2001 CENSUS: LEVEL, MAIN FIELD AND YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION

(Census Paper No. 03/09)

2001 CENSUS:
LEVEL, MAIN FIELD AND YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION

Anne Travis

Population Census Evaluation

October 2003

INQUIRIES

For further information about this paper, contact the Assistant Director, Census Evaluation by telephone: (02) 6252 5611 or email: joanne.healey@abs.gov.au.

SUMMARY OF FINDINGS

This 2001 Census Paper evaluates the data quality of questions relating to highest non-school qualifications of Australians. Overall, high non-response and incomplete response are the greatest data quality issues. The main findings of this paper are:

- The non-response rate for the non-school qualification indicator question increased marginally from 1996. Increasing evidence showed that this question is confusing to respondents and needs further attention. For example, a high majority of respondents who failed to complete this question, but answered the questions relating to level, main field and year of completion of highest non-school qualification were found to hold a formal qualification (such as a Bachelor degree).

- The non-response rate for year of completion of highest non-school qualification approximately doubled between Censuses. This increase is most likely to be associated with a change in questionnaire design from a selection list response to a write-in box.

- While the non-response rate for level of highest non-school qualification halved between Censuses, the proportion of responses which were coded to 'inadequately described' increased. This paper argued that the space on the 2001 Census form (three rows of boxes for level of highest non-school qualification as opposed to four rows of boxes for main field of highest non-school qualification), as well as the quality of information provided by respondents were contributing factors. This will be further investigated by the ABS following the August 2003 Census test, which has assigned four rows of boxes to both questions.

- Non-school qualifications which were obtained some time ago (such as 'General nursing' qualifications) were particularly difficult to code. A high proportion of these responses were coded as 'inadequately described'.

- Level of highest non-school qualification presented more coding problems than main field of highest non-school qualification, both in terms of the proportion of records which were 'inadequately described' and query rates measured at the DPC. These problems were more prevalent within certain main fields, including 'Health' and 'Management and commerce'.

- The predominant data quality issue associated with main field of highest non-school qualification was found to be a high proportion of records in some main fields (such as 'Information technology', 'Natural and physical sciences' and 'Engineering and related technologies') which could only be coded at a very broad (2-digit) level.

- Comparisons with the 2001 Survey of Education and Work revealed undercounts for both level and main field of highest non-school qualification. The greatest concern was the severe undercount for highest non-school qualifications at the Certificate I/II level.

- Changes to the validation and derivation systems will be considered for the 2006 Census, to ensure full comparability between the Census and other statistical collections.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

1. INTRODUCTION

1.1 About Census Papers

The Australian Bureau of Statistics (ABS) has a stated, corporate objective to provide the means for informed and increased use of statistics. This paper is one of a series produced after each census by the Australian Bureau of Statistics' Population Census Evaluation Team, whose role is to review the data quality of the 5 yearly Census of Population and Housing. The aim of a Census Paper is to inform users of issues that have been identified as impacting on the quality of the census data, to be considered when utilising the data. Analyses such as this are a critical factor in the continuous quality improvement of the Census Program. The ABS welcomes your feedback and suggestions.

1.2 This Paper

This Census Paper evaluates the quality of questions relating to the highest non-school qualifications of Australians collected in the 2001 Census, including the effects of respondent error and non-response as well as information on the development and processing of the Census form. Details of these questions, including images from the 2001 Census Form, are discussed in *Section 2, Question Design.*

Section 3 (Collection of Data) and *Section 4 (Processing at the Data Processing Centre)* contains information about how various stages in the development and processing of the 2001 Census Forms might have affected the data. The previous Census Working Paper: *1996 Census Data Quality: Qualification Level and Field of Study* (00/2), identified the following data quality issues for the questions on highest post-school qualification collected from the 1996 Census:

- a large percentage of responses could not be coded to a detailed field due to insufficient detail in their responses;
- coders had difficulty coding a qualification level for vocational qualifications; and
- the broad fields: 'Business administration' and 'Society and culture', were frequently miss-allocated and classified under other field codes in the data processing stages.

Section 5 (Sample Data Analysis) provides analyses of a 2 per cent sample of Collection Districts (CDs) from each State and Territory in Australia of the non-school qualification indicator question, focussing on cases of multiple marks, the frequency of respondents who indicated that they did not have a non-school qualification in the non-school qualification indicator but responded to either the level or main field of highest non-school qualification questions, and instances where respondents aged under 15 years incorrectly completed questions relating to non-school qualifications.

Section 6 (Final Data Analysis) discusses data quality issues relating to respondent error. The most serious errors were high non-response rates for all questions relating to highest non-school qualifications. Analyses to support the hypotheses that most non-respondents did not hold a qualification, as well as demographics (such as age and labour force status) are included. Other errors identified and discussed include insufficient detail provided by respondents and errors reported on the Census form.

Section 6 also contains a time series analysis between the 1996 and 2001 Censuses and comparisons with data collected from the 2001 ABS Survey of Education and Work (SEW), with particular attention to data quality issues and comparability.

Finally, *conclusions and recommendations* are made so that users of Census data can be fully informed about the quality and usefulness of statistics produced. Findings will also help the ABS to identify particular groups that could benefit from increased attention in the public relations campaign for the 2006 Census and improve form design, processing, as well as the content of guides such as *Household Guide: How to Complete your Census Form*.

1.3 Background

Statistics about the outcomes from Australia's education system, including levels of highest non-school qualifications, are important to society. Education and training help people to develop knowledge and skills that may be used to enhance their own living standards and those of the broader community. For an individual, educational attainment is widely seen as a factor contributing to a rewarding career and is related to income growth by providing people with specialised skills. For the nation as a whole, having a skilled workforce is vital to supporting ongoing economic development and improvements in living conditions.

Indicators on levels of highest non-school qualifications completed by society tend to be related to increasing levels of participation in education and training, but participation and attainment among identified target groups remain well below the overall population. Such indicators are therefore key inputs into Government policy development as well as performance monitoring. Statistical data are used to address social issues such as:

- monitoring and improving levels of highest non-school qualifications for target groups;
- promoting both equal access to education and equity in educational outcomes; and
- ensuring educational resources are sufficiently and appropriately distributed.

1.4 Comparability with the 1996 Census

Since the 1996 Census, the provision of education and training has undergone significant structural change, including:

- the boundaries between the three major sectors (schools, Vocational Education and Training (VET) and higher education) have become less distinct with developments such as a few universities offering programmes under the National Training Framework (NTF), VET programmes being offered in schools and some Bachelor programmes being offered by VET institutions;

- an increase in the number of multi-sector institutions;

- greater emphasis on appropriate skills formation and entry level training, particularly in the VET sector; and

- industry involvement in the development of work based competency based training programs (including the introduction of the Australian Qualifications Framework by the Ministerial Council on Education, Employment, Training and Youth Affairs).

These developments substantially reduced the usefulness of the *Australian Bureau of Statistics Classification of Qualifications* (ABSCQ), which was used for the processing of the 1996 Census. After a lengthy review process, the ABS developed and implemented a new national standard classification, the *Australian Standard Classification of Education* (ASCED). ASCED has a significantly broader scope than the ABSCQ because it is not directed specifically at only post-school qualifications study. Therefore, 2001 Census statistics compiled using the ASCED should be compared to 1996 Census Statistics compiled using the former ABSCQ with extreeme caution. This is discussed in more detail in *Section 6.5 (Classification Issues)* and *Section 6.6 (Intercensal Comparisons)*.

2. QUESTION DESIGN

2.1 Main questions in the 2001 Census Forms

The nature and content of questions on highest non-school qualifications included in the Census have undergone continuous improvement over time, in line with changes to classification systems (such as the implementation of ASCED), user requirements, and attempts to minimise respondent errors, (such as non-response, incomplete information, omissions and incorrect reporting). Changes have also been driven by improved form design and processing techniques such as Optical Mark Recognition (OMR) and Intelligent Character Recognition (ICR).

An education question in which respondents reported their highest level of achievement was first included in the 1911 Census. However, a question directly asking about educational qualifications was not included until 1966 when respondents were asked to provide the name of the qualification and the institution at which it was obtained. Since 1971, people aged 15 years or over have been asked whether they obtained a qualification and if so, information on the name of the highest qualification, the field of study, the institution from which it was granted, and the year of completion have been obtained. This enabled coding to produce a detailed classification of level and field of study. The year in which a qualification was obtained was coded separately from 1981.

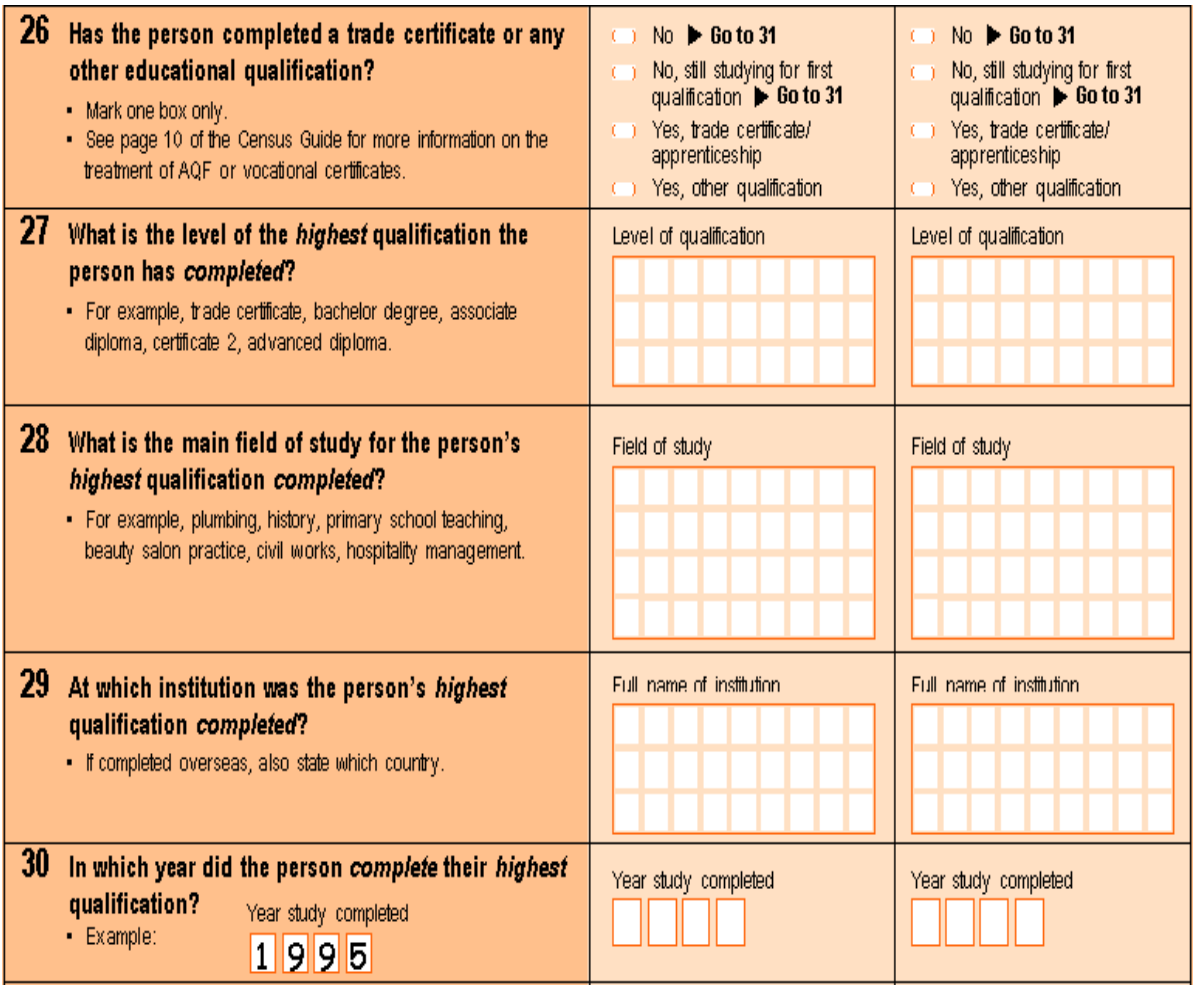
The Census of Population and Housing has different form types to accomodate differences in the living arrangements of Australians (share accommodation dwellings and family dwellings) and to ensure indigenous Australians are enumerated correctly.

The data relating to highest non-school qualifications is collected on the Household Form (HHF), Personal Form (PF) and Special Indigenous Personal Form (SIPF). The HHF and the PF are nearly identical in content: the HHF is designed to collect data for up to six people living in private dwellings; whereas the PF is designed for individuals or families spending Census Night away from their usual residence, or people who did not want their personal details disclosed to other members of the dwelling (such as those in group houses). The SIPF is designed to facilitate interviewer-based enumeration of Indigenous Australians in identified communities.

The 2001 Census had five questions on non-school qualifications (refer to Figure 1). The first question, generally referred to as the Qualification Indicator question, required a mark-box response (completion of trade certificate or any other educational qualification). The remaining four questions required a write in response (shown in the Census form as highest qualification, field of study, institution at which qualification was completed and year of completion).

The same questions relating to non-school qualifications were included in the PF, with the minor difference that the instruction for the question on the year in which the highest qualification was completed (Q31 in the PF) did not include an example.

FIGURE 1: NON-SCHOOL QUALIFICATIONS QUESTIONS, 2001 CENSUS HOUSEHOLD FORM



The SIPF did not differ much from the HHF and PF on basic content, however, the wording on questions and instructions in the SIPF were significantly different from the HHF and PF (refer to Figure 2). The SIPF is collected by an interviewer, whereas the HHF and the PF are self-enumerated. These differences may affect the quality of data collected from the SIPF, given that approximately 19% of Census data on Indigenous Australians aged 15 and over is collected via the SIPF (refer to Table 1).

The wording of question 23 on the SIPF to include only non-school qualifications since leaving school is inconsistent with the broader scope of ASCED and question 26 of the HHF, which includes all non-school qualifications irrespective of the respondents schooling status. As a consequence, non-school qualifications within the scope of ASCED which can be obtained prior to leaving school (predominantly, Certificate Levels I and II) may be understated for Census data collected using SIPFs.

In addition, the wording of question 23 would appear to exclude qualifications obtained from private organisations. There are large numbers of private organisations within the VET sector, particularly in the areas of English language studies, secretarial, data management and computer studies and business management.

FIGURE 2: NON-SCHOOL QUALIFICATIONS QUESTIONS, 2001 CENSUS SPECIAL INDIGENOUS PERSONAL FORM

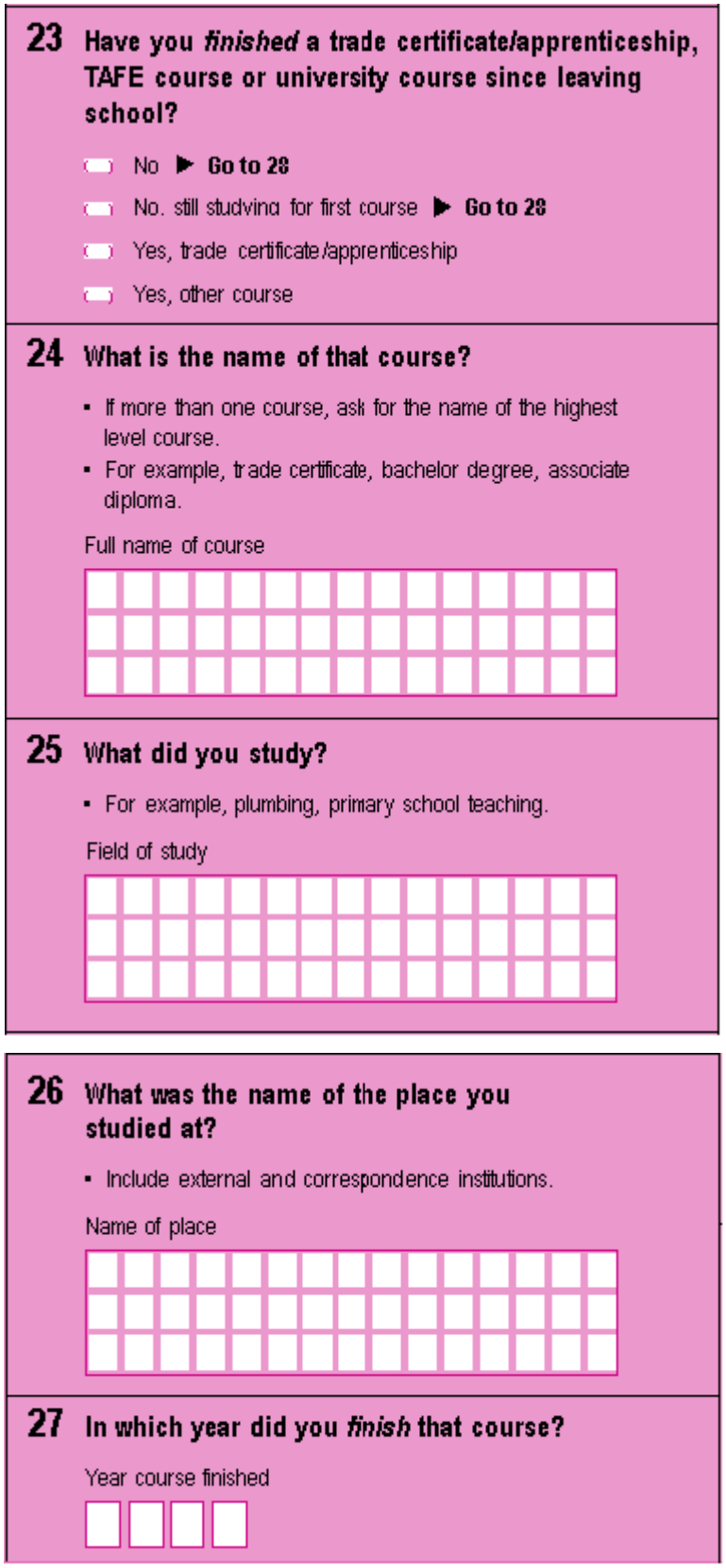


TABLE 1: INDIGENOUS STATUS OF PERSONS AGED 15 YEARS OR OVER, BY FORM TYPE (a) - 2001 CENSUS

Indigenous Status	*Personal Form*	*Household Form*	*Special Indigenous Personal Form*	*Total*
Non-Indigenous	461,993	13,526,056	1,240	13,989,289
Indigenous	12,825	184,256	46,089	243,890
Not Stated	42,974	209,165	448	252,587
Total	**517,792**	**13,919,477**	**47,777**	**14,485,766**

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

The following three questions relating to education collected from the 2001 Census are outside the scope of this Census Paper:

- Is the person attending a school or any other educational institution? If attending:
    - What type of educational institution is the person attending?; and
- What is the highest level of primary or secondary school the person has completed.

2.2 Intercensal changes to questions

2.2.1 Question content, instructions and examples

Structural changes (such as the increased tendency for high school students to obtain VET qualifications as part of their school curriculum) to the education and training sector since 1996 and the resulting replacement of ABSCQ with ASCED, had a large impact on the design and wording of the 2001 Census Form. This lead to the dropping of the phrase *'since leaving school'* from the qualification indicator and level questions (Questions 26 and 27 shown in Figure 1), thereby allowing all persons aged 15 years or more (including school students) to complete these questions for the first time. Comparability between the 1996 and 2001 Census arising from these structural changes are discussed in more detail in *Section 6.6 (Intercensal Comparisons)*.

It should be noted that the wording of the 2001 Census questions on non-school qualifications differs slightly from the formal variable definitions assigned by ASCED. For example, main field of study of highest qualification (which is the terminology used in the 2001 Census form), is refered to in subsequent sections of this paper as main field of highest non-school qualification (which is the ASCED standard definition).

Other changes to the 2001 Census questions on non-school qualifications were inclusion of:

- a bullet point instruction for the highest non-school qualification indicator question to encourage those who may have completed a VET qualification (which were predominantly out of scope of the 1996 Census) to refer to the Census guide and complete the question correctly;

- the word 'level' to the question relating to the level of the highest qualification the person has completed, to provide greater clarity to the question and may have contributed to the halving in the non-response rate for this question between Censuses; and

- a broader range of examples of fields (beauty salon practice, civil works and hospitality management), designed to address the inability to adequately code approximately 20 per cent of the responses to the *'Business and administration'* category in the 1996 Census.

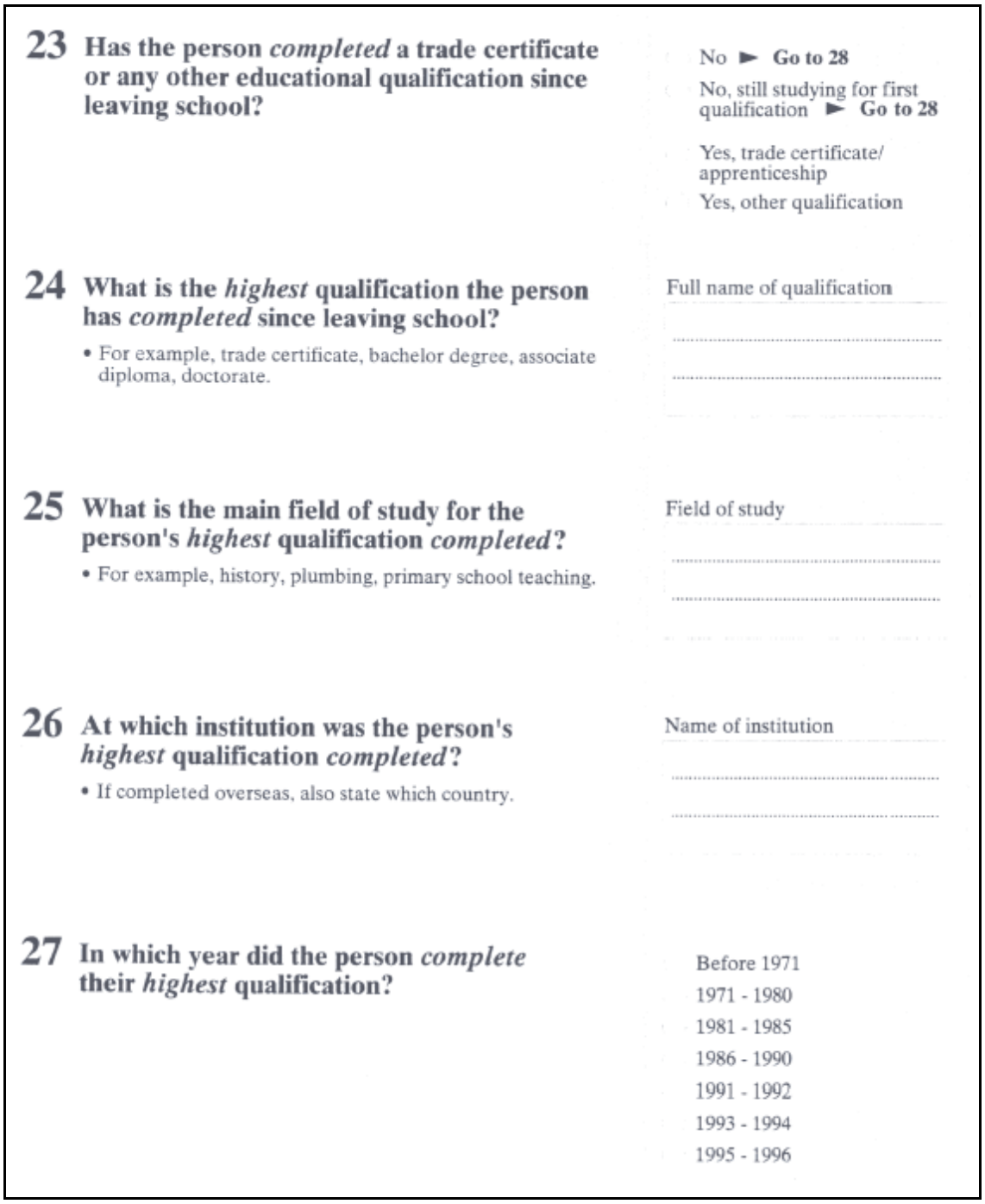
2.2.2 Question layout

A highly visible change to the form design for the 2001 Census was the replacement of dotted lines with boxes for respondents to answer the highest qualification level, field, name of institution and year of completion questions (refer to Figure 3 for the 1996 version). The change was primarily implemented to enable ICR and to facilitate automatic coding. However, this change also required the use of block letters and would have discouraged the use of cursive writing, which was used by many respondents in answering the questions on the 1996 Census Form.

This change may have affected data quality if there were insufficient boxes for respondents to adequately complete the question, resulting in responses being coded to *'inadequately described*'. However other factors, such as the quality of information provided by respondents, were also contributing factors. For more analysis, refer to *Section 6.3 - Analysis of 'Inadequately Described' Responses* and *Section 6.4 - 'Inadequately Described' and 'Not-Stated' Responses to Level of Highest Non-School Qualification*.

The layout of the question on the year of highest qualification completed changed from a selection list and mark box response in 1996 to a write-in response in 2001, which saved space on the Census form and also utilised ICR. However, this change would have been a significant contributor to the approximate doubling of non-response rates between Censuses (7.8 per cent in 2001, compared to 4.4 per cent in 1996). This increase in non-response is discussed in more detail in *Section 6.2.2 (Non-Response Analysis of Year of Completion of Highest Non-School Qualification)*.

FIGURE 3: POST-SCHOOL QUALIFICATIONS QUESTIONS, 1996 CENSUS HOUSEHOLD FORM



23 **Has the person *completed* a trade certificate or any other educational qualification since leaving school?**

- No ► **Go to 28**
- No, still studying for first qualification ► **Go to 28**
- Yes, trade certificate/ apprenticeship
- Yes, other qualification

24 **What is the *highest* qualification the person has *completed* since leaving school?**

- For example, trade certificate, bachelor degree, associate diploma, doctorate.

Full name of qualification

25 **What is the main field of study for the person's *highest* qualification *completed*?**

- For example, history, plumbing, primary school teaching.

Field of study

26 **At which institution was the person's *highest* qualification *completed*?**

- If completed overseas, also state which country.

Name of institution

27 **In which year did the person *complete* their *highest* qualification?**

- Before 1971
- 1971 - 1980
- 1981 - 1985
- 1986 - 1990
- 1991 - 1992
- 1993 - 1994
- 1995 - 1996

2.3 Possible list effect

The non-school qualification indicator question (completion of trade certificate or any other educational qualification) required a mark box response. For this type of question, there can be a possible bias in self-coded responses, known as a "list effect".

The list effect may not be significant, given that there were only four responses to select from, the list was fully exhaustive and results shown in *Section 5, Sample Data Analysis*, indicate that multiple marking occurred in less than 1% of cases.

3. COLLECTION OF THE DATA

During the collection phase of the 2001 Census, collectors reported increased difficulty contacting some householders. Access to secure small and large apartment buildings, gated communities, and growing community concerns about security, made it increasingly difficult to judge whether the residents of a dwelling were absent or not.

System Created Records (SCRs) were created during Census processing for people for whom a Census form had not been received but where the collector believed the dwelling to be occupied on census night. SCRs had values imputed for age, sex, marital status and usual residence only, and values for other variables set to 'not stated' or 'not applicable', depending on the imputed value for age.

An increase in non-response rates was apparent for many census variables in the 2001 Census. Most of the change was attributed to the increase in the proportion of SCRs. A Fact Sheet, *Effect of Census Processes on Non-response Rates and Person Counts* has been produced and stored on the ABS web site, which discusses the factors that may have contributed to the increase in SCRs for 2001, and the percentage of records affected by state.

After elimination of the impact of SCRs, the non-response rate for the year of completion of the highest non-school qualification doubled when compared to the 1996 Census. This large increase is most likely to be attributable to changes in question design. The non-response rates of the remaining questions on highest non-school qualifications either remained relatively constant or decreased. Analysis of the non-response rates for the questions relating to highest non-school qualifications are found in *Section 6.2 (Non-Response Analysis*).

4. PROCESSING AT THE DATA PROCESSING CENTRE (DPC)

4.1 Background

2001 Census forms were processed in the ABS' DPC in Sydney. After receipt and precapture checks (such as ensuring key fields are completed), the forms are then prepared for the first stage of processing, Data Capture (DC). Following DC, all Automatic Coding (AC) and Computer Assisted Coding (CAC) takes place. Quality control checks are constantly undertaken for coding accuracy and all census data are extensively validated before data is released.

Responses to non-school qualification indicator were coded using DC and CAC. Responses to level and main field of highest non-school qualification were coded using a combination of AC and CAC; 70.1% of these responses were coded using AC.

4.2 Data Capture

DC is the process of scanning the forms into the image and text files that are used for all subsequent processes. At this stage, mark box responses are captured and coded, and text responses are translated into machine readable symbols (through a process that assigns percentages of surety for each individual character) which are examined for their fitness for AC. Where the degree of tolerance was so low that automatic coding was not possible, the field was sent to CAC.

4.3 Automatic Coding

Automatic coding is the process of computer matching the captured text responses to entries on an index for that topic. If no match is made during AC, the response is sent to an operator for CAC.

In this second stage of processing, the Automatic Coder attempted to match the textualised ICR version of a response to an entry in the Coding Index. A table of tolerances was created to provide a framework for operation. Using its own in-house developed system, the ABS was one of the first international statistical agencies to utilise such technology to process Census forms.

4.4 Coding Procedures

CAC is the process of using procedures and rules to allow a coder to match the image of the text responses to entries on an index for that topic. If no match can be made, the response may be 'dump' coded to a less specific index entry or to 'inadequately described'. The operators also confirm if there is no response to the question.

4.4.1 Coding of Not-Stated Responses

In the CAC stage of processing, a bulletin was issued to DPC coders in February 2002, setting out procedures to be followed in the coding of 'not-stated' responses to the qualification indicator, level and main field of highest non-school qualification questions. The content of this bulletin is shown in Figure 4.

FIGURE 4: WHEN TO USE 'NOT-STATED' FOR QUALIFICATIONS, 2001 DPC CODING BULLETIN - 2001 CENSUS

If you are coding the 'qualification indicator question', (Q26), and Q26 is empty:

- but there is level information, such as 'Diploma', 'Bachelor' or 'Year 10', in either the level or field snippets (Q27 or Q28), you should select 'Yes, other qualification' for Q26.
- The exception is if the respondent has written 'Apprenticeship' or 'Trade Certificate' at Q27 or Q28. Then you should select 'Yes, trade certificate/apprenticeship' for Q26.
- Only if there is no level information in the field or level snippets, then you can code Q26 to 'Not-Stated'.

If field snippet (Q28) is empty, but there is a response such as 'Diploma of Visual Arts' in the level snippet (Q27), use this information to code qualification field (eg 'Visual arts') and then Qualification Level (eg 'Diploma').

If the level snippet (Q27) is empty:

- but there is a response such as 'Bachelor of Science' in the field snippet (Q28), you should use this information to code qualification field (eg, 'Science') and then Qualification Level (eg, 'Bachelor').
- or, if the 'Yes, trade certificate/apprenticeship' box has been marked in the 'qualification indicator' question (Q26), use this information to code Qualification Level.

If there is institution information, but NO qualification level or field information, then code to 'not-stated'.

Table 2 provides an analysis of 'not-stated' responses to the non-school qualification indicator question, which were processed using CAC procedures. There were a total of 175,320 instances where level of highest non-school qualification was stated and hence the first two coding rules outlined in Figure 4 could have applied, but were not correctly executed by DPC coders. In the majority of these cases, the respondents completed the year of completion of highest non-school qualification. Therefore, it is highly likely that these non-respondents have completed a non-school qualification.

Similarly, analysis of 'not-stated' responses to the question on level of highest non-school qualification indicator question revealed that there were 16,678 instances where the non-school qualification indicator question was stated by respondents as 'Yes, trade certificate/apprenticeship' and hence the second coding rule for level of highest non-school qualification could have applied. The majority of these records (16,671) were coded using CAC processes.

TABLE 2: NON-RESPONSES TO THE NON-SCHOOL QUALIFICATION INDICATOR QUESTION, BY LEVEL AND YEAR OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 2001 CENSUS

Level of Highest Non-School Qualification	*Response to Year of Completion of Highest Non-School Qualification*		
	'Stated'	*'Not-Stated'*	*Total*
Postgraduate degree	10,876	600	11,476
Graduate diploma / Graduate certificate	7,800	369	8,169
Bachelor degree	56,792	3,798	60,590
Advanced diploma / Diploma	38,239	2,998	41,237
Certificate III / IV	26,397	4,211	30,608
Certificate I / II	17,422	781	18,203
Certificate not further defined	3,075	1,962	5,037
Total	**160,601**	**14,719**	**175,320**

(a) Table excludes: Persons aged under 15 years, persons without a non-school qualification, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

Given the above mentioned pattern of 'not-stated' responses to the non-school qualification indicator and level of highest non-school qualification questions, non-response could be substantially lowered (refer to Table 3) in the future by implementing the derivation rules outlined in Section 6.5.3 for the 2006 Census.

TABLE 3: NON-RESPONSE RATES FOR HIGHEST NON-SCHOOL QUALIFICATION QUESTIONS AFTER THE IMPLEMENTATION OF REVISED CODING PROCEDURES (a) - 2001 CENSUS

		Persons for whom there was no response		*Non-response Rate (%)*	
Highest Non-School Qualification Question	*Persons for whom Questions were relevant*	*2001 Census*	*Revised Coding Procedures*	*2001 Census*	*Revised Coding Procedures*
Indicator	14,485,766	1,030,409	869,808	7.1	6.0
Level	5,455,279	300,636	283,958	5.5	5.2

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors. Refer to Section 6.2.1 for the methodology relating to the calculation of non-response rates.

4.4.2 'Not Further Defined' Coding of Main Field of Highest Non-School Qualification

For main field of highest non-school qualification, some responses only provided enough information to enable coding at a relatively broad level (2-digit or 4-digit level), instead of the most detailed classification under ASCED possible (defined as the 6-digit level, for example Astronomy 010303). In such instances, a 'NFD' (Not Further Defined) code was allocated and either of the following options were taken:

- responses were coded to the NFD category of the Narrow Field, called the 4-digit level (for example, Physics and Astronomy NFD 0103); or

- responses were coded to the NFD category of the Broad Field, called the 2-digit level (for example, Natural and Physical Sciences NFD 01).

The frequency of NFD coding, also known as 'dump coding', needs to be considered by users of Census data in determining the most appropriate level of detail to analyse main field of highest non-school qualification statistics. The frequency of 'not-stated' and 'inadequately described' responses also needs to be considered by users of Census data. These issues are discussed in *Section 6 (Final Data Analysis)*.

Table 4 shows that three-quarters of the responses in the 'Information technology' main field of highest non-school qualification could only be classified at a very broad level (2-digit). Other main fields which had a high proportion of responses coded at this very broad level include 'Natural and physical sciences' and 'Engineering and related technologies'. This indicates that additional guidance in definition booklets, such as a broader range of examples in the Census Form and Census Guide, should be concentrated within these main fields.

On the other hand, the main fields of highest non-school qualification 'Agriculture, environmental and related studies', 'Health' and 'Food, hospitality and personal services' had a very high proportion of responses coded to the most detailed level possible (6-digit level) and therefore do not present a problem to coders.

TABLE 4: DISTRIBUTION OF 'NOT FURTHER DEFINED RESPONSES', BY MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 1996 AND 2001 CENSUS

Main Field of Highest Non-School Qualification	*% of responses coded to Broad Field (2-digit code)*	*% of responses coded to Narrow Field (4-digit code)*	*% of responses coded to Detailed Field (6-digit code)*	*Total*
2001 Census:				
Natural and physical sciences	29.4	18.8	51.8	192,088
Information technology	75.1	17.9	7.0	153,183
Engineering and related technologies	15.8	17.1	67.0	1,270,721
Architecture and building	0.1	16.9	83.0	406,049
Agriculture, environmental and related studies	0.2	0.5	99.3	139,748
Health	4.3	1.9	93.8	592,009
Education	6.3	23.1	70.6	508,781
Management and commerce	5.6	16.7	77.7	1,053,735
Society and culture	5.7	22.5	71.8	562,311
Creative arts	8.7	16.9	74.4	198,511
Food, hospitality and personal services	0.1	0.7	99.2	318,253
Mixed field programmes	2.8	8.9	88.2	4,598
Field 'inadequately described'	. .	. .	. .	70,402
2001 Census - Total	**10.1**	**14.1**	**75.7**	**5,399,987**
1996 Census (b) - Total	**7.0**	**14.2**	**82.8**	**5,834,776**

. . Not applicable.

(a) Table excludes: Persons aged under 15 years, persons without a non-school qualification, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) VET Certificate Level 1 qualifications were outside the scope of the 1996 Census.

4.5 Indexes used in coding

All coding of responses is done by matching to index entries that map to the standard output classification for the topic. Indexes are constantly updated during the processing phase, in response to the type of answers respondents have provided. All additions to the index must be mapped to a category in the standard output classification and are done so with the assistance and approval of the ABS' classification experts. Index updates are requested by the coders to allow them to better code frequently occurring responses, and by the teams looking at the data throughout processing, such as in response to discrepancy reports.

Responses to level and main field of highest non-school qualification were classified according to ASCED (refer to *Section 6.5 - Classification Issues* for more details). The greatest limitation in 2001 with the index was that most queries were readily coded in terms of main field of highest non-school qualification, but matching on level of highest non-school classification was a difficulty. This resulted in a query rate for level of highest non-school qualification of 21.6%.

Analysis of 'inadequately described' responses to level of highest non-school qualification (refer to Section 6.4 for a detailed analysis) revealed that this difficulty was concentrated within a few main fields (such as 'Secretarial and clerical studies' and 'General nursing'). In these instances respondents used general terminology (such as 'Trade Certificate') or used terminology (such as 'Trained Nurse') which was applicable to nursing qualifications obtained prior to 1990.

Improvements to the coding indexes, particularly in relation to level of highest non-school qualification, as well as providing additional guidance to respondents in known problem areas, will improve data accuracy as well as improving the efficiency and work flow of processing of the 2006 Census.

4.6 Edits applied to the data

The ABS Census program has a minimalist editing approach, with most data output as reported on census forms. However, editing is the systematic way of altering data to ensure that it is:

- more complete. For example, if the basic demographic variables of age, sex or usual residence are not stated, they are imputed based on known distributions;

- socially consistent to some extent. For example, age edits do not allow five year olds to be attending high school; and

- consistent with ABS classifications used in other ABS collections. For example, Census labour force status is derived using the same derivation used in the Labour Force Survey, to allow clients to more accurately compare data.

For the processing of questions relating to the highest non-school qualification of persons aged 15 years and over, edits were applied to:

- edit level, main field and year of completion of highest non-school qualification to 'not applicable' in instances where respondents had answered 'No' or 'Still studying for first qualification' to the non-school qualification indicator question;

- eliminate non-school qualifications reported which were outside the scope of the Census (including those with a code of 610, 611, 613 and 621, which were secondary school qualifications and were edited to 'not applicable'); and

- derive a more detailed level for highest non-school qualifications at the diploma and certificate levels, following the implementation of the Australian Qualifications Framework and ASCED (refer to Table 5 below).

TABLE 5: EDIT RULES, LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION - 2001 CENSUS

Edit Specification
If ASCED level = 400 & year obtained < 1998, then new ASCED code = 411
If ASCED level = 400 & year obtained > 1997, then new ASCED code = 421
If ASCED level = 500 & ABSCQ level code = 7 & year obtained < 1998, then new ASCED code = 520
If ASCED level = 500 & ABSCQ level code = 6 & year obtained < 1998, then new ASCED code = 510

The edits listed in Table 5 did not apply to any of the 'not stated' responses to year of completion of highest non-school qualification. For example, 24.1% of responses to level of highest non-school qualification which were coded to 500 (Certificate not further defined), did not provide a response to year of completion of highest non-school qualification (refer to Table 6). *Section 6.2.2 (Non-Response Analysis of Year of Completion of Highest Non-School Qualification)* provides a detailed analysis of 'not stated' responses to year of completion of highest non-school qualification.

TABLE 6: NON-RESPONSE RATES FOR YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION, BY LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 2001 CENSUS

Level of Highest Non-School Qualification Variable	*Persons for whom Questions were relevant*	*Persons for whom there was no response*	*Non-response Rate (%)*
Postgraduate degree	257,566	5,297	2.1
Graduate diploma / Graduate certificate	195,759	4,075	2.1
Bachelor degree	1,385,353	35,471	2.6
Advanced diploma / Diploma	851,122	30,417	3.6
Certificate III / IV	1,907,882	224,846	11.8
Certificate I / II	312,043	7,314	2.3
Certificate not further defined	68,168	16,452	24.1
Level 'inadequately described'	176,750	15,807	8.9
Level 'not stated'	300,636	84,352	28.1
Total	**5,455,279**	**424,031**	**7.8**

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

4.7 Quality management

A Quality Management (QM) system was established to identify systematic discrepancies in processing, provide feedback to coders on discrepancies, and produce and analyse discrepancy rates by topics.

4.7.1 The QM Process

QM processing takes a sample of each coder's work, plus samples of codes resulting from data capture and automatic coding, for duplicate coding by a second coder. When the original code and second code differ, both outcomes are written to a mismatch file; these mismatches are then recoded for a third time, by an adjudicator, who determines which is the correct code. When the adjudicator determines a code that differs from the original and/or second coder, a discrepancy is recorded for that source; in some cases the adjudicator may determine both are incorrect, and both will have a discrepancy recorded. A report of these discrepancies is fed back to the relevant coder, or process, so that retraining can be done, or systems updates can be made.

4.7.2 Discrepancy Rates

In the majority of cases, the data is not corrected as a result of this sampling: the aim is to improve the coder or process so that such errors do not reoccur. Discrepancy Rates therefore show error rates that are very close to those existing in the final data. However, in extreme cases the production data is recoded. The discrepancies are also aggregated into the Management Information System reports which provide data on the types and frequencies of coding errors over time.

The QM system in place during processing allowed the detection of discrepancies and the calculation of a crude discrepancy rate. This crude discrepancy rate differs from a true discrepancy rate for the following reasons:

- a higher proportion of ‘poor’ coders’ work was included in the quality monitoring sample;
- the QM check coders could make the same mistake as the original coder and therefore an error would not be detected;
- there is not always an absolutely correct code for every response; and
- discrepancies were recorded for any difference in coding between the Quality Management coder and the original coder.

The DPC routinely reviewed between 10% and 50% of automatic and manual coding. This practice was ongoing, though, particularly with a ‘human’ coder, the percentage chosen for review varied depending on their performance. In this way a measure of quality could be made, and extra training or ongoing support provided if a staff member was having continuing problems. Automatic processes were also continuously monitored.

4.7.3 Discrepancy Rates in final data

There would invariably have been errors that coders or systems would have made that were repeated by the QM coders - therefore ensuring that further review of Adjudication never occurred. Such occurrences, however, would have been small - no doubt less than the confirmed Discrepancy Rate. Balancing out this aspect was the greater scrutiny of coders experiencing difficulty.

Generally, the Discrepancy Rates outlined below can be presumed to be close to the error rate in final data. The final Discrepancy Rates, broken down by each coding process, for level and main field of highest non-school qualification questions are presented in Table 7.

TABLE 7: DISCREPANCY RATES FOR LEVEL AND MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION, BY CODING PROCESS - 2001 CENSUS

Highest Non-School Qualification Variable	*Automatic Coding*	*Computer Assisted Coding*	*Combined*
2-Digit Level			
Level	1.4%	1.4%	1.4%
Main Field	0.6%	0.9%	0.7%
Detailed Level			
Level	1.5%	1.4%	1.5%
Main Field	1.3%	1.9%	1.6%

These results were very close to the 1% expected error rate had all the records been coded manually. The most significant difference in Discrepancy Rates was the difference between coding at the 2-Digit and detailed levels, for field of highest non-school qualification. The reason for the difference is that responses were often ambiguous and judgment calls by coders and adjudicators were made.

4.7.4 Discrepancies

For level of highest non-school qualification, the following discrepancies were identified during the QM process:

- 30% of AC discrepancies were due to adjudicators determining the correct code should have been 514 (Certificate III). In two-thirds of these cases, coders coded to ‘not stated’, because they did not see the trade certificate box marked; and

- one-quarter of CAC discrepancies were due to adjudicators determining the correct code should have been 514 (Certificate III), while coders coded the responses to 500 (Certificate not further defined), 524 (Certificate I) or 521 (Certificate II).

For main field of highest non-school qualification, the highest proportion of both AC and CAC discrepancies were due to adjudicators determining the correct code should have been within the 08 (management and commerce) group, but were coded to a different code within the 08 group. For example, 11% of AC discrepancies were due to adjudicators determining the correct code to be 080000, but one-quarter of these cases were coded to 080101.

4.8 Validation

The role of validation in the processing system is to ensure that the data produced, and released, meets the requirements of the users. This role is carried out by checking the data produced by the system to ensure that it meets the stated output requirements, and identifying and correcting the errors that occur. When the source of the error is identified, that part of the system that is generating the error is reviewed for the most suitable method of correction. In some cases, a procedural correction may be more appropriate than a system update.

4.8.1 The Validation Process

The main objectives of the validation process are to:

- identify system procedures that affect data (such as edits and legal values) and provide specifications for updating the system as part of the continuous data quality improvement strategy of the Census program; and

- ensure data quality for Census output by checking all variables for accuracy and consistency, so that final data from the DPC meets the requirements of all relevant stakeholders.

Before data is released from the DPC, all output variables are validated, including a check for intercensal comparability between 1996 and 2001, with investigation of any changes outside specified tolerance levels. There are a number of cases where census variables are subject to imputation when no response is given by the respondent (including Age, Marital Status, State, Statistical Local Area of Usual Residence on Census Night and Labour Force Status).

The variables level, main field and year of completion of highest non-school qualification were output variables and were subject to this validation strategy. However, the non-school qualification indicator question was not identified as an output variable and was therefore not validated. This variable was retained for internal use only for subsequent census phases such as the evaluation and development program for the 2006 Census.

As discussed in more detail in Section 6.5.3 *(Measurement of persons holding a non-school qualification),* it is recommended that the 2006 Census Validation Strategy change the non-school qualification indicator question to an output variable primarily because reliance on level of highest non-school qualification alone significantly understates the proportion of persons holding a non-school qualification. This is because a number of respondents completed the non-school qualification indicator question, but either did not respond to level of highest non-school qualification, or provided a response which could only be coded to 'inadequately described'. This information could be used to provide a derivation of *'level not determined'* instead of 'Not-Stated'; or provide further information to get a better outcome than 'Inadequately Described'.

5. SAMPLE DATA ANALYSIS

A 2% sample statistically derived of CDs (approximately 760) from each State and Territory in Australia, representing a range of urban and rural CDs; and two smaller samples, focused on Indigenous, and Homeless populations, were identified for 2001. Using these samples, Data Quality Investigation (DQI) tasks were carried out at the 2001 DPC, directly related to the areas for which in-depth investigations were planned. The resulting data quality information is made available to clients in Census Papers and other related publications, and through analysis provided via the Census query service.

This section provides analyses the data quality of the non-school qualification indicator, focusing on:

- the frequency of multiple marks for the non-school qualification indicator question, which had a selection list and mark box format; and

- the response type (stated or not-stated) to level and main field of highest non-school qualifications of those that indicated in the non-school qualification indicator that they did not hold a non-school qualification or were still studying for their first qualification.

The remaining three questions discussed in this paper (level, main field and year of completion of highest non-school qualification) will not be analysed in detail because these were only assigned a stated/not stated data label in the DQI dataset because the format for these questions was a write-in response. Consequently, no further insights can be obtained beyond those already addressed in *Section 6 - Final Data Analysis* of this paper.

Given the close proximity of highest non-school qualification questions to the age banner (question 24 on the form) and the instructions to continue completing the remaining questions of the 2001 Census Form for persons aged 15 years or more, the proportion of stated responses to questions on highest non-school qualifications which are completed for persons who are under the age of 15 will be presented.

5.1 Multiple Marks: Non-School Qualification Indicator Question

Table 8 shows that based in the DQI sample, multiple marking of question 26 was found in approximately 0.6% of the DQI records.

TABLE 8: NUMBER AND FREQUENCY OF SINGLE AND MULTIPLE MARKS FOR THE NON-SCHOOL QUALIFICATION INDICATOR QUESTION, based on 2001 DQI sample (a)

	Frequency	*Percentage (%)*
Single Mark	266,573	92.0
Multiple Mark	1,724	0.6
Non Response	21,418	7.4
Total	**289,715**	**100.0**

(a) Table excludes: Persons aged under 15 years and those that did not state their age.

Table 9 shows that of the three-quarters of the multiple marks related to those that reported that they had both a 'trade certificate/apprenticeship' and 'other qualification'. During the scanning and DC of this question, the question was imaged and the first mark (ie, 'trade certificate/ apprenticeship') was captured by DC. Although the percentage of records for this anomaly was relatively small (0.4% of total records), there are some data quality issues if the respondent actually holds a non-school qualification at a higher level than a 'trade certificate/apprenticeship' (such as a 'Bachelor degree').

TABLE 9: BREAKDOWN OF MULTIPLE MARKS FOR THE NON-SCHOOL QUALIFICATION INDICATOR QUESTION, based on 2001 DQI sample (a)

	Frequency	*Percentage of Respondents Aged 15 Years or Over (%)*	*Percentage of Multiple Marks (%)*
'No' and 'No, still studying for first qualification'	109	0.0	6.3
'No' and 'Yes, trade certificate/apprenticeship'	64	0.0	3.7
'No' and 'Yes, other qualification'	191	0.1	11.1
'No, still studying for first qualification' and 'Yes, trade certificate/apprenticeship'	22	0.0	1.3
'No, still studying for first qualification' and 'Yes, other qualification'	38	0.0	2.2
'Yes, trade certificate/apprenticeship' and 'Yes, other qualification'	1,294	0.4	75.1
'No', 'Yes, trade certificate/apprenticeship' and 'Yes, other qualification'	1	0.0	0.0
'No, still studying for first qualification', 'Yes, trade certificate/apprenticeship' and 'Yes, other qualification'	2	0.0	0.0
'No', 'No, still studying for first qualification', 'Yes, trade certificate/apprenticeship' and 'Yes, other qualification'	3	0.0	0.0

(a) Table excludes: Persons aged under 15 years and those that did not state their age.

5.2 Non-Response to Level and Main Field of Highest Non-School Qualifications against the response options of the Non-School Qualification Indicator Question

Section 6.2 (Non-Response Analysis) includes at Table 18 an analysis of the frequency of response patterns to the non-school qualification indicator question against the response patterns of level, main field and year of completion of highest non-school qualification questions. This analysis concentrates on 'Yes' and 'Not-Stated' responses to the non-school qualification indicator question because the remaining response options ('No' and 'Still studying for first qualification') had values for level, main field and year of completion set to 'not applicable'.

The response characteristcs of persons who selected either 'No' or 'Still studying for first qualification' for the non-school qualification indicator question can be analysed using the 2001 DQI sample to provide some indication on whether:

- the sequencing instructions (not to complete the remaining questions relating to highest non-school qualifications) are being followed by respondents; and

- there could be potential for situations where respondents who correctly indicated in the level question that they hold a non-school qualification but responded incorrectly to the non-school qualification indicator question, resulting in their non-school qualification being not counted in the 2001 Census.

Respondents who did not provide a response to the year of completion of highest non-school qualification have been eliminated from analyses in this section on the assumption that while they may have provided a response to level of highest non-school qualification, it is likely that they have not completed their first non-school qualification.

Table 10 shows that between 2.2% and 3.1% of persons who indicated that they did not hold a non-school qualification also provided a response to level and year of completion of highest non-school qualification. Slightly lower percentages were found to provide a response to main field of highest non-school qualification.

TABLE 10: RESPONSE FREQUENCY TO LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION FOR RESPONDENTS WHO INDICATED IN THE INDICATOR QUESTION THAT THEY DID NOT HOLD A NON-SCHOOL QUALIFICATION, BY AGE - based on 2001 DQI sample(a)

	Response to Level of Highest Non-School Qualification						
	'Stated'					*Percentage of 'Stated' respondents (b)*	*'Stated' Respondents as a % of Age Level (b)*
Age group (years)	*Year 'Stated'*	*Year 'Not-Stated'*	*Total*	*'Not-Stated'*	*Total*		
15-24	573	154	727	25,481	26,208	2.2	1.1
25-34	688	132	820	21,112	21,932	3.1	1.3
35-44	662	155	817	24,227	25,044	2.6	1.2
45-54	687	161	848	23,340	24,118	2.8	1.4
55-64	526	133	659	18,895	19,554	2.7	1.5
65-74	336	132	468	14,661	15,129	2.2	1.3
75-84	254	64	318	9,224	9,542	2.7	1.6
85 or older	71	15	86	3,001	3,087	2.3	1.4
Total	**3,797**	**946**	**4,743**	**139,941**	**144,684**	**2.6**	**1.3**

(a) Table excludes: Persons aged under 15 years and those that did not state their age.
(b) Excludes respondents that did not respond to the question: year of completion of highest non-school qualification.

Table 11 shows that, overall, 4.2% of persons who indicated that they were still studying for their first non-school qualification also provided a response to level and year of completion of highest non-school qualification. While this percentage is higher than for respondents who indicated that they did not hold a non-school qualification, the percentage, expressed as a total of all respondents in each age group, is much lower.

TABLE 11: RESPONSE FREQUENCY TO LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION FOR RESPONDENTS WHO INDICATED IN THE INDICATOR QUESTION THAT THEY WERE STILL STUDYING FOR THEIR FIRST QUALIFICATION, BY AGE - based on 2001 DQI sample(a)

	Response to Level of Highest Non-School Qualification						
	'Stated'					*Percentage of 'Stated' respondents (b)*	*'Stated' Respondents as a % of Age Level (b)*
Age group (years)	*Year 'Stated'*	*Year 'Not-Stated*	*Total*	*'Not-Stated'*	*Total*		
15-24	366	190	556	9,443	9,999	3.7	0.7
25-34	77	28	105	1,276	1,309	5.9	0.2
35-44	37	11	48	613	661	5.6	0.1
45-54	23	6	29	257	286	8.0	0.1
55-64	8	4	12	56	68	11.8	0.0
65-74	3	3	6	29	35	8.6	0.0
75-84	5	0	5	24	29	17.2	0.0
85 or older	1	0	1	13	14	7.1	0.0
Total	**520**	**242**	**762**	**11,711**	**12,401**	**4.2**	**0.2**

(a) Table excludes: Persons aged under 15 years and those that did not state their age.
(b) Excludes respondents that did not respond to the question: year of completion of highest non-school qualification.

The findings shown in Table 10 and 11 indicate that while there may be some confusion about the sequencing instructions and/or the question content of the non-school qualification indicator question, this data anomaly was found for a relatively small proportion of respondents. Further analysis of this question, including assertions that it is confusing to respondents, is provided in *Section 6.2.3 (Non-Response to the Non-School Qualification Indicator against the response patterns of the other Highest Non-School Qualification Questions)*.

5.3 Responses by persons aged under 15 years

Table 12 shows that the majority of persons aged under 15 years did not respond to the questions on highest non-school qualifications. This problem was most associated with the HHF, with 99% of erroneous reporting occuring on this form type. While 17.2% of persons aged under 15 years did respond to the non-school qualification indicator, the majority indicated that they did not hold a qualification.

TABLE 12: FREQUENCY OF PERSONS AGED UNDER 15 COMPLETING THE NON-SCHOOL QUALIFICATION QUESTIONS, based on 2001 DQI sample(a)

Highest Non-School Qualification Question	*Persons aged under 15 years*	*Persons for whom there was a response*	*Percentage (%)*
Indicator	75,660	12,987	17.2
Level	75,660	363	0.5
Main Field	75,660	285	0.4

(a) Table excludes: Persons aged 15 years or over and those that did not state their age.

Table 13 shows that the proportion of respondents under the age of 15 incorrectly completing the questions relating to highest non-school qualifications is lower than questions occurring on subsequent pages of the 2001 Census Form (which do not include an age banner). Consistent with Table 12, this was mostly associated with the HHF. On this basis, the age banner seems to have a mixed impact on reporting errors, with the greatest impact being for the questions located on the same page as the age banner.

TABLE 13: FREQUENCY OF PERSONS AGED UNDER 15 COMPLETING QUESTIONS ON THE NATURE OF JOBS UNDERTAKEN, based on 2001 DQI sample (a)

Question	*Persons aged under 15 years*	*Persons for whom there was a response*	*Percentage (%)*
Full-time/Part-time Job	75,660	21,024	27.8
Job Last Week	75,660	609	0.8

(a) Table excludes: Persons aged 15 years or over and those that did not state their age.

6. FINAL DATA ANALYSIS

6.1 Overview of Data Analysis Methodology

6.1.1 Introduction

The following section contains an analysis of the 2001 Census questions on highest non-school qualifications. Evaluation will focus on:

- non-response and insufficient response;
- implementation of ASCED;
- intercensal changes and comparability; and
- comparisons with results of other statistical collections.

6.1.2 Removal of selected records

SCRs, which have values imputed for age, sex, marital status and usual residence have been eliminated from all analyses of 1996 and 2001 Census data because values for questions on highest non-school qualifications were set to either not stated or not applicable, depending on the imputed value for age. The removal of SCRs means that:

- some tables using 1996 Census data, such as the non-response rate tables, are not comparable with the previous Census Working Paper: *1996 Census Data Quality: Qualification Level and Field of Study* (00/2), because SCRs were not excluded;
- comparative non-response rates differ from those published in the *2001 Census of Population and Housing Fact Sheet: Non-Response Rates, 2001* which calculate non-response rates including SCRs; and
- population counts, and therefore, proportion of persons with a non-school qualification differ from those published in *Census of Population and Housing, Selected Education and Labour Force Characteristics, Australia, 2001* (cat. no. 2017.0) which included SCRs.

Other record types which have been eliminated from all analyses include:

- overseas visitors; and
- forms where questions relating to highest non-school qualifications were not asked, including Substitute forms, Summary forms, Special Short forms and E-files.

Additional records, such as those persons still attending school, but holding a non-school qualification, have been eliminated from intercensal analyses (for more detail, refer to *Section 6.6.2 - Comparison Methodology)* to reduce the population to a similar scope to that of the ABSCQ. Similarly, comparisons with the SEW have eliminated records (such as persons aged 65 years and over) from Census data to reduce the Census Population to the same scope as the SEW (for more detail, refer to *Section 6.7.2 - Reconciliation of 2001 Census Data with the SEW*).

6.1.3 Calculation of proportions

Section 6.2 highlights the fact that the 2001 Census data on level, main field and year of completion of highest non-school qualifications have high non-response rates (between 3.8% and 7.8%) when compared to other Census variables and other education surveys, such as the SEW (refer to *Section 6.7 - Comparisons with the ABS 2001 Survey of Education and Work*).

The inclusion of a large number of 'not-stated' responses in the population base of calculations such as the proportion of persons holding a non-school qualification significantly underestimates these calculations. Therefore, in the following sections of this paper, 'not-stated' responses have been removed from the population base of all calculations involving proportions. The effect of excluding a large number of 'not-stated' responses in such calculations is shown in Table 14 below, which utilises data presented in Appendix D2.

TABLE 14: PROPORTION OF PERSONS AGED 15-64 WITH A NON-SCHOOL QUALIFICATION: INCLUDING AND EXCLUDING 'NOT-STATED' RESPONSES (a), 2001 CENSUS

	Excluding 'not-stated' responses			*Including 'not-stated' responses*			*Difference*
State/ Territory	*Persons with non-school qualification ('000) (b)*	*Total Persons ('000) (c)*	*Per cent*	*Persons with non-school qualification ('000) (b)*	*Total Persons ('000) (d)*	*Per cent*	*Percentage points*
NSW	1,767.0	3,766.2	46.9	1,767.0	3,981.6	44.4	2.5
Vic.	1,246.6	2,807.2	44.4	1,246.6	2,973.9	41.9	2.5
Qld	871.2	2,146.7	40.6	871.2	2,254.8	38.6	1.9
SA	354.6	881.9	40.2	354.6	922.2	38.5	1.8
WA	484.8	1,094.2	44.3	484.8	1,145.5	42.3	2.0
Tas.	103.5	271.4	38.1	103.5	285.9	36.2	1.9
NT	40.9	88.6	45.7	40.9	95.0	43.1	2.6
ACT	104.1	191.9	54.3	104.1	200.8	51.8	2.4
Aust.	**4,972.9**	**11,249.1**	**44.2**	**4,972.9**	**11,861.8**	**41.9**	**2.2**

(a) Table excludes: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) Derived using either of the 'Yes' responses to the non-school qualification indicator question.

(c) Total excludes 'not stated' responses to the non-school qualification indicator in the population base.

(d) Persons for whom the non-school qualification indicator was relevant.

6.2 Non-Response Analysis

6.2.1 Initial Analysis

The high non-response rates for the questions relating to the highest non-school qualification have a serious bearing on overall data quality. Table 15 contains the 1996 and 2001 non-response rates for the highest non-school qualification questions for persons for whom questions were relevant, after the elimination of SCRs. The non-school qualification indicator question was relevant to all persons aged 15 years of more, whereas questions on level, main field and year of completion of highest non-school qualification were deemed to be relevant to persons aged 15 years or more except those who either did not respond to the non-school qualification indicator question or indicated that they did not have a non-school qualification.

TABLE 15: NON-RESPONSE RATES FOR HIGHEST NON-SCHOOL QUALIFICATION QUESTIONS (a) - 1996 AND 2001 CENSUS

	2001 Census			*1996 Census (c)*
Highest Non-School Qualification Question	*Persons for whom Questions were relevant (b)*	*Persons for whom there was no response*	*Non-response Rate (%)*	*Non-response Rate (%)*
Indicator	14,485,766	1,030,409	7.1	6.5
Level	5,455,279	300,636	5.5	10.9
Main Field	5,455,279	207,773	3.8	3.9
Year of Completion	5,455,279	424,031	7.8	4.4

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.
(b) For level, main field and year of completion of highest non-school qualification, persons who indicated that they held a non-school qualification (ie, responded 'Yes, trade certificate/apprenticeship' or 'Yes, other qualification' to the non-school qualification indicator question).
(c) VET Certificate Level 1 qualifications were outside the scope of the 1996 Census.

Table 15 shows that 7.1 % of respondents aged 15 years or over failed to provide a response on whether or not they held a non-school qualification, a small increase from the 1996 Census. The non-response rate for main field of highest non-school qualification remained unchanged, however the non-response rate for year of completion of highest non-school qualification nearly doubled, increasing from 4.4 % to 7.8 %. While the non-response rate for level of highest non-school qualification halved between Censuses, this was partially offset by an increase in the proportion of responses which could only be coded to 'inadequately described' (refer to Section 6.3 for further details).

Non-response rates for questions relating to highest non-school qualifications are significantly higher than labour force status, which had a non-response rate of 2.0 % and highest level of schooling completed, which had a non-response rate of 5.4 % (refer to Appendix A).

The exclusion of non-responses to the non-school qualification indicator from Table 15 tends to underestimate non-response rates because some of these persons may have held a non-school qualification.

The calculation of adjusted non-response rates, which includes not only those that responded to the non-school qualification indicator question, but also those that did not respond to the non-school qualification indicator question reveals significantly larger non-response rates for level, main field and year of completion of highest non-school qualification (refer to Table 16). This increase in non-response is significantly higher than equivalent calculations for questions (such as Occupation) which rely on the completion of the question on whether the person had a full-time or part-time job of any kind in the previous week (refer to Appendix A).

TABLE 16: ADJUSTED NON-RESPONSE RATES FOR HIGHEST NON-SCHOOL QUALIFICATION QUESTIONS (a) - 1996 AND 2001 CENSUS

	2001 Census			*1996 Census (d)*
Highest Non-School Qualification Question	*Persons for whom Questions could have been relevant (b)*	*Persons for whom there was no response (c)*	*Adjusted Non-response Rate (%)*	*Adjusted Non-response Rate (%)*
Level	6,485,688	1,142,144	17.6	23.7
Main Field	6,485,688	1,015,299	15.7	17.4
Year of Completion	6,485,688	1,223,760	18.9	17.4

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.
(b) Includes all responses to non-school qualification indicator except those who indicated that they did not have a non-school qualification plus those that did not respond to the non-school qualification indicator question.
(c) Persons who did not respond to level, main field or year of completion of highest non-school qualification and either indicated that they held a non-school qualification or did not respond to the non-school qualification indicator question.
(d) VET Certificate Level 1 qualifications were outside the scope of the 1996 Census.

Approximately one in five respondents who may have held a non-school qualification failed to provide information on the year of completion, an increase when compared to the 1996 Census. Reasons for this increase are explored in the following section.

6.2.2 Year of Completion of Highest Non-School Qualification

Changes to Questionnaire design from a selection list to a write-in box is likely to be the main contributor which lead to the approximate doubling in non-response rates between the 1996 and 2001 Census for the question year of completion of highest non-school qualification. This is because respondents (particularly those in higher age groups) may have experienced increased difficulties in recalling the precise year in which they completed their highest non-school qualification.

This assertion is supported in Table 17, which reveals that only 12 per cent of relevant respondents (those respondents that indicated that they held a non-school qualification) who failed to respond to the year of completion of highest non-school qualification also failed to respond to both the level and main field of highest non-school qualification questions. Therefore, the majority of these non-respondents could complete at least one of the two questions on level and main field of highest non-school qualification, but could not complete the question on year of completion of highest non-school qualification.

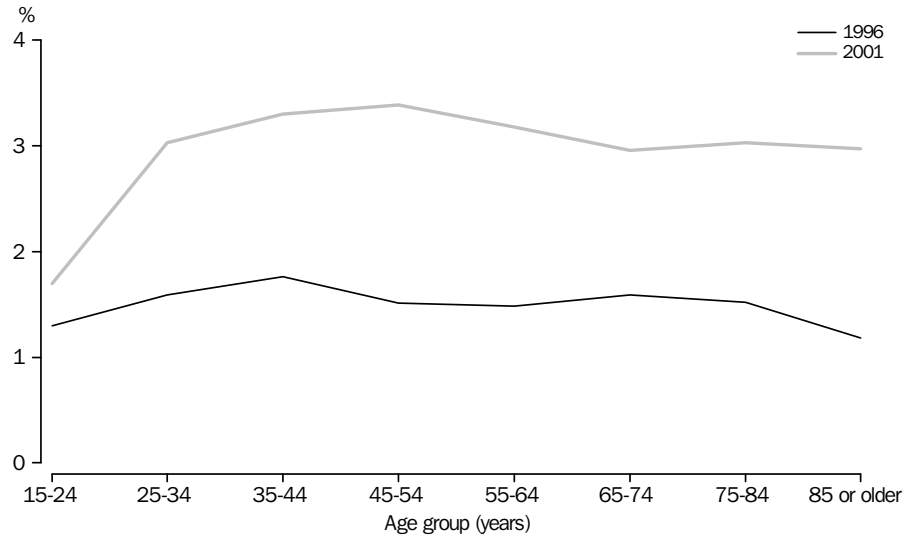
TABLE 17: DISTRIBUTION OF NON-RESPONSES, YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 2001 CENSUS

Respondent's responses to Level and Main Field of Highest Non-School Qualification (b)	*Non-response Count*	*Proportion of Total Non-Response (%)*
Both Level and Main Field stated	257,362	60.7
Level stated but Main Field not stated	82,317	19.4
Main Field stated but Level not stated	33,650	7.9
Neither Main Field or Level Stated	50,702	12.0
Total Non-Response	**424,031**	**100.0**

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.
(b) Includes all respondents who failed to complete the question on year of completion of highest non-school qualification and who indicated that they held a non-school qualification.

The hypothesis that respondents in higher age groups experienced increased difficulty in completing this question on the 2001 Census Form is supported by Figure 5, which shows an increased tendency towards non-response between the 1996 and 2001 Census as age increases.

FIGURE 5: PROPORTION OF THE POPULATION THAT DID NOT RESPOND TO YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION, BY AGE (a) - 1996 (b) AND 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.
(b) VET Certificate Level 1 qualifications were outside the scope of the 1996 Census.

6.2.3 Non-Response of the Non-School Qualification Indicator against the response patterns of the other Highest Non-School Qualification questions

Table 18 shows the frequency of 'Yes' and 'Not-Stated' responses to the non-school qualification indicator question against the response patterns of level, main field and year of completion of highest non-school qualification questions. Detailed analysis of 'No' and 'Still studying for first qualification' responses are provided in *Section 5.2*.

Consistent with the 1996 Census, if a respondent completed the non-school qualification indicator question, there was a high probability (87.2%) that all questions on level, main field and year of completion of highest non-school qualification were answered.

TABLE 18: FREQUENCY OF 'STATED' RESPONSES FOR LEVEL, MAIN FIELD AND YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION, BY RESPONSE TO THE NON-SCHOOL QUALIFICATION INDICATOR (a) - 2001 CENSUS

	Response to Non-School Qualification Indicator			
Number of 'Stated' Responses to Other Highest Non-School Qualification Variables	*Yes*	*% of 'Yes' Respondents*	*Not Stated*	*% of 'Not Stated' Respondents*
1996 Census (b):				
0	95,549	2.0	762,434	85.2
1	80,834	1.7	39,849	4.5
2	463,153	9.8	28,679	3.2
3	4,109,527	86.5	63,993	7.2
Total	**4,749,063**	**100.0**	**894,955**	**100.0**
2001 Census:				
0	50,702	0.9	768,137	74.5
1	131,249	2.4	48,412	4.7
2	517,836	9.5	47,528	4.6
3	4,755,492	87.2	166,332	16.1
Total	**5,455,279**	**100.0**	**1,030,409**	**100.0**

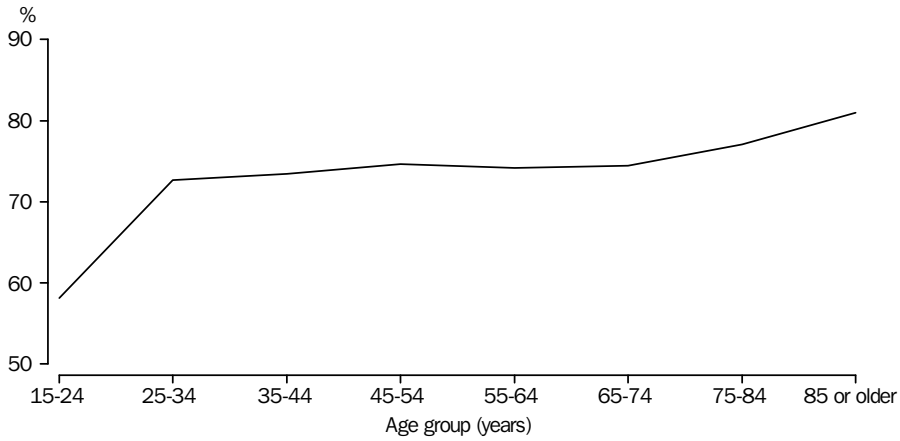
(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) VET Certificate Level 1 qualifications were outside the scope of the 1996 Census.

However, there was a doubling between Censuses in the percentage of respondents that failed to answer the non-school qualification indicator question but completed all questions relating to level, main field and year or completion of highest non-school qualification. On this basis, the DPC coded all responses to level, main field and year of completion of highest non-school qualification, irrespective of whether the non-school qualification indicator question had been answered.

This question *('has the person completed a trade certificate or any other educational qualification?')* may be interpreted by respondents as seeking information primarily about trade certificates only and not other non-school qualifications, such as Bachelor degrees. This hypothesis is supported in Figure 6, which shows that the majority of respondents that failed to answer the non-school qualification indicator question but responded to level, main field and year of completion of highest non-school qualification had completed a formal qualification (such as a Bachelor degree).

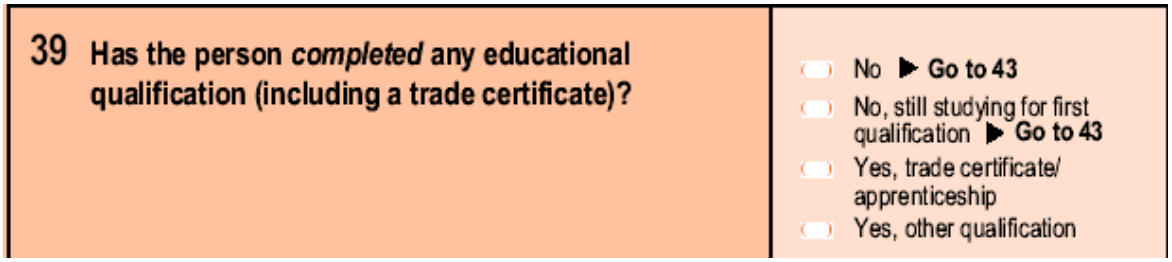
FIGURE 6: PROPORTION OF NON-RESPONSES TO THE NON-SCHOOL QUALIFICATION INDICTOR ONLY WHO HELD A DIPLOMA OR HIGHER, BY AGE (a) - 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

Given this pattern, it is considered that question design changes would improve response rates for this question. One alternative question wording, which is currently being tested in the August 2003 Census Test, is *'Has the person completed any educational qualification (including a trade certificate)?'* (refer to Figure 7).

FIGURE 7: 2003 AUGUST CENSUS TEST HOUSEHOLD FORM (NON-SCHOOL QUALIFICATION INDICATOR QUESTION)



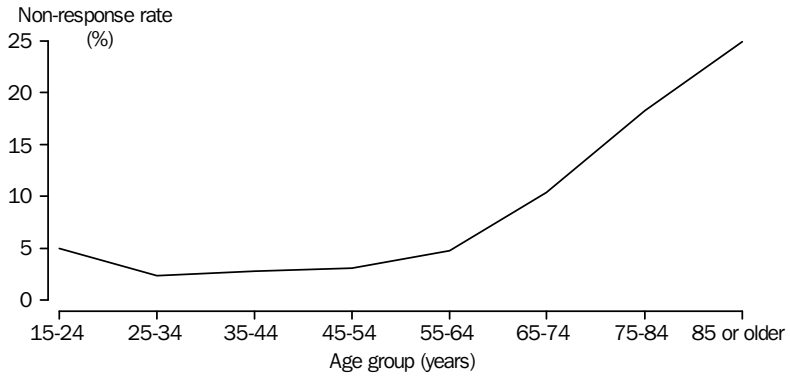
An alternative proposal would be to use the standard question module for interviewer-based surveys as the question wording for the Census, instead of the current use of the self-enumerated module. This standard wording is: *"Has the person completed a trade certificate, diploma, degree or any other educational qualification?'*.

6.2.4 Characteristics of Not-Stated Responses

A more detailed analysis of the characteristics of respondents that did not respond to any non-school qualification question is warranted to examine the correlation between non-response and the probability of sub-groups of the population holding a non-school qualification. If an inverse correlation exists, it can be assumed that the majority of non-respondents did not hold a formal qualification and, therefore, there is a large over-correction in the calculation of adjusted non-response rates.

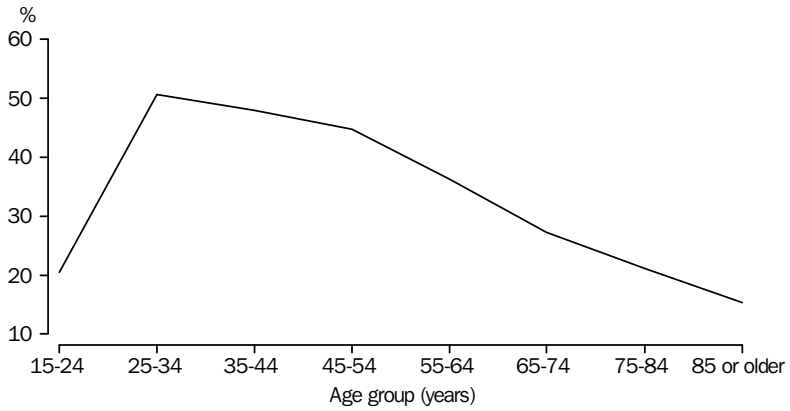
Figures 8 and 9 show that non-response increases with the age of the population, while the proportion of the population with a non-school qualification decreases.

FIGURE 8: NOT-STATED RESPONSES FOR NON-SCHOOL QUALIFICATION QUESTIONS, BY AGE (a) - 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

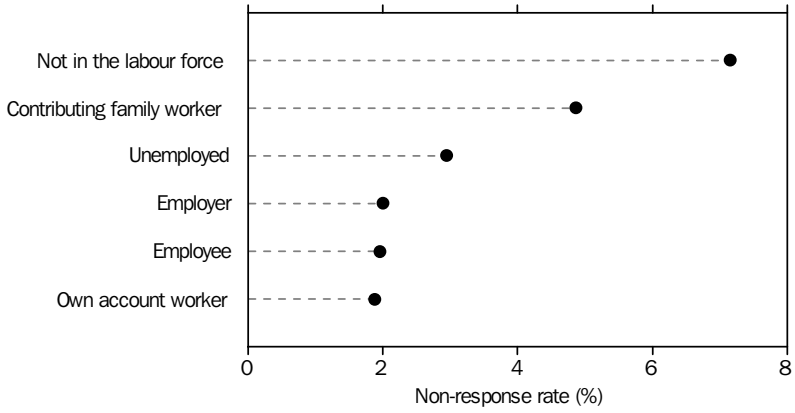
FIGURE 9: PROPORTION OF PERSONS WITH A NON-SCHOOL QUALIFICATION, BY AGE (a) - 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors. Population base excludes 'not stated' responses to non-school qualification indicator.

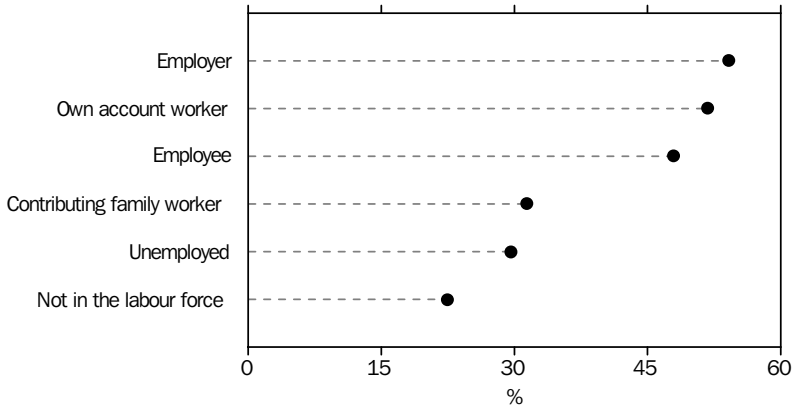
Figures 10 and 11 show that persons not in the labour force had the highest non-response of all persons who stated their labour force status and the lowest proportion of persons with a non-school qualification.

FIGURE 10: NOT-STATED RESPONSES FOR NON-SCHOOL QUALIFICATION QUESTIONS, BY LABOUR FORCE STATUS (a) - 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

FIGURE 11: PROPORTION OF PERSONS WITH A NON-SCHOOL QUALIFICATION, BY LABOUR FORCE STATUS (a) 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors. Population base excludes 'not stated' responses to non-school qualification indicator.

Given that there is a large over-correction in the calculation of adjusted non-response rates, an alternative method to calculate adjusted non-response rates is presented, which increases the non-response rate by an estimate of the likely number of relevant responses (termed 'likely non-response rates'), based on the proportion of the stated population holding a non-school qualification. This provides a more realistic estimate of total non-response and the degree of potential understatement in the Census figures attributable to non-response.

Table 19 shows the calculation of 'likely non-response rates' for the level and main field of highest non-school qualification, based on age. Using this approach, a non-response rate of 8.7% for level of highest non-school qualification was obtained, whereas a smaller rate (7.1%) was obtained for main field of highest non-school qualification.

TABLE 19: LIKELY NON-RESPONSE FOR LEVEL AND MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION, BY AGE (a) - 2001 CENSUS

Age group (years)	*Persons for whom Questions were relevant (b)*	*Relevant persons for whom there was no response (c)*	*Persons for whom there could have been a response (d)*	*Proportion of persons holding a non-school qualification (e)*	*Persons for whom there would have likely been a response (d*e) (f)*	*Estimated non-response (c + f)*
Level of Highest Non-School Qualification:						
15-24	658,141	28,762	132,400	20.5	27,078	55,840
25-34	1,443,095	49,541	74,366	50.7	37,718	87,259
35-44	1,471,928	60,727	88,409	48.0	42,455	103,182
45-54	1,252,733	60,174	87,601	44.7	39,162	99,336
55-64	743,026	42,844	93,191	36.3	33,790	76,634
65-74	485,749	31,289	139,546	27.3	38,037	69,326
75-84	325,729	21,300	158,333	21.1	33,432	54,732
85 or older	105,287	5,999	67,662	15.3	10,356	16,355
Total	**6,485,688**	**300,636**	**841,508**	. .	**262,028**	**562,664**
Main Field of Highest Non-School Qualification:						
15-24	658,141	22,798	129,462	20.5	26,477	49,275
25-34	1,443,095	35,254	70,647	50.7	35,832	71,086
35-44	1,471,928	42,456	83,709	48.0	40,198	82,654
45-54	1,252,733	40,473	82,535	44.7	36,897	77,370
55-64	743,026	27,639	87,839	36.3	31,850	59,489
65-74	485,749	21,172	134,121	27.3	36,558	57,730
75-84	325,729	14,695	153,176	21.1	32,343	47,038
85 or older	105,287	3,286	66,037	15.3	10,107	13,393
Total	**6,485,688**	**207,773**	**807,526**	. .	**250,262**	**458,035**

. . Not applicable.

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) Includes all persons except those who indicated that they did not have a non-school qualification.

(c) Persons who indicated that they held a non-school qualification, but did not respond to level or main field of highest non-school qualification.

(d) Persons who did not respond to both non-school qualification indicator and either level or main field of highest non-school qualification.

(e) Population base excludes 'not stated' responses to non-school qualification indicator in the denominator.

6.3 Analysis of 'Inadequately Described' Responses

A number of responses to level and main field of highest non-school qualification were classified to 'inadequately described' because respondents provided insufficient information to enable full coding. Table 20 shows that the percentage of responses coded to 'inadequately described' is higher for level of highest non-school qualification than main field of highest non-school qualification.

TABLE 20: RESPONSES TO LEVEL AND MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION WHICH WERE CODED TO 'INADEQUATELY DESCRIBED' (a) - 1996 AND 2001 CENSUS

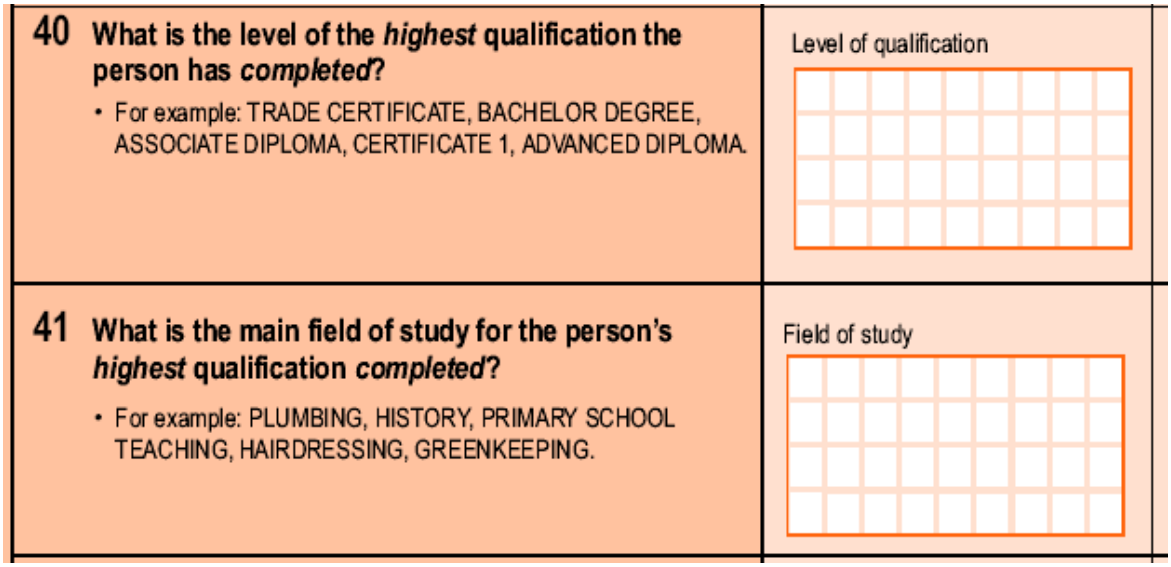
Highest Non-School Qualification Question	*Persons for whom Questions was relevant (b)*	*Responses coded to inadequately described*	*Percentage 'inadequately described' (%)*
Level			
1996 Census (c)	4,748,780	121,385	2.5
2001 Census	5,455,279	176,750	3.2
Main Field			
1996 Census (c)	4,748,780	53,177	1.1
2001 Census	5,455,279	57,880	1.1

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.
(b) Persons who indicated that they held a non-school qualification (ie, responded 'Yes, trade certificate/apprenticeship' or 'Yes, other qualification' to the non-school qualification indicator question).
(c) VET Certificate Level 1 qualifications were outside the scope of the 1996 Census.

The allotted space of three rows of boxes may be one factor that contributed to the higher percentage of responses to level of highest non-school qualification which could only be coded to 'inadequately described' (four rows of boxes were assigned to main field of highest non-school qualification question). However other factors, such as the quality of information provided by respondents, also were contributing factors (refer to Section 6.4 for more details).

The ABS will further analyse the effect of such space considerations on data quality using the results obtained from the August 2003 Census Test, which has four rows of boxes for both questions (refer to Figure 12).

FIGURE 12: 2003 AUGUST CENSUS TEST HOUSEHOLD FORM (LEVEL AND MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION)



6.4 'Inadequately Described' and 'Not-Stated' Responses to Level of Highest Non-School Qualification

Table 21 shows responses to the main field of highest non-school qualification question that either did not respond or provided a response to level of highest non-school qualification which could only be coded to 'inadequately described'. Some main fields, including 'Health' and 'Management and commerce' are highly problematic for coders and have a high proportion of responses to level of highest non-school qualification which were only coded to 'inadequately described'.

TABLE 21: FREQUENCY OF 'INADEQUATELY DESCRIBED' AND 'NOT-STATED' RESPONSES TO LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION, BY MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 1996 AND 2001 CENSUS

	Response to Level of Highest Non-School Qualification			
	'Inadequately Described'		*'Not-Stated'*	
Main Field of Highest Non-School Qualification	*Number of Responses*	*Percentage of Responses*	*Number of Responses*	*Percentage of Responses*
2001 Census:				
Natural and physical sciences	2,960	1.5	4,087	2.1
Information technology	4,003	2.6	6,784	4.4
Engineering and related technologies	19,534	1.5	38,231	3.0
Architecture and building	6,213	1.5	10,375	2.6
Agriculture, environmental and related studies	3,027	2.2	7,121	5.1
Health	42,870	7.2	36,881	6.2
Education	7,806	1.5	15,641	3.1
Management and commerce	63,033	6.0	97,881	9.3
Society and culture	16,791	3.0	20,708	3.7
Creative arts	7,337	3.7	14,474	5.3
Food, hospitality and personal services	3,076	1.0	15,175	4.8
Mixed field programmes	910	19.8	1,765	38.4
2001 Census - Total	**177,560**	**3.3**	**265,123**	**4.9**
1996 Census (b) - Total	**120,847**	**2.6**	**392,969**	**8.5**

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) VET Certificate Level 1 qualifications were outside the scope of the 1996 Census.

In the previous Census Working Paper: *1996 Census Data Quality: Qualification Level and Field of Study* (00/2), it was hypothesised that the high prevalence of responses to field of study of highest post-school qualification which either had a 'not-stated' or 'inadequately described' response to level of highest post-school qualification was caused by *'respondents reporting qualifications* (such as first aid/medical certificates, typing/keyboard courses and basic computing/word processing courses*) which do not lie within the ABSCQ definition of a post-school qualification'*.

A detailed breakdown of responses using 2001 Census data revealed that similar main fields of highest non-school qualification had the highest frequency of 'not-stated' and 'inadequately described' responses to level of highest non-school qualification (refer to Table 22).

TABLE 22: TOP 10 FREQUENCY OF 'INADEQUATELY DESCRIBED' AND 'NOT-STATED' RESPONSES TO LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION, BY DETAILED MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION (a) 2001 CENSUS

'Inadequately Described' Response to Level of Highest Non-School Qualification		*'Not-Stated' Response to Level of Highest Non-School Qualification*	
Detailed Main Field of Highest Non-School Qualification	*Number of Responses*	*Detailed Main Field of Highest Non-School Qualification*	*Number of Responses*
Secretarial and clerical studies	26,237	Secretarial and clerical studies	45,190
General nursing	23,239	General nursing	24,008
Business and management, nfd	8,358	Accounting	11,697
Accounting	7,415	Business and management, nfd	9,293
Dental assisting	4,410	Teacher education: Primary	5,864
Real estate	3,741	Information technology, nfd	5,513
Children's services	3,093	Hospitality	5,382
Security services	3,068	Management and commerce, nfd	4,832
Information technology, nfd	2,703	Engineering and related technologies, nfd	4,415
Plant and machine operations	2,507	Teacher education, nfd	4,146

(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

A detailed case study of data and images from the ACT was conducted during Second Release Processing at the DPC and disproved the previous hypothesis made about the likely reporting of qualifications outside the scope of the Census. For example, nearly all of the Secretarial and clerical studies responses to level of highest non-school qualification which were coded as 'inadequately described' were described by respondents as 'Trade Certificate'.

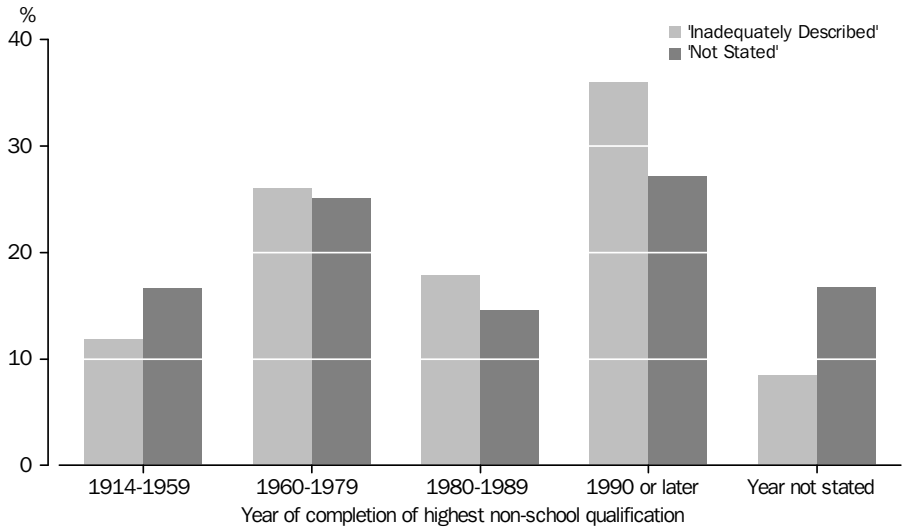
Secondly, the following characteristics were found for 'General nursing' responses to main field of highest non-school qualification which had a level of highest non-school qualification coded to 'inadequately described':

- most described their qualification level as either 'Trade Certificate' or 'Trained Nurse'; and

- nearly all responses stated a hospital as the non-school qualification institution and all were obtained prior to 1990.

The high prevalence of highest non-school qualifications obtained prior to 1990 which were sufficiently described to field of highest non-school qualification but were either 'inadequately described' or 'not stated' to level of highest non-school qualification was common among many main fields (refer to Figure 13).

FIGURE 13: FREQUENCY OF 'INADEQUATELY DESCRIBED' AND 'NOT-STATED' RESPONSES TO LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION RESPONSES AND 'STATED' RESPONSES TO MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION, BY YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

In conclusion, a number of factors have contributed to the high prevalence of responses to main field of highest non-school qualification which either had a 'not-stated' or 'inadequately described' response to level of highest non-school qualification, including

- poor quality of information being provided by some respondents;
- the space assigned to level of highest non-school qualifications (three rows of boxes);
- illegal combinations for some main field and level of highest non-school qualifications being provided by respondents (for example, 'Trade Certificate' and 'Secretarial and clerical studies') which do not exist in current coding indexes; and
- insufficient guidance being provided to respondents (such as in the Census Guide) who have specific types of non-school qualifications (eg, 'General nursing') which are not easily coded.

6.5 Classification Issues

6.5.1 Implementation of a New Classification System, ASCED

After a lengthy review process, the ABSCQ, which was used to classify level and field of highest post-school qualification in the 1991 and 1996 Censuses, was replaced by the ASCED which was implemented for the 2001 Census. ASCED has a significantly broader scope than the former classification, and addressed significant structural changes which had occurred in the provision of education and training since 1996, including:

- the boundaries between the three major sectors (schools, VET and higher education) have become less distinct with developments such as a few universities offering programmes under the NTF, VET programmes being offered in schools and some Bachelor programmes being offered by VET institutions;

- an increase in the number of multi-sector institutions;

- greater emphasis on appropriate skills formation and entry level training, particularly in the VET sector; and

- industry involvement in the development of work based competency based training programs (including the introduction of the Australian Qualifications Framework by the Ministerial Council on Education, Employment, Training and Youth Affairs).

ASCED was developed as a national framework which also addressed a lack of comparability between data collected and used by the different education and training sectors and by the ABS.

ASCED has been introduced into the ABS collections the SEW, Survey of Education, Training and Information Technology, and the 2001 Census of Population and Housing from 2001 to classify level of education and field of education. ASCED has also been introduced by government agencies other than the ABS, enabling comparability of data across many collections. Qualification data classified according to ASCED can also be used for international comparisons, because of its alignment, as far as possible, with the International Standard Classification of Education.

6.5.2 Consistency and breaks in series

ASCED has a much broader scope than the ABSCQ because it is not directed specifically at only post-school qualifications. This has enabled the question modules for the 2001 Census and Education Surveys to be expanded to include persons still at school. In addition, the scope of ASCED includes Certificate Level 1, whereas ABSCQ excluded one semester courses such as Certificate Level 1. As a result, there are major series breaks between what was previously collected under ABSCQ and what is now collected under ASCED.

To enable meaningful time series comparisons, the SEW, Survey of Education, Training and Information Technology and the 2001 Census of Population and Housing have coded qualification level questions to both ASCED and ABSCQ.

The implementation of ASCED and the scope inclusion of persons still at school has brought about two distinct measurements of level of education: highest level of educational attainment and highest level of non-school qualification. The 2001 Census measures level of highest non-school qualification only, as opposed to highest level of educational attainment, which is only measured by surveys such as SEW and the Survey of Education, Training and Information Technology in addition to measuring level of highest non-school qualification.

In these above mentioned surveys, it is possible for qualification statistics to be derived differently when measured according to the standard classifications of highest level of educational attainment and highest level of non-school qualification. This is particularly evident in the case of Certificates I and II where level of highest educational attainment may be derived as Year 12 if that person has completed Year 12. However, as highest non-school qualification relates only to non-school qualifications, highest non-school qualification would be derived as Certificate I and II (refer to Figure 14).

FIGURE 14: ASCED DECISION TABLE

AUSTRALIAN STANDARD LEVEL OF EDUCATION (ASCED) CODES							
	Certificate not further defined (500)	**Certificate III or IV not further defined (510)**	**Certificate IV (511)**	**Certificate III (514)**	**Certificate I or II not further defined (520)**	**Certificate II (521)**	**Certificate I (524)**
Secondary Education not further defined (600)	Certificate not further defined	Certificate III or IV not further defined	Certificate IV	Certificate III	Certificate I or II not further defined	Certificate II	Certificate I
Senior Secondary Education not further defined (610)	Senior Secondary not further defined	Certificate III or IV not further defined	Certificate IV	Certificate III	Senior Secondary not further defined	Senior Secondary not further defined	Senior Secondary not further defined
Year 12 (611)	Year 12	Certificate III or IV not further defined	Certificate IV	Certificate III	Year 12	Year 12	Year 12
Year 11 (613)	Year 11	Certificate III or IV not further defined	Certificate IV	Certificate III	Year 11	Year 11	Year 11
Junior Secondary Education not further defined (620)	Certificate not further defined	Certificate III or IV not further defined	Certificate IV	Certificate III	Certificate I or II not further defined	Certificate II	Certificate I
Year 10 (621)	Year 10	Certificate III or IV not further defined	Certificate IV	Certificate III	Year 10	Certificate II	Year 10
Year 9 (622)	Certificate not further defined	Certificate III or IV not further defined	Certificate IV	Certificate III	Certificate I or II not further defined	Certificate II	Certificate I
Year 8 (623)	Certificate not further defined	Certificate III or IV not further defined	Certificate IV	Certificate III	Certificate I or II not further defined	Certificate II	Certificate I
Year 7 (624)	Certificate not further defined	Certificate III or IV not further defined	Certificate IV	Certificate III	Certificate I or II not further defined	Certificate II	Certificate I

(a)

Matrix is sourced from ABS publication, 'Education and Work', (Cat. No. 6227.0), May 2002.

This relationship emphasises that great care should be taken when comparing Census data with ABS survey data, because the two measurements yield different statistics (refer to Table 23). In particular, Census data should not be compared to ABS survey data derived using highest level of educational attainment. Therefore, in this paper, comparisons with the SEW have been restricted to data on level and main field of highest non-school qualification.

TABLE 23: LEVEL OF HIGHEST EDUCATIONAL ATTAINMENT AND LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64 ('000), AUSTRALIA (a) - 2001 SEW

Qualification Level	*Educational Attainment*	*Highest Non-School Qualification*
Postgraduate degree	283.9	283.9
Graduate diploma / Graduate certificate	300.3	300.3
Bachelor degree	1,595.5	1,595.5
Advanced diploma / Diploma	854.6	854.6
Certificate III / IV	1,798.6	1,798.6
Certificate I / II	121.8	873.6
Certificate not further defined	13.7	192.1

(a) Statistics are sourced from the 2001 Survey of Education and Work.

Further details about the implementation of ASCED, including a full concordance between ASCED and ABSCQ, can be found in *Australian Standard Classification of Education (ASCED), 2001* (Cat. no. 1272.0).

6.5.3 Measurement of persons holding a non-school qualification

Wherever possible, data should be captured, processed and validated to enable the greatest flexibility for the output of data and facilitate comparisons with other surveys (for example, SEW).

ASCED recommends that the following standard output categories for '*Level of highest non-school qualification'* be used:

- Persons with a non-school qualification:
    - Postgraduate degree;
    - Graduate diploma / Graduate certificate;
    - Bachelor degree;
    - Advanced diploma / Diploma;
    - Certificate III/IV;
    - Certificate I/II;
    - Certificate not further defined; and
    - Level not determined.
- Persons without a non-school qualification.

The category '*level not determined*', which is currently an output category in the SEW publication, includes those that responded to the non-school qualification indicator question, indicating they held a non-school qualification, but either:

- did not respond to the question on level of highest non-school qualification; or
- provided a response which could only be coded to the category 'inadequately described'.

However, the standard output categories listed for the 2001 Census in ASCED excludes the category *'level not determined'* from Census Output, and therefore the calculation of persons holding a non-school qualification. This exclusion significantly underestimates the proportion of persons with a non-school qualification (refer to Table 24) and causes a lack of comparability between other calculations (such as the SEW). On this basis, it is recommended that consideration be given to changing:

- ASCED to include the category *'level not determined'* as an output category for all statistical collections (including the Census); and
- the validation strategy for the 2006 Census to include the non-school qualification indicator question as an output field to facilitate publication of the category *'level not determined'* in future Census data.

Given that a high percentage of respondents who failed to complete the non-school qualification indicator question but completed the questions on level, main field and year of completion of highest non-school qualification held a non-school qualification (such as a Bachelor degree) (refer to Figure 6 in Section 6.2.3), this validation strategy would allow the non-school qualification indicator question to be derived as:

- 'Yes, other qualification' where the non-respondents indicated that they have a non-school qualification in the level question (such as a Bachelor Degree); or
- 'Yes, trade certificate/apprenticeship' where non-respondents indicated that they have a certificate qualification in the question on level of highest non-school qualification.

The derivations listed above should only be applied in cases where respondents have completed the year of completion of highest non-school qualification question, to exclude the potential for inclusion of persons who have not yet completed their first non-school qualification.

Given that there were also 16,678 instances where level of highest non-school qualification was not stated and the non-school qualification indicator was stated as 'Yes, trade certificate/apprenticeship', the strategy could also use this information to code level of highest non-school qualification.

TABLE 24: MEASUREMENT OF THE PROPORTION OF PERSONS WITH A NON-SCHOOL QUALIFICATION (a) - 2001 CENSUS

Level of Highest Non-School Qualification	*2001 Census Output*	*ASCED Standard Output Categories*
Persons with a non-school qualification:	269,042	269,042
Postgraduate degree		
Graduate diploma / Graduate certificate	203,928	203,928
Bachelor degree	1,445,943	1,445,943
Advanced diploma / Diploma	892,359	892,359
Certificate III / IV	1,938,490	1,938,490
Certificate I / II	330,246	330,246
Certificate not further defined	73,205	73,205
Level not determined	. .	477,386
Level not stated	. .	300,636
Level inadequately described	. .	176,750
Total persons with a non-school qualification	5,153,213	5,630,599
Total population	14,485,766	14,485,766
Proportion of persons with a non-school qualification	**35.6**	**38.9**

. . Not applicable

(a) Table excludes: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

6.6 Intercensal Comparisons

6.6.1 Introduction

The purpose of this section is to compare the change and state/territory characteristics of non-school qualifications held by the Australian population between the 1996 and the 2001 Censuses. This section also highlights any data quality issues arising from this comparison, such as changes to questionnaire formats.

6.6.2 Comparison Methodology

Comparison of the 1996 and the 2001 Census estimates on level and year of completion of highest non-school qualification has been compiled at a very broad level as a result of the implementation of ASCED and a lack of comparability between the ABSCQ and ASCED for specific clasification levels (refer to Figure 15). In addition, persons with a Certificate 1 non-school qualification and persons holding a non-school qualification and were still attending school were eliminated from the 2001 Census estimates in this section to enable, as far as possible, close comparability with the 1996 Census and the ABSCQ.

FIGURE 15: ABSCQ-ASCED CORRESPONDENCE TABLE, 2001 SEW

ABSCQ Level of Attainment		ASCED Level of Education		% of ABSCQ Level at each ASCED Level
Code	Name	Code	Name	
1	Higher degree	110–114	Doctoral degree	28
		120–123	Master degree	72
2	Postgraduate diploma	210–213	Graduate diploma	88
		220–222	Graduate certificate	11
3	Bachelor degree	310–312	Bachelor degree	100
4	Undergraduate diploma	411	Advanced diploma	77
		310–312	Bachelor degree	7
		420–421	Diploma	16
5	Associate diploma	420–421	Diploma	50
		511	Certificate IV	49
6	Skilled vocational qualification	514	Certificate III	99
		500	Certificate n.f.d.	1
7	Basic vocational qualification	500	Certificate n.f.d.	16
		511	Certificate IV	3
		521	Certificate II	80

(a)

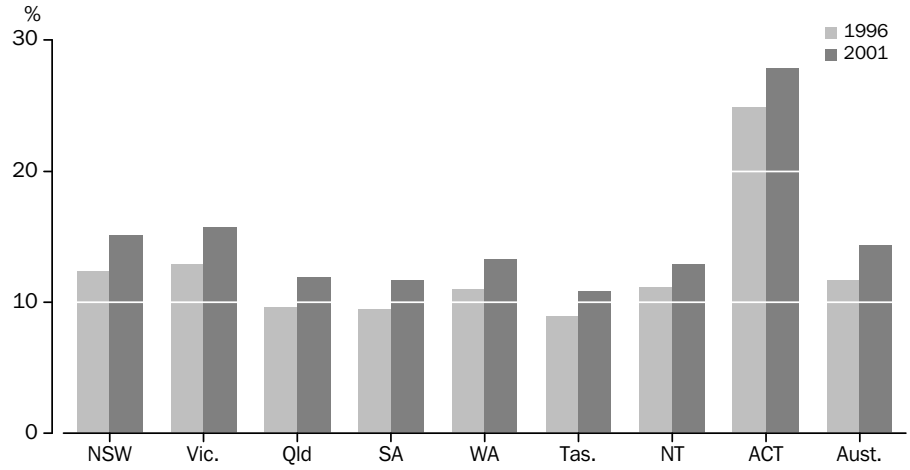
Correspondence table is sourced from ABS publication, 'Education and Work', (Cat. No. 6227.0), May 2001.

6.6.3 Broad Level of Highest Non-School Qualification

As shown in Appendix C, in 2001, there were 5,141,541 persons with a non-school qualification, accounting for 38.4% of all persons aged 15 years or over. Between 1996 and 2001, the number of persons holding a non-school qualification in Australia increased from 4,179,092 (33.7% of persons aged 15 years and over). High proportions of persons aged 15 years or over with a non-school qualification were recorded in the Australian Capital Territory (49.8%), New South Wales (40.4%), Western Australia (38.9%) and Victoria (38.6%). These same areas also recorded the highest increase in persons with a non-school qualification between Censuses.

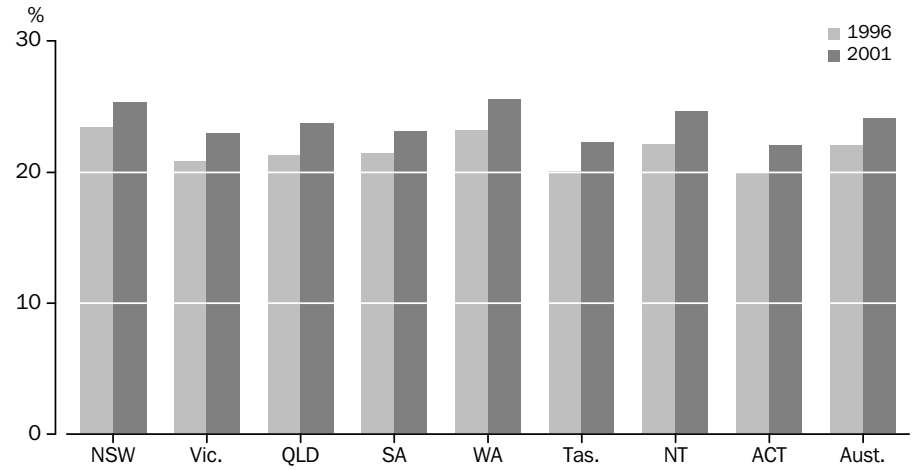
Significant increases were recorded for the proportion of persons holding a Bachelor degree or higher (refer to Figure 16) between 1996 and 2001. The Australian Capital Territory had the highest proportion of persons holding a Bachelor degree or higher in 2001 (27.8%), while Tasmania had the lowest proportion of persons holding a Bachelor degree or higher in 2001 (10.9%). Significant increases were also recorded in the proportion of persons holding an Advanced diploma or below, although the differences between States and Territories was much smaller (refer to Figure 17).

FIGURE 16: PROPORTION OF PERSONS WITH A BACHELOR DEGREE OR HIGHER, BY STATE (a) - 1996 AND 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors. 2001 Census figures excludes persons with a Certificate Level I qualification and persons still attending school. Population base excludes 'not stated' responses to the non-school qualification indicator question.

FIGURE 17: PROPORTION OF PERSONS WITH AN ADVANCED DIPLOMA OR BELOW, BY STATE (a) - 1996 AND 2001 CENSUS



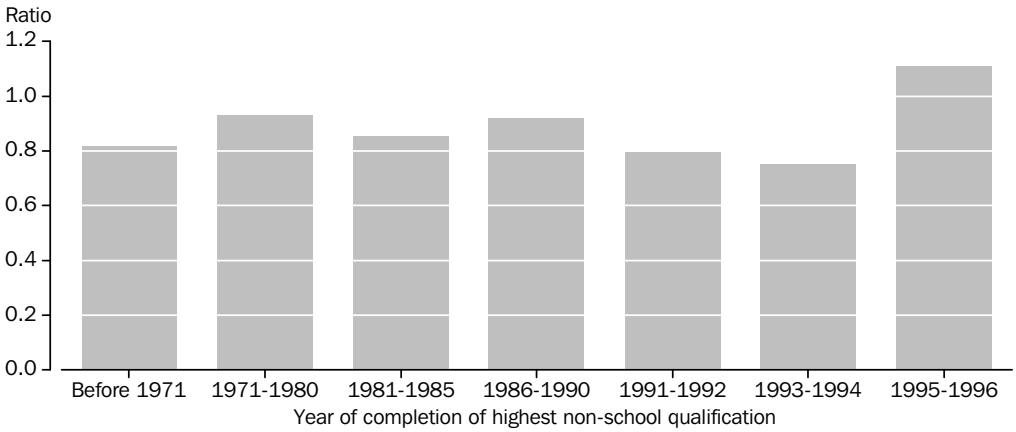
(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors. 2001 Census figures excludes persons with a Certificate Level I qualification and persons still attending school. Population base excludes 'not stated' responses to the non-school qualification indicator question.

6.6.4 Year of Completion of Highest Non-School Qualification

Section 6.2.2, which analysed the doubling in the non-response rate for year of completion of highest non-school qualification, proved that switching from a selection list to a write-in box for this question had a negative impact on data quality. This increase in non-response caused a consequential decrease in the 2001 Census figures for year of completion of highest non-school qualification for non-school qualifications which were in the scope of both Censuses (ie, those qualifications which were completed prior to August 1996).

The magnitude of the likely understatement for year of completion of highest non-school qualification has been quantified in Figure 18. Undercount ratios are presented using the age brackets used in the selection list of the 1996 Census, and dividing the 2001 Census figure by the 1996 Census figure. If an undercount ratio of 0.80 is obtained, the 2001 Census figure was 20% smaller than the 1996 Census figure.

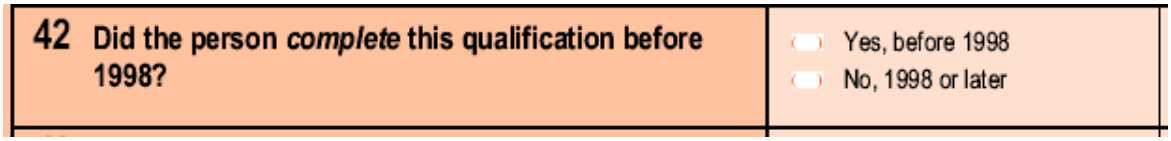
FIGURE 18: UNDERCOUNT RATIO, YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 2001 CENSUS



(a) Figure excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors. 2001 Census figures excludes persons with a Certificate Level I qualification and persons still attending school.

High undercounts were found for highest non-school qualifications which were completed before 1971, between 1991 and 1992 or between 1993 and 1994. Given these undercounts are significant and exist for non-school qualifications which have been completed relatively recently (as late as 1994), the ABS is proposing changing this question from a write-in box to a mark box response for future Censuses. Such a change is being trialled for the August 2003 Census Test (refer to Figure 19).

FIGURE 19: 2003 AUGUST CENSUS TEST HOUSEHOLD FORM (YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION)



6.7 Comparison with the ABS 2001 Survey of Education and Work

6.7.1 Introduction

The purpose of this section is to compare the collection of the Census and the SEW, having regard to both quantifiable (such as inclusion or exclusion of subsets of the population) and non-quantifiable (such as self-enumeration versus an interview approach) differences.

6.7.2 Reconciliation of 2001 Census Data with the SEW

Although the Census and ABS population surveys (such as the SEW) collect similar data, they are not strictly comparable due to differences in scope, coverage, timing, measurement of underlying concepts and collection methodology. While the Census population can be reduced to a similar population as the SEW, non-quantifiable differences mean that the Census and the SEW are still not strictly comparable. These differences include:

- under-enumeration in the Census for which Census data are not adjusted;
- the use in SEW of population benchmarks derived from incomplete information about population change;
- differing treatment for non-response; and
- sampling variability.

Self-enumeration in the Census versus the interview approach using any responsible adult in the household adopted in the SEW is another difference. The extent to which the SEW estimate for Graduate diploma / Graduate certificate is influenced by the prompting of the interviewer where the initial response was just ‘certificate’ or ‘diploma’ is unknown.

In this Census Paper, to enable analyses of Census data and the SEW, their scopes have been reduced to a common population. Therefore, the following population groups have been excluded from census data (refer to Table 25):

- members of the Australian permanent defence forces, civilian personnel employed by the defence forces, personnel employed by government units mainly engaged in defence administration (ANZSIC Classification 8200 - Defence);
- persons aged under 15 and over 64;
- overseas visitors;
- residents of Jervis Bay Territory and external territories;
- persons living in remote and sparsely settled parts of Australia;
- patients in hospitals and sanatoriums; and
- inmates in reformatories, goals, etc.

Boarding school pupils aged 15 years or more were within the scope of the SEW. However, limited information was collected for these students, in particular, details are not available for level of highest non-school qualification. Therefore, comparisons in this section have excluded boarding school pupils from statistics on specific levels and main field of highest non-school qualifications.

TABLE 25: ADJUSTMENTS MADE TO THE 2001 CENSUS TO DERIVE A COMMON POPULATION FOR DATA ON HIGHEST NON-SCHOOL QUALIFICATIONS (a) - 2001 CENSUS

Population Group	*Number*	*Treatment*
Defence Force Personnel	62,081	Deducted
Overseas visitors	157,954	Deducted
Jervis Bay Territory and external territories	1,587	Deducted
Persons living in remote and sparsely settled areas of Australia	134,035	Deducted
Patients in hospitals and sanatoriums, inmates in reformatories, goals	105,986	Deducted
Persons aged 65 years or over	2,319,247	Deducted
Boarding school pupils	13,627	Deducted

(a) Table excludes: Persons aged under 15 years, SCRs and forms where highest non-school qualification questions were not asked.

6.7.3 Sampling Variability and Comparisons with Census data

Estimates derived from the SEW are based on information obtained from occupants from a sample of dwellings, whereas the Census aims to collect information from all dwellings in Australia. A measure to quantify this likely difference (known as the standard error or SE) is presented in this paper. There are 19 chances in 20 that the difference between sampling a population and obtaining information from the entire population will be less than 2 SEs. Therefore, if the Census figure is within the ranges of the SEW estimate plus or minus 2 SEs, it can be assumed that there are no other fundamental differences between the two collections. However, if the Census figure is significantly outside this range, the data quality of both collections needs to be further investigated.

Due to space limitations, it is impractical to print the SE of each estimate for some detailed tables (for example, level of highest non-school qualification by State/Territory) in this paper. A table of SEs is provided in Appendix D to enable readers to determine the SE for an estimate from the size of that estimate.

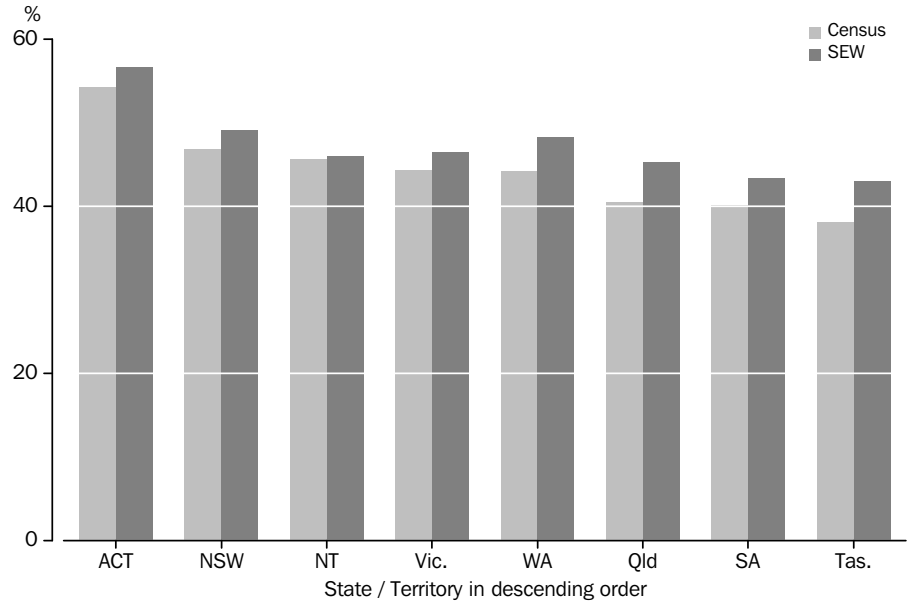
6.7.4 Proportion of persons with a non-school qualification

Comparison of the Census and SEW data reveals that the proportion of persons with a non-school qualification is lower for the Census than the SEW. As shown in Appendix D, these differences are much greater than 2 SEs of the SEW, and therefore indicate a data quality concern. For example, the difference between the Census and the SEW figure for New South Wales and Queensland is approximately 9 times 2 SEs of SEW. Figure 20 reveals that:

- the greatest difference between the Census and the SEW were for Western Australia, Queensland, South Australia and Tasmania; and

- differences in interstate relativities, such as the difference between Victoria and Queensland, were much greater for the Census than the SEW.

FIGURE 20: PROPORTION OF PERSONS AGED 15-64 YEARS WITH A NON-SCHOOL QUALIFICATION, BY STATE - 2001 CENSUS (a) AND 2001 SEW (b)



(a) Census figures exclude: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors. Proportion is derived using either of the 'Yes' responses to the non-school qualification indicator question and excludes 'not stated' responses in the population base.

(b) Statistics are sourced from the 2001 Survey of Education and Work.

6.7.5 Level of Highest Non-School Qualification

Table 26 shows that the 2001 Census figures have a significant undercount of the number of persons holding a Certificate I and II, Bachelor degree and Graduate diploma or Graduate certificate. These undercounts are several times 2 SEs of the SEW estimates and therefore represent a data quality issue. One contributing factor to these undercounts is the substantially higher non-response and inadequate response for the Census as opposed to the SEW. These undercounts, analysed by age (refer to Figures 21 and 22), mean that:

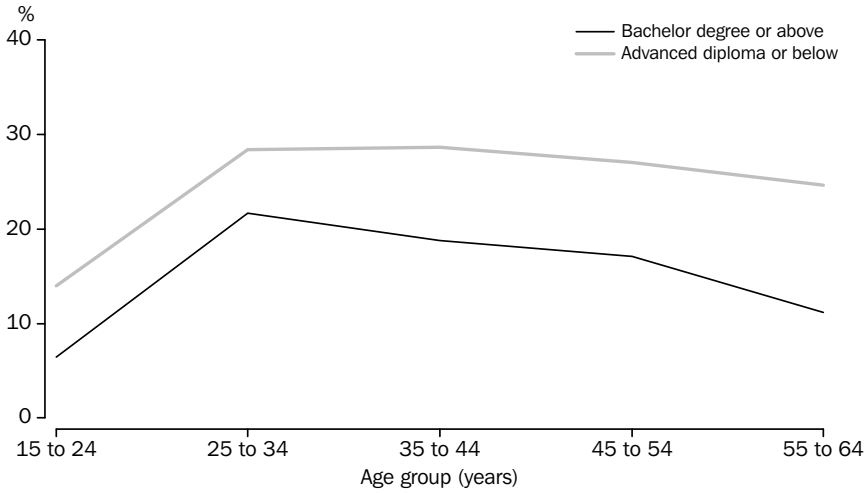
- due to undercount across all qualification levels, the proportions calculated are much smaller for the Census as opposed to the SEW;
- the relative increments of persons with a Bachelor degree or above as opposed to an Advanced diploma or below is much lower for the Census as opposed to the SEW, particularly for those aged 35 years or over; and
- due to the undercount in Graduate diplomas and Graduate certificates, the proportion of persons with a Bachelor degree or above calculated from the Census decreases much more sharply for those aged 45 or over.

TABLE 26: COMPARISON OF THE 2001 CENSUS WITH THE 2001 SEW: LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64

	Census (a)	*SEW (b)*	*Comparative Statistics*	
Level of Highest Non-School Qualification	*Persons ('000)*	*Persons ('000)*	*Census under-count (c)*	*2 SEs of SEW*
Persons with a non-school qualification:				
Postgraduate degree	244.4	283.9	39.5	18.7
Graduate diploma / Graduate certificate	189.4	300.3	110.9	15.6
Bachelor degree	1,334.7	1,595.5	260.8	38.3
Advanced diploma / Diploma	781.3	854.6	73.2	32.5
Certificate III / IV	1,672.5	1,798.6	126.1	43.2
Certificate I / II	298.5	873.6	575.0	33.2
Certificate not further defined	68.2	192.1	123.9	15.0
Level not determined	383.8	142.7	(241.1)	14.6
Level not stated	236.3	114.2	(122.1)	11.6
Level inadequately described	147.5	28.5	(119.0)	7.4
Total with non-school qualifications	*4,972.9*	*6,041.2*	*1,068.3*	*60.4*
Persons without a non-school qualification (d)	6,424.7	6,747.1	322.4	67.5
Non-school qualification indicator and level not stated	464.2	. .	. .	. .
Total	**11,861.8**	**12,788.3**	**925.5**	**102.3**

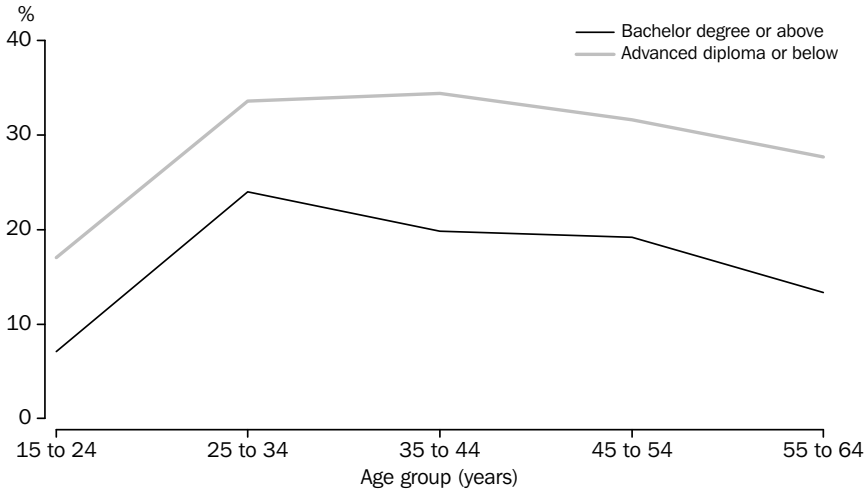
. . Not applicable
(a) Census figures exclude: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.
(b) Statistics are sourced from the 2001 Survey of Education and Work.
(c) Undercount is derived by subtracting the Census figure from the SEW figure.
(d) Includes boarding school pupils.

FIGURE 21: LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64 (a) - 2001 CENSUS



(a) Figure excludes: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors. Proportion excludes 'not stated' responses to non-school qualification indicator and level of highest non-school qualification in the population base.

FIGURE 22: LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64 (a) - 2001 SEW



(a) Statistics are sourced from the 2001 Survey of Education and Work.

6.7.6 Main Field of Highest Non-School Qualification

Significant undercounts in the 2001 Census data exist for all main fields of highest non-school qualification, with the exception of 'Education'. Table 27 shows that the highest discrepancies were found for the following main fields of highest non-school qualification, where differences between the Census and SEW were several multiples of 2 SEs of the 2001 SEW estimate:

- 'Engineering and related technologies';
- 'Health';
- 'Management and commerce';
- 'Society and culture'; and
- 'Food, hospitality and personal services'.

TABLE 27: COMPARISON OF THE 2001 CENSUS WITH THE 2001 SEW: MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64

	Census (a)	*SEW (b)*	*Comparative Statistics*	
Main Field of Highest of Non-School Qualification	*Persons ('000)*	*Persons ('000)*	*Census under-count (c)*	*2 SEs of SEW*
Persons with a non-school qualification:				
Natural and physical sciences	174.3	222.2	47.9	14.7
Information technology	149.0	181.6	32.6	14.2
Engineering and related technologies	1,066.0	1,289.1	223.1	30.9
Architecture and building	351.0	413.5	62.6	21.5
Agriculture, environmental and related studies	126.1	162.2	36.1	12.6
Health	514.8	622.4	107.6	23.7
Education	450.9	448.4	(2.5)	23.3
Management and commerce	954.8	1,321.9	367.1	31.7
Society and culture	513.8	710.6	196.8	27.0
Creative arts	180.9	240.0	59.1	15.8
Food, hospitality and personal services	293.2	384.7	91.5	20.0
Mixed field programs	3.8	5.4	1.7	2.7
Field not determined	658.7	39.2	(619.5)	8.2
Field inadequately described	598.4	3.8	(594.6)	2.3
Field not stated	60.3	35.4	(25.0)	7.4
Total (d)	**5,437.1**	**6,041.2**	**604.1**	**60.4**

(a) Census figures exclude: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.
(b) Statistics are sourced from the 2001 Survey of Education and Work.
(c) Undercount is derived by subtracting the Census figure from the SEW figure.
(d) For the Census data, total does not match level of highest non-school qualification because a number of respondents failed to respond to the non-school qualification indicator and level of highest non-school qualification questions, but completed the main field of highest non-school qualification question.

7. CONCLUSIONS

This section summarises the main data quality implications raised in this paper. Detailed recommendations to improve data quality are discussed in detail in Section 8 of this paper.

7.1 Non-School Qualification Indicator

Overall, the non-response rate for this question increased marginally from 6.5% in 1996 to 7.1% in 2001. However, a breakdown of non-responses to this question revealed that there was a doubling in the percentage of non-respondents who failed to answer this question, but could complete the questions relating to level, main field and year of completion of highest non-school qualification.

This paper highlighted that the emphasis on completion of a trade certificate in this question could have confused respondents who may think that their level of non-school qualification is not relevant to this question. A high percentage of respondents who failed to complete the non-school qualification indicator question only held a formal qualification (such as a Bachelor degree).

Recommendations to improve data quality focus on changes to questionnaire design as well as strategies to address situations where respondents have failed to complete the non-school qualification indicator question, but completed the question on level of highest non-school qualification.

7.2 Level of Highest Non-School Qualification

This question had a higher percentage of both 'not-stated' and 'inadequately described' responses, when compared to main field of highest non-school qualification. One contributing factor may have been allotted space on the 2001 Census form (there were four rows of boxes for main field of highest non-school qualification, but only three rows for level of highest non-school qualification).

Other factors, such as the quality of information provided by respondents, could also contributed to larger percentages being coded to *'inadequately described*'. This paper has argued that providing additional information to respondents in main fields where this calculation was high (such as 'Health' and 'Management and commerce') as well as updating coding indexes, instructions and training will improve data quality.

Insufficient space on the 2001 Census form may have contributed to undercounts (compared to the 2001 SEW) in some levels of highest non-school qualification (particularly Certificate I/II qualifications). Further analysis of space allotted to these two questions will be undertaken by the ABS after completion of the August 2003 Census test, which has four rows of boxes for both level and main field of highest non-school qualification.

7.3 Main Field of Highest Non-School Qualification

The question on main field of highest non-school qualification had the lowest non-response and inadequate response rates. However, undercounts (compared to the 2001 SEW) were still prevalent for most main fields.

The high frequency of dump coding for some main fields, such as 'Information technology', 'Natural and physical sciences' and 'Engineering and related technologies' indicates that additional guidance, such as a broader range of examples in the Census Form and Census Guide, should be concentrated within these main fields.

7.4 Year of Completion of Highest Non-School Qualification

A change from a selection list question in 1996 to a write-in box in 2001 adversely affected data quality in a number of ways:

- the calculated non-response has increased from 4.4% in 1996 to 7.8% in 2001;
- of those respondents that failed to complete this question, 88% could complete either or both of the other qualification questions analysed in this paper (level and main field of highest non-school qualification); and
- significant undercounts exist in this variable when compared to the ranges (with the exception of 1995 to 1996) used for the mark boxes in the 1996 Census.

Recommendations to improve data quality for this question strongly argue that this question should be changed back to a selection list for future Censuses.

8. RECOMMENDATIONS

8.1 Questionnaire Design

- The non-school qualification indicator question should be changed to reduce the confusing nature of this question.

- The same space should be assigned to both level and main field of highest non-school qualifications (four rows of boxes).

- Year of completion of highest non-school qualification should be changed from a write-in box to a mark box response.

8.2 Instructions provided to respondents

- Additional guidance on main field of highest non-school qualification should be concentrated within main fields (such as 'Information technology', 'Natural and physical sciences' and 'Engineering and related technologies') which have a high proportion of responses which could only be coded at the broad (2-digit) level.

- Additional guidance needs to be provided to respondents who completed their qualifications some time ago (such as 'Secretarial and clerical studies' and 'General nursing' qualifications), which caused increased problems for coders.

8.3 Coding

- Responses to level, main field and year of completion of highest non-school qualification that have a non-response to the non-school qualification indicator question should countinue to be coded.

- Improvements to coding indexes, particularly for level of highest non-school qualification, need to be made to improve efficiency in coding processes and data quality.

- Changes to coding indexes should address particular qualifications (such as those in the main fields of 'Secretarial and clerical studies' and 'General nursing') which have prevalent coding problems. Further investigation and analysis of these responses by the ABS will also improve coding practices in the future.

8.4 Derivation and Validation Strategy

- Consideration should be given to changing ASCED to include the category *'level not determined'* as an output category for all statistical collections (including the Census).

- The Validation for the 2006 Census should be changed to include the non-school qualification indicator question as an output field.

- Derivations (as detailed in section 6.5.3) in situations where respondents have failed to complete the non-school qualification indicator question but have completed the question on level and year of completion of highest non-school qualification should be implemented to improve data quality.

- Where level of highest non-school qualification was not stated and the non-school qualification indicator was stated as 'Yes, trade certificate/apprenticeship', the derivation strategy should use this information to code level of highest non-school qualification.

APPENDIX A - NON-RESPONSE RATES FOR SELECTED VARIABLES, 2001 CENSUS

TABLE A1: NON-RESPONSE RATES FOR SELECTED VARIABLES (a) - 2001 CENSUS

Variable	*Persons for whom Questions were relevant*	*Persons for whom there was no response*	*Non-response Rate (%) 2001*
Labour Force Status	14,485,766	284,114	2.0
Individual Income (weekly)	14,485,766	767,616	5.3
Highest Level of Schooling Completed	14,485,766	784,221	5.4
Occupation	8,298,606	98,829	1.2
Industry of Employment	8,298,606	144,613	1.7
Industry Sector	8,298,606	202,177	2.4

(a) Table excludes: Persons aged under 15 years, SCRs, forms where specified questions were not asked and overseas visitors.

TABLE A2: ADJUSTED NON-RESPONSE RATES FOR LABOUR FORCE RELATED VARIABLES (a) - 2001 CENSUS

Variable	*Persons for whom Questions could have been relevant (b)*	*Persons for whom there was no response (c)*	*Non-response Rate (%) 2001*
Occupation	8,582,720	382,943	4.5
Industry of Employment	8,582,720	428,727	5.0
Industry Sector	8,582,720	486,291	5.7

(a) Table excludes: Persons aged under 15 years, SCRs, forms where labour force related questions were not asked and overseas visitors.
(b) Includes all responses to question *'Last week' did the person have a full-time or part-time job of any kind?'* who indicated that they had a job last week plus those that did not respond to this question.
(c) Persons who did not respond to the labour force related variables above and either indicated that they had a full-time or part-time job last in the previous week or did not respond to the question: *'Last week' did the person have a full-time or part-time job of any kind?'*.

APPENDIX B - NON-STATED RESPONSES, 2001 CENSUS

TABLE B1: NOT-STATED RESPONSES FOR QUESTIONS RELATING TO HIGHEST NON-SCHOOL QUALIFICATION, BY AGE (a) - 2001 CENSUS

Age group (years)	*Number of Topic Non-Respondents*	*% of Topic Non-Respondents*	*Non-Respondents as a % of Age Level*
15-24	123,653	16.1	5.0
25-34	64,788	8.4	2.4
35-44	77,401	10.1	2.8
45-54	76,816	10.0	3.1
55-64	83,293	10.8	4.8
65-74	129,386	16.8	10.4
75-84	148,333	19.3	18.3
85 or older	64,467	8.4	24.9
Total	**768,137**	**100.0**	. .

. . Not applicable.
(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

TABLE B2: NOT-STATED RESPONSES FOR QUESTIONS RELATING TO HIGHEST NON-SCHOOL QUALIFICATION, BY LABOUR FORCE STATUS (a) - 2001 CENSUS

Labour Force Status	*Number of Topic Non-Respondents*	*% of Topic Non-Respondents*	*Non-Respondents as a % of Labour Force Status*
Employee	133,813	17.4	2.0
Employer	11,558	1.5	2.0
Own account worker	15,743	2.1	1.9
Contributing family worker	2,930	0.4	4.9
Not in the labour force	374,798	48.8	7.2
Not Stated	209,884	27.3	73.9
Unemployed	19,411	2.5	2.9
Total	**768,137**	**100.0**	**. .**

. . Not applicable.
(a) Table excludes: Persons aged under 15 years, SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

APPENDIX C - INTERCENSAL COMPARISONS

TABLE C1: HIGHEST LEVEL OF NON-SCHOOL QUALIFICATION OF PERSONS AGED 15 YEARS OR OVER, BY STATE (a) - 1996 AND 2001 (b) CENSUS

State/ Territory	*Bachelor degree or above*	*Advanced diploma or below (c)*	*Total persons with a non-school qualification*	*Total Population (d)*	*Proportion of persons with a non-school qualification*
1996					
NSW	513,198	972,185	1,485,383	4,158,457	35.7
Vic.	392,653	635,998	1,028,651	3,053,489	33.7
Qld	223,199	493,424	716,623	2,316,570	30.9
SA	97,026	217,959	314,985	1,017,723	30.9
WA	131,198	276,448	407,646	1,189,926	34.3
Tas.	28,182	63,311	91,493	315,636	29.0
NT	13,589	27,150	40,739	122,747	33.2
ACT	51,582	41,299	92,881	206,918	44.9
Aust. (e)	**1,450,842**	**2,729,250**	**4,179,092**	**12,383,558**	**33.7**
2001					
NSW	676,245	1,130,025	1,806,270	4,466,096	40.4
Vic.	519,225	754,076	1,273,301	3,298,254	38.6
Qld	305,191	606,457	911,648	2,559,254	35.6
SA	125,235	248,219	373,454	1,073,176	34.8
WA	173,700	332,747	506,447	1,302,833	38.9
Tas.	35,156	71,931	107,087	322,297	33.2
NT	17,294	33,094	50,388	134,412	37.5
ACT	62,788	49,535	112,323	225,527	49.8
Aust. (e)	**1,915,047**	**3,226,494**	**5,141,541**	**13,384,031**	**38.4**

(a) Table excludes: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) 2001 Census Figures excludes persons still at school

(c) Excludes persons with a Certificate Level I qualification.

(d) Excludes 'not stated' responses to level of highest non-school qualification.

(e) Australia figures include data for External Territories.

TABLE C2: YEAR OF COMPLETION OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 1996 AND 2001 CENSUS

Year of completion of highest non-school qualification	*1996 Census*	*2001 Census (b)*
Before 1971	1,251,461	1,020,694
1971-1980	929,354	865,773
1981-1985	631,384	538,293
1986-1990	657,633	604,877
1991-1992	380,837	303,474
1993-1994	447,996	337,261
1995-1996	360,797	400,645
After 1996	. .	1,169,189
Not Stated	984,233	1,182,115
Total	**5,643,695**	**6,422,321**

. . Not applicable.

(a) Table excludes: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) Excludes persons with a Certificate Level I qualification and persons still attending school.

APPENDIX D - COMPARISON BETWEEN THE 2001 SURVEY OF EDUCATION AND WORK AND THE 2001 CENSUS

TABLE D1: STANDARD ERRORS OF ESTIMATES (a) - 2001 SEW

									Aust.	
Size of estimate (no.)	*NSW (no.)*	*Vic. (no.)*	*Qld (no.)*	*SA (no.)*	*WA (no.)*	*Tas. (no).*	*NT (no.)*	*ACT (no.)*	*SE (no.)*	*RSE (%)*
100	100	. .	. .	. .	. .	. .	. .	. .	80	80.0
200	170	180	. .	. .	. .	150	150	170	140	70.0
300	230	240	. .	270	280	180	180	190	200	66.7
500	340	340	420	330	350	220	220	230	290	58.0
700	430	420	490	380	410	250	260	250	370	52.9
1,000	550	430	580	440	480	290	290	280	470	47.0
1,500	720	670	690	520	570	340	350	330	610	40.7
2,000	860	790	790	590	650	380	390	360	730	36.5
2,500	1,000	900	850	650	700	400	400	400	850	34.0
3,000	1,100	1,000	950	700	750	450	450	400	950	31.7
3,500	1,200	1,050	1,000	750	800	500	500	450	1,050	30.0
4,000	1,300	1,150	1,100	800	850	500	500	450	1,100	27.5
5,000	1,450	1,250	1,200	850	950	550	550	500	1,250	25.0
7,000	1,700	1,500	1,400	1,000	1,100	650	650	600	1,550	22.1
10,000	2,050	1,750	1,600	1,150	1,250	700	700	650	1,850	18.5
15,000	2,450	2,100	1,900	1,350	1,500	850	850	800	2,250	15.0
20,000	2,800	2,350	2,200	1,500	1,650	950	950	900	2,600	13.0
30,000	3,300	2,750	2,600	1,800	1,950	1,000	1,050	1,050	3,150	10.5
40,000	3,650	3,100	2,900	2,000	2,200	1,250	1,200	1,150	3,550	8.9
50,000	3,950	3,300	3,200	2,200	2,350	1,350	1,300	1,300	3,900	7.8
100,000	4,950	4,200	4,250	2,900	3,050	1,750	1,650	1,750	5,100	5.1
150,000	5,600	4,850	5,050	3,400	3,500	2,000	1,900	2,100	5,900	3.9
200,000	6,150	5,450	5,650	3,800	3,900	2,250	2,100	2,400	6,550	3.3
300,000	7,200	6,450	6,650	4,450	4,450	2,600	. .	2,850	7,650	2.6
500,000	8,900	8,100	8,150	5,450	5,300	3,100	. .	. .	9,300	1.9
1,000,000	12,450	11,350	10,700	7,150	6,600	. .	. .	. .	12,150	1.2
2,000,000	18,300	16,450	13,950	9,350	8,150	. .	. .	. .	16,050	0.8
5,000,000	32,850	28,350	19,650	. .	. .	. .	. .	. .	24,600	0.5
10,000,000	. .	. .	. .	. .	. .	. .	. .	. .	43,150	0.4

. . Not applicable.

(a) Standard errors are sourced from the 2001 Survey of Education and Work.

TABLE D2: COMPARISON OF THE 2001 CENSUS WITH THE 2001 SEW: PROPORTION OF PERSONS AGED 15-64 WITH A NON-SCHOOL QUALIFICATION

	Census (a)			*SEW (d)*			*Comparative Statistics*	
State/ Territory	*Persons with non-school qualification ('000) (b)*	*Total Persons ('000) (c)*	*Per cent*	*Persons with non-school qualification ('000)*	*Total Persons ('000)*	*Per cent*	*Census under-count (e)*	*2 SEs of SEW*
NSW	1,767.0	3,766.2	46.9	2,114.6	4,300.3	49.2	347.5	38.7
Vic.	1,246.6	2,807.2	44.4	1,499.6	3,224.1	46.5	253.0	34.0
Qld	871.2	2,146.7	40.6	1,082.8	2,386.1	45.4	211.5	23.2
SA	354.6	881.9	40.2	423.7	974.5	43.5	69.1	12.6
WA	484.8	1,094.2	44.3	621.8	1,284.2	48.4	137.0	13.2
Tas.	103.5	271.4	38.1	129.9	302.1	43.0	26.4	4.5
NT	40.9	88.6	45.7	46.5	101.0	46.0	5.6	2.8
ACT	104.1	191.9	54.3	122.4	216.0	56.7	18.3	4.3
Aust.	**4,972.9**	**11,249.1**	**44.2**	**6,041.2**	**12,788.3**	**47.2**	**1068.3**	**60.4**

(a) Table excludes: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) Derived using either of the 'Yes' responses to the non-school qualification indicator question.

(c) Total excludes 'not stated' responses to the non-school qualification indicator in the population base.

(d) Statistics are sourced from the 2001 Survey of Education and Work.

(e) Undercount is derived by subtracting the Census figure from the SEW figure.

TABLE D3: LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64, BY STATE (a) - 2001 CENSUS, ('000)

Level of Highest Non-School Qualification	*NSW*	*Vic.*	*Qld*	*SA*	*WA*	*Tas.*	*NT*	*ACT*	*Aust.*
Persons with a non-school qualification:									
Postgraduate degree	101.0	60.0	35.1	13.4	18.7	3.8	1.9	10.5	244.4
Graduate diploma / Graduate certificate	56.5	61.4	29.4	13.3	16.3	3.4	1.8	7.3	189.4
Bachelor degree	470.6	366.1	215.5	87.4	123.0	24.6	9.9	37.5	1,334.7
Advanced diploma / Diploma	276.2	199.3	135.5	54.7	79.4	15.0	6.3	14.9	781.3
Certificate III / IV	568.3	395.9	330.0	131.6	171.0	40.1	14.8	20.9	1,672.6
Certificate I / II	122.0	61.8	47.8	22.1	30.1	6.5	2.5	5.8	298.5
Certificate not further defined	26.5	14.4	11.8	5.4	7.0	1.3	0.6	1.2	68.2
Level not determined	145.9	87.8	66.1	26.7	39.4	8.8	3.2	5.9	383.8
Level not stated	89.1	57.3	39.3	15.3	24.6	5.2	1.9	3.6	236.3
Level inadequately described	56.7	30.5	26.8	11.4	14.8	3.6	1.3	2.3	147.5
Total with non-school qualifications	*1,767.1*	*1,246.6*	*871.2*	*354.6*	*484.8*	*103.5*	*40.9*	*104.1*	*4,972.9*
Persons without a non-school qualification (b)	2,054.1	1,602.3	1,297.6	536.3	622.7	170.6	49.9	91.1	6,424.7
Non-school qualification indicator and level not stated	160.5	125.0	86.0	31.1	40.0	11.8	4.2	5.6	464.2
Total	**3,981.7**	**2,973.9**	**2,254.8**	**922.0**	**1,147.5**	**285.9**	**95.0**	**200.8**	**11,861.8**

(a) Table excludes: SCRs, forms where non-school qualification questions were not asked and overseas visitors.

(b) Includes boarding school pupils.

TABLE D4: LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64, BY STATE (a) - 2001 SEW, ('000)

Level of Highest Non-School Qualification	*NSW*	*Vic.*	*Qld*	*SA*	*WA*	*Tas.*	*NT*	*ACT*	*Aust.*
Persons with a non-school qualification:									
Postgraduate degree	104.3	76.8	46.1	13.7	25.7	4.2	2.1	10.9	283.9
Graduate diploma / Graduate certificate	94.1	102.4	38.9	20.1	25.7	4.3	3.7	11.0	300.3
Bachelor degree	570.8	434.5	256.3	99.5	156.4	26.6	10.1	41.4	1,595.5
Advanced diploma / Diploma	304.3	220.8	139.9	59.8	85.3	19.3	8.0	17.1	854.6
Certificate III / IV	584.7	411.4	384.2	134.4	198.8	46.6	15.9	22.6	1,798.6
Certificate I / II	324.7	182.5	163.2	69.4	93.9	21.9	4.2	13.7	873.6
Certificate not further defined	66.2	44.9	36.5	18.2	18.7	3.7	0.6	3.2	192.1
Level not determined	65.4	26.2	17.7	8.5	17.3	3.1	1.9	2.5	142.7
Level not stated	58.3	16.9	12.9	7.2	12.8	2.5	*1.2	2.4	114.2
Level inadequately described	7.1	9.4	*4.8	1.3	4.5	*0.7	*0.7	**0.1	28.5
Total with non-school qualifications	*2,114.6*	*1,499.6*	*1,082.8*	*423.7*	*621.8*	*129.9*	*46.5*	*122.4*	*6,041.2*
Persons without a non-school qualification (b)	2,185.7	1,724.5	1,303.3	550.9	662.4	172.2	54.5	93.6	6,747.1
Total	**4,300.3**	**3,224.1**	**2,386.1**	**974.5**	**1,284.2**	**302.1**	**101.0**	**216.0**	**12,788.3**

* estimate has a relative standard error of between 25% and 50% and should be treated with caution.

** estimate has a relative standard error greater than 50% and is considered too unreliable for general use.

(a) Statistics are sourced from the 2001 Survey of Education and Work.

(b) Includes boarding school pupils.

TABLE D5: LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64, BY AGE (a) - 2001 CENSUS, ('000)

	Age group (years) at August 2001					
Highest Level of Non-School Qualification	*15 to 24*	*25 to 34*	*35 to 44*	*45 to 54*	*55 to 64*	*Total*
Persons with a non-school qualification:						
Bachelor degree or above	145,694	542,171	496,721	405,415	178,448	1,768,449
Advanced diploma or below	318,848	711,010	756,173	641,412	392,143	2,820,586
Level not determined	42,766	79,689	100,537	95,982	65,072	383,846
Total	508,308	1,332,870	1,353,431	1,142,609	635,663	4,972,881
Total Persons (b)	**2,285,074**	**2,499,158**	**2,645,885**	**2,373,763**	**1,593,677**	**11,397,557**

(a) Table excludes: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

(b) Total includes boarding school pupils but excludes 'not stated' responses to non-school qualification indicator and level of highest non-school qualification in the denominator.

TABLE D6: LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64, BY AGE (a) - 2001 SEW, ('000)

	Age group (years) at May 2001					
Highest Level of Non-School Qualification	*15 to 24*	*25 to 34*	*35 to 44*	*45 to 54*	*55 to 64*	*Total*
Persons with a non-school qualification:						
Bachelor degree or above	191,535	686,198	572,429	498,780	230,701	2,174,644
Advanced diploma or below	459,422	959,383	997,747	819,551	482,712	3,721,815
Level not determined	14,067	36,003	37,511	31,825	23,326	142,731
Total	665,024	1,681,585	1,607,685	1,350,156	736,740	6,041,190
Total Persons (b)	**2,703,776**	**2,854,571**	**2,987,097**	**2,592,300**	**1,740,587**	**12,788,330**

(a) Statistics are sourced from the 2001 Survey of Education and Work.

(b) Total includes boarding school pupils.

TABLE D7: MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64, BY LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 2001 CENSUS, ('000)

Main Field of Non-School Qualification	*Postgraduate degree*	*Graduate diploma/ Graduate certificate*	*Bachelor degree*	*Advanced diploma/ Diploma*	*Certificate III or IV*	*Certificate I or II*	*Certificate not further defined*
Persons with a non-school qualification:							
Natural and physical sciences	34.3	3.8	106.3	15.8	3.0	4.7	0.6
Information technology	9.7	8.7	51.8	34.6	16.1	13.2	4.5
Engineering and related technologies	20.6	4.2	109.6	95.5	763.5	25.3	4.8
Architecture and building	3.2	2.3	27.4	20.0	279.6	3.7	1.1
Agriculture, environmental and related studies	5.6	2.5	23.7	22.6	48.1	12.6	2.5
Health	27.3	22.9	226.4	106.0	51.4	14.2	4.0
Education	27.8	78.7	204.2	116.1	5.5	1.5	1.1
Management and commerce	52.5	31.4	236.7	199.7	112.3	168.3	26.8
Society and culture	52.2	27.9	248.6	78.2	42.9	22.7	9.3
Creative arts	6.8	5.3	64.7	50.4	28.0	9.3	3.1
Food, hospitality and personal services	_	0.1	2.0	27.0	226.0	17.3	4.8
Mixed field programmes	_	_	0.1	0.1	0.4	0.8	0.5
Field not determined	4.4	1.6	33.3	15.5	95.9	4.9	5.1
Field inadequately described	2.1	0.9	25.8	6.4	10.1	2.0	0.5
Field not stated	2.3	0.7	7.5	9.0	85.7	2.9	4.6
Total	**244.4**	**189.4**	**1,334.7**	**781.3**	**1,672.6**	**298.5**	**68.2**

_ null cells

(a) Table excludes: SCRs, forms where highest non-school qualification questions were not asked and overseas visitors.

TABLE D8: MAIN FIELD OF HIGHEST NON-SCHOOL QUALIFICATION OF PERSONS AGED 15-64, BY LEVEL OF HIGHEST NON-SCHOOL QUALIFICATION (a) - 2001 SEW, ('000)

Main Field of Non-School Qualification	*Postgraduate degree*	*Graduate diploma/ Graduate certificate*	*Bachelor degree*	*Advanced diploma/ Diploma*	*Certificate III or IV*	*Certificate I or II*	*Certificate not further defined*
Persons with a non-school qualification:							
Natural and physical sciences	38.5	9.1	135.9	17.3	*4.0	12.2	*1.7
Information technology	8.6	10.3	60.6	36.4	20.4	27.4	14.5
Engineering and related technologies	21.2	8.2	148.6	95.9	856.2	114.4	19.0
Architecture and building	*3.2	*3.8	29.6	18.0	322.2	21.3	5.9
Agriculture, environmental and related studies	*5.5	*4.6	32.1	26.0	41.0	38.2	8.1
Health	43.0	43.3	277.5	123.8	58.8	49.7	10.6
Education	23.1	104.9	198.0	97.8	11.2	*5.4	*3.4
Management and commerce	62.4	56.6	277.5	232.5	137.8	441.9	71.3
Society and culture	67.3	48.9	333.9	100.1	76.8	51.4	20.7
Creative arts	9.2	8.5	82.1	66.2	30.6	24.7	10.0
Food, hospitality and personal services	**0.1	**0.3	4.2	32.0	236.0	81.1	24.4
Mixed field programmes	_	**0.3	_	**0.1	**0.5	*2.3	*1.9
Field not determined	*1.6	*1.4	15.3	8.5	*3.0	*3.7	**0.6
Field inadequately described	**0.7	*1.0	15.3	8.0	*2.0	*2.8	**0.6
Field not stated	**0.9	**0.5	_	**0.5	*1.0	*1.0	_
Total	**283.9**	**300.3**	**1,595.5**	**854.6**	**1,798.6**	**873.6**	**192.1**

***** estimate has a relative standard error of between 25% and 50% and should be treated with caution.
** estimate has a relative standard error greater than 50% and is considered too unreliable for general use.
_ nil or rounded to zero (including null cells)
(a) Statistics are sourced from the 2001 Survey of Education and Work.

ACRONYMS

ABS	Australian Bureau of Statistics
ABSCQ	Australian Bureau of Statistics Classification of Qualifications
AC	Automatic Coding
ASCED	Australian Standard Classification of Education
CAC	Computer Assisted Coding
CD	Collection District
DC	Data Capture
DPC	Data Processing Centre
DQI	Data Quality Investigation
HHF	Household Form
ICR	Intelligent Character Recognition
NFD	Not Further Defined
NTF	National Training Framework
OMR	Optical Mark Recognition
PF	Personal Form
QM	Quality Management
SCRs	System Created Records
SE	Standard Error
SEW	Survey of Education and Work
SIPF	Special Indigenous Personal Form
VET	Vocational Education and Training

GLOSSARY

Australian Standard Classification of Education (ASCED) - The ASCED is a new national standard classification which includes all sectors of the Australian education system; that is, schools, Vocational Education and Training and Higher education. From 2001, ASCED replaces a number of classifications used in administrative and statistical systems, including the ABSCQ. The ASCED comprises two classifications: Level of Education and Field of Education. Further details can be found in *Australian Standard Classification of Education (ASCED), 2001* (Cat. no. 1272.0).

Automatic coding - The matching of textual responses (as interpreted by ICR) to the Index, without manual intervention.

Census Guide - An explanatory booklet that provides advice and background information on how to complete a Census Form. A Guide was distributed with each form.

Census Inquiry Service - A phone-based (13 number) facility set up to provide translation and other information services relating to the 2001 Census.

Certificate not further defined - Survey responses are coded to Certificate not further defined where there is not enough information to code them to Certificate I, II, III or IV in the ASCED.

Collection District - The smallest geographic area defined in the Australian Standard Geographical Classification. It has been designed for use in the Census of Population and Housing as the smallest unit for collection, processing and output of data.

Completion - A completing student is one who has completed all academic requirements for the conferring of an award from an institution.

Data Capture - The process that ensures that marks on the Census Form (mark box or writing) are reproduced on an image. DC registers and codes mark box responses.

Data Processing - Completed census forms are delivered to the DPC as soon as possible after Census Night. They are then put though a number of processes, including precapture, data capture and AC and CAC.

Data Processing Centre - The DPC is the central site where the capture and data entry of census forms takes place. After census forms have been completed on Census Night, they are collected and returned to the State distribution offices or sent directly to the DPC.

Data Quality Investigation - A DQI Team operated at the DPC, conducting additional coding exercises to uncover data quality issues.

Discrepancy Rate - The rate at which Quality Management and subsequent Adjudication coding differed from that of an individual human or system coding. It is expressed as a percentage and is regarded as the error rate within final data.

Higher education - Education offered by a university or other recognised higher education institution, leading to the award of a degree or higher level qualification.

Index - The listing of valid responses to a Census question or topic.

Indigenous Status -'Indigenous' refers to people who identify as being of Aboriginal and/or Torres Strait Islander origin.

Intelligent Character Recognition - The system used to convert any mark box or hand-written responses found on an image into machine processable information.

Level of highest non-school qualification - Level of highest non-school qualification identifies the highest non-school qualification a person has completed in any area of study. It is categorised according to the ASCED Level of Education classification.

Level not determined - Level not determined includes inadequately described responses or where no responses were given.

Main field of highest non-school qualification - Main field of highest non-school qualification is defined as the subject matter of an educational activity. It is categorised according to the ASCED Field of Education classification.

Mark boxes - Boxes that invite the respondent to place a dash on one of a possible series of selection boxes on the Census Form. The ICR system then identified marked boxes during DC.

Non-school qualification - Non-school qualifications are awarded for educational attainments other than those of pre-primary, primary or secondary education. They include qualifications at the Post Graduate Degree Level, Master Degree Level, Graduate Diploma and Graduate Certificate Level, Bachelor Degree Level, Advanced Diploma and Diploma and Certificates I, II, III and IV levels. Non-school qualifications may be obtained concurrently with school qualifications.

Quality Management - The process of regular review of a percentage of coding work, though also a term for broader DPC-wide ongoing reviews.

Second Release Processing - Responses to the more complex Census topics, such as highest non-school qualifications, were processed within this second phase.

Self-enumeration - Self-enumeration is the term used to describe the way census data are collected. The census forms are generally completed by householders (or individuals in non-private dwellings) rather than by interviewers, although interviewers are available in some areas (such as Indigenous communities).

System Created Record - A record created during census processing for a person for whom a census from has not been received but where a collector believed the dwelling was occupied on census night. These records have values imputed for age, sex, marital status and usual residence only. Values for other variables are set to 'not stated' or 'not applicable', depending on the imputed value for age.

Validation - The checking of all Census variables for signs of any remaining or emerging system problems. This was undertaken by the DPC-based Validation Team.

Vocational Education and Training - Post-compulsory education and training, excluding degree and higher level programs delivered by higher education institutions, which provides people with occupational or work-related knowledge and skills. Vocational education and training also includes programs which provide the basis for subsequent vocational programs.

Write-in Response Boxes - A response box on the Census Form requiring a written response. It was generally coded using ICR and then AC.

REFERENCES

Australian Bureau of Statistics, 1993, *ABS Classification of Qualifications, ABSCQ,* (cat. no. 1262.0), ABS, Canberra

Australian Bureau of Statistics, 2000, *Information Paper, 2001 Census of Population and Housing, Nature and Content,* (cat. no. 2008.0), ABS, Canberra

Australian Bureau of Statistics, 2000, *How Australia Takes a Census,* (cat. no. 2903.0), ABS, Canberra

Australian Bureau of Statistics, 2001, *Australian Standard Classification of Education (ASCED),* (cat. no. 1272.0), ABS, Canberra

Australian Bureau of Statistics, 2001, *Measuring Wellbeing: Frameworks for Australian Social Statistics*, (cat. no. 4160.0), ABS, Canberra

Australian Bureau of Statistics, 2002, *Measuring Australia's Progress,* (cat. no. 1370.0), ABS, Canberra

Australian Bureau of Statistics, 2002, *Education and Training Indicators, Australia, 2002,* (cat. no. 4230.0), ABS, Canberra

Australian Bureau of Statistics, 2002, *Education and Work, Australia, May 2001,* (cat. no. 6227.0), ABS, Canberra

Australian Bureau of Statistics, 2002, *Education and Work, Australia, May 2002,* (cat. no. 6227.0), ABS, Canberra

Australian Bureau of Statistics, 2002, *Education and Training Experience, 2001,* (cat. no. 6278.0), ABS, Canberra

Australian Bureau of Statistics, 2002, *2001 Census of Population and Housing Fact Sheet, Effect of Census Processes on Non-response Rates and Person Counts,* ABS, Canberra

Australian Bureau of Statistics, 2002, *2001 Census of Population and Housing Fact Sheet, Non-response Rates, Australia,* ABS, Canberra

Australian Bureau of Statistics, 2003, *Census of Population and Housing, Selected Education and Labour Force Characteristics, Australia, 2001*, (cat. no. 2017.0), ABS, Canberra

LIST OF CENSUS PAPERS

2001 Census Papers:

03/09 *2001 Census: Level, Main Field and Year of Completion of Highest Non-School Qualification*
03/06 *2001 Census: Occupation*
03/05 *2001 Census: Labour Force Status*
03/04 *2001 Census: Income*
03/03 *2001 Census: Computer and Internet Use*
03/02 *2001 Census: Housing*
03/01b *2001 Census: Ancestry - Detailed Paper*
03/01a *2001 Census: Ancestry - First and Second Generation Australians*
02/03 *2001 Census: Form Design Testing*
02/02 *Report on Testing of Disability Questions for Inclusion in the 2001 Census*
02/01 *2001 Census: Digital Geography Technical Information Paper*

1996 Census Working Papers:

00/4 *1996 Census Data Quality: Income*
00/3 *1996 Census Data Quality: Industry*
00/2 *1996 Census Data Quality: Qualification Level and Field of Study*
00/1 *1996 Census Data Quality: Journey to Work*
99/6 *1996 Census Data Quality: Occupation*
99/4 *1996 Census: Review of Enumeration of Indigenous Peoples in the 1996 Census*
99/3 *1996 Census Data Quality: Housing*
99/2 *1996 Census: Labour Force Status*
99/1 *1996 Census: Industry Data Comparison*
97/1 *1996 Census: Homeless Enumeration Strategy*
96/3 *1996 Census of Population and Housing: Digital Geography Technical Information Paper*
96/2 *1996 Census Form Design Testing Program*

A range of 1991 Census Working Papers, from 93/1 to 96/1 are also available.

These Papers can be accessed on the ABS web site at <http://www.abs.gov.au>. From the ABS home page, select **Census -> (Census Information) Fact Sheets and Census Papers -> (Fact Sheets and Information Papers) Census Papers.**

If you have further data quality queries, please contact the Assistant Director, Census Evaluation by telephone: (02) 6252 5611 or email: <joanne.healey@abs.gov.au>.