the classifications, along with the variables, is the overseas-born population.

CODING PROCEDURES

The Year of Arrival in Australia classification stores the four digit calendar year response to the question (i.e. the classification has a four digit code structure which directly relates to calendar years). Special codes are reserved for those people born in Australia (code '0000') and for not stated and inadequately described responses (code '9999').

YEAR OF ARRIVAL IN AUSTRALIA

CODING PROCEDURES
continued

The two digit code structure for the Period of Residence in Australia classification is the set {00, 01, 02, ... , 99}, which relates to the two digit period of residence data expressed in completed years. Code '00' represents the response where the year of arrival is the same as the year of collection. Code '01' represents a one year period of residence for the respondent, and so on up to code '97' which represents a period of residence of 97 years *or more.* Code '98' is reserved for 'Born in Australia' and code '99' for 'Not stated/Inadequately described'. If a value of '–1' is derived after subtracting all significant absences from Period of Residence in Australia data for a person who has had multiple arrivals, then code '00' (i.e. zero years of residence) should be substituted.

OUTPUT CATEGORIES

The standard structure of output categories for Year of Arrival in Australia depends on the year of collection. The basic output structure consists of the year of collection, the previous four single years, a time period covering the years between the last single year listed and the previous census year (e.g. 1991–1994), two standard five-year intercensal time periods, and an open-ended time period (e.g. 'Arrived before 1981'). Output categories for 'Born in Australia' and 'Not stated/Inadequately described' are also required. For example, the standard output categories for data collected in 1999 would be:

Arrived before 1981
Arrived 1981–1985
Arrived 1986–1990
Arrived 1991–1994
Arrived 1995
Arrived 1996
Arrived 1997
Arrived 1998
Arrived 1999
Born in Australia
Not stated/Inadequately described

These output categories can be adjusted for a particular collection year. For example, if a collection takes place in 2001, the output categories would be:

Arrived before 1986
Arrived 1986–1990
Arrived 1991–1995
Arrived 1996
Arrived 1997
Arrived 1998
Arrived 1999
Arrived 2000
Arrived 2001
Born in Australia
Not stated/Inadequately described

OUTPUT CATEGORIES
continued

The standard structure of output categories for Period of Residence in Australia data is simpler than for Year of Arrival in Australia as it does not vary with the year of collection. It aggregates the single year categories into five-year groupings up to 20 years of residence. For a collection in any year, the following output categories should be standard:

0–4 years of residence
5–9 years of residence
10–14 years of residence
15–19 years of residence
20 or more years of residence
Born in Australia
Not stated/Inadequately described

Since the focus of analytical interest will frequently be on persons who have a short period of residence, the following presentation of output may be used.

0–4 years of residence
  Less than 1 year of residence
  1 year of residence
  2 years of residence
  3 years of residence
  4 years of residence
5–9 years of residence
10–14 years of residence
15–19 years of residence
20 or more years of residence
Born in Australia
Not stated/Inadequately described

CHAPTER 4 QUESTION MODULES

A question module is a set of questions, with response categories and associated sequence guides, designed to collect data for the measurement of a particular variable or group of related variables. More than one question module may be developed for a topic (such as cultural and language diversity) and each of the modules is designed for a specific purpose.

When designing the question module, there are several factors which need to be taken into consideration. These include:

- the collection methodology (how the questionnaire will be administered);
- analytical requirements (the information you want to obtain from the data);
- time, space and cost constraints; and
- provider or respondent load.

Two question modules have been developed for the measurement of cultural and language diversity.

- The Question Module for the Minimum Core Set of Cultural and Language Indicators comprises four questions. This is the minimum question module recommended for obtaining data on cultural or language diversity.
- The Question Module for the Standard Set of Cultural and Language Indicators contains all of the standard questions which can be used to measure cultural and language diversity.

It should be noted that it is not necessary that *all* of the full set of questions be used in any one questionnaire. Rather a combination of variables from the standard set, based on the factors that need to be taken into consideration which are itemised above, can be added to the Minimum Core Set of questions to create a tailored question module.

The content of the question module chosen depends on how the data will be used, the collection methodology, and competing claims for interviewer time and space on the questionnaire. As far as possible, conceptual and operational consistency is maintained across all of the modules in order to maximise comparability. Inevitably, however, data collected using different question modules will have some limitations in terms of comparability due to differences in the degree of precision and the level of detail collected.

This chapter presents the question modules for the Minimum Core Set and the Standard Set of Cultural and Language Indicators. An example of a tailored set of questions, and its sequencing order, is also given.

COLLECTION METHODOLOGY

Collections can be administered using several different methodologies. The questions asked and the way in which they are worded depends on the collection methodology chosen. Australian Bureau of Statistics (ABS) household collections generally use one of the following methodologies:

- Personal interview (PI), where a trained interviewer personally interviews each person selected;
- Any responsible adult (ARA), where a trained interviewer asks one person in the household (a usual resident aged 18 or over) for information about all persons in the household (Both PI and ARA surveys can be conducted face-to-face or by telephone); or
- Self-enumeration, where questionnaires left at the household may be completed either by each person or by one person on behalf of all household members.

Following is a guide to the relative efficiency of each methodology.

Personal Interview (PI)

The PI methodology enables collection of very complex and detailed information, offering great scope and flexibility in manipulating, compiling and analysing the survey data. It usually achieves high response rates and high quality data. However, the costs involved in employing, training and managing interviewers makes this method of data collection expensive. Also, for more sensitive topics, respondents might not be inclined to reveal private information using this method.

Any Responsible Adult (ARA)

The ARA methodology is used when time and cost constraints prevent the use of personal interview. The ABS Monthly Population Survey, which administers the Labour Force Survey and one or more Supplementary Surveys to 35,000 households each month, uses the ARA methodology. A limitation of this methodology for collecting data on cultural and language diversity is the subjectivity of responses. Certain questions rely on the respondent's opinion, or the respondent may not be aware of some of the information required for variables such as Country of Birth of Father, Country of Birth of Mother or First Language Spoken. This could result in larger non-response rates or guessing on behalf of the person who is answering the questions. Cultural diversity data collected by this method may therefore be less precise, less detailed and may support fewer analytical applications than data collected by personal interview.

Self-Enumeration

Self-enumeration questionnaires such as those used in the Census of Population and Housing must be simple and self-explanatory. The complex sequencing and detailed questions often used in interviewer-administered questionnaires are not feasible in self-enumerated collections. The data quality is dependent on the respondent reading and understanding all of the instructions without the prompts and assistance that an interviewer can provide. This is the least expensive method of data collection, although information obtained from such an enumeration strategy is less precise.

DETAILED AND MINIMUM DATA QUESTIONS

As noted in Chapter 3, most cultural and language variables contain two question options: at least one detailed data question and a minimum data question. The question used will depend upon the type of information required from each question. Detailed questions elicit more complete information.

There are usually two forms of detailed question: one that contains a tick box list and an 'Other—please specify' response category, and a short form that allows the respondent to write in their response. This second form of detailed question has higher coding costs than the tick box option. For example, the full detailed data question for Main Language Other Than English Spoken at Home is:

[Do you] [Does the person] [Does (name)] speak a language other than English *at home?*
(If more than one language, indicate the one that is spoken most often.)

- No, English only ☐
- Yes, Italian ☐
- Yes, Greek ☐
- Yes, Cantonese ☐
- Yes, Mandarin ☐
- Yes, Arabic ☐
- Yes, Vietnamese ☐
- Yes, German ☐
- Yes, Spanish ☐
- Yes, Tagalog (Filipino) ☐
- Yes, other—please specify: _____

The reduced form of this question for detailed data is:

[Do you] [Does the person] [Does (name)] speak a language other than English *at home?*
(If more than one language, indicate the one that is spoken most often.)

- No, English only ☐
- Yes, other—please specify: _____

Each of these questions elicits the same information (i.e. the language which a person speaks at home). The question used will depend upon survey design requirements and budget. Tick box responses are 'self-coded' which reduces the amount of manual or office coding needed, as the majority of responses are captured by the list. However, if there is a need for detailed information but space constraints are an overriding consideration, then the reduced form of the question for detailed data could be used. This is a more expensive and time consuming option as each response other than English will need to be manually coded to a classification category using the coding index.

DETAILED AND MINIMUM DATA QUESTIONS *continued*

When there is no requirement for detailed data, the standard question for minimum data should be used. This question differs from the reduced standard question for detailed data in that the 'Other' response category does not include the 'please specify' prompt. In the case of Main Language Other Than English Spoken at Home (example given below), it is primarily recommended for collections that seek to determine whether or not a person only speaks English in the home. Similarly, for the Country of Birth variables, the question should be asked when the only information needed is whether or not the person, their mother, or their father, is a migrant to Australia.

[Do you] [Does the person] [Does (name)] speak a language other than English *at home?*
(If more than one language, indicate the one that is spoken most often.)

No, English only..... ☐
Yes, other..... ☐

The question modules for the Minimum Core Set and Full Standard Set provided in this chapter contain the detailed standard questions for each variable. The tailored set includes examples of both the reduced format of detailed questions and minimum data questions.

NAVIGATING THE QUESTION MODULES

Question modules in this chapter are displayed with suggested sequencing. Text in **bold type** on the right hand side of responses are instructions for sequencing past the following inapplicable questions, or are used to indicate that no more questions are required to be answered. The instructions given only apply to the response they are next to. For any response that does not have a sequencing instruction, the respondent should go on to the next question in the sequence.

Example responses, prompts and additional information about how to answer the question are given in *italics* directly beneath the relevant question. These are important for the correct interpretation of the question and should always be included on self-enumerated questionnaires. On questionnaires being administered by interviewers, the interviewer should use this information when clarification is necessary.

The way in which a question is worded will depend on the collection methodology used. These variations are in square brackets: [ ]. Only the option that relates to the methodology being used should be included when designing the questionnaire. For example, if the collection is a personal interview, where an interviewer is directly questioning the respondent, a direct address should be used (i.e. 'Were you', 'Are you', 'Do you'). If the collection is self-enumerated, where any person in the household could answer for others in the household, the question wording 'Does the person', 'Is the person', etc. is used.

TAILORING RESPONSE CATEGORIES

As detailed in Chapter 3, it is possible to 'tailor' tick boxes to suit the needs of the collection. The tick box options provided for responses to each of the questions have been included primarily on the basis of their statistical significance in the Australian context. However, different regions may differ in certain characteristics. In these cases, it may be more cost effective to 'tailor' the tick boxes to better represent the population being surveyed using ABS census data. This can be obtained by contacting the ABS Consultancy Services section in your State or Territory as outlined on the back page of this publication.

QUESTION MODULE FOR THE MINIMUM CORE SET

The following question module is used to collect data on the Minimum Core Set of variables for the measurement of cultural and language diversity. This is the minimum set of questions which should be used when measuring cultural and language diversity.

Q1 In which country [were you] [was the person] [was (name)] born?

- Australia ☐
- England ☐
- New Zealand ☐
- Italy ☐
- Viet Nam ☐
- Scotland ☐
- Greece ☐
- Germany ☐
- Philippines ☐
- Netherlands ☐
- Other—please specify: ☐

Q2 [Do you] [Does the person] [Does (name)] speak a language other than English *at home*?

(If more than one language, indicate the one that is spoken most often.)

- No, English only ☐ >**Go to Q4**
- Yes, Italian ☐
- Yes, Greek ☐
- Yes, Cantonese ☐
- Yes, Mandarin ☐
- Yes, Arabic ☐
- Yes, Vietnamese ☐
- Yes, German ☐
- Yes, Spanish ☐
- Yes, Tagalog (Filipino) ☐
- Yes, other—please specify: ☐

Q3 How well [do you] [does the person] speak English?

- Very well ☐
- Well ☐
- Not well ☐
- Not at all ☐

Note: Question 3 requires a modification when being asked in an interview situation. When collecting data via interview, the following question should be used:

QUESTION MODULE FOR THE MINIMUM CORE SET
continued

Q3 Do you consider [you speak] [(name) speaks] English very well, well, or not well?

- Very well ☐
- Well ☐
- Not well ☐
- Not at all ☐

Q4 [Are you] [Is the person] [Is (name)] of Aboriginal or Torres Strait Islander origin?

(For persons of both Aboriginal and Torres Strait Islander origin, mark both 'Yes' boxes.)

- No ☐
- Yes, Aboriginal ☐
- Yes, Torres Strait Islander ☐

QUESTION MODULE FOR THE STANDARD SET

The following question module is used to collect data on the full Standard Set of variables for the measurement of cultural and language diversity. The questions are presented in the suggested order with sequencing instructions included. However, it is recognised that all questions would rarely be required in any one collection. This is especially true for the questions relating to languages, where it is likely that collections would require only one or two language variables rather than all four. It is suggested that the variable or combination of variables which best suits information needs be used.

Q1 In which country [were you] [was the person] [was (name)] born?

- Australia ☐ >**Go to Q3**
- England ☐
- New Zealand ☐
- Italy ☐
- Viet Nam ☐
- Scotland ☐
- Greece ☐
- Germany ☐
- Philippines ☐
- Netherlands ☐
- Other—please specify: ☐

Q2 In what year did [you] [the person] [(name)] first arrive in Australia to live here for one year or more?
(Write in the calendar year of arrival or mark the box if here less than one year.)

☐☐☐☐

Will be here less than one year ☐

QUESTION MODULE FOR THE STANDARD SET
continued

Note: Any of Questions 3, 5 or 6 can be used as a filter for Question 4. See Chapters 2 and 3 for further information on the relative efficiencies of each language variable. When using the full set, it is unlikely that all language questions would be needed. Rather the variable, or combinations of variables, that best suit the needs of the collection should be used.

Q3 [Do you] [Does the person] [Does (name)] speak a language other than English *at home*?
(If more than one language, indicate the one that is spoken most often.)

- ☐ No, English only ..... >**Go to Q5**
- ☐ Yes, Italian .....
- ☐ Yes, Greek .....
- ☐ Yes, Cantonese .....
- ☐ Yes, Mandarin .....
- ☐ Yes, Arabic .....
- ☐ Yes, Vietnamese .....
- ☐ Yes, German .....
- ☐ Yes, Spanish .....
- ☐ Yes, Tagalog (Filipino) .....
- Yes, other—please specify: ☐

Q4 How well [do you] [does the person] speak English?

- ☐ Very well .....
- ☐ Well .....
- ☐ Not well .....
- ☐ Not at all .....

QUESTION MODULE FOR THE STANDARD SET
continued

AND/OR

Q5 Which language [did you] [did the person] [did (name)] *first* speak as a child?
(Mark only one box.)

- English ☐
- Italian ☐
- Greek ☐
- Cantonese ☐
- Mandarin ☐
- Arabic ☐
- Vietnamese ☐
- German ☐
- Spanish ☐
- Tagalog (Filipino) ☐
- Other—please specify: ☐

AND/OR

Q6 Which language [do you] [does the person] [does (name)] mainly speak *at home*?
(If more than one language, indicate the one that is spoken most often.)

- English ☐
- Italian ☐
- Greek ☐
- Cantonese ☐
- Mandarin ☐
- Arabic ☐
- Vietnamese ☐
- German ☐
- Spanish ☐
- Tagalog (Filipino) ☐
- Other—please specify: ☐

QUESTION MODULE FOR THE STANDARD SET
continued

AND/OR

Q7 Which language or languages [do you] [does the person] [does (name)] speak *at home*?
(Please indicate all languages spoken.)

- English ☐
- Italian ☐
- Greek ☐
- Cantonese ☐
- Mandarin ☐
- Arabic ☐
- Vietnamese ☐
- German ☐
- Spanish ☐
- Tagalog (Filipino) ☐
- Other—please specify: ☐

Q8 [Are you] [Is the person] [Is (name)] of Aboriginal or Torres Strait Islander origin?
(For persons of both Aboriginal and Torres Strait Islander origin, mark both 'Yes' boxes.)

- No ☐
- Yes, Aboriginal ☐
- Yes, Torres Strait Islander ☐

QUESTION MODULE FOR THE STANDARD SET
continued

Q9 In which country was [your] [the person's] [(name's)] mother born?

- Australia ☐
- England ☐
- Italy ☐
- New Zealand ☐
- Scotland ☐
- Greece ☐
- Viet Nam ☐
- Netherlands ☐
- Lebanon ☐
- Philippines ☐
- Other—please specify: ☐

Q10 In which country was [your] [the person's] [(name's)] father born?

- Australia ☐
- England ☐
- Italy ☐
- New Zealand ☐
- Scotland ☐
- Greece ☐
- Netherlands ☐
- Germany ☐
- Viet Nam ☐
- Lebanon ☐
- Other—please specify: ☐

Note: If only one question on birthplace of parents can be accommodated, Country of Birth of Mother (Q9) should be collected.

QUESTION MODULE FOR THE STANDARD SET
continued

Q11 What is [your] [the person's] [(name)'s] Ancestry?
(For example: Hmong, Kurdish, Australian South Sea Islander, Maori.)
(Provide more than one ancestry if necessary.)

- English ☐
- Irish ☐
- Italian ☐
- German ☐
- Greek ☐
- Chinese ☐
- Vietnamese ☐
- Dutch ☐
- Filipino ☐
- Australian ☐
- Other—please specify: ☐

Q12 What is [your] [the person's] [(name)'s] religion?
(Answering this question is OPTIONAL.)
(For example: Salvation Army, Hinduism, Judaism or Humanism.)
(If no religion, mark last box.)

- Catholic (not Eastern Churches) ☐
- Anglican (Church of England) ☐
- Uniting Church ☐
- Presbyterian ☐
- Greek Orthodox ☐
- Baptist ☐
- Lutheran ☐
- Islam ☐
- Buddhism ☐
- Other—please specify: ☐
- No religion ☐

EXAMPLE OF A TAILORED QUESTION MODULE

The following question module is an example of tailoring the set of cultural and language diversity variables to suit particular data requirements. In addition to the questions contained in the Minimum Core Set, it contains Year of Arrival in Australia, Country of Birth of Mother and First Language Spoken.

Year of Arrival in Australia is a useful addition to the Minimum Core Set as it can be used as an indication of the respondent's likelihood of having adapted to Australian society. Users interested in determining which individuals or community groups have the most difficulty, or take the longest time to adapt to Australian society, may wish to include this question in their question module.

The combination of First Language Spoken and Main Language Other Than English Spoken at Home provides a good measure of changes in language usage over time. Users interested in identifying those communities or individuals that tend to retain their community language may be interested in the inclusion of both of these language variables. This information can be used to assist in identifying individuals and community groups that may require particular services. Either Main Language Other Than English Spoken at Home or First Language Spoken can be used as a filter for Proficiency in Spoken English.

Country of Birth of Mother can be used in conjunction with the Minimum Core Set to provide a finer measure of ethnic or cultural identity (than does Country of Birth of Person used alone) and to give an indication of the retention of a person's mother's culture and language. Users with an interest in broad measures of cultural diversity may wish to include this variable in their question module.

EXAMPLE OF A TAILORED QUESTION MODULE
continued

Q1 In which country was the person born?

Australia	☐	**>Go to Q3**
England	☐	
New Zealand	☐	
Italy	☐	
Viet Nam	☐	
Scotland	☐	
Greece	☐	
Germany	☐	
Philippines	☐	
Netherlands	☐	

Other—please specify: ☐

Q2 In what year did the person first arrive in Australia to live here for one year or more?
(Write in the calendar year of arrival or mark the box if here less than one year.)

☐☐☐☐

Will be here less than one year ..... ☐

Q3 Does the person speak a language other than English *at home*?
(If more than one language, indicate the one that is spoken most often.)

No, English only	☐
Yes, Italian	☐
Yes, Greek	☐
Yes, Cantonese	☐
Yes, Mandarin	☐
Yes, Arabic	☐
Yes, Vietnamese	☐

Yes, other—please specify: ☐

EXAMPLE OF A TAILORED QUESTION MODULE
continued

Q4 Which language did the person *first* speak as a child?
(Mark only one box.)

English	☐	>**Go to Q6**
Italian	☐	
Greek	☐	
Cantonese	☐	
Mandarin	☐	
Arabic	☐	
Vietnamese	☐	
German	☐	
Spanish	☐	
Tagalog (Filipino)	☐	
Other—please specify:		

Q5 How well does the person speak English?

Very well	☐
Well	☐
Not well	☐
Not at all	☐

Q6 Is the person of Aboriginal or Torres Strait Islander origin?
(For persons of both Aboriginal and Torres Strait Islander origin, mark both 'Yes' boxes.)

No	☐
Yes, Aboriginal	☐
Yes, Torres Strait Islander	☐

Q7 In which country was the person's mother born?

Australia	☐
Other country	☐

GLOSSARY

Administrative collection (data)

This refers to data obtained from existing records and documents such as: hospital admissions forms; employment records; births, deaths and marriages records; trade union membership records, etc. Much social and economic research can be done using such existing data without the need for costly and time consuming surveys.

Any responsible adult (ARA) interview

This method of data collection involves asking questions of the first responsible adult (usual resident aged 18 or over) contacted by the interviewer. This person is required to answer the questions on behalf of either a single member or all members of the household. This person is generally related to the people/person they are answering for or knows them very well. However, this method may result in lower response rates and data quality compared to a personal interview, where the respondent is directly questioned, as the ARA may not know all details about the respondent or may answer the questions subjectively.

Classification

A classification can be defined as a set of categories structured in a way that allows each unit in a population to be assigned unequivocally to only one category in the set according to characteristics of the unit.

Categories in a classification must be mutually exclusive and jointly exhaustive of the population of units under consideration. Classifications must also be useful (e.g. relevant) and realistic.

Classifications often have hierarchical structures with two, three, four or more hierarchical levels. A classification with only one level is termed *non-hierarchical* or *flat*. There may need to be rules to ensure categories are not too small, nor too large. There will always be a conflict between keeping a classification up to date with reality and maintaining unbroken time series.

A standard classification can generally be applied to a number of different variables. For example, the *Australian Standard Classification of Languages (ASCL)* (Cat. no. 1267.0) can be used to classify data collected by the four language variables: Main Language Other Than English Spoken at Home, First Language Spoken, Main Language Spoken at Home, and Languages Spoken at Home.

Coding index

A coding index enables descriptive information such as the responses to one or more survey questions to be coded to the appropriate category in a classification. The coding index is generally a comprehensive, alphabetical list of all the probable responses, usually derived from previous survey responses, together with the appropriate numerical code.

Coding procedures

Coding procedures are the rules whereby data collected for a variable (and related question) are assigned to categories of the classification. This coding is generally facilitated by a coding index. Standard coding procedures promote consistency across collections and over time.

Indicator

An indicator is an attribute of a person or set of attributes of a population (defined and measured by a statistical variable) that indicates possible aspects of their situation in Australian society. For instance, the variables Income and Occupation can be used as indicators of socioeconomic status. Similarly, Country of Birth of Person and Main Language Spoken at Home can be used as indicators of a likely connection to a particular cultural or ethnic group. If Proficiency in Spoken English is added to the indicators, broad conclusions about socioeconomic status and the possible degree of advantage or disadvantage in Australian society can also be drawn.

Whereas a variable is an objective measure of some attribute of a person (or group of people), an indicator uses the information supplied by a variable to indicate or imply, to a greater or lesser extent, a possibility in regard to the person's (or group of persons') status or functioning.

Interview-based collection

This method of data collection involves an interviewer asking questions of the respondent in a personal interview or an ARA interview. The costs involved in employing, training and managing interviewers make this method of data collection more expensive than self-enumeration. Also, for more sensitive topics, respondents might not be inclined to reveal private information using this method. Interview-based collections may be conducted face-to-face or using the telephone.

Personal interview (PI)

Personal interviews involve an interviewer directly asking the questions of a respondent and recording the responses. Another person does not answer the questions on behalf of the respondent. This method usually achieves high response rates and high quality data.

Question module

A question module is a set of standard questions, relating to particular variables, designed to collect a range of information. For example, the question module for the Minimum Core Set of Cultural and Language Indicators consists of four standard questions (for the variables Country of Birth of Person, Main Language Other Than English Spoken at Home, Proficiency in Spoken English and Indigenous Status), and is designed to collect fundamental information about a person's cultural and language diversity.

Self-enumerated collection

Self-enumerated collections are those in which respondents are required to complete the survey or administrative questions for themselves. Respondents usually feel more willing to supply personal or sensitive data with this method, compared to the personal interview method, however, there may be some confusion and misinterpretation of questions unless appropriate instructions are included. All question wording, instructions and sequencing of responses must be simple and self-explanatory.

Statistical standard for a variable

A statistical standard for a variable is defined as:

> '*… a set of components which, when used together, produce consistent and high quality statistical output (about concepts which underpin the cultural and language variables) across collections and over time.'*

The essential set of components that comprise and specify a statistical standard for a variable are as follows:

- Standard Name of the Variable;
- Standard Definition of the Variable;
- Standard Question(s);
- Standard Classification;
- Standard Coding Procedures; and
- Standard Output Categories.

Statistical standards are designed to add quality to the data produced by improving the accuracy, reliability, relevance and timeliness of data. Statistical standards involve thorough and rigorous development of concepts, definitions, questions, classifications, and processing and dissemination procedures. This rigorous development process is combined with consideration of practical issues such as statistical feasibility and extensive consultation with users of the standard and of the resulting data, and usually leads to long term cost savings by providing an 'off the shelf solution' for most statistical collections.

Standards improve data comparability and comprehensibility. As Australian Bureau of Statistics standards are designed to harmonise as far as possible with established Australian and international practices, comparability applies internationally, across collections, across time, across agencies and within a given subject area. Comprehensibility means the ability to be understood, by users of the standard and by respondents to administrative and statistical collections. It involves clarity of definitions, realism in the sense of modelling the real world, and providing logical and coherent structure. Standards also provide an integrated statistical picture of society and the economy. It means drawing all the data about a particular topic, variable or population together—in a meaningful way—from the full range of statistical sources.

Variable A variable, in simple terms, is a concept which can be measured. More precisely, for Australian Bureau of Statistics purposes, a variable is an attribute of a statistical unit whose value can, in theory, be measured in a statistical collection. In this context (i.e. for the Standard Set of Cultural and Language Indicators), each (statistical) variable names and defines a particular attribute relating to the origins of each person questioned. The variable also describes and itemises the set of attributes of a population of people. It is called a variable because the defined attributes may vary from person to person. It is a 'statistical' variable because it is used to collect, classify and aggregate responses relating to a defined attribute of each person in a statistical collection.

For instance, First Language Spoken is a statistical variable which attributes a defined element of language usage to each person (the first language they spoke as a child) and itemises the total range of languages first spoken by a population of people. Main Language Spoken at Home and Main Language Other Than English Spoken At Home are other language variables. Language itself is not used as a statistical variable because it does not closely define a particular aspect of language usage or refer to a particular entity being counted, and is not statistically useful. All of the variables recommended in the Standard Set of Cultural and Language Indicators relate to people, but it should be noted that statistical variables can also relate to families, households, activities, businesses, etc., as well as to people.

BIBLIOGRAPHY

Australian Bureau of Statistics 1996, *1996 Census Dictionary,* Cat. no. 2901.0, ABS, Canberra.

Australian Bureau of Statistics 1996, *1996 Census Directory of Classifications,* Cat. no. 2904.0, ABS, Canberra.

Australian Bureau of Statistics 1996*, ABS Views on Census Classification,* Cat. no. 2012.0, ABS, Canberra.

Australian Bureau of Statistics 2000, *Australian Standard Classification of Cultural and Ethnic Groups (ASCCEG),* Cat. no. 1249.0, ABS, Canberra (to be released mid 2000).

Australian Bureau of Statistics 1997*, Australian Standard Classification of Languages (ASCL)*, Cat. no. 1267.0, ABS, Canberra.

Australian Bureau of Statistics 1996, *Australian Standard Classification of Religious Groups (ASCRG),* Cat. no. 1266.0, ABS, Canberra.

Australian Bureau of Statistics 1998, *Standard Australian Classification of Countries (SACC),* Cat. no. 1269.0, ABS, Canberra.

Australian Bureau of Statistics 2000, *Statistical Concepts Library* (CD-ROM), Cat. no. 1361.0.30.001, ABS, Canberra (to be released early 2000).

Australian Bureau of Statistics 1986, *The Measurement of Ethnicity in the Australian Census of Population and Housing: Report to the Australian Statistician by the 1986 Population Census Ethnicity Committee,* Cat. no. 2172.0, ABS, Canberra.

CONTACT DETAILS

For more information on the implementation of the standard variables, or any aspect of the collection and use of cultural and language diversity data, please contact the ABS. Feedback or comments on this publication are welcome.

Postal address:

Australian Bureau of Statistics
PO Box 10
Belconnen ACT 2616

Telephone:

(02) 6252 5736—Social Classifications hotline

Fax:

(02) 6252 5281

Email:

social.classifications@abs.gov.au

FOR MORE INFORMATION...

SELF-HELP ACCESS TO STATISTICS

CPI INFOLINE — For current and historical Consumer Price Index data, call 1902 981 074 (call cost 75c per minute).

DIAL-A-STATISTIC — For the latest figures for National Accounts, Balance of Payments, Labour Force, Average Weekly Earnings, Estimated Resident Population and the Consumer Price Index call 1900 986 400 (call cost 75c per minute).

INTERNET — www.abs.gov.au

LIBRARY — A range of ABS publications is available from public and tertiary libraries Australia-wide. Contact your nearest library to determine whether it has the ABS statistics you require.

WHY NOT SUBSCRIBE?

PHONE — +61 1300 366 323

FAX — +61 03 9615 7848

CONSULTANCY SERVICES

ABS offers consultancy services on a user pays basis to help you access published and unpublished data. Data that is already published and can be provided within 5 minutes is free of charge. Statistical methodological services are also available. Please contact:

INQUIRIES

City	*By phone*	*By fax*
Canberra	02 6252 6627	02 6207 0282
Sydney	02 9268 4611	02 9268 4668
Melbourne	03 9615 7755	03 9615 7798
Brisbane	07 3222 6351	07 3222 6283
Perth	08 9360 5140	08 9360 5955
Adelaide	08 8237 7400	08 8237 7566
Hobart	03 6222 5800	03 6222 5995
Darwin	08 8943 2111	08 8981 1218

POST — Client Services, ABS, PO Box 10, Belconnen ACT 2616

EMAIL — client.services@abs.gov.au