

# Content Analysis of the April 2017 Census Test data

## Table of Contents

Section 1: Summary of key findings – by variable .....	6
Overall comments.....	6
Key issues by variable .....	7
Population, Social and Identity.....	7
Age.....	7
Sex .....	7
Ethnicity .....	7
Māori descent.....	8
Iwi affiliation .....	8
Birthplace.....	8
Languages spoken.....	8
Number of children born .....	8
Religious affiliation .....	9
Relationship to reference person .....	9
Living arrangements .....	9
Unpaid activities .....	10
Health Variables.....	10
Disability/activity limitations .....	10
Cigarette smoking behaviour.....	10
Location Variables .....	10
Years at usual residence .....	10
Years since arrival in New Zealand .....	11
Person record type, usual residence address and census night address.....	11
Usual residence one year ago.....	11
Education Variables .....	12
Highest secondary school qualification .....	12
Highest post-school qualification .....	12
Highest qualification (Derivation).....	12
New Zealand/Overseas post-school qualification indicator .....	12
Field of study .....	12
Study participation .....	12

Income Variables .....	13
Total Income .....	13
Sources of income .....	13
Tenure of Household and associated variables .....	13
Dwelling owned or in family trust .....	13
Mortgage payments .....	13
Rent Indicator .....	14
Weekly rent paid by household .....	14
Sector of landlord .....	14
Tenure of household.....	14
Individual home ownership .....	14
Occupied Dwelling Type and associated variables .....	14
Dwelling description .....	14
Dwelling joined or separate.....	15
Number of storeys .....	15
Occupied dwelling type .....	15
Unoccupied dwelling type .....	15
Number of rooms .....	15
Number of bedrooms .....	15
Numbers of other room types .....	15
Housing Quality Variables.....	16
Main types of heating .....	16
Dwelling mould indicator.....	16
Dwelling dampness indicator .....	16
Access to basic amenities .....	16
Access to Telecommunication systems .....	16
Transport Variables .....	17
Main means of travel to work .....	17
Workplace address (analysis of non-response only) .....	17
Travel to education.....	17
Educational institution location.....	17
Number of motor vehicles.....	18
Work Variables .....	18
Work and Labour Force Status .....	18
Job (Work) Indicator .....	18
Hours Worked in Employment per Week.....	18
Occupation.....	18
Industry.....	18

Workplace Address .....	18
Job Search/Availability for Work .....	18
Census Variables Not Included in the Report: .....	19
Section 2: Detailed Analysis of Variables .....	20
Individual Variables – Social, Population, Identity .....	20
Age .....	20
Sex .....	22
Ethnic Group .....	24
Māori Descent .....	26
Iwi .....	28
Religious affiliation .....	31
Birthplace.....	33
Languages spoken.....	34
Number of Children Born .....	35
Relationship to Reference Person .....	37
Living Arrangements .....	40
Unpaid activities .....	43
Health .....	44
Disability/activity limitations .....	44
Cigarette Smoking Behaviour .....	47
Location variables .....	49
Years at Usual Residence .....	49
Years since Arrival in New Zealand .....	49
Person Record Type, Usual Residence Address and Census Night Address .....	50
Usual Residence One Year Ago .....	52
Education and Training Variables .....	53
Study Participation .....	53
Highest Secondary School Qualification .....	55
Highest Post-School Qualification .....	57
Highest Qualification – Derivation 54 .....	59
Overseas post-school qualification indicator .....	61
Field of Study .....	61
Income variables.....	63
Total personal income .....	63
Sources of personal income.....	66
Tenure of household and associated variables .....	68
Dwelling owned or in family trust .....	69
Mortgage payments .....	70

Rent Indicator .....	71
Weekly rent paid by household.....	72
Sector of landlord .....	74
Tenure of household (derived output variable) .....	75
Individual Home Ownership .....	77
Occupied dwelling type and associated variables .....	78
Dwelling description (question two).....	79
Dwelling joined or separate (question three).....	81
Number of storeys (question four).....	82
Occupied dwelling type .....	83
Unoccupied dwelling type .....	84
Rooms .....	85
Number of bedrooms .....	85
Numbers of other room types .....	86
Number of rooms .....	87
Main types of heating.....	89
Housing quality.....	93
Dwelling mould indicator.....	94
Dwelling dampness indicator .....	98
Access to basic amenities .....	100
Access to telecommunication systems.....	102
Transport variables .....	104
Main means of travel to work and workplace address .....	104
Main means of travel to work .....	105
Workplace address (analysis of non-response only) .....	106
Travel to education.....	107
Main means of travel to education .....	108
Educational institution location.....	110
Number of motor vehicles.....	111
Work variables .....	112
Work and Labour Force Status .....	112
Job (Work) Indicator .....	114
Hours Worked in Employment per Week.....	115
Status in Employment.....	116
Occupation.....	116
Industry.....	119
Workplace Address.....	120
Job Search/Availability for Work .....	121



## Section 1: Summary of key findings – by variable

### Overall comments

- There is a common pattern of lower non-response rates for online forms compared with paper forms. Greater use of online forms will reduce item non-response. Use of online forms should be promoted as much as possible.
- Non-response rates for the bilingual paper forms were much higher for most questions than English paper forms. This issue needs to be resolved.
- Lower non-response for online forms may be due to several factors:
  - differences between the characteristics of respondents who use the online form and the characteristics of respondents who use the paper form
  - the online form takes respondents to the next question relevant to them whereas those using the paper form need to use the instructions on the form to find the next relevant question
  - mandatory questions must be answered online before the respondent moves to the next page of the questionnaire
  - the one-by-one presentation of questions that occurs when using the online forms may encourage a response to each individual question. When using the paper forms respondents can see all the questions all at once. For some respondents, using paper forms may lead to a tendency to pick and choose which questions they answer.
- A pattern of respondents putting responses in the wrong place (ie not in the response ovals) which are not captured by scanning. This has contributed to higher than acceptable levels of non-response for a majority of paper forms. This was much worse on bilingual paper forms but also occurred on English forms.
- The most common error in bilingual form completion is marking a new dash to the right of the English question version instead of the circle in the middle. A tick box to the right of text is probably what English completers would expect to see given the current layout of the form. Two instructions may confuse respondents – Māori is printed on the left and English on the right, and the 'how to complete' example illustrates to mark answers to the *right* of the text.
- Some redesign of the forms is necessary to improve the completion of paper forms, particularly the bilingual form (eg instructions, layout, use of colour density/contrast). It would be best to avoid manual grooming of paper forms as this would be very time consuming.
- There is some evidence that responses are not always being captured in scanning. Numeric recognition is not always accurate. Further checking of the processing system is needed.
- There appears to be a drop in response rates for the questions on the last page of the individual form (page 4), possibly related to age. Further analysis may be required.
- Further checking of the online form is needed to make sure that it is working as it should. (eg children under 1 not being routed away from address one year ago, rent amount) – see information on these instances in the summaries below)

#### Notes:

1. Percentages are calculated out of the total stated for all stated responses, and out of the overall total for residual categories, as would be done for output.
2. Images can only be checked for paper forms.
3. The Census test was undertaken in the Whanganui district only. Comparative data used in this report is the 2013 Census Whanganui territorial authority population, not the New Zealand population, unless stated otherwise.
4. The 2018 Census population is likely to have different characteristics to the Whanganui test sample of voluntary participants and an older age group profile.

5. Expectation reports were prepared prior to the April test for Age, Sex, Tenure of Household, Ethnicity, Māori descent, Iwi affiliation and Disability. Warrants of Fitness were completed for Age, Sex and Tenure of Household. These were completed to try and test the process, end-to-end. However, due to tight timeframes, parts of the processing system not being in place, delays in processing etc, these reports are short and very high level. There will need to be more work done to develop the templates for 2018.
6. The incomplete processing system, delays in processing the data – especially the paper forms - and the conflicting objectives of the test – processing were checking the systems and content was checking data quality – have had an impact on the completeness of the analysis.
7. There are a few variables that have not been analysed to date, as the data is not ready for analysis.
  - Absentees
  - Family Type
  - Extended Family Type
  - Household Composition

## Key issues by variable

### Population, Social and Identity

#### Age

- The age profile of the Whanganui sample is biased towards older age groups 55 and over. This is likely to be because the test is voluntary and it was difficult to engage young people with the census.
- Non-response was minimal and not possible online. Around [REDACTED] responses to age, sex and/or census night address were imputed because of missing or incorrect data.
- Age data from the Census Test is not fit for use, as it is not a representative sample of the population. There are no obvious issues with the question itself and if the uptake of internet forms is higher than in 2013, 2018 data should be high quality.
- Our recommendation is to retain the question in its current format.

#### Sex

- Non response is not possible online. Non-response for the bilingual paper form was high but this was a global issue across the variables.
- The proportion of male responses were lower than expected and vice versa for females, probably because the test was voluntary.
- There were [REDACTED] intersex responses overall, or 0.1% of the total stated responses. In a follow up survey [REDACTED] respondents were asked about the third sex category and all said they were comfortable with the question being included on the forms, and understood the concept, although they were not asked if they did identify as intersex.
- [REDACTED] paper responses of 'intersex', multiple response or not stated were imputed.
- Overall the third response category of 'intersex' does not appear to impact on the male and female distribution of responses, but is too small a population to produce a quality population estimate for the intersex population.
- The 2018 Census programme is no longer recommending a sex question with three response options.

#### Ethnicity

- Non-response was low as online responses were mandatory. Paper non-responses were also low apart from on bilingual forms.
- Ethnic group compositions reflected the age biases in the sample.

- Less than 1% of respondents gave the 'New Zealander' response.
- Our recommendation is to retain the question in its current format.

### Māori descent

- Non-response online was not possible as the question was mandatory, but non-response was 14% overall for paper, with the bilingual form non-response rate double that of the English paper form. Overall non-response rate was low compared with 2013.
- Only 2% of respondents answered Don't know to Māori Descent and of these, only 5% said that they knew their iwi.
- Some common paper response issues were:
  - skipping the Māori Descent and Iwi indicator question if not relevant to respondent
  - skipping if already identified as Māori ethnic group and writing in one or more iwi
  - skipping country of birth, Māori Descent and iwi, but continuing with the form
  - putting a dash or tick to the right of the question, not in the box provided
- Our recommendation is to retain the question in its current format. A decision will need to be made on the output of the iwi data for respondents who answer 'don't know' to Māori Descent.

### Iwi affiliation

- Iwi data is not fit for use and analysis **should only be used to indicate possible trends**.
- There was an issue with the capture of iwi data which was not picked up until near the end of the test. Online forms did not capture the iwi of a significant proportion of respondents.
- Some respondents (■) did not answer the iwi indicator but gave a valid iwi on the paper forms.
- When responses were captured correctly, ■ out of ■ respondents named one iwi only.
- Indicative data shows that the iwi question should be retained in its current form with improvements to the data capture of responses.

### Birthplace

- Overall non-response and unidentifiable rates were low.
- Birthplace patterns fit with expectations.
- Note that online there are 2 options, NZ born and overseas, with a drop down box for the latter, whereas on paper a list of most common countries is provided.
- Recommendation is to retain the question in its current format.

### Languages spoken

- Overall non-response was low, and non-response for this question was lower than most other questions on the bilingual form.
- Language responses are within expectations for this variable. Analysis of the 'other' responses indicates that all 'other' responses had a written response in the text box. There were also 7 cases of written responses with no tick boxes marked, slightly lowering the residuals for this question.
- Overall this variable performed well, with slightly higher than ideal non-response on the English paper form. The write-in component of the question is working successfully and the AYT was not in place on the online form so for 2018, which should increase the quality of the data.

### Number of children born

- Frequencies for number of children born are consistent with 2013 data.
- ■ (paper) respondents outside the subject population responded to this question. Some of these responses appear to be due to other related variables (e.g. sex, age) being incorrect for various reasons. Online non-response was very low, paper was higher than acceptable.
- Some quality issues with the subject population, some respondents are wrongly included or excluded because of scanning errors or imputation. Consistency edits and checks are required.
- Issue with text responses for number of children (e.g. TWO).
- Recommendation is to retain the question in its current format, but to drop the term 'alive' which has been an outstanding issue and could upset parents of stillborn children.



## Religious affiliation

- Paper non-response rates were high for this variable.
- Some respondents are skipping question 16 all together and giving a response to question 17 only.
- ‘No religion’ responses are within expectations. Higher levels of ‘no religion’ responses online may be because ‘no religion’ paper respondents skipped the question altogether.
- Proportions of Christian, Buddhism, Hinduism, Islam and Judaism response rates were within expectations. ‘Other religion’ responses were higher than expected.
- Overall results indicate that the question is successfully collecting information on the first level of the religious classification for the categories with tick boxes, but is not collecting good data on lower level categories.
- Category 6 - Māori Religions, Beliefs and Philosophies and Category 7 – Spiritualism and New Age Religions are not being collected by this question.
- Recommend major change - question to no longer include any tick boxes religious groupings, for both level one groupings within the classification (Christianity, Buddhism, Hinduism), and Christian-based religions (e.g. Anglican, Catholic etc.). Instead all respondents who indicate they have a religion will be asked to write in their religion. This may affect the counts of religious groupings, with possible impacts from proposed changes to the online question design, relating to the ‘No Religion’ and ‘Object to Answer’ other fields, autofill and drop down response options

## Relationship to reference person

- Note that this variable comes from two questions on the dwelling form, but is included in the individual (and absentee) datasets, as it is an individual variable. This requires successful linking of individual and dwelling forms (8% overall could not be linked in the test).
- Overall non-response rate is 1.4%. If unlinked numbers are included, non-response increases to 9.3% overall. Therefore the quality of this data is heavily dependent on successful linking of records.
- Results were generally comparable with 2013 data, although there was a noticeable increase in the percentage of reference people and partners, and a decrease in the percentage of children and flatmates, but this fits with the living arrangements data in the test.
- Fewer guests/visitors/inmates/residents in the 2017 data –visitors may be less likely to take part in a voluntary test.
- Higher proportions of reference people in the paper data (indicating smaller households), and higher proportions of children in the online data.
- This data is low quality due to the high percentage of records that could not be linked from the dwelling form to the individual dataset.

## Living arrangements

- Overall non-response rate was low (higher for paper).
- Results were generally comparable with 2013 data, except for a noticeable increase in the percentage of reference people and partners, and a decrease in children and flatmates. Also a higher proportion of those living alone completing paper forms (age effect), and higher families completing online.
- Errors in coding rest home residents to living arrangements; they should be excluded from the dataset as non-private dwelling (NPD) residents.<sup>1</sup> Rest homes are classified as NPD, but ‘residents’ living independently eg townhouse, villa etc. are in private dwellings.
- Data is possibly more accurate than in 2013 because the textbox is being coded. Splitting of coding of terms such as Dad’s wife is a minor issue.
- Edits could eliminate issues such as multiple responses from people living alone.
- “Foster children” is being coded to “Sibling”.

---

<sup>1</sup> A rest home (including any rest home serviced apartments) counts as one NPD. Each villa/townhouse/unit in which residents live fully independently is a private dwelling (usual tenure is licence to occupy). We do not produce family and household data from NPDs.

- “Parents in law” needs to be added to the synonym list.

### Unpaid activities

- Non-response for both paper forms is quite high and does not appear to be directly related to low response rates to the work questions preceding it. Of the ■■■ cases of non-response on the English paper forms, 73% were 65 years or older. This indicates that older respondents are disproportionately not responding to this section of the paper form. This is consistent with ■■■ out of ■■■ non-responses to the work indicator question being respondents of 65 years or older.
- The percentage totals for all of the categories are comparable to 2013 results. No issues were identified with the counts for the activities.
- Overall this variable performed well, with the exception of high non-response to both English and bilingual paper forms. These issues are identified as being wider than this variable, and as such this question format is recommended for the 2018 Census.

## Health Variables

### Disability/activity limitations

- Overall, non-response rates for the question set were acceptable. The data produced for the overall disability indicator should be of suitable quality for output. Unidentifiable responses were relatively few.
- Non-response rates to individual questions within the question set were acceptable. Non-response was highest for the ‘hearing’ question; feedback from cognitive testing indicated some respondents were finding the wording confusing, as it includes a reference to using a hearing aid.
- Overall level of 9.2% disabled population is within the expected range. Paper responses for disability were higher, probably reflecting the tendency for older people to respond on paper.
- The individual question with the highest contribution to the disability indicator is the walking question. This is consistent with the July 2016 Test results.
- Overall the results for this variable indicate that the question set is working successfully in producing the overall disability indicator. While non-response to individual questions on the paper forms is higher than ideal, this is not impacting on the ability to output this variable.

### Cigarette smoking behaviour

- Non response for this variable overall is acceptable, other than the bilingual form. A small number of respondents who answered ‘no’ to current smoking behavior (Q25) are skipping the second question on ex-smoking on both paper and online formats.
- The totals for the regular smoker and ex-smoker categories are quite low compared to 2013. While smoking rates are likely to be trending down, this shift is likely to be related to the older respondent population to the 2017 test, as overall smoking prevalence peaks in the age bracket of 25-34.
- Overall this variable performed well, with some incidence of non-response across the two questions. It is recommended to check if we are using any previous census data to inform the ex-smoker output, as this may help with data quality for the ex-smoker counts.

## Location Variables

### Years at usual residence

- Overall, non-response rate was acceptable
- 61% of all respondents have been at their current usual residence for less than 10 years, similar to 2013.

- Recommendation is to retain the question in its current format.

#### Years since arrival in New Zealand

- Some paper respondents born in NZ (N = ■) did not follow the routing and answered the question.
- Non-response rate for overseas-born New Zealand residents was above acceptable levels (32%), but the majority of these were paper respondents who did not answer the filter question of country of birth. Only ■ people overseas-born New Zealand resident who responded to country of birth did not respond to the arrival question.
- Data fits within expectations overall.
- The question does not work well for paper respondents who do not answer country of birth. Could routing be improved to make it clear to overseas-born that they should answer this question?

#### Person record type, usual residence address and census night address

- Adults and children had been processed and coded correctly into person record type and residence status.
- Ratio of adults to children higher than in 2013 (80:20) but reasonable given the low test response rate. Children were more likely than adults to complete online.
- Coding of overseas and New Zealand residents using the tick boxes in Q4 (where do you usually live) showed 23 errors (see table below in red), which were mostly respondent error on the paper forms.
- ■ paper records did not have a census night address or a usual residence address. Some of these were overseas-born students who did not understand the concept of usual residence. Others did not fill in an address in either place but ticked the boxes.
- Around half of all paper respondents did not tick either of the usual residence tick boxes provided (in New Zealand or Overseas).
- The derived usual residence indicator showed a majority (close to 90%) were at their usual residence on census night.
- Recommendation that paper forms are checked manually when responses to usual residence and/or census night address are inconsistent, for example respondents who say they live overseas but give a New Zealand address for their usual residence, or overseas students who live in boarding schools. Additional guide notes may assist respondents to work out what is their 'usual address'.

#### Usual residence one year ago

- This is a new question so we do not have any comparative data from 2013.
- Non-response was low (3%), and unidentifiable responses were mainly respondents ticking 2 boxes.
- Children under 1 year old were not routed away from the question online. This error in the ICS needs to be corrected for 2018.
- Responses look feasible, apart from the not born 1 year ago category which was too low, as noted above.
- The "at my census night address" tick box confuses paper respondents. It was used by less than 1% of online respondents, and most of those that did tick it had the same census night and usual residence address. Very few respondents actually lived at their census night address one year ago unless it was also their usual residence.
- Only a small number of respondents (■) were at their usual address one year ago and also said they had been at their usual residence 'less than one year'. ■ respondents who said they were at their usual residence one year ago gave a numerical response of less than one year for this question.
- Recommend the tick box 'at my census night address given in 6' is removed from the forms, and the space is allocated to collecting the country of residence for those who were overseas 1 year ago.

## Education Variables

### Highest secondary school qualification

- Non-response to this question was acceptable overall; paper unidentifiable and not stated higher than online.
- Data comparable to 2013, with a decrease in respondents with no secondary school qualification. This fits with the trend for increasing formal qualifications over time.
- There was a direct relationship between the level of qualification and the mode of completion which is probably age-related; respondents with low level school qualifications (generally older respondents) were more likely to complete paper forms.
- Our recommendation is to retain the question in its current format.

### Highest post-school qualification

- Note that there has been a change from an open write-in box to a tick box format in 2017.
- Both internet and paper samples had a high proportion of non-response. Unidentifiable responses accounted for 60% of all residuals, mostly vague responses that could not be coded– for example occupations or job titles instead of actual qualifications.
- One in four respondents who ticked Yes to the post school qualification indicator did not tick a qualification level but wrote in the text box below. Around 7% of all text box responses were unable to be coded from the codefile. The As-You-Type suggestions may have contributed to the lower non-response rate online.
- Respondents are more likely to answer the question than to skip it, but around 1 in 4 respondents are writing a response into the text box rather than marking a qualification level.
- Test data indicates that this question is causing some issues for respondents in selecting a qualification level, especially on paper, but the new question format will reduce the manual coding burden. It is recommended that this question be included in census. Improvements in the codefile may assist in more accurate coding of the variable.

### Highest qualification (Derivation)

- Derived from highest secondary school and post-school qualifications
- High proportion of unidentifiable paper responses, because of high non-response to the base questions.
- Data fits with expectations

### New Zealand/Overseas post-school qualification indicator

- New question for 2018 Census
- Sample falls within acceptable non-response rate
- Findings fit with expectations from the July 2016 test data
- No major issues identified, question should produce fit for purpose data

### Field of study

- Overall the non-response rate was within the acceptable range
- Data compares well with 2013 with a similar pattern of high-level qualification categories.
- Detailed level field of study responses do not show any unexpected trends
- Test data indicates that this question working acceptably and would likely produce fit for purpose data. Our recommendation is to retain the question in its current format.

### Study participation

- Subject population now all New Zealand residents, previously NZ adults only; filter question for the new travel to education question.
- Non-response was within acceptable range. Only █ people gave a multiple response (e.g. full and part-time)

- A decision on how to code multiple responses is needed. Household surveys recode a person who ticks full-time AND part-time study to full-time (code 1), processing codes it to unidentifiable (code 7) and the draft derivation codes multiple responses to code 3 (part-time and full-time study).
- Note that the draft derivation needs to be corrected - 'not studying' output code should be 4 not 3.
- Recommendation is to retain the question in its current format. Decision on multiple responses to be made.

## Income Variables

### Total Income

- Not stated non-response for paper forms was unacceptably high, particularly for bilingual forms. This has not been investigated to see if it was responses in the wrong place on the bilingual form or true non-response.
- The data looks sensible when compared with 2013 Census and expected increases in income levels over time. This limited analysis shows no new issues with this data.
- Although increases in online responses may improve data quality in 2018, bias may still be an issue – the test data shows paper form respondents tend to have lower income levels and higher levels of non-response than online respondents.

### Sources of income

- Overall non-response was low (well within the acceptable range) and lower than national non-response in the previous three censuses
- The distribution of responses looked sensible – mostly similar to previous census data, but with some differences due to the skewed nature of the sample for this voluntary test. There do not appear to be any issues arising from the changes to some category names.
- It looks like greater use of online forms may be effective in lowering non-response (and increasing data quality) for this variable.

## Tenure of Household and associated variables

### Dwelling owned or in family trust

- Overall non-response was within the acceptable range, but non-response for paper forms was quite high and mostly true non-response
- Some paper respondents skipped all tenure-related questions, then completed the rest of the form, or answered certain tenure-related questions only
- Multiple response of owned and family trust should be coded to family trust for tenure of household.
- Test data is skewed toward home owners, probably due to the voluntary nature of the test and characteristics of those who responded. Family trust responses were as expected.
- Suitable for inclusion in 2018. For online forms, need a help note explaining how to answer if the dwelling is partly owned and partly in a family trust (mark family trust).

### Mortgage payments

- Overall non-response was within the acceptable range
- Paper non-response was too high. Respondents were not following routing instructions correctly. A common pattern was for homeowners to miss the “go to 9” instruction in Q5 and answer the rent questions.

- Our recommendation is to review the placement of the “go to 9” instruction on the paper form.

### Rent Indicator

- Overall non-response was within the acceptable range
- Higher proportion than expected of respondents saying they did not pay rent, but some of these appeared incorrect. In census proper further investigation would be done to deal with this.
- This question is suitable for inclusion in the 2018 Census in this format.
- Use of online forms (which present questions one at a time) will be a major factor in minimising non-response and ensuring good quality output data for households who rent their home.

### Weekly rent paid by household

- Overall non-response is around the acceptable range, but higher than in 2013.
- Some odd results for online forms – most online respondents coded to ‘not stated’ answered rent period without giving a rent amount. Online form should be checked for possible errors in display or capture.
- Some respondents put answers in the wrong place.
- Minimal cases of very high, incorrect, rent amounts – were due to respondents not using the pre-printed decimal place and cents or crossing out the first two boxes.

### Sector of landlord

- Overall non-response was in the acceptable range.
- No evidence of major issues with the new categories, but some evidence of people answering ‘other community provider’ when it is a licence to occupy dwelling in a retirement village.
- Numbers in new categories are likely to be small, so any error affecting them could have a significant effect on data quality. In 2018, responses in these categories should be checked to make sure they seem correct.

### Tenure of household

- Overall percentage of not stated is acceptably low, but data quality for paper forms is lower – reflects higher non-response to input questions used to derive tenure of household.
- Test data skewed toward homeowners, reflecting biased sample.
- Some respondents didn't follow questionnaire routing correctly – review of questionnaire design recommended.
- Some evidence that people with a licence to occupy do not answer correctly and/or are confused by the tenure-related questions – strategies to help them answer correctly would be useful (eg flyers in retirement villages on how to answer these questions, group census form filling sessions with help provided).
- This variable appears to be working sufficiently well. It is difficult to tell how much the more condensed questionnaire design has improved data quality.

### Individual home ownership

- Non-response is remarkably low – maybe due to the more compliant nature of respondents in this voluntary test, and to test data being skewed toward home-owners  
No evidence of any major problems with this new style of question.

## Occupied Dwelling Type and associated variables

### Dwelling description

- Overall non-response was in the acceptable range.
- Responses look sensible. Most respondents appear to be making good (appropriate) use of the new tick boxes in this simplified dwelling description question.
- Some respondents seem to be giving a written response because the term they use to describe their home is not on the form eg villa, flat.

- In general this question appears to be working well and suitable for inclusion in this format in the 2018 Census.

#### Dwelling joined or separate

- Overall non-response was in the acceptable range.
- This question appears to be working well and suitable for inclusion in this format in the 2018 Census.

#### Number of storeys

- Overall the non-response rate was acceptable. Most non-response for paper forms was on English forms.
- This question appears to be working satisfactorily and suitable for inclusion in this format in the 2018 Census.
- Unfortunately test data from Whanganui does not tell us if this question works for respondents in high-rise apartments because Whanganui doesn't have this type of housing.

#### Occupied dwelling type

- Data as expected for the area where this test was done.
- Some evidence of improved data quality compared with 2013. The new questionnaire design may be contributing to this.
- The questions used to derive this variable appear to be working well and suitable for inclusion in this new format in the 2018 Census.
- The occupied dwelling type derivation needs to be checked as a separate exercise to make sure it is working correctly.

#### Unoccupied dwelling type

- There was a low number of unoccupied dwellings in this test. The data looks different than 2013 but may not be comparable with 2013 due to de-scoping of non-response follow-up.
- As far as is possible to tell from this test, it appears that we will still be able to produce data on unoccupied, empty dwellings and unoccupied, residents away dwellings, but it is difficult to tell how good the quality of this test data is.

#### Number of rooms

- New question style appears to be mostly working well and suitable for inclusion in the 2018 Census.
- Overall not stated is higher than desirable. Several causes: true non-response, insufficient response, marking a box instead of giving a number, responses in wrong place, or responses not being captured.
- Version A of the derivation has been used for this test analysis. The results using version A need to be compared with version B and a decision on which version will be used for 2018 needs to be made.

#### Number of bedrooms

- High not stated for paper forms – often seems to be true non-response.
- Some respondents mark the box instead of giving a number.
- New questionnaire design might be negatively affecting data quality for bedrooms data – it is recommended that minor questionnaire design changes be considered eg move bedrooms response option to the top, put “count” in bold.
- No evidence of numeric misrecognition for high bedroom counts

#### Numbers of other room types

- Non-response for some room types was high (eg dining rooms) or extremely high (conservatories and studies) but the derivation allows for this.
- Overall non-response for kitchens, living rooms (and bedrooms) was acceptably low, so it should still be possible to produce total rooms data that is of good enough quality to be output.
- Respondents often leave boxes blank if they don't have that type of room, as in previous tests.
- Some respondents answered with a dash instead of a number. These respondents still usually put numbers for bedrooms – suggest moving bedrooms to top of list.

## Housing Quality Variables

### Main types of heating

- Overall non-response is low. It was also within the acceptable level for those using paper forms.
- Many multiple responses, as expected.
- Very few inconsistent multiple responses (heating and no heating)
- Responses looked sensible, and there were few cases where fuel type could not be coded from heating type.
- Written responses were relatively low (5%); adding or changing response categories used in the question could reduce written responses further.
- The question used in this test is performing well both for online and paper and is suitable for inclusion in this format in the 2018 Census.

### Dwelling mould indicator

- Non-response is acceptably low overall. Low level of 'don't know' responses.
- The data looks sensible overall and shows the expected patterns.
- There do not appear to be any issues with respondents of particular household tenures or landlord types being unwilling to answer.
- The test data indicates that this question is working acceptably well and is suitable for inclusion in this format in the 2018 Census.

### Dwelling dampness indicator

- Overall non-response is acceptably low. Low level of 'don't know' responses.
- There don't appear to be any issues with respondents of particular household tenures or landlord types being unwilling to answer this question.
- The data looks sensible overall, and generally as expected
- The test data indicates that this variable and question is suitable for inclusion in the 2018 Census.

### Access to basic amenities

- Overall non-response is acceptably low.
- The data looks sensible and as expected. The vast majority of dwellings have all amenities listed in the question.
- The amenities most likely to be missing were: drinkable tap water, a fridge. The follow-up survey indicated some concern about water quality in this area. This may be related to the flooding in Whanganui at the time of the test.
- The follow-up survey also provided some evidence that respondent error to this question has contributed to the number of dwellings that apparently lack one or more amenities.
- This variable and question appear to be suitable for inclusion in the 2018 Census. However, given the element of respondent error identified in the follow-up survey, it is recommended that checks be done during evaluation to try to identify such errors and caution be applied when outputting this data and drawing conclusions from it.

## Access to Telecommunication systems

- Non-response was acceptably low, both on paper and online.
- Only a low amount of apparent inconsistent multiple responses.
- The data looks sensible with increases in internet and cellphones, decreases in landlines, and a very low proportion with no access to any telecommunication systems.
- The question is suitable for inclusion in the 2018 Census in this format.
- Note this variable was coded incorrectly in the test – this needs to be fixed for census proper.



## Transport Variables

### Main means of travel to work

- Non-response to this question appears to be exceptionally low, even on paper.
- The relative frequencies of the different modes look sensible.
- The number of 'other' responses was relatively low, suggesting that the selection of response boxes provided worked well.
- The new (changed) question appears to be working very well and is suitable for inclusion in 2018 in this format.
- One aspect that may need discussion is whether the motorbike category should be reinstated.

### Workplace address (analysis of non-response only)

- Overall non-response was fairly acceptable. For the paper form, complete non-response was higher than acceptable and non-response to individual parts of the question was very high, particularly for building name and suburb.
- It appears that some respondents do not know some details of their workplace address, have privacy concerns, or don't see the relevance of the question.
- This (unchanged) question remains suitable for inclusion.
- Promoting online completion may be particularly useful for ensuring the quality of this data.

### Travel to education

- There was almost no non-response, even for those using paper forms.
- There was a data quality issue for paper forms – respondents giving a multiple response that may be modes used for journeys on different days.
- Minor issue of children driving themselves to school – easily fixed with an edit for census proper.
- The data looks mostly sensible and as expected overall. The number of 'other' responses is quite low which suggests that the selection of response boxes provided worked well.
- In general this question appears to be working well and is suitable for inclusion in this format for the 2018 Census.
- Most respondents in each age group appear to be answering correctly.

### Educational institution location

- Non-response for educational institution name was acceptably low and non-response for city was fairly acceptable, but non-response for suburb was very high.
- Usually it should still be possible to determine the geographic location of the educational institution if suburb is missing but the other information has been provided.
- Some respondents put answers in a different box to that intended – having the processing system allow for this would help maximise data quality.
- This variable is suitable for inclusion in the 2018 Census in the current format. Although many respondents do not answer it fully, it is expected that the information they do provide will still be sufficient to produce good quality data.
- It appears that this information can be successfully collected for pre-schoolers as well as for school and tertiary students. These results suggest that data quality issues are more likely to affect the data for tertiary students.
- For the online form, it is recommended that questionnaire design/presentation of this question be reviewed to see if the quality of responding from tertiary students can be improved.
- To achieve the best data quality possible, automatic coding processes may need to allow for responses being in a different place than intended.

### Number of motor vehicles

- The overall level of non-response is acceptable
- The overall distribution of responses looks sensible.
- Checking responses for three or more vehicles showed no evidence of numeric misrecognition and little evidence of obvious respondent error.
- This question appears to be working well and is suitable for inclusion in 2018 in this format.
- The planned change to produce more detailed output (including exact vehicle number) should be possible to implement.

## Work Variables

### Work and Labour Force Status

- Issue with the capture of hours worked (online forms), which may have impacted on the data by inflating the proportion of full-time workers and vice versa for part-time. Needs to be corrected for 2018.
- Relatively high level of non-response to the work questions on paper meant 24% of responses were imputed.

### Job (Work) Indicator

- Paper non response rate was relatively high, particularly the 65 and over age group, who may not see the relevance in answering the question. Online response rates were good.
- Data looks sensible when analysed by age- reflects bias towards older respondents on paper.

### Hours Worked in Employment per Week

- Issue with capture of hours worked online, which impacted on the data. A work-around was used to populate the hours worked in “all other jobs”, namely that job hours 1 = job hours 2, which must be addressed in 2018.
- 65 respondents answered 0 hours to main job; responses were coded to 888 (out of scope). Some of these respondents worked, others were on a benefit or retired, there was no clear pattern to the responses.
- Analysis of hours worked in ‘main job’ correlates with the 2013 Census data. Non-response was low for main job, unable to comment on ‘other jobs’.

### Occupation

- Occupation coding appears to have worked very well, probably due in part to prompts from the As-You-Type lists. The number of occupations that could only be coded to Not Elsewhere Classified were relatively small.

### Industry

- There was a large proportion of unidentifiable responses (over 1 in 3), due to the low processing priority given to this variable in the test.

### Workplace Address

- Physical workplace address was not analysed for this report.
- Most respondents worked away from home in their main job.

### Job Search/Availability for Work

- The majority of respondents who should have answered this question gave valid responses.
- Most respondents were not looking for paid work in the past 4 weeks. 2/3 of respondents not seeking work and unavailable for work were aged 65 and over.
- Of those respondents who were actively seeking work, the most common method used was looking at job advertisements.

### **Census Variables Not Included in the Report:**

- Absentees
- Family Type
- Extended Family Type
- Household Composition

The above 2017 census test family and household variables will not be analysed in this report.

## Section 2: Detailed Analysis of Variables

### Individual Variables – Social, Population, Identity

#### Age

**2** When were you born?

day (eg 28)	month (eg 2)	year you were born (eg 1984)
<input type="text"/>	<input type="text"/>	<input type="text"/>

Subject population: total census night population (after imputation)

#### Comments:

There are some paper responses not being captured, for example a 2-digit year is not recognised, or a written form of a day or month (e.g. Nov), but not widespread. The raw data shows a few responses where raw age was blank and the imputation did not fit with respondent characteristics (e.g. gets super at age 19, married at age 2), which may affect analysis of other variables by age but is a minor issue.

*Non-response* -The paper sample non-response was 111 (0.9%) and there was 1 internet non-response (mandatory question). Missing age responses will be imputed.

*Response unidentifiable* - There were a few “odd” responses in the paper forms, probably facetious.

#### Paper and Online responses

Graph below shows that fewer respondents aged 65+ completed their forms online; 1/3 of older respondents used paper forms, compared with around 1/5 of younger respondents. The table below shows the data prior to imputation.

Age Group	Online		Paper		Total	
	Number	%	Number	%	Number	%
No valid age given		0.9		99.1		100
0-14		84.1		15.9		100
15-39		79.4		20.6		100
40-64		81.6		18.4		100
65+		65.8		34.2		100
Total		76.4		23.6		100

The age profile of the Whanganui population completing the test was biased towards older age groups, particularly the 65 and over group. This is likely to reflect the voluntary nature of the test and the difficulty in getting young people engaged with the census.

Age Group Distribution	2017 Test	Expectation (based on 2013 data)
0 to 14 years (child)	15%	19 - 21%
15 to 39 years	19%	24 - 26%
40 to 64 years	36%	33 - 35%

65 and over

30%

21 - 23%

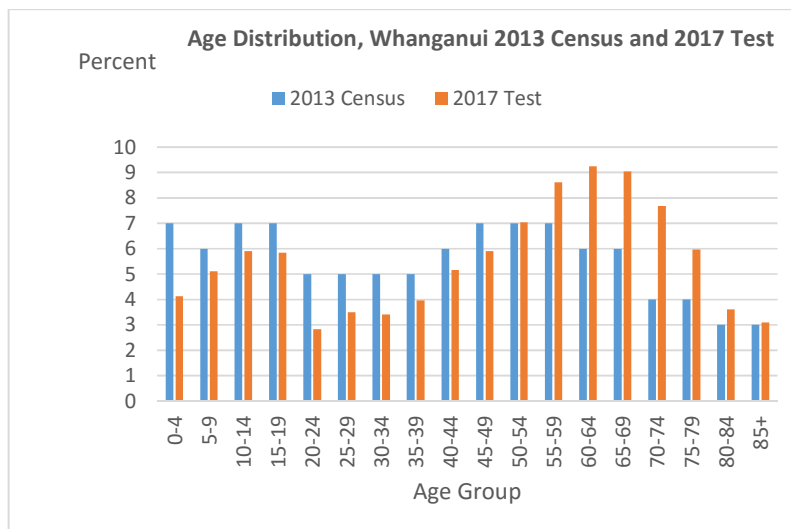
Detailed Age Groups by Mode, Census 2017 Test

Age Group	Online	%	Paper	%	Total	%
0-4		4.6		3.0		4.2
5-9		6.0		2.1		5.1
10-14		5.9		5.6		5.9
<b>0-14</b>		<b>16.6</b>		<b>10.7</b>		<b>15.2</b>
15-19		5.3		7.5		5.8
20-24		3.0		2.1		2.8
25-29		3.7		2.9		3.5
30-34		3.8		2.0		3.4
35-39		4.4		2.6		4.0
<b>15-39</b>		<b>20.2</b>		<b>17.1</b>		<b>19.4</b>
40-44		5.9		2.7		5.1
45-49		6.5		4.0		5.9
50-54		7.4		5.6		7.0
55-59		8.9		7.6		8.6
60-64		9.5		8.4		9.2
<b>40-64</b>		<b>38.1</b>		<b>28.4</b>		<b>35.8</b>
65-69		9.1		8.9		9.0
70-74		7.2		9.3		7.7
75-79		4.9		9.6		6.0
80-84		2.5		7.4		3.7
85+		1.5		8.6		3.1
<b>65+</b>		<b>25.1</b>		<b>43.8</b>		<b>29.5</b>
TOTAL		100.0		100.0		100.0

Note: Data 31 May, shows dataset after age imputation.

Analysis of responses

A more in-depth analysis of age reinforces the skewed age distribution of the Whanganui test sample towards older respondents 55+.



*Imputation:* Just over [redacted] responses to age, sex and/or census night address were imputed because of missing or incorrect data as shown in the table below.

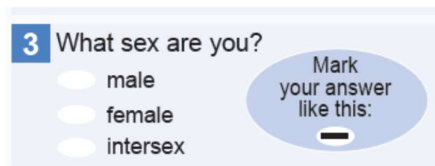
Type of Imputation	Online form	Paper form	Total
bday			
bday,bmonth			
bday,bmonth,byear			
bday,byear			
bmonth			
bmonth,byear			
byear			
cn_address			
sex			
sex,bday			
sex,bday,bmonth			
sex,bday,bmonth,byear			
sex,bday,bmonth,byear,cn_address			
sex,bmonth			
sex,byear			
TOTAL IMPUTED			
Not imputed			

*Assessment*

Age data from the Whanganui Census Test is not fit for use, as it is not a representative sample of the population. There are no obvious issues with the question itself and if the uptake of internet forms is higher than in 2013, 2018 data should be of higher quality.

Our recommendation is to retain the question in its current format.

**Sex**



Subject population: total census night population

*Comments:*

*Non-response* - Non response is not possible online. Non-response for the bilingual form was high at 14.7%, but consistent with other variables when the bilingual form was used. It is not an issue specific to the sex variable (however interest in the term used for 'intersex' on the Te Reo side of the online form should be noted).

The non-response for the English paper form was 2.0%. This is higher than the overall expectation of less than 1%. It is possible that the three response options increased the level of non-response to this variable. Non-

response for the three response option sex question in the July 2016 Census Test was 0.8%. This should be compared against non-response levels for the English paper form overall to assess whether non-response levels are comparatively high. As a priority one variable, higher non-response to sex is of some concern.

*Response unidentifiable* - There were two cases of multiple response on the paper forms.

*Analysis of Responses*

April 2017 Census Test – total New Zealand population								
Sex	Internet form		English paper form		Bilingual paper form		Grand Total	
	Number	%	Number	%	Number	%	Number	%
Male		45.7		44.6		47.7		45.6
Female		54.2		55.4		52.1		54.3
Intersex		0.1		0.0		0.2		0.1
Total Stated		100.0		100.0		100.0		100.0
Responses unidentifiable				0.0		0.1		0.0
Not Stated				2.0		14.7		1.2
Residual Codes				2.0		14.8		1.2
Grand Total		100.0		100.0		100.0		100.0

*Male Responses* – The overall levels of male responses at 45.6% is slightly lower than expected, with an expected range of 46.5-48.5%. This is likely due to the nature of the test as voluntary.

*Female responses* – The overall levels of male responses at 54.3% is slightly higher than expected, with an expected range of 51.5-53.5%. This is likely due to the nature of the test as voluntary.

*Intersex responses* – There were [redacted] intersex responses overall. At 0.1% of the total stated responses, this falls in line with expectations of less than 0.5%. Compared to the July 2016 Census test which had 0.2% of ‘indeterminate’ responses of which a high number were facetious or made in error, this indicates the term ‘intersex’ may be working more effectively, or respondents for this test were less likely to respond facetiously. Three respondents who stated ‘intersex’ were successfully called back as part of the follow up survey. All three of these respondents indicated they were comfortable with the question being included on the forms, and understood that the concept meant something other than male or female. While there was no direct question about whether the respondent identified as intersex, none of the respondents indicated that they were intersex.

*Edits and imputation*

- [redacted] records of ‘intersex’, multiple response or not stated were imputed.
- [redacted] intersex responses were coded to [redacted] males and [redacted] females
- [redacted] responses unidentifiable were coded to [redacted] males
- [redacted] response unidentifiable was male and intersex / 1 was female and intersex (CHECK: is this following the imputation rules?)
- [redacted] records with sex not stated were coded to 88 males and 97 females

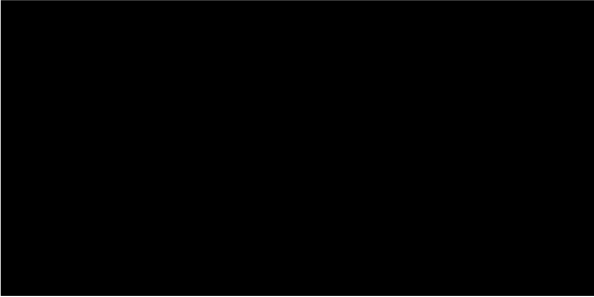
*Summary comments*

Overall the results for this variable indicate that while the third response category of ‘intersex’ is not impacting on the male and female distribution of responses, the responses to this category are not likely to produce a quality population estimate for the intersex population. This is due to the very small size of the population and the error which occurs in a self-completed questionnaire which is relatively large by comparison.





*Total Response Data at Level 1*

<b>Ethnic Group (Total Response Data)</b>	<b>Internet %</b>	<b>Paper %</b>	<b>Total</b>	<b>Expectation (based on 2013 data)</b>
European				81-83%
Māori				24-26%
Pacific				3%
Asian				4%
MELAA				0.6%
Other				
<i>Not stated/residual</i>				2-3%

*Detailed Analysis of Ethnic Group Response Types.*

Note that the following analysis only includes categories of 20 or more responses. Further analysis may be possible if derivations are created, as this analysis is only combination responses.

The most common response combination was NZ European, which fits with expectations. Māori/European and Māori only were the next most common responses. Indian and Chinese were the most common Asian ethnic groups, and Samoan was the only Pacific group with 20+ responses. Less than 1% of respondents gave a 'New Zealander' response.

The pattern of ethnic responses also fits with expectations from the Whanganui profile (2013 Census).

<b>Ethnic Group Response</b>	<b>Percent of respondents (where number of responses = 20 or more)</b>
NZ European	73.2
Māori/European	6.6
Māori	6.4
English	1.6
Indian	1.2
No response	1.0
Chinese	0.7
New Zealander	0.7
Australian	0.5
Other European	0.4
Dutch	0.3
German	0.3
Samoan	0.3
South East Asian	0.3
British NFD	0.3
American	0.3
Scottish	0.2
NZ European/Dutch	0.2
NZ European/Samoan	0.2
Korean	0.2
Japanese	0.2

Canadian	0.1
European NFD	0.1

*Assessment* – The data indicates that the current ethnic question would likely produce higher quality data in the 2018 Census than in 2013, as non-response should be lower with higher completion of internet forms.

Our recommendation is to retain the question in its current format.

## Māori Descent

11 Are you descended from a Māori (that is, did you have a Māori birth parent, grandparent or great-grandparent, etc)?

yes → go to 12

don't know → go to 12

no → go to 14

Subject population: total New Zealand resident population (overseas residents are excluded from output and analysis)

Comments:

*Non-response* – The paper sample non-response was 13.6%; there was only one internet non-response. The impact of the change to a Priority One variable with a mandatory response required online was positive. The overall non-response rate of 3.2% was an improvement from 2013 and an acceptable response level, but paper non-response is high. However, in the 2013 Census, a higher proportion (15.6% of Whanganui residents) did not respond to the Māori Descent question.

*Response unidentifiable* – There was one unidentifiable paper response. Unidentifiable responses were not possible on the internet form.

*Bilingual and English paper forms* - As might be expected, the proportion of people with Māori descent was much higher in the bilingual form respondent group. However, non-response for the bilingual forms was twice the rate of English forms (22% compared with 11%), highlighting the difficulty respondents had in completing the question in the current format.

*Don't Know responses* – 2.3% of respondents did not know whether they had Māori descent (and ticked Don't Know), similar to 2013. Respondents who ticked 'Don't Know' were able to answer the iwi indicator question on all forms - 5.2% of the respondents who answered the filter question said that they 'Know their Iwi', a decrease from 2013 (22%). The high uptake of internet forms may have contributed to the decrease.

*Overseas residents* – 37 paper respondents (17.8%) did not follow routing and completed the question.

*Māori Descent responses* – 16.6% was lower than expected (26% of the Wanganui population as Māori Descent in 2013), but acceptable given the voluntary nature of the test and the bias towards older respondents (Māori descent population is relatively youthful compared with non-Māori). There was also a lower than expected proportion of Māori ethnic group respondents in the test.

*Non-Māori Descent responses* – as above, higher than expected, but within an acceptable range for the test.

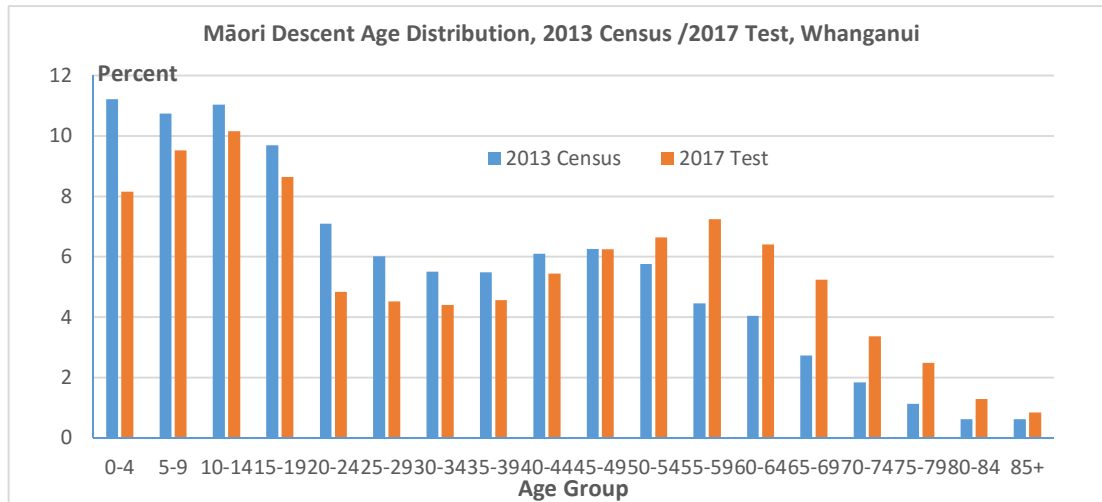
**April 2017 Census Test - NZ Residents, Whanganui, Māori Descent**

Māori Descent	Internet form	% stated	Paper form	% stated	Grand Total	% stated
Yes		16.6		18.2		16.9
No		81.2		79.0		80.7
Don't know		2.2		2.9		2.3
Yes and No		0.0		0.0		0.0
Total Stated		100		100		100
Not Stated		0.0		13.6		3.2
TOTAL						

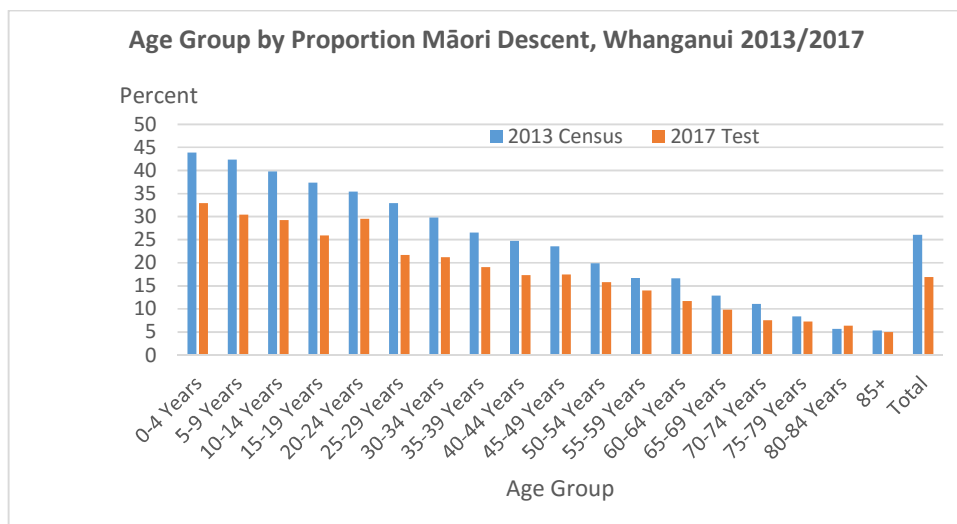
**Iwi Indicator Responses, Māori Descent Respondents**

	Yes	%	No	%	Don't Know	%
Iwi Known		78.6		2.9		5.2
Iwi Not Known		21.4		97.1		94.8

*Age profile* – The Māori descent age distribution profile of Whanganui 2017 Test respondents is biased towards the 50 and over group, compared with the 2013 Census (see graph below).



The next graph shows the proportion of people with Māori descent in each age group compared with the 2013 Census, confirming the bias towards older Māori descent respondents. This may help explain the lower response rate for Māori descent in the test.



*Other comments* – There seems to be a few patterns to the paper non-responses.

- Skipping the question and the following iwi questions if it is not relevant to the respondent (i.e. non-Māori)
- Skipping the question (and the iwi indicator) but identify as Māori ethnic group and write in one or more iwi
- Skipping country of birth, MD and iwi, but continuing on the next page of the form
- Filling out the question but draw a dash or tick to the right of the question rather than in the box provided

*Assessment* – This test data indicates that a Māori descent question in this format for the New Zealand population would likely produce high quality data. We would expect the non-response rate to drop significantly if a majority of respondents complete the form online.

Our recommendation is to retain the question in its current format.

Note: If the current routing from Māori descent to iwi indicator is retained for the 2018 Census (Yes and Don't Know can answer), a decision will need to be made on the output of this data, which traditionally has not been output in the past.

## Iwi

Notes: There was an issue with the capture of iwi data which was not picked up until near the end of the test. Online forms only captured the iwi of the first person in the household to complete the form. Results were distorted as it looked as if nearly half of all Māori descent respondents did not name an iwi in the test.

Results may also be impacted by the targeted approach to the distribution of paper bilingual forms and the relatively poor completion of these forms.

Iwi data is not fit for use and analysis of iwi groups **should only be used to indicate possible trends**.



Iwi responses	Online	%	Paper	%	Total	%
Name of iwi(s) given*		47.9		88.7		55.2
Unidentifiable		9.4		8.3		9.2
Out of Scope		0.4		1.5		0.6
Not stated		42.3		1.5		35.0
Grand Total		100.0		100.0		100.0

\*excludes a small number of responses not given in first row of grid

#### *Iwi responses (indicative only)*

Five iwi fields can be filled out online and 4 on paper. Of those whose responses to iwi were captured in the test, 2/3 named one iwi only. The proportion of respondents giving multiple responses to the iwi question in the test was lower than in the 2013 Census.

Common Iwi groups for Whanganui district were consistent with the 2013 Census results. The proportions of responses in each group were not the same, which is not unexpected given the voluntary nature of the test, and problems with the data, but the top 10 iwi named were the same as in 2013, as follows (in 2017 order):

#### **Common Iwi as a proportion of all iwi named by respondents (Whanganui District) 2013/2017**

<i>Iwi</i>	<i>2013 Census</i>	<i>2017 Census Test</i>
<i>Percent of respondents</i>		
Te Ati Haunui-a-Pāpārangi	31	28
Ngāpuhi	9	7
Ngā Rauru	9	6
Ngāi Tahu / Kāi Tahu	5	5
Ngāti Porou	7	5
Te Atiawa (Taranaki)	3	4
Ngāti Tūwharetoa	10	4
Ngāti Apa (Rangitīkei)	8	4
Ngāti Maniapoto	4	3
Ngāti Raukawa, region unspecified	4	3
Ngāti Kahungunu ki Te Wairoa	3	2
Tūhoe	3	2

#### **Māori Descent Respondents - Common Iwi Responses (>20 respondents)\***

<i>Iwi</i>	<i>Know Iwi</i>	<i>Don't Know Iwi</i>	<i>Not Stated</i>	<i>TOTAL</i>
Te Ati Haunui-a-Pāpārangi				
Ngāpuhi				
Ngā Rauru				
Ngāi Tahu / Kāi Tahu				
Ngāti Porou				
Te Atiawa (Taranaki)				
Ngāti Tūwharetoa				
Ngāti Apa (Rangitīkei)				
Ngāti Maniapoto				
Ngāti Raukawa, region unspecified				
Ngāti Kahungunu ki Te Wairoa				
Tūhoe				
Unidentifiable				
Response Outside Scope				
Not Stated				

\* Does not include responses where the respondent did not write in the first response box on the form

*Recommendation* – indicative data shows that the iwi question should be retained in its current form with improvements to the data capture of responses.

## Religious affiliation

### First question – Question 16

**16** What is your religion?

no religion → go to **18**

Christianity

Buddhism

Hinduism

Islam

Judaism

other religion → go to **17**

or  object to answering → go to **18**

Subject population: New Zealand usual residents

April 2017 Census Test – total New Zealand population								
Religious affiliation	Internet form		English paper form		Bilingual paper form		Grand Total	
	Number	%	Number	%	Number	%	Number	%
No Religion		48.2		38.5		40.7		46.5
Christianity		44.5		54.8		57.1		46.4
Buddhism		0.7		0.9		1.0		0.7
Hinduism		0.5		0.5		0.0		0.5
Islam		0.2		0.5		0.1		0.3
Judaism		0.1		0.4		0.1		0.1
Other religion		5.8		7.3		6.7		6.1
Total Stated								
Object to answering		4.8		6.5		5.1		5.1
Responses unidentifiable								
Not Stated		0.4		<b>8.9</b>		<b>14.7</b>		3.0
Residual Codes		5.2		15.5		24.9		8.1
Grand Total								

Please note this question is multiple response and therefore totals will not always add to 100 percent

Comments:

#### *Non-response*

Non-response rates for both of the paper forms is high for this variable. The level of non-response on the bilingual form is consistent with non-response across all variables for the bilingual forms.

For the English paper forms, the level of non-response is higher than expected at 8.9%. 168 of the 238 records of non-response on the English paper forms had a written response to question 17. This indicates that some respondents are skipping question 16 all together and giving a response to question 17 only. If the current

question format is to stay similar, this edit for this variable will need to ensure that this type of response behavior results in responses to question 17 only being coded.

Non-response on the online forms is at an acceptable level.

#### *No religion responses*

Levels of 'no religion' responses are within expectations and consistent with previous collections. The levels of 'no religion' response from the online forms are higher than from the paper forms. This is possibly due to the lower levels of non-response on the online forms, with those with no religion more likely to not respond on the paper forms.

#### *Christian responses*

46.4% of Christian responses overall is within expectations and follows the overall slowly declining trend for Christian affiliation observed over the past censuses. Levels of Christian affiliation are lower from online respondents, which is possibly due to the higher levels of residual categories on the paper forms being respondents who are not likely to indicate Christian affiliation.

#### *Buddhism, Hinduism, Islam, Judaism responses*

Responses to these categories are within expectations – no specific comments.

#### *Other religion*

Levels of 'other religion' responses are higher than from previous censuses. Of the 866 'other religion' responses, 233 of these had a written response in question 17. The largest of these were responses indicating 'Ratana' or a synonym, with 67 responses.

*Response unidentifiable* – There were no responses unidentifiable for this question, given that it is multiple response. There were no occurrences of multiple response from the online form (CHECK: recheck to see that the online form allowed multiple response?).

#### *Summary comments*

Overall the results for this question indicate that this question is successfully collecting information on the first level of the religious classification for the categories with tick boxes. Category 6 - Māori Religions, Beliefs and Philosophies and Category 7 – Spiritualism and New Age Religions are not being collected by this question.



## Birthplace

Subject population: total New Zealand population. Table below excludes countries with less than 50 responses.

**8** Which country were you born in?

New Zealand → go to **10**

Australia

England

China (People's Republic of)

India

South Africa

Samoa

Cook Islands

other. Print the country name:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Country of Birth	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
New Zealand		84.0		84.4		84.1
England		4.5		4.9		4.6
Australia		1.6		1.8		1.7
South Africa		1.2		0.5		1.0
United Kingdom NFD		1.1		0.0		0.9
India		0.8		0.6		0.7
Scotland		0.7		0.9		0.7
United States		0.5		0.6		0.6
Netherlands		0.4		0.8		0.5
Germany		0.4		0.6		0.4
China		0.5		0.0		0.4
Canada		0.5		0.2		0.4
Fiji		0.4		0.3		0.4
Total Stated		100.0		100.0		100.0
Unidentifiable (Code 0)		0.1		0.3		0.1
Not Stated		0.3		5.0		1.4
Total						

### Comments:

*Non-response* – The paper sample non-response was 5.0% and internet non-response was 0.3%. The overall non-response rate of 1.4% was acceptable. Non-response for paper bilingual forms was high compared with paper English forms – 11% compared with 3%. This has been flagged as an issue with the layout and design of the bilingual form.

*Response unidentifiable* – There was only Code 0 (inadequately described) responses. The majority of these responses were facetious.

*Response Categories* – 84.1% of all respondents (stated) were born in New Zealand. The largest group not born in New Zealand were English-born at 4.6%, followed by Australian (1.7%). Pattern was similar to the 2013 Census (Whanganui) with 88% NZ-born, 3.9% English and 1.4% Australian.

*Internet v paper responses* – The small number of responses in most categories makes it difficult to comment, but it was interesting that all Chinese-born respondents completed online forms. German and Dutch born respondents were least likely to complete online forms (around completed on paper).

*Comment on question design* – Note that the internet form does not have a list of main categories like the paper form. Online there are two options, NZ born and overseas, and overseas-born respondents choose from a drop-down box of countries once they have ticked overseas.

*Countries listed in the question* – a comparison of 2013 data with the countries listed on the 2017 form shows that the top six countries listed were the same in 2013. Samoa had a lower count than Fiji in 2013, and the Cook Islands did not feature in the top 10 countries, but the inclusion of these countries are related to other issues, which are still relevant. The drop-down As-You-Type box should assist online respondents to complete the question.

*Assessment* – Our recommendation is to retain the question in its current format.

## Languages spoken

**15** In which language(s) could you have a conversation about a lot of everyday things?  
 Mark the space or spaces which apply to you.

Remember to mark English if you can have a conversation in English.

English  
 Māori  
 Samoan  
 New Zealand Sign Language  
 other language(s), for example GUJARATI, CANTONESE, GREEK.  
 Print the language(s):


or  none (for example too young to talk)

Subject population: New Zealand usual residents

April 2017 Census Test – New Zealand usual residents								
Languages spoken	Internet form		English paper form		Bilingual paper form		Grand Total	
	Number	%	Number	%	Number	%	Number	%
English								98.3
Maori								3.1
Samoan								0.3
New Zealand Sign Language								0.5
Other								5.7

Total Stated						
Responses unidentifiable						
Not Stated		0.3		5.1		11.3
Residual Codes				5.1		11.3
Grand Total		100.0		100.0		100.0

This analysis is of the tick boxes in question 25 only. Analysis of written responses is included below.

This question is multiple response therefore totals will not add up to 100%.

Comments:

*Non-response*

Non response across the form types is acceptable other than the bilingual form non-response which is an issue across all variables on the bilingual form. Notably, the non-response for this question is lower on the bilingual form than others. Non-response for paper is slightly higher than expected.

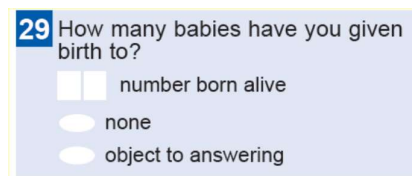
*Responses to language categories*

Language responses are within expectations for this variable. Analysis of the 'other' responses indicates that all 'other' responses had a written response in the text box. There were also 7 cases of written responses with no tick boxes marked, slightly lowering the residuals for this question.

*Summary comments*

Overall this variable performed well, with slightly higher than ideal non-response on the English paper form. The write-in component of the question is working successfully.

## Number of Children Born



Subject population -Female NZ Usual Resident Adult (≥15 years old) respondents.

Frequency of responses

The below table shows the frequency and percentage of responses at the lowest level of the classification. Percentages are calculated out of the Total Stated for all stated responses, and out of the overall Total for residual categories.

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
No Children		22.4		23.4		22.6
One Child		11.1		6.7		10.1

Two Children		30.2		25.2		29.1
Three Children		20.9		20.7		20.9
Four Children		8.6		12.2		9.4
Five Children		2.7		4.3		3.1
Six Children		1.1		2.1		1.3
Seven Children		0.3		1.5		0.6
Eight Children		0.2		0.3		0.2
Nine Children		0.2		0.4		0.3
Ten or More Children		0.1		0.7		0.2
<i>Don't Know</i>						
Object to Answering		2.3		2.5		2.3
<b>Total Stated</b>						
Response Unidentifiable		0.0		0.1		0.0
<i>Response Outside Scope</i>						
Not Stated		0.7		11.5		3.3
<b>Total</b>						

Frequencies for this question appear to be in line with the 2013 data for the Whanganui Territorial Authority in SuperCross. There is a slightly lower percentage of those who object to answering this question in the test data than there was in 2013 (3.4% in 2013), but this may reflect the voluntary nature of the Census Test.

There is a slight, but noticeable difference between the frequency distributions for online and paper forms, with higher proportions of respondents with few (1-2) births in the online data, and higher proportions of respondents with many (4+) births in the paper data. There is little difference between online and paper forms for respondents with no children born, or those who objected to answering the question.

#### Responses outside subject population

█ respondents outside the subject population responded to this question. Some of these seem to be respondents whose age or country of usual residence were incorrectly coded. Others may have been introduced by sex imputation.

Except for intersex respondents (which were later imputed to male or female), responses from those outside the subject population were only available for paper forms, as the online form routes respondents away from this question as appropriate.

#### *Residuals*

##### Non-response

The internet sample has a very low non-response rate of 0.7%. **The paper sample is outside the acceptable non-response level at 11.5%**, but a large proportion of this (4.1% of total respondents) are respondents who have had sex imputed to female. The overall non-response rate is 3.3%.

Two respondents coded to Not Stated had in fact written a number of children, in text (“one” and “three”), outside the text box. This was picked up by scanning, but not coded.

Overall the non-response rate for the Test data shows improvement over the non-response rate for Whanganui in 2013 (8.1%), although this is possibly due to the voluntary nature of the Census Test.

Response unidentifiable

The level of response unidentifiable on the paper form is 0.1%. There is no response unidentifiable on the online form (as there were possibly rules in place to prevent multiple response).

Response outside scope

No respondents were coded to Response Outside Scope for the Test data (there were 3 in Whanganui in 2013). It is unclear if this category is being used for coding in the Test.

Assessment

There are some quality issues with the subject population for this variable – some respondents seem to have been included/excluded due to scanning errors or imputation. This will inflate the non-response rate, as respondents who have had children may be imputed to “male”, and respondents that can’t have children may be imputed to “female”. These issues will presumably be improved by improvements in scanning and imputation, or by the implementation of consistency edits and checks.

Coding from the paper form needs to be improved to include text responses for number of children (“one”, “two”, etc.).

---

## Relationship to Reference Person

Note that this variable comes from two questions on the dwelling form, but is included in the individual (and absentee) datasets, as it is an individual variable. This requires successful linking of individual and dwelling forms.

Subject population

NZ Usual Resident respondents in private dwellings.

Notes for this analysis:

- All data referenced below (2017 and 2013) excludes absentees.
- **2017 Test data excludes respondents who could not be linked to their dwelling form (0.5% of online, 32.9% of paper, and 8.0% overall of usual residents)**

Frequency of responses

The below table shows the frequency and percentage of responses at the lowest level of the classification. Dotted lines show groupings for higher levels of the classification. Percentages are calculated out of the Total Stated for all stated responses, and out of the overall Total for residual categories.

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Reference person		47.1		58.7		49.0
Partner <sup>a</sup>		25.1		21.8		24.5
Child <sup>b</sup>		22.3		13.7		20.9
Parent <sup>b</sup>		1.4		0.6		1.3
Other relative, not further defined		0.0				0.0
Cousin		0.0		0.0		0.0
Sibling <sup>b</sup>		0.6		0.4		0.6
Grandparent <sup>b</sup>		0.1				0.1
Great-grandparent <sup>b</sup>						
Grandchild <sup>b</sup>		1.3		2.3		1.5
Great-grandchild <sup>b</sup>		0.0				0.0
Aunt/uncle		0.0				0.0
Great aunt/uncle						
Niece/nephew		0.2		0.0		0.2
Great niece/nephew						
Parent in law		0.1				0.1
Son/daughter in law		0.2		0.3		0.2
Sibling in law		0.1		0.1		0.1
Other relative not elsewhere classified						
Non relative not further defined						
Flatmate		0.9		1.3		0.9
Child of flatmate		0.0				0.0
Employer, partner of employer						
Child of employer						
Employee, partner of employee						
Child of employee						
Boarder		0.3		0.3		0.3
Child of boarder						
Lodger/roomer		0.0				0.0
Landlord/landlady						
Non-relative, not elsewhere classified		0.0				0.0

Guest/visitor/inmate/patient/ resident		0.1	0.4	0.1
<i>Don't know</i>				
<i>Refused to answer</i>				
<b>Total Stated</b>				
Response unidentifiable		0.1	0.0	0.1
<i>Response outside scope</i>				
Not stated		1.0	3.2	1.4
<b>Total</b>				

<sup>a</sup>Spouse/civil union partner/partner/de facto/boyfriend or girlfriend

<sup>b</sup>Birth/biological, adopted, step or other

**Note that the analysis below excludes respondents who could not be linked from the dwelling from to the individual dataset. Thus it is strongly biased towards online data.**

Frequencies for this question appear to be roughly in line with the values for Whanganui in 2013, although there is a significant increase in the percentage of reference people and partners, and a decrease in the percentage of children and flatmates. This agrees with living arrangements data, which shows more people living alone or with partners than in 2013, and fewer people living with parents and flatmates. This implies the Test is biased towards couples and small family households.

There is also a dramatic decrease in the number of guests/visitors/inmates/residents in the 2017 data (0.1%, compared with 4.7% in 2013 for Whanganui). It's unclear whether this is due to synonym list changes, or because those with visitors were unlikely to take part in the Test.

As seen for the living arrangements variable, there is a noticeable difference between the frequency distributions for online and paper forms, with a higher proportion of reference people in the paper data (indicating smaller households), and higher proportions of children in the online data.

- "Don't have wife. No children." coded to child.

There are at least ■ records in the Test data where living arrangements are "Live alone" but relationship to reference person is not "Reference person", "Visitor" or residual. This is a conservative estimate, limited to usual residents without multiple responses to living arrangements. It is likely that there are also other inconsistencies between these questions, but this is difficult to analyse without the family coding household matrix.

One specific issue found with the data was a case where the text response "Don't have wife. No children." was coded to Child. It may be worth investigating a method to identify negative statements like this.

Responses outside subject population

This isn't relevant for this question – technically all respondents should have a response for this, but only usual residents are used for family coding.

Residuals

## Non-response

Ignoring all those who could not be linked from the dwelling form to the individual dataset, the internet and paper samples have non-response rates of 1.0% and 3.2% respectively. The overall non-response rate is 1.4%. **If unlinked numbers are included, the non-response increases to 1.6% for online, 35.1% for paper and 9.3% overall.** Therefore the quality of this data is heavily dependent on successful linking of records.

All of these values are much higher than the non-response rate for Whanganui in 2013 (0.3%).

## Response unidentifiable

Of the ■ cases of Response Unidentifiable, many could be coded correctly if synonyms were added to coding (eg, “my wife”, “renter”, “au pair”).

## Assessment

This data is low quality due to the high percentage of records that could not be linked from the dwelling form to the individual dataset. Because of this it is difficult to accurately assess any other issues the data may have.

Coding could be improved with more synonyms and some way to handle negative statements (“don’t have children”). There is some inconsistency between this variable and living arrangements that needs further investigation.

## Living Arrangements

**18** Who lives with you at your usual address?  
 Mark the space or spaces which apply to you.

- I live alone
- my wife or husband, partner or de facto
- my mother and/or father
- my daughter(s) and/or son(s)
- my sister(s) and/or brother(s)
- my grandparent(s)
- my grandchild(ren)
- my flatmate(s)
- other, for example STEP-SON, MOTHER-IN-LAW etc.  
Please state:


## Subject population

NZ Usual Resident respondents in private dwellings.



## Frequency of responses

The below table shows the frequency and percentage of responses at the lowest level of the classification. Dotted lines show groupings for higher levels of the classification. Percentages are calculated out of the Total Stated for all stated responses, and out of the overall Total for residual categories.

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Live alone		12.6		27.8		15.8
Partner <sup>a</sup>		55.5		43.5		52.9
Child <sup>b</sup>		24.2		14.1		22.1
Parent <sup>b</sup>		25.0		18.7		23.6
Other relative, not further defined		0.1		0.2		0.1
Sibling <sup>b</sup>		16.6		13.7		16.0
Sibling-in-law		0.2		0.1		0.2
Grandparent <sup>b</sup>		1.8		0.3		1.4
Great-grandparent <sup>b</sup>		0.0		0.0		0.0
Grandchild <sup>b</sup>		1.8		0.2		1.5
Great-grandchild <sup>b</sup>		0.0		0.0		0.0
Aunt/uncle		0.6		0.2		0.5
Great-aunt/great-uncle						
Niece/nephew		0.7		0.2		0.6
Great-niece/nephew						
Daughter/son-in-law		0.5		0.3		0.5
Parent-in-law		0.2		0.3		0.2
Cousin		0.4		0.1		0.3
Other relative, not elsewhere classified						
Non-relative, not further defined						
Flatmate		2.2		6.0		3.0
Child of flatmate		0.0		0.0		0.0
Employer						
Child of employer						
Employee		0.0		0.4		0.1
Child of employee						
Partner/spouse of employee						
Boarder		0.6		1.0		0.6

Child of boarder				
Lodger/roomer		0.0	0.0	0.0
Landlord/landlady		0.0	0.3	0.1
Non-relative, not elsewhere classified		0.1	0.5	0.2
Guest/visitor/inmate/patient/resident		0.3	1.9	0.6
<i>Don't know</i>				
<i>Refused to answer</i>				
<b>Total Stated</b>				
Response unidentifiable		0.4	3.5	1.0
Response outside scope		0.0	0.2	0.1
Not stated		0.2	7.1	1.7
<b>Total</b>				

<sup>a</sup>Spouse/civil union partner/partner/de facto/boyfriend or girlfriend

<sup>b</sup>Natural, step, adopted or foster

Frequencies for this question appear to be roughly in line with the values for Whanganui in 2013, although there is a noticeable increase in the proportion of people living alone or with partners, and a decrease in those living with parents and flatmates. This agrees with the relationship to reference person data, which shows more reference people and partners than in 2013, and fewer children and flatmates. This implies the Test is biased towards couples and small family households.

As seen for the relationship to reference person data, there is a noticeable difference between the frequency distributions for online and paper forms, with a higher proportion of those living alone in the paper data, and higher proportions of family living arrangements in the online data.

One major issue is the number of people coded to Guest/visitor/inmate/patient/resident. No one should be coded here for this question, as it is asking for living arrangements at the respondent's usual, private address. In almost all cases this code was used for rest home residents, who should have been excluded as in an NPD. Others were coded here for variations on "family friend", which should have been coded elsewhere.

Aside from this issue, data is possibly more accurate than in 2013 because the textbox is being coded. There are many cases of respondents ignoring the checkboxes and writing in "mother", "partner", etc. into the textbox – these responses would have been lost in 2013.

The following specific issues were found:

- [REDACTED] respondents were coded for "Live Alone" and another living arrangement. These were a mixture of scanning error and multiple response on paper forms. Suggest an edit (enforced on the online form) to check this situation manually for 2018.
- There are issues with coding not recognizing possessives in the freetext field – eg. "my dads wife" gets coded for "dad" and "wife" separately, and "mother and her partner" codes for "mother" and "partner". This could be a matter of adding these synonyms to the coding (some other examples, eg. "my mums partner", did code correctly).
- For some reason "foster children" seems to be coded to "Sibling". This should be changed. There are also a few cases of religious brothers and sisters being coded as siblings.

- “Parents in law” needs to be added to the synonym list (coded to “Parent in law”).
- In one instance “grandson” was coded to “grandparent” and “child”.

#### Residents and NPDs

It seems that many people coded as guest/visitor/inmate/resident are part of an aged care facility which should have been coded as an NPD (and excluded from this subject population). It is unclear how to exclude NPDs in the current dataset.

#### Responses outside subject population

A lot of rest home residents were coded for this question when they should have been excluded as an NPD. It also seems many overseas residents responded to this question, but these can be filtered by person record type.

#### Residuals

##### Non-response

The internet sample has a very low non-response rate of 0.2%. **The paper sample is outside the acceptable non-response level at 7.1%.** The overall non-response rate is 1.7%.

All of these values show improvement over the non-response rate for Whanganui in 2013 (8.1%), although this is possibly due to the voluntary nature of the Census Test.

##### Response unidentifiable

## Unpaid activities

**51** In the last 4 weeks, which of these have you done, **without pay**?

*Mark the space or spaces which apply to you.*

- household work, cooking, repairs, gardening, etc, for my household
- looked after a child who is a member of my household
- looked after a member of my household who is ill or has a disability
- looked after a child (who **does not** live in my household)
- helped someone who is ill or has a disability (who **does not** live in my household)
- other help or voluntary work for or through any organisation, group or marae
- or  none of these

Subject population: New Zealand usual resident adults

April 2017 Census Test – New Zealand usual resident adults

Unpaid activities	Internet form		English paper form		Bilingual paper form		Grand Total	
	Number	%	Number	%	Number	%	Number	%
Household work								87.1
Child in household								22.0
Ill in household								7.9
Child not in household								16.6
Ill not in household								11.8
Voluntary work								20.6
None of these								11.1
Total Stated								
Responses unidentifiable								0.9
Not Stated		0.7		13.4		13.3		3.8
Residual Codes								1.2
Grand Total		100.0		100.0		100.0		100.0

Please note 'responses unidentifiable' have been indicated where a respondent has indicated one or more unpaid activities as well as 'none of these'. These have not been removed from the total stated population.

This question is multiple response therefore totals will not add up to 100%.

Comments:

*Non-response*

Non-response for both paper forms is quite high. For the English paper form this is higher than for earlier variables in the form. Analysis of page 4 (all work questions) indicates that all of page 4 was not responded to by [redacted] respondents out of the [redacted] who did not respond to the work indicator question at the start of page 4. This indicates that this question is not causing respondents to not respond specifically. Of the 233 cases of non-response on the English paper forms [redacted] out of the [redacted] were 65 years or older. This indicates that older respondents are disproportionately not responding to this section of the paper form. This is consistent with [redacted] out of 358 non-responses to the work indicator question being respondents of 65 years or older.

*Responses to unpaid activities categories*

The percentage totals for all of the categories are comparable to 2013 results. No issues were identified with the counts for the activities.

*Response unidentifiable*

Response unidentifiable is not possible for this question online. [redacted] people gave an unidentifiable response to this question on paper.

*Summary comments*

Overall this variable performed well, with the exception of high non-response to both English and bilingual paper forms. These issues are identified as being wider than this variable, and as such this question format is recommended for the 2018 Census.

## Health

### Disability/activity limitations

Subject population: New Zealand usual resident population aged 5 and over

**23** This question is about difficulties you may have doing certain activities because of a health problem.

Do you have difficulty with any of the following:

- seeing, even if wearing glasses?
  - no difficulty
  - some difficulty
  - a lot of difficulty
  - cannot do at all
- hearing, even if using a hearing aid?
  - no difficulty
  - some difficulty
  - a lot of difficulty
  - cannot do at all
- walking or climbing steps?
  - no difficulty
  - some difficulty
  - a lot of difficulty
  - cannot do at all
- remembering or concentrating?
  - no difficulty
  - some difficulty
  - a lot of difficulty
  - cannot do at all
- washing all over or dressing?
  - no difficulty
  - some difficulty
  - a lot of difficulty
  - cannot do at all

*Comments:*

*Non-response* – Non-response rates for the question set overall were acceptable, outside of non-response to the bilingual forms which is an issue with all variables on the bilingual form. Overall non-response to the English paper form is acceptable at 3.4%. This indicates that data produced for the overall disability indicator will be of suitable quality for output.

Non-response rates to individual questions within the question set were acceptable, ranging from 2.2-2.9% overall, which is higher comparatively than the overall non-response of 1.5%. Removing the 1.5% of overall non-response gives each question a rate of 0.7-1.4% of respondents who responded to at least one other question, but did not respond to this individual question. This rate is highest for the 'hearing' question; feedback received in cognitive testing indicated that some respondents were finding the wording confusing, as it includes a reference to using a hearing aid. Non-response to the 'hearing' question on the English paper forms is 5.3%, this is lower than 5.9% for the paper sample in the July 2016 Census Test. This is higher than ideal, but not impacting on overall data quality.

**Overall disability indicator**

**April 2017 Census Test – New Zealand usual residents aged 5 or over**

Disability Indicator	Internet form		English paper form		Bilingual paper form		Grand Total	
	Number	%	Number	%	Number	%	Number	%
Not Disabled		93.2		82.8		81.4		90.8
Disabled		6.8		17.2		18.6		9.2
Total Stated		100.0		100.0		100.0		100.0
Responses unidentifiable				0.0		0.1		0.0
Not Stated		0.0		3.4		15.4		1.5
Residual Codes		0.0		3.4		15.5		1.5
Grand Total		100.0		100.0		100.0		100.0

### Individual questions

April 2017 Census Test – New Zealand usual residents aged 5 or over						
	Seeing		Hearing		Walking	
	Number	%	Number	%	Number	%
Not Disabled		97.9		97.6		95.1
Disabled		2.1		2.4		4.9
Total Stated		100.0		100.0		100.0
Responses unidentifiable		0.0		0.0		0.0
Not Stated		2.2		2.9		2.4
Residual Codes		2.3		2.9		2.5
Grand Total		100.0		100.0		100.0

April 2017 Census Test – New Zealand usual residents aged 5 or over						
	Remembering		Washing		Communicating	
	Number	%	Number	%	Number	%
Not Disabled		97.4		98.0		98.8
Disabled		2.6		2.0		1.2
Total Stated		100.0		100.0		100.0
Responses unidentifiable		0.0		0.0		0.0
Not Stated		2.6		2.8		2.5
Residual Codes		2.7		2.8		2.5
Grand Total		100.0		100.0		100.0

*Response unidentifiable* – Occurrences of responses unidentifiable were low overall, and across the individual questions, which is good, given that there was some concern over respondents having issues deciding between two response categories for some questions.

*Disabled/Not disabled responses* – The overall level of 9.2% disabled population is within the expected range. This is higher than the result from the July 2016 Census Test results which was 6.2%, but this was considered a low level. The overall levels of the indicator for the disabled population vary largely between the online and paper forms. This is possibly as a result of older respondents being more likely to respond using the paper forms and more likely to be disabled.

The individual question with the highest contribution to the disability indicator is the walking question. This is consistent with the July 2016 Test results.

*Recommendation*

Overall the results for this variable indicate that the question set is working successfully in producing the overall disability indicator. While non-response to individual questions on the paper forms is higher than ideal, this is not impacting on the ability to output this variable.

## Cigarette Smoking Behaviour

**25** Do you smoke cigarettes regularly (that is, one or more a day)?

Don't count pipes, cigars or e-cigarettes.  
Count only tobacco cigarettes.

yes → go to **27**

no

**26** Have you ever been a regular smoker of one or more cigarettes a day?

yes

no

Question 25 – current smoking

Subject population: New Zealand usual resident adults

April 2017 Census Test – New Zealand usual resident adults								
Current smoking	Internet form		English paper form		Bilingual paper form		Grand Total	
	Number	%	Number	%	Number	%	Number	%
Regular smoker (yes)								10.2
Not regular smoker (no)								89.8
Total Stated								100.0
Responses unidentifiable								0.9
Not Stated		0.5		4.2		14.8		3.8
Residual Codes								1.2
Grand Total		100.0		100.0		100.0		100.0

Question 26 – ex smoking

<b>April 2017 Census Test – New Zealand usual resident adults who did not answer 'yes' to question 25</b>
---

Ex smoking	Internet form		English paper form		Bilingual paper form		Grand Total	
	Number	%	Number	%	Number	%	Number	%
Ex-smoker (yes)								10.2
Never regular smoker (no)								89.8
Total Stated								100.0
Responses unidentifiable								0.0
Not Stated		0.5		6.3		18.2		3.0
Residual Codes		0.5		6.3		18.2		3.0
Grand Total		100.0		100.0		100.0		100.0

Comments:

*Non-response*

Not response for this variable overall is acceptable, other than the bilingual form. Because the non-respondent population for question 25 has been included in the data for question 26, these are slightly higher. However, there is an indication that a small number of respondents who answered 'no' to question 25 are skipping question 26 on both paper and online formats.

*Responses to smoking categories*

The totals for the regular smoker and ex-smoker categories are quite low compared to 2013 census data for the Wanganui region, which had were 18% and 24% respectively. While smoking rates are likely to be trending down, this shift in the data is likely to be a result of the older respondent population to the 2017 Census Test, as overall smoking prevalence peaks in the age bracket of 25-34 and trending downwards. While the results are low, this is not of concern.

*Response unidentifiable*

There was only one case of multiple response to these questions.

*Summary comments*

Overall this variable performed well, with some incidence of non-response across the two questions. It is recommended to check if we are using any previous census data to inform the ex-smoker output, as this may help with data quality for the ex-smoker counts.



## Location variables

### Years at Usual Residence

**5** How long have you lived at the address or country given in 4?

less than one year or   years

Subject population: New Zealand residents (overseas residents excluded from analysis)

#### Comments

*Not stated* – Originally there were no residual codes in the dataset, as all blank responses were coded to 0. This needed further investigation, and a new variable has been created for analysis.

New data shows the non-response rate as 3.6% overall, with a 15.2% rate for paper and 3.0% for online data. There were 11 responses coded as N/A, which were unidentifiable write-in answers like “all my life” or “six”.

*Analysis* – 61% of all respondents have been at their current usual residence for less than 10 years, which compares well with the 2013 data for Whanganui (63%).

Around 1000 respondents said they have lived in the same residence for their lifetime (age = years at usual residence); it is possible that some respondents misunderstood the question intent.

Years at Usual Residence	Online	%	Paper	%	Total	%
less than 10 years		61.5		54.6		61.4
10-49 years		37.4		40.2		38.1
50 years or more		1.1		5.2		2.0
Total Stated		100.0		100.0		100.0
Not Stated /NA		3.0		15.2		3.6

### Years since Arrival in New Zealand

**9** When did you first arrive to live in New Zealand?

month if known (eg 4)      year (eg 1985)

*Subject Population* – the subject population is the census night population, but the output population is usually NZ residents who were not born in New Zealand. Output is by number of years since arrival in NZ. Note that a small number of paper respondents born in NZ did not follow the routing and answered the question.

*Not stated* – the not-stated response rate for overseas-born New Zealand residents was 9.6% which is relatively high. Paper non-response was unacceptably high and made up the bulk of non-response; 32.3% of paper respondents did not answer the question. However, 70% of these respondents did not answer either country of birth or years since arrival, which has inflated the non-response. Only one person overseas-born New Zealand resident who responded to country of birth actually gave no response to the arrival question.

In contrast, online non-response rates were very low (1.9%).

*Response unidentifiable* – There were only 5 unidentifiable responses which were wrongly coded to the 0 years category. Around 43% of overseas visitors answered the question. Only 61 respondents said they had lived at their usual residence longer than they had been in New Zealand.

*Frequency of responses* – Analysis of the data for overseas born residents only shows that around half of respondents (50.9%) arrived less than 20 years ago. This compares favourably with the 2013 Census data for Whanganui (52.8%).

Years Since Arrival in NZ	Online	%	Paper	%	Total	%
0-9 years		27.8		21.9		26.7
10-19		23.1		15.9		21.7
20-29		10.4		7.5		9.9
30-39		6.9		8.9		7.3
40-49		12.8		14.5		13.1
50-59		12.0		14.7		12.5
60-69		6.4		14.0		7.9
70-79		0.5		1.9		0.8
80-89		0.0		0.9		0.2
No response		1.9		32.3		9.6
Grand Total						

*Recommendation* – The question does not work well for paper respondents, but the main problem is when residents do not answer either country of birth or years since arrival. Could routing be improved so that it is clear to respondents born overseas that they should answer this question?

## Person Record Type, Usual Residence Address and Census Night Address

*Subject Population* – total New Zealand census night population

*Analysis* – a cross check of age by record type showed that adults and children had been processed correctly according to their age.

The proportion of adults to children in the 2017 Whanganui test was a little higher than in the 2013 Census (80%/20%), but a reasonable representation given the response rate in 2017. As expected, and in line with 2013, children were more likely to complete online.

Person Type	Person Record Type, 2017 Census Test		Grand Total
	Online	Paper	
New Zealand Adult	82.8%	86.8%	83.8%
New Zealand Child	16.5%	9.7%	14.9%
Overseas Adult	0.6%	2.9%	1.1%
Overseas Child	0.1%	0.5%	0.2%
Grand Total	100.0%	100.0%	100.0%

It is difficult to check whether the usual residence address indicator and census night address indicator are correct as there are a large number variables in the dataset that contribute to these indicators.

Coding of overseas and New Zealand residents using the tick boxes in Q4 (where do you usually live) showed errors (see table below in red), which were mostly respondent error on the paper forms. For example,

people ticking 2 boxes on the usual resident question, or writing in their residential address and then writing in the overseas box as well. However, about half of all paper respondents did not tick either of the usual residence tick boxes provided (in New Zealand or Overseas).

Usual Residence Indicator					
Person Type	New Zealand	NZ and Overseas	Overseas	Not Stated	Total
NZ Adult	██████████	██████████	██████████	██████████	██████████
NZ Child	██████████	██████████	██████████	██████████	██████████
Overseas Adult	██████████	██████████	██████████	██████████	██████████
Overseas Child	██████████	██████████	██████████	██████████	██████████
TOTAL	██████████	██████████	██████████	██████████	██████████

There were █████ paper records which did not have a census night address or a usual residence address. A spot check of these records showed that some of these respondents were overseas-born students who did not understand the concept of usual residence, so ticked overseas but also ticked 'at usual residence' for Census Night Address. Some respondents did not fill in an address in either place but ticked the boxes. The guide notes direct secondary students at boarding school who are overseas students to put their home address country, which is not their usual New Zealand residence.

**4 Which address do I give?**

- If you live in New Zealand or are staying here for 12 months or more, excluding short overseas trips, follow these guidelines to give the right address:
- If you are a primary or secondary school student at boarding school, give your home address.
  - If you are a tertiary student, give the address where you live during term.

The derived usual residence indicator showed 85.8% of adults and 89.2% of children were at their usual residence on census night (see table below for numbers). The equivalent proportion in 2013 was 96%.

Usual Residence Indicator	NZ Adult	NZ Child	Overseas Adult	Overseas Child	Grand Total
Same as Census Night Address	██████████	██████████	██████████	██████████	██████████
Elsewhere in NZ	██████████	██████████	██████████	██████████	██████████
Overseas	██████████	██████████	██████████	██████████	██████████
Grand Total	██████████	██████████	██████████	██████████	██████████

*Recommendation* – paper forms may need to be checked manually when responses are inconsistent. Also, could more information be added to the form to assist respondents to work out what is their usual address?

## Usual Residence One Year Ago

**14** Where did you usually live 1 year ago, on 4 April 2016?

not born 1 year ago

at my usual address, given in **4**

at my census night address, given in **6**

at another New Zealand address:

street number      flat number

street name

suburb or rural locality

city, town or district

or  I lived overseas 1 year ago

Subject population: New Zealand resident population

*Comments:*

This is a new question so we do not have any comparative data from 2013.

Note: coding of this variable did not match the classification Census Usual Residence One Year Ago Indicator V1.0, but processing has now corrected this error.

*Not stated* – The not-stated response rate was 2.8%, which is below the acceptable 5% non-response rate.

*Response unidentifiable* – There were [redacted] responses (0.2%) that would fit into this category, these were mainly multiple responses (e.g. ticked both usual address and census night address).

Children under 1 year old – should be routed away from the question online. However, this did not happen and parents got to see the question without the ‘not born 1 year ago’ option, so they filled out other options instead (see below).

Usual Residence 1 Year Ago (Children under 1 year	Online	Paper	Grand Total
Not Born 1 Year Ago	[redacted]	[redacted]	[redacted]
At Usual Address	[redacted]	[redacted]	[redacted]
At Census Night Address	[redacted]	[redacted]	[redacted]
At Another NZ Address	[redacted]	[redacted]	[redacted]
Overseas	[redacted]	[redacted]	[redacted]
Not Stated	[redacted]	[redacted]	[redacted]
Grand Total	[redacted]	[redacted]	[redacted]

*Frequency of responses* – An analysis of the raw tick box data shows that responses look feasible, apart from the not born 1 year ago category which was too low, as noted above. The following analysis excludes children under 1 and overseas visitors (both of whom should not answer this question).

Usual Residence 1 Year Ago, NZ residents aged 1 and over

Address 1 Year Ago	Online	%	Paper	%	TOTAL	%
At Usual Address	[redacted]	84.1	[redacted]	86.0	[redacted]	84.5
At Census Night Address	[redacted]	0.2	[redacted]	3.4	[redacted]	0.9

At Another NZ Address		14.1		9.6		13.2
Overseas		1.6		1.1		1.5
Valid Responses		100		100		100
Not Born 1 Year Ago						
Not Stated						
Multiple Responses/NA						
Grand Total						

*Census Night Address Tick Box* – The “at my census night address” tick box is not working, and potentially confuses paper respondents. It was used by only 0.2% of respondents on the internet form. Analysis of the 3.4% of respondents who ticked the box on paper shows that the majority of their responses were not valid, as at least half of respondents who ticked at my census night address gave their census night address in Q6 as their usual residence (at the address in Q4). Another contributing factor was people not completing the census night address question properly or at all. Overall very few respondents actually lived at their census night address one year ago unless it was also their usual residence.

*Other analysis* – Comparison of usual residence one year ago with years at usual residence showed that only a very small proportion (█ respondents) said they were at their usual address one year ago, but also ticked the box for ‘less than one year’ at their usual residence. A further █ respondents who said they were at their usual residence one year ago gave a numerical response of less than one year for this question. Overall this error was only a minor issue.

*Recommendation:* the tick box ‘at my census night address given in 6’ is removed from the forms, and the space is allocated to collecting the country of residence for those who were overseas 1 year ago. Population Statistics (SMA) supports this change.

## Education and Training Variables

### Study Participation

**19** Are you attending, studying or enrolled at school or any other place of education?

yes – full-time (20 hours or more a week) → go to **20**

yes – part-time (less than 20 hours a week) → go to **20**

or  no, neither of these → go to **22**

Subject population: New Zealand resident population

Comments:

This question was asked of all New Zealand residents in the 2017 test as it is a filter question for the new travel to education question. In prior censuses it was only asked of New Zealand adults.

*Not stated* – The not-stated response rate was 2.5% which is below the acceptable 5% non-response rate.

*Response unidentifiable* – Only [redacted] responses were unidentifiable or out of scope, one person ticked full time study and part time study, the other ticked all three boxes.

*Frequency of responses* – Analysis is more meaningful by age group, but the data below generally seems to fit with expectations by age.

### Study Participation Whanganui Census Test 2017

Age Group	Full-time	Part-time	Total studying	Not studying	Not stated (% total)
0-4	25.2	10.5	35.7	64.3	3.2
5-9	91.9	0.6	92.5	7.5	0.8
10-14	95.3	0.7	96.0	4.0	1.5
15-19	77.0	2.5	79.5	20.5	1.3
20-24	16.5	6.7	23.2	76.8	1.5
25-29	6.5	7.0	13.5	86.5	0.8
30-34	4.7	6.8	11.4	88.6	0.2
35-39	5.5	6.4	11.9	88.1	1.3
40-44	2.3	5.9	8.2	91.8	1.7
45-49	1.9	5.7	7.7	92.3	1.6
50-54	1.9	4.3	6.2	93.8	1.6
55-59	0.7	2.2	2.9	97.1	2.5
60-64	0.9	1.7	2.6	97.4	2.2
65-69	0.1	0.8	1.0	99.0	3.2
70-74		0.4	0.4	99.6	5.5
75-79		0.8	0.8	99.1	6.4
80-84		1.0	1.0	99.0	10.7
85+		0.7	0.7	99.3	12.6
Total	17.4	3.1	20.5	79.5	3.0

Census 2013 data is not strictly comparable as the question was asked of a different subject population, but across selected age groups below, the data looks comparable apart from the 15-24 year age group, which has increased full-time participation from 37% to 57%. The likely reason for this is that students are staying on longer at school, and taking up tertiary study in greater numbers than before.

Age Group	15-24 years		25-44 years		45-64 years	
	2013	2017	2013	2017	2013	2017
Full time study	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Part-time study	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Not studying	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Not stated	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

*Other comments:* Online forms do not allow for multiple responses but they are still possible on paper. Only [redacted] respondents did not answer correctly in the test – [redacted] gave full time/part time as an answer the other ticked all 3 boxes, and these were coded to 7 = unidentifiable.

There is an issue with the coding of multiple responses. Household surveys recode a person who ticks full-time AND part-time study as studying full-time to full-time (code 1). Processing has coded this combination to unidentifiable (code 7) and the derivation has been drafted to recode multiple responses to code 3 (part-time and full-time study).

Also note that the draft derivation for this variable incorrectly recodes the third tick box (not studying) to output code 3 which is not correct (should be output code 4).

*Assessment* – Our recommendation is to retain the question in its current format. A decision on processing multiple paper responses (Full-time and Part-time) is pending.

## Highest Secondary School Qualification

Subject population: New Zealand residents aged 15 and over

**31** What is your highest **secondary school** qualification?

none

one of:

- NZ School Certificate in one or more subjects, *or*
- National Certificate level 1, *or*
- NCEA level 1

one of:

- NZ Sixth Form Certificate in one or more subjects, *or*
- National Certificate level 2, *or*
- NZ UE in one or more subjects before 1986, *or*
- NCEA level 2, *or*

one of:

- NZ Higher School Certificate, *or*
- Higher Leaving Certificate, *or*
- NZ University Bursary/Scholarship, *or*
- National Certificate level 3, *or*
- NCEA level 3, *or*
- NZ Scholarship

other secondary school qualification **gained in NZ:**

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

or  other secondary school qualification **gained overseas**

**Highest Secondary School Qualifications by Level**

Qualification	Online		Paper		Grand Total	
	Number	Percent	Number	Percent	Number	Percent
None		26.5		47.1		30.9
Level 1 Certificate		22.1		21.3		21.9
Level 2 Certificate		21.1		14.9		19.7
Level 3 Certificate		19.3		10.5		17.5
Level 4 Certificate		0.0		0.0		0.0
Overseas sec school qualification		11.1		7.0		10.2

Total Stated		100		100		100
Don't Know		0.0		0.0		0.0
Unidentifiable		0.3		4.4		1.3
Outside scope		0.3		0.9		0.4
Not Stated		1.6		11.6		4.5
Residual Codes		2.2		16.9		6.2
Grand Total		100.0		100.0		100.0

Comments:

*Non-response* – Not stated responses to this question were unacceptably high for paper (11.6%), but online non-response was only 2.2%. The overall non-response rate of 4.5% is acceptable.

*Other residual codes* – Low overall, with a high of 4.4% unidentifiable responses on paper. Overall, the proportion of residual code responses was 16.9 for paper and 2.2 for internet forms. The low level of non-response for internet forms has reduced the total residual code responses from 12.6% in the 2013 Census to 6.2% in the test.

*Response Categories* – 30.7% of respondents had no secondary school qualification (cf. 38% in 2013 Census). Comparisons with the 2013 Whanganui census results show a rise in secondary school qualification levels since 2013 as shown below.

Qualification	2013	2017
None	38.4%	30.9%
Level 1	22.8%	21.9%
Level 2	18.2%	19.7%
Level 3 or 4 certificate	13.5%	17.5%
Overseas Secondary School Qualification	7.1%	10.2%
<i>Total Residuals</i>	12.6%	6.2%

Data from the 2013 Census shows that qualification levels are age-related; 58% of the 65+ age group had no school qualifications in 2013 compared with only 27% of 15-39 year olds. In the Census Test results, 18% of 15-39 year olds had no secondary qualification, compared with 47% of the 65+ age group. This result fits well with what we might expect in the 2018 Census, given the trend towards increasing formal qualifications over time.

*Internet v paper responses* – There was a direct relationship between the level of qualification and the mode of completion; respondents with low level school qualifications were more likely to complete paper forms than those with higher qualifications. This is likely to be age-related – older respondents are generally less qualified and were more likely to complete on paper.

**Secondary School Qualifications – by mode of completion**

Qualification	Online		Paper		Total %
	Number	Percent	Number	Percent	
None		68.1		31.9	100
Level 1 Certificate		79.6		20.4	100
Level 2 Certificate		84.1		15.9	100



Level 3 Certificate		87.4		12.6	100
Level 4 Certificate		100.0		0.0	100
Overseas qualification		85.5		14.5	100
Total Stated		78.9		21.1	100
<i>Residual Codes</i>		28.9		71.1	100
Grand Total					100

*Comment on classification question design* – Note that there are no Level 4 secondary school qualifications (formerly bursary) and this category is obsolete.

*Assessment* – Our recommendation is to retain the question in its current format.

## Highest Post-School Qualification

Subject population: NZ Usual Resident Adult respondents who indicate they have completed a qualification other than a secondary school qualification

*Note that there has been a change from an open write-in box to a tick box format in 2017*

### Filter Question

**32** Apart from secondary school qualifications, do you have another completed qualification?

**Don't count** qualifications that take less than 3 months of full-time study to get.

yes → go to **33**

no → go to **36**

Response		
Yes		47.9%
No		52.1%
Unidentifiable		0.0%
Not Stated		7.8%

Comparison with 2013 Census data is not straight-forward as the question (and the output) has changed, and it is not clear how “post-school qualification level of attainment” output was constructed in 2013. Around 60% of Whanganui adults who gave a valid answer to the question had no post-school qualifications in 2013, and it is likely that the proportion would have dropped in the past four years, given the increase in tertiary study in recent years.

**33** What is your highest qualification?

- Level 1 Certificate
- Level 2 Certificate
- Level 3 Certificate
- Level 4 Certificate
- Level 5 Diploma
- Level 6 Diploma
- Bachelor Degree or Level 7 Qualification
- Bachelor Hons or Postgraduate Certificate/Diploma
- Masters Degree
- PhD
- other qualification. Print what it is:


The table below analyses the responses to Q33, for respondents who said Yes to the filter question (Q32) above.

Highest Post School Qualification						
Qualification Level	Online	%	Paper	%	Grand Total	%
Level 1 Certificate		2.8		2.5		2.8
Level 2 Certificate		3.0		4.1		3.2
Level 3 Certificate		7.2		8.4		7.3
Level 4 Certificate		18.7		20.2		18.9
Level 5 Diploma		10.1		8.0		9.8
Level 6 Diploma		13.1		16.7		13.5
Bachelor Degree/Level 7		25.2		24.8		25.2
Bachelor Honours Degree/Post Grad Cert/Diploma		12.8		11.0		12.6
Masters Degree		6.0		3.6		5.7
PhD		1.0		0.8		1.0
<b>Total Stated</b>		<b>100</b>		<b>100</b>		<b>100</b>
Response Unidentifiable		5.6		13.2		6.6
Response Outside Scope		0.3		0.1		0.3
Not Stated		4.0		3.8		4.0
No Qualification*		0.1		0.0		0.1
<b>Residuals</b>		<b>10.1</b>		<b>17.1</b>		<b>11.0</b>
<b>Grand Total</b>						

\*people who said Yes to Post –School Qual then wrote in None

Comments:

*Not stated/unidentifiable* – Both the internet and paper samples have a high proportion of residual responses – 10% and 17% respectively. Unidentifiable responses accounted for 60% of all residual responses, mostly vague responses that could not be coded– for example occupations or job titles instead of actual qualifications.

In the 2017 test, around ██████ respondents (1/4) who ticked Yes to the post school qualification indicator did not tick a qualification level but wrote in the text box below. Around 7% of all text box responses were unable

to be coded from the codefile. The As-You-Type suggestions may have contributed to a lower response rate for the internet forms, as only 5.6% of internet responses were unidentifiable compared with 13.2% of paper responses. There was a similar level of not stated responses for paper and internet forms.

To summarise, respondents are more likely to answer the question than to skip it, but around 1 in 4 respondents are writing a response into the text box rather than ticking a qualification level.

*Frequency of responses* – Proportions of responses to this question are comparable to the 2013 Census (Whanganui District), apart from Level 4 Certificates and post-graduate/honours degrees, which are both considerably lower in 2017. This could be due to a number of reasons including the makeup of the sample and the new question format gathering responses more accurately or inaccurately than the 2013 Census responses which had a large level of manual coding. No manual coding of written responses has been done for the test.

	2013 Census	2017 Test
Level 1, 2 or 3 Certificate	10.4	13.3
Level 4 Certificate	31.9	18.9
Level 5 Diploma	10.6	9.8
Level 6 Diploma	14.1	13.5
Bachelor Degree and Level 7 Qualification	23.5	25.2
Post-graduate and Honours Degrees	5.4	12.6
Masters Degree	3.4	5.7
Doctorate Degree	0.7	1.0
	100%	100%

*Assessment* – Test data indicates that this question is causing some issues for respondents selecting a qualification level on the paper form. There is also significant variation from the 2013 Census data for this variable in some of the qualification levels indicated by respondents. Despite this, the change to the 2013 question format will reduce the manual coding burden associated with this question. Some of the differences between 2013 and 2017 could also be related to whether the codefile is able to allocate a code to the written responses.

It is recommended that this question be included in the March 2017 Census Test with solutions to the issues identified above to be considered before finalising the question format.

---

## Highest Qualification – Derivation 54

Subject population: NZ usual resident adults

*General* – Highest qualification is a derived variable which includes school and post-school qualifications and combines information from Q31-33 of the Census Test.

*Non-response* – The paper sample non-response was 9.2% and the internet non-response was 0.6%. The overall proportion of not-stated responses was 6.1%.

*Response unidentifiable* – A high proportion of unidentifiable responses on paper (17.5%) which correlates to the highest school and post-school qualification questions from which this question is derived. Internet responses were much better, with only 2.4% of unidentifiable responses, probably because of the As You Type lists provided for post-school qualification. As this output is a derivation of Q31-33 it is difficult to assess how each question contributes to the overall residuals rates.

*Analysis* – Overall, the data fits with expectations, when compared with 2013 data. The differences between online and paper respondents in 2017 also fits with the age data analysis which shows that older people are

much more likely to complete on paper than younger people, and are also more less likely to have higher qualifications.

#### Highest Qualification Gained by Mode - 2017 test data

Highest Qualification	Online	Paper	Total
00 No Qualification	19.2	32.6	21.9
01 Level 1 Certificate Gained at School	14.3	17.9	15.0
02 Level 2 Certificate Gained at School	10.3	10.1	10.3
03 Level 3 or 4 Certificate Gained at School	5.2	3.5	4.9
04 Overseas Secondary School Qualification	3.8	3.2	3.6
05 Level 1 Certificate Gained Post-School	1.2	0.8	1.1
06 Level 2 Certificate Gained Post-School	1.6	1.8	1.7
07 Level 3 Certificate Gained Post-School	3.4	2.7	3.3
08 Level 4 Certificate Gained Post-School	8.9	6.6	8.4
09 Level 5 Diploma	4.8	3.0	4.5
10 Level 6 Diploma	6.1	5.1	5.9
11 Bachelor Degree and Level 7 Qualifications	11.8	8.2	11.1
12 Post-Graduate and Honours Degree	6.0	3.3	5.4
13 Masters Degree	2.8	1.0	2.4
14 Doctorate Degree	0.5	0.3	0.4
<b>Total Responses</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
97 Response Unidentifiable	2.4	17.5	6.1
99 Not Stated	0.6	9.2	2.7

Comparison of the 2013 Census data with the test data below shows good consistency across the 2 data points. Note that the output categories differ between 2013 and 2017. Residuals are higher for paper in 2017 but the proportion of paper responses will be lower in the 2018 Census.

#### Highest Qualification by Mode, Whanganui adults, 2013 Census and 2017 Test

	Internet		Paper form	
	2013	2017	2013	2017
No Qualification	19.2	19.2	33.1	32.6
Level 1 Certificate	15.5	20.4	16.4	18.7
Level 2 Certificate	12.0	11.9	11.3	11.9
Level 3/4 Certificate	20.7	17.5	7.6	12.8
Overseas Secondary School Qualification	3.9	3.8	2.7	3.2
Level 5 Diploma	4.7	4.8	3.5	3.0
Level 6 Diploma	5.9	6.1	4.7	5.1
Bachelor Degree and Level 7 Qualification	12.3	11.8	6.9	8.2
Post-graduate and Honours Degrees	3.1	6.0	1.5	3.3
Masters Degree	2.1	2.8	0.8	1.0
Doctorate Degree	0.5	0.5	0.1	0.3
Not Elsewhere Included	3.3	3.0	17.8	26.7

## Overseas post-school qualification indicator

**34** Was your highest qualification gained in New Zealand or overseas?

- New Zealand  
 overseas

Subject population: NZ usual resident adult respondents who indicate they have completed a qualification other than a secondary school qualification

Qualification Gained	Internet Responses		Paper Responses		TOTAL	
	Number	%	Number	%	Number	%
New Zealand		85.3		85.5		85.3
Overseas		14.7		14.5		14.7
Total Stated		100		100		100
Unidentifiable		0		0		
Not Stated (% of total)		0.9		3.1		1.2
TOTAL (eligible respondents)						

Comments:

*Non-response* – The internet sample falls within the acceptable non-response rate of 5% or less. Paper forms had a slightly higher non-response rate of 3% compared with the internet forms (0.9%)

*Response unidentifiable* – The level of response unidentifiable was negligible.

*Frequency of responses* – No difference across the modes. This is a new question for 2018, so there is no Whanganui data for comparison, but in the July 2016 test the proportion of respondents with New Zealand qualifications was 75% (internet) and 83% (paper), which is in the same range as the 2017 data..

*Assessment* – Test data indicates that this question would likely produce fit for purpose data.

## Field of Study

Subject population: NZ usual resident adult respondents who indicate they have completed a qualification other than a secondary school qualification

Field of Study	Percent of subject responses	2013 Census (Whanganui)
Health	17.3	15.4
Society and Culture	15.1	13.3
Management and Commerce	14.9	12.4
Engineering and Related Tech	14.6	18.2
Education	14.2	11.3
Architecture and Building	5.8	8.8
Creative Arts	5.0	5.6
Food, Hospitality and Personal Services	4.2	6.6
Natural and Physical Sciences	4.1	2.5

Agriculture, Environmental and related Information Technology	3.0	4.2
Information Technology	2.0	1.7
Residual Codes (DK/Unidentifiable/Not Stated)	6.4	

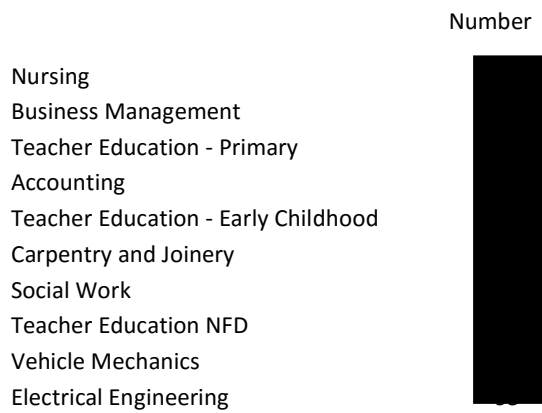
Comments:

*Not stated*– The non-response rate was 4.5%, which is below the acceptable 5% non-response rate.

*Response unidentifiable* – Around 1.8% of responses were unidentifiable or out of scope.

*Frequency of responses* – The data looks acceptable – when compared with 2013 Census data for Whanganui, with a similar pattern of responses at the higher level of the qualification.

A brief examination of detailed level field of study responses does not show any unusual patterns, with the top 10 responses as follows:



*Assessment* – Test data indicates that this question working acceptably and would likely produce fit for purpose data. Our recommendation is to retain the question in its current format.

## Income variables

### Total personal income

**37** From all the sources of income you marked in **36**, what will the total income be:

- that you yourself got
- before tax or anything was taken out of it
- in the 12 months that ended on 31 March 2017

- loss
- zero income
- \$1 – \$5,000
- \$5,001 – \$10,000
- \$10,001 – \$15,000
- \$15,001 – \$20,000
- \$20,001 – \$25,000
- \$25,001 – \$30,000
- \$30,001 – \$35,000
- \$35,001 – \$40,000
- \$40,001 – \$50,000
- \$50,001 – \$60,000
- \$60,001 – \$70,000
- \$70,001 – \$100,000
- \$100,001 – \$150,000
- \$150,001 or more

Subject population: usual residents aged 15 years and over

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Loss		0.4		0.7		0.5
Zero income		5.5		5.4		5.5
\$1-\$5,000		3.6		6.7		4.2
\$5,001-\$10,000		3.4		3.6		3.4
\$10,001-\$15,000		8.7		11.4		9.3
\$15,001-\$20,000		11.7		16.3		12.7
\$20,001-\$25,000		9.6		15.0		10.8
\$25,001-\$30,000		7.0		8.7		7.3
\$30,001-\$35,000		5.7		6.4		5.9
\$35,001-\$40,000		6.1		4.8		5.8
\$40,001-\$50,000		10.2		6.7		9.5
\$50,001-\$60,000		8.7		5.0		7.9
\$60,001-\$70,000		5.9		3.1		5.3
\$70,001-\$100,000		8.5		4.7		7.7
\$100,001-\$150,000		3.2		1.6		2.8
\$150,001 or more		1.8		0.0		1.4
Total stated						
Response unidentifiable						
Not stated						
NA						
Total						
Percentage not stated		2.2		16.0		5.6

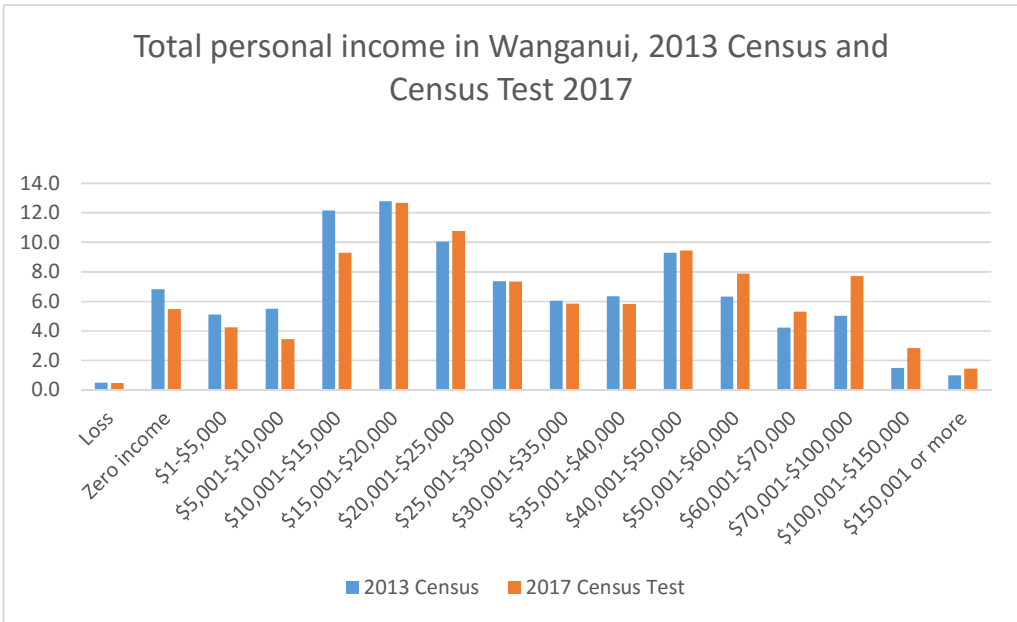
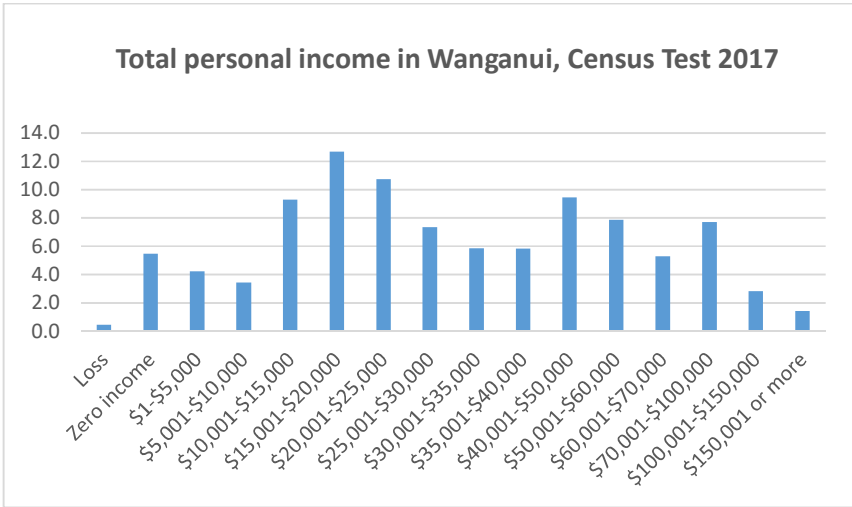
Comments:

*Non-response* – The overall level of not stated was fairly acceptable, and not stated for those using online forms was low. However, the percentage of not stated for paper forms was unacceptably high, and those using bilingual forms were more likely (at 17.1 percent) to be coded to not stated than those using English forms (4.9 percent). This has not been investigated further to see if it was due to responses being put in the wrong place on the bilingual form or true non-response.

*Response unidentifiable* – A low number of responses were coded to response unidentifiable.

*Frequency of responses* – The data looks sensible. The distribution of the data shows a similar pattern to that for the 2013 Census for Wanganui, and the expected pattern of increases in income levels over time.





As expected, those filling it in online tended to have higher personal income levels than those filling it in on paper.

*Assessment* – This limited analysis does not show any new issues with this data.

It appears that greater use of online forms may result in improved quality for this data in 2018 than in previous censuses, due to lower non-response overall. However, it seems likely that there will still be an issue with bias (skew) in the data due to lower representation of those on relatively low incomes. This test data shows that those using paper forms tend to have lower income levels and higher levels of non-responses.

## Sources of personal income

**36** Mark as many spaces as you need to show all the ways you yourself got income in the 12 months ending today.

Don't count loans because they are not income.

- wages, salary, commissions, bonuses, etc, paid by my employer
- self-employment, or business I own and work in
- interest, dividends, rent, other investments
- regular payments from ACC or a private work accident insurer
- NZ Superannuation or Veteran's Pension
- other superannuation, pensions or annuities (other than NZ Superannuation, Veteran's Pension or war pensions)
- Job Seeker Support
- Sole Parent Support
- Supported Living Payment
- Student Allowance
- other government benefits, government income support payments, war pensions, or paid parental leave
- other sources of income, counting support payments from people who do not live in my household

or  no source of income in that time

Subject population: usual residents aged 15 years and over

	Number	Percent
No source of income during that time		5.8
Wages, salary, commissions, bonuses etc paid by my employer		50.5
Self-employment or business I own and work in		13.4
Interest, dividends, rent, other investments		23.8
Regular payments from ACC or a private work accident insurer		1.8
New Zealand Superannuation or Veteran's Pension		32.7
Other superannuation, pensions or annuities (other than NZ Superannuation, Veteran's Pension or war pensions)		5.1
Jobseeker Support		3.4
Sole Parent Support		1.6
Supported Living Payment		2.7
Student Allowance		1.4
Other government benefits, government income support payments, war pensions or paid parental leave		4.4
Other sources of income, including support payments from people who do not live in my household		1.4
Total stated		
Not Stated		2.4
Total		

Note: this analysis has not been done split by mode (online, paper) due to the complicated nature and amount of work involved in dealing with a multiple response variable with many categories.

Comments:

*Non-response* – Overall non-response was well within the acceptable range at 2.4 percent. This is lower than national non-response in the previous three censuses. Breaking the test data down by mode showed very low non-response for online forms (0.6 percent) and a non-response rate that was a bit higher than desirable for paper forms (7.7 percent).

*Response unidentifiable* – No responses were coded to response unidentifiable for this test.

*Frequency of responses* – The distribution of the data looks sensible. The distribution is mostly fairly similar to national results from the previous three censuses, except for a much higher percentage of people receiving NZ Superannuation and lower percentages of people receiving supported living payments (previously this was sickness and invalids benefits) or a student allowance. These differences reflect the skewed nature of the sample for this voluntary test. It is skewed toward older age groups and people on supported living payments are probably less likely to be included in a voluntary test than in census proper.

*Assessment* – This question appears to be working well and – as previously – is suitable for inclusion in the 2018 Census in this format. There do not appear to be any issues arising from the changes to some category names. It looks like greater use of online forms may be effective in lowering non-response (and increasing data quality) for this variable.

## Tenure of household and associated variables

Note: The analysis below is indicative because it has not been possible to restrict the data to the correct subject population of households. The derived variables needed to identify households are not available. The analysis has been done for all dwellings in the dwelling dataset. Generally these are private, but it includes a few previously unidentified non-private dwellings and may also include some visitor-only private dwellings. This is not expected to have much effect on the results. It may slightly decrease data quality.

**5** Do you, or anyone else who lives here:

hold this dwelling in a family trust?

own or partly own this dwelling, with or without a mortgage? → go to **9**

or  neither of these? → go to **6**

See the Guide Notes for help

**6** If nobody who lives here owns this dwelling, who owns it?

private person, trust or business

local authority or city council

Housing New Zealand Corporation

iwi, hapū or Māori land trust

other community housing provider

other state-owned corporation or state-owned enterprise, or government department or ministry

don't know



respondents find these questions harder than other dwelling form questions, are reluctant to answer them, or feel that some of these questions are not relevant to them. Respondents putting responses in the wrong place has also contributed to this high percentage for not stated.

*Response unidentifiable* – Only a small number of multiple responses on paper forms were coded to response unidentifiable, which is good for data quality. (Note: no multiple response is possible on online forms.) These were a mixture of owned and in family trust, which is a possible situation, and so not a data quality issue, and inconsistent responses (eg owned and don't own) which would be a data quality issue if there were many of these responses. A multiple response of owned and in a family trust on paper forms will get coded to a family trust category by the tenure of household derivation.

*Frequency of responses* – The test data is somewhat skewed toward home owners, with a percentage of home owners that is around 10 percent higher than expected (around half was expected), and lower representation of those who do not own their home or have it in a family trust (mostly renters). This is probably just due to the voluntary nature of the test and characteristics of those who responded rather than indicative of problems with this question. The percentage in a family trust is as expected and follows the previous trend of a small increase.

*Assessment* – The question itself generally appears to be working well and suitable for inclusion in this format in the 2018 Census. However, there seems to be greater reluctance or less interest in answering it compared with other dwelling form questions. Maximising use of online forms may be the best strategy for ensuring good quality data.

For online forms, it would be useful to have a help note explaining how to answer if the dwelling is partly owned and partly in a family trust (mark family trust).

## Mortgage payments

Subject population: Households in which the respondent said they or someone else living there owned the dwelling or had it in a family trust. However, this analysis has been done for all dwellings in the dwelling dataset where this response was given rather than for households only because the derived variables to identify households were not available. A small number of non-private dwellings and visitor-only private dwellings may be included in the analysis. This should make very little, if any, difference to the results. It may have slightly decreased data quality.

A few cases where there was a multiple response of owned and in a family trust have been omitted from the analysis because these were coded to response unidentifiable for this test, along with other types of inconsistent multiple response.

Mortgage payments by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Mortgage payments made		47.0		33.2		44.9
Mortgage payments not made		52.1		65.8		54.2
Don't know		0.9		1.0		0.9
Total stated						
Not stated		0.9		17.3		3.7
Total						

*Non-response* – Overall non-response is within the acceptable level, and non-response was very low for those using online forms, but unacceptably high for those using paper forms.

Checking some images showed that much of the non-response on paper forms is due to respondents not following routing instructions correctly. There is a common pattern of marking own or family trust in question five, then answering ‘no’ for the rent indicator question, (and sometimes answering sector of landlord too), and then skipping to question 10 rooms. They appear not to have seen the “go to 9” instruction in question five, perhaps because it is quite far over toward the right side. The dwelling forms for these respondents were otherwise filled in very well. A less common pattern was respondents filling in their dwelling form fully except for question 9 and one case of contradictory answers (own and pay rent) and an “n/a” answer in the wrong place.

*Response unidentifiable* – No responses were coded to response unidentifiable. There do not appear to have been any multiple responses or responses that were unclear as to intent.

*Frequency of responses* – The data looks sensible in that the split between making mortgage payments and not making them is reasonably even. However, it is somewhat contrary to expectations. The percentage making mortgage payments is about 10 percent lower than expected and the percentage not making mortgage payments is about 10 percent higher than expected. This may be due to the test sample being skewed toward older respondents who are more likely to have paid off a mortgage, with lower representation of younger respondents who are more likely to still have a mortgage.

The percentage of ‘don’t know’ answers was very low, which is good for data quality.

*Assessment* – As far as is possible to tell from these test results, this question appears to be working in a satisfactory way and appears to be suitable for inclusion in this format in the 2018 Census. Previously there has been some concern that respondents could have difficulty answering ‘yes’ or ‘no’ to a two-part question. To properly explore this would require cognitive testing.

It is recommended that the placement of the “go to 9” instruction on the paper form be reviewed to see if it can be moved to a more visible location. However, promoting use of online forms may be the best strategy for ensuring good quality data on mortgage payments.

## Rent Indicator

Subject population: those households who said they did not own the dwelling or have it in a family trust. However this analysis has been done for all dwellings where this response was given rather than households because the derived variables to identify households were not available. A small number of non-private dwellings and visitor-only private dwellings are likely to be included in the analysis. This should make very little, if any, difference to the results. It may have slightly decreased data quality.

Rent indicator by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Rent paid		84.0		75.3		81.9
Rent not paid		16.0		24.7		18.1
Total stated						
Not stated		0.6		6.4		2.1
Total						

Comments:

*Non-response* – Non-response for paper forms is a little higher than desirable but the very low rate for online forms brings it down to an acceptably low level overall. It appears that greater use of online forms (which present questions one at a time) will be a major factor in minimising non-response. Other factors contributing to this low non-response could be the compliant nature of test participants and the more condensed questionnaire design for the tenure-related questions.

Checking images for paper forms where rent indicator was coded to not stated showed some cases of respondents answering the question in the wrong place or in a different way than intended (eg outside the checkbox area, by underlining the text, or to the right of the text, even on the English form). Solving these issues would reduce the amount of data coded to not stated.

*Response unidentifiable* – There appear to have been no multiple responses to this question in this test. This seems to be an easy question to answer.

*Frequency of responses* – As expected, most respondents who needed to answer this question indicated that they paid rent ie were renting. However, the percentage who said they did not pay rent seems quite high. (A slightly rough comparison using 2013 output data for tenure of household for Wanganui gives a figure of 11 percent.)

Some of these no rent paid responses appear to be incorrect, based on answers to other questions (eg sector of landlord, rent amount) so the true percentage occupying the dwelling rent-free is likely to be lower than this and more similar to the 2013 figure. In census proper, further work would be done to investigate this and corrective action would be taken where necessary.

























*Assessment* – This question generally appears to be working well and is suitable for inclusion in the 2018 Census in this format. It appears that greater use of online forms will help to ensure low non-response and good quality output data for households who rent their home.

## Weekly rent paid by household

Subject population: households who rent (ie category 21 in tenure of household), but the analysis below is for all dwellings in the dwelling dataset where “yes” was marked for rent indicator. Generally these are private dwellings but there may be some previously unidentified non-private dwellings and it may include some private dwellings that do not contain a household (visitor-only private dwellings).

The analysis has been done this way because a household indicator is not available and to avoid complications from any issues with the tenure of household derivation. It is not expected to have made much difference to the results. It may have slightly decreased data quality.

Weekly rent paid by household by mode

Weekly Rent (banded)	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Under \$50		0.3		1.9		0.7
\$50-79		2.6		7.2		3.8
\$80-99		1.4		5.3		2.4
\$100-124		8.3		19.2		11.2
\$125-149		5.2		10.4		6.6
\$150-174		12.7		16.4		13.7
\$175-199		13.0		14.5		13.4
\$200-249		25.6		13.5		22.4



\$250-299	■	19.4	■	5.0	■	15.6
\$300-349	■	6.9	■	1.6	■	5.5
\$350 or more	■	4.6	■	5.0	■	4.7
<b>Total stated</b>	■		■		■	
Response unidentifiable	■	0.1		0.0	■	0.1
Not stated	■	3.7	■	9.1	■	5.2
<b>Total</b>	■		■		■	

Comments:

*Non-response* – Overall non-response is around the acceptable range, but higher than in 2013 for this area (3.6 percent). Non-response for online forms is within the acceptable range, but non-response for paper forms is unacceptably high.

Checking images for paper forms showed that factors contributing to this are a pattern of respondents answering either rent amount or rent period, but not both (this occurred on bilingual and English forms), and respondents putting responses in the wrong place, for example:

- rent amount inside rent period box (two cases seen, one a bilingual form, one English)
- rent amount to the right of the “.00” in the box (bilingual form)
- a response of “6000 per year” in the rent amount box – the rent period wasn’t captured from this.

In census proper, responses such as these should (ideally) be brought up to operators so they can get coded correctly at that point. Otherwise they will need to be identified via evaluation checks and then corrective action will need to be taken to get them coded correctly. This is more time consuming and less effective than fixing during processing.

The data for online respondents coded to not stated showed that most answered rent period without giving a rent amount, which is odd. Possibly the question on the online form is not always being displayed correctly or online responses to this question are not being captured correctly.

*Response unidentifiable* – The only unidentifiable response occurred on the online form. This was a text entry of “2weeks”.

*Frequency of responses* – In general, the distribution looks sensible. It is similar to the 2013 Census data for Whanganui.

As expected, there is a general pattern of slight decreases for lower rent categories and slight increases for higher rent categories compared with 2013. This probably reflects rent increases since 2013. However, there is a dramatic decrease in households paying \$80-99, and similar increase in households paying \$100-124. Looking at sector of landlord for these households suggests that a local authority has put its rent up since the last census, pushing a large group of households up a rent bracket.

There were only a few strangely high rent amounts. Checking images for these showed that this was partly caused by respondents putting in decimal points and amounts for cents and not using the pre-printed “.00” for cents (eg an apparent response of 5,000 a week was really \$50.00 and an apparent response of 10350 a week was really 103.50). In census proper, such responses could probably be mostly fixed by dividing the rent amount by 100. In another case the respondent also appears to have not given the correct number of zeros for the dollar amount (they gave a response of 10.02 which was recognised as 1002.) Another factor contributing to these very high rents is respondents crossing out the first two boxes with diagonal lines, resulting in a five-figure rent amount starting with \$11.... (eg a rent amount captured as “11230” when the intended response was \$230).

Checking images of very low rent amounts did not reveal any error, introduced or otherwise. Even when the respondent wrote a low amount in the wrong boxes ie starting on the far left side (96,— or 50, followed by

three empty boxes) the intended response (\$96, \$50) was recognised correctly. It appears that some respondents pay very low/nominal rent amounts (eg those on low incomes in Housing New Zealand homes).

*Assessment* – In general this question appears to be working acceptably well and suitable for inclusion in this format in the 2018 Census.

However, to help ensure good data quality it is recommended that further testing of the online form be done to check whether responses are being captured correctly and whether the rent amount question is being displayed correctly.

## Sector of landlord

**6** If nobody who lives here owns this dwelling, who owns it?

- private person, trust or business
- local authority or city council
- Housing New Zealand Corporation
- iwi, hapū or Māori land trust
- other community housing provider
- other state-owned corporation or state-owned enterprise, or government department or ministry
- don't know

Subject population: households who rented their home (ie the renting category from tenure of household), but the analysis below is for all dwellings in the dwelling dataset for which the derived tenure was renting because no household identifier variable was available. These dwellings are generally private, but there may be a few previously unidentified non-private dwellings and some visitor-only private dwellings. This is expected to have made little, or no, difference to the results. It may have slightly decreased data quality.

### Sector of landlord by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Private person, trust, or business		84.6		70.1		80.3
Local authority or city council		3.5		15.9		7.1
Housing New Zealand Corporation		5.4		10.0		6.8
Iwi, hapū, or Māori land trust		0.2		0.5		0.3
Other community housing provider		2.0		2.1		2.1
Other state-owned corporation or state-owned enterprise, or government department or ministry		0.9		0.2		0.7
Don't know		3.4		1.2		2.7
Total stated						
Not stated		0.3		6.1		2.1
Total						

Comments:

*Non-response* – Overall non-response is acceptably low, but non-response for paper forms is a little higher than desirable.

*Response unidentifiable* – No responses were coded to response unidentifiable.

*Frequency of responses* – The distribution looks sensible, with the majority renting from a private person, trust, or business. As expected, the percentages in social housing were relatively low, and the number of responses to the new categories (iwi, hapu, or Maori land trust; other community housing provider) was low. Almost all responses to the new categories were for “other community housing provider”.

One response of “other community housing provider” seen appears incorrect, as the respondent said the owner is “Summerset”. This is probably a licence to occupy dwelling in a retirement village. (This respondent also appears to have answered question five, dwelling owned or in family trust, incorrectly.)

Responses of don’t know were quite low overall, but higher for online forms than paper forms. Possibly people using online forms tended to still respond (with a don’t know answer) if they did not know who the landlord was whereas those using paper forms tended to skip the question altogether (and get coded to not stated) if they didn’t know who the landlord was.

*Assessment* – This question generally appears to be working satisfactorily and is suitable for inclusion in the 2018 Census. There is no evidence of major issues with the new categories. As the numbers in the new categories are likely to be small, any error affecting these categories could have a significant effect on data quality. For census proper, it will be important to have a close look at the responses for these categories to make sure they appear correct.

## Tenure of household (derived output variable)

Subject population: all households, but this analysis is for all dwellings in the dwelling dataset because the derived variables to identify households were not available. Generally these are private dwellings, but it includes a few previously unidentified non-private dwellings and may also include some visitor-only private

dwellings. This is expected to have made very little, or no, difference to the results. It may have slightly decreased data quality.

Tenure of household by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Dwelling owned or partly owned, mortgage payments not further defined		0.8		8.3		2.3
Dwelling owned or partly owned, mortgage payments made		32.6		15.2		29.1
Dwelling owned or partly owned, mortgage payments not made		30.3		28.9		30.0
Dwelling not owned and not held in a family trust, rental arrangements not further defined		0.1		0.7		0.2
Dwelling not owned and not held in a family trust, rent payments made		18.3		32.0		21.1
Dwelling not owned and not held in a family trust, rent payments not made		1.0		1.5		1.1
Dwelling held in a family trust, mortgage arrangements not further defined		0.6		3.4		1.1
Dwelling held in a family trust, mortgage payments made		5.0		2.9		4.6
Dwelling held in a family trust, mortgage payments not made		11.2		7.0		10.3
Total stated						
Not stated		3.3		10.8		4.9
Total						

Comments:

*Non-response* – The overall percentage not stated (which means insufficient information to classify tenure) is acceptably low, and a small improvement compared with 2013. Factors contributing to this improvement might include: greater use of online forms with associated low non-response, the more condensed style of questioning, changes to the derivation (which has been completely revised), and the voluntary nature of the test and relatively compliant nature of test respondents.

The percentage not stated for paper forms is higher than acceptable, reflecting higher non-response on paper forms to the input questions used to derive tenure of household.

*Response unidentifiable* – No responses were coded to response unidentifiable. Checking of the derivation (which is being done as a separate exercise) would be useful to see if any combinations of responses were made that should have been coded to response unidentifiable.

Checking images for paper forms showed that respondents occasionally gave inconsistent or unnecessary responses to the tenure-related questions and did not always follow the questionnaire routing. For example, there were cases of respondents marking the family trust response option and then answering the sector of landlord and rent questions. Four respondents unnecessarily answered rent amount despite marking “no” for rent indicator. Two of these were text responses of “no rent” and “I own this home”, the latter of which was

from a respondent who answered “neither of these” to the question about whether the dwelling was owned or in a family trust.

*Frequency of responses* – The data generally looks sensible but is skewed toward home owners and away from those who do not own their home or have it in a family trust (mostly renters), which probably reflects the characteristics of the people who participated in this test.

The percentages in “not further defined” categories are low, which is good for data quality.

Checking forms showed some evidence of people in licence to occupy dwellings not answering correctly (see the sector of landlord section). Evidence of respondents in this situation answering incorrectly or being confused about the tenure-related questions was also found in the follow-up survey.

In the WOF for tenure of household for this test, it was noted that a small bug was found in the derivation logic. This incorrectly assigned some households to category 21 (renting) instead of category 31 (in family trust, mortgage payments made) or 32 (in family trust, no mortgage payments made). This error needs to be addressed when the derivation is checked and tested.

*Assessment* – This variable appears to be working sufficiently well. It is difficult to gauge whether, or to what degree, the more condensed style of questioning has improved data quality.

Some tweaks to questionnaire design for the paper form are worth considering, given the issue with questionnaire routing not being followed. However, promoting use of online forms may be the most effective strategy for ensuring good quality data for tenure of household. Strategies to help respondents in licence to occupy dwellings (eg flyers on retirement village notice boards?) answer correctly would be helpful for interpreting this data.

## Individual Home Ownership

Subject population: New Zealand residents aged 15 years and over

**30** Thinking about the dwelling that you usually live in, do you yourself:

- hold it in a family trust?
- own or partly own it, with or without a mortgage?
- or  neither of these

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Hold in a family trust	█	14.9	█	12.1	█	14.3
Own or partly own	█	55.8	█	45.0	█	53.4
Do not own and do not hold in a family trust	█	29.3	█	42.9	█	32.4
Total stated	█		█		█	

Response unidentifiable	0	0.8	0.2
Not stated	0.6	5.8	1.8
<b>Total</b>			

Comments:

*Non-response* – Non-response is well within acceptable levels, and remarkably low overall, given the issues with non-response to this question identified during evaluation in 2013. This may be due to the more compliant nature of respondents in this voluntary test, and to the test data being skewed toward home-owners. Home owners might be more interested in answering this question than non-home owners (eg because non-home owners feel that it is not relevant to them.)

Response unidentifiable – The amount of unidentifiable responses is very low, which is good for data quality.

*Frequency of responses* – The relative figures for owning, not owning, and family trust generally look sensible, reflecting the somewhat skewed nature of the sample (skewed toward older age groups and home owners).

Comparing the test data results with the 2013 Census results for Wanganui shows the expected differences, given the skewed nature of the test sample. However the data is not completely comparable because of the introduction of a separate response category for family trust. It is not known how well respondents in 2013 followed the questionnaire instruction to mark ‘yes’ if they had their dwelling in a family trust.

#### Individual home ownership in Wanganui, 2013 Census and Census Test 2017

	2013 Census	Census Test 2017
Own or partly own usual residence	54.0	53.4
Hold in a family trust*	-	14.3
Do not own usual residence	46.0	32.4
Not elsewhere included	5.9	2.0

\*not asked as a separate category in 2013. Respondents were asked to mark own.

*Assessment* – There is no evidence of any major problems with this new style of question about individual home ownership. It appears to be working well and is suitable for inclusion in this format for the 2018 Census. The introduction of a separate family trust category may help with interpretation of this data as well as provide a better respondent experience.

## Occupied dwelling type and associated variables

The following analysis is restricted to occupied private dwellings. No data for non-private dwellings was available.



Dwelling description by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
House		88.7		75.0		85.8
Townhouse		4.6		7.9		5.3
Unit		5.9		15.0		7.8
Apartment		0.8		1.5		0.9
Mobile dwelling not in a motor camp		0.1		0.1		0.1
Improvised dwelling or shelter		0.0		0.1		0.0
Non-private dwelling		0.0		0.3		0.1
Total stated						
Response unidentifiable		0		1.0		0.2
Not stated		2.8		5.9		3.5
Total						

*Non-response* – Overall non-response is acceptably low. As for many other questions, it is a bit higher than desirable for paper forms. Checking some images for paper forms showed that respondents putting responses in the wrong place (particularly on bilingual forms, but also on English forms) has contributed to this.

*Response unidentifiable* – Very few responses (all on paper forms) were coded to response unidentifiable, which is good from a data quality point of view.

*Frequency of responses* – The data looks sensible. As expected for this area, ‘house’ was by far the most common dwelling description. Units and townhouses were much less common, and apartments were relatively rare. Those using the paper form were more likely to be in units or townhouses than those using the online form. Dwellings that were joined were commonly described as a ‘unit’ for dwelling description, which seems sensible.

*Written responses* – There were ■ written responses (0.6 percent of those who stated a response). Some respondents gave a written response as well as marking a tick box.

This seems like quite a low amount of written responses. Most respondents appear to be making good (appropriate) use of the new tick boxes in this simplified dwelling description question.

Some of these were genuine ‘other’ responses for which a written response is necessary, but there were others that were unnecessary or did not provide the type of information required (eg indicating things such as the materials used for the dwelling, number of bedrooms, or tenure). As previously, some respondents seem to be giving a written response because the term they use to describe their home is not on the form eg villa, flat.

There were a few written responses indicating that the dwelling was really non-private, not private. In census proper, these would need to be put back through the processing system and reclassified as private. Other dwelling form questions (eg relationship to reference person, rooms) and the activity limitation questions on the individual form may also be useful for identifying dwellings that are really non-private. (For example, for residential and community care facilities, relationship to reference person could be clients, the activity limitation questions may show instances of people who have a lot of difficulty doing certain activities or cannot do them at all, and the number of bedrooms/rooms may be higher than usual.)

It is difficult to gauge whether this new questionnaire design has reduced the amount of unnecessary or otherwise non-useful written responses compared with 2013. Many respondents gave unnecessary written responses in 2013 but the number is not available.



A few written responses to this question identified the dwelling as being in a retirement village. This information is not needed for classifying occupied dwelling type, but could be useful for checking how respondents in this situation answer tenure-related questions.

*Assessment* – In general this question appears to be working well and suitable for inclusion in this format in the 2018 Census.

Although it is difficult to determine whether this new questionnaire design has reduced the level of unnecessary or non-useful responses, it is hoped that it will help to reduce this and will make processing easier and faster and contribute to better data quality for the occupied dwelling type variable for 2018.

*Non-private dwellings* – No data on non-private dwellings was available so it is not possible to assess how well the enumeration/classification of non-private dwellings worked in this test. Some information relating to non-private dwellings is available in reports from field operations and infrastructure, but these focus on field processes and do not provide much insight into aspects relating to counting or classifying dwellings. There are some indications of issues with private dwellings located at non-private complexes. At some point, the process of feeding through the non-private dwelling information into the processing system and into the occupied dwelling type classification will need to be tested to make sure it works. The process being used for 2018 is new.

### Dwelling joined or separate (question three)

Subject population: All private dwellings, but the analysis below includes all dwellings in the dwelling dataset. Essentially these are private dwellings with a few previously unidentified non-private dwellings, so this should make no or very little difference to the results.

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Joined		8.0		16.1		9.7
Separate		92.0		83.9		90.3
Total stated		100.0		100.0		100.0
Not stated		2.5		6.0		3.2
Total						

*Non-response* – As for some other questions, non-response is acceptably low overall, but a little higher than desirable for paper forms.

*Response unidentifiable* – No responses have been coded to response unidentifiable, so there appear to be no multiple responses, which is good for data quality.

*Frequency of responses* – The distribution of responses is as expected for this area – the vast majority of private dwellings are separate, and only a small percentage are joined.

*Assessment* – This question appears to be working well and suitable for inclusion in this format in the 2018 Census. Having this aspect of the dwelling type information split out into a simple, separate question looks promising for helping to raise the quality of the output data on occupied dwelling type.

## Number of storeys (question four)

Subject population: All private dwellings, but the analysis below includes all dwellings in the dwelling dataset. Essentially these are private dwellings with a few previously unidentified non-private dwellings, so this should make no or very little difference to the results.

	online		paper		Total	
	Number	Percent	Number	Percent	Number	Percent
One storey		83.2		91.4		84.9
Two or three storeys		16.7		8.5		15.0
Four to six storeys		0.1		0.1		0.1
Seven to nine storeys		0.0		0.0		0.0
Ten or more storeys		0.0		0.0		0.0
Total stated						
Response unidentifiable						
Not stated		2.7		7.9		3.8
Total						

Comments:

*Non-response* – The level of non-response was acceptably low overall, and low for those using the online form, but higher than desirable for those using paper forms. The majority of this non-response (over two thirds) was on English forms.

Checking some images for paper forms showed that respondents putting their response in the wrong place (ie not in the response oval), and this not being captured during scanning, has contributed to this higher than desirable not stated percentage. This happened on bilingual forms and (maybe to a lesser extent) on English forms. Often the response was on the right side of the response options. Sometimes it was beside the question on the right side. Some respondents continued in this way throughout the form. Others seem to have realized their mistake part-way through filling in the form. There is an instruction on the paper form indicating where answers should go. It appears that some respondents did not see this.

*Response unidentifiable* – No responses have been coded to response unidentifiable. There appear to be no cases of multiple response, which is good for data quality.

*Frequency of responses* – The frequency distribution is as expected for Wanganui, with most private dwellings being of one storey/in a one-storey building, and a relatively low percentage with two or three storeys (15.0 percent). There were very few with four or more storeys, one of which appears to be an unidentified non-private dwelling, based on responses to number of rooms.

*Assessment* – As far as is possible to tell from this test data, this question appears to be working satisfactorily and suitable for inclusion in this format in the 2018 Census.

Unfortunately, because this test was done in Wanganui only, this test data does not indicate how well this question will work for respondents in large, multi-storey apartment complexes. Given the issue in 2013 with data being unusable due to respondent error, for the 2018 Census it will be important to check that respondents in high-rise apartments have answered this question correctly.

To ensure good data quality, it is recommended that steps be taken to help prevent respondents putting responses in the wrong place, or make sure that these responses are captured if this cannot be prevented. Changes to questionnaire design (eg greater colour contrast between response ovals and the surrounding form, changing the layout of the bilingual form) may help prevent this problem. If it is not possible to prevent

this or allow for it during scanning, manual checking of paper forms before scanning for responses in the wrong place and transcribing of forms may be needed.

## Occupied dwelling type

Subject population: All occupied dwellings

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Occupied private dwelling not further defined		3.1		7.9		4.1
Separate house no storey information		0.5		2.0		0.8
Separate house with one storey		74.3		69.8		73.3
Separate house with two or more storeys		14.7		5.8		12.8
Joined dwelling no storey information		0.0		0.3		0.1
Joined dwelling in a one-storey building		5.9		12.5		7.3
Joined dwelling in a two- or three storey building		1.5		1.3		1.5
Joined dwelling in a ten- or more storey building		0.0		0.0		0.0
Mobile dwelling not in a motor camp		0.1		0.1		0.1
Improvised dwelling or shelter		0.0		0.1		0.0
Hotel, motel or guest accommodation		0.0		0.2		0.0
Total		100.0		100.0		100.0

Note: this variable does not have residual categories. All occupied dwellings are at least classified as private not further defined or non-private not further defined.

Comments:

*Frequency of responses* – The data generally looks sensible and as expected given the area in which this test was done, with most dwellings being separate houses of one storey, and joined dwellings in one-storey buildings being more common than those in two- or three storey buildings. The dwelling coded as joined in a ten-or storey building appears to be due to a facetious response, given the responses made elsewhere by this respondent.

These test results provide some evidence of improved data quality. The percentage of private dwellings coded to 'not further defined' is lower overall (4.1 percent) than in the final output data for 2013 (5.8 percent), even without the additional work done to improve data quality for data which is to be output. As the percentage in this category was lower for forms filled in online, online completion may also be contributing to improving data quality.

The percentage of joined dwellings for which there was no storeys information is lower (0.6 percent) for this test than in 2013 (0.8 percent), again even without the additional work that would be done to improve data quality for data being output from a census. The percentage of separate dwellings with no storey information was actually higher in this test (0.9 percent) than in 2013 (0.4 percent), but again, this is without the benefit of the additional work that would be done to increase data quality for data that is to be output.

One of the dwellings identified as non-private in dwelling description has somehow disappeared from the non-private dwelling categories for this derived variable. How this has happened needs to be investigated when the derivation for occupied dwelling type is checked.







*Assessment* – The questions used to derive this variable appear to be working well and suitable for inclusion in this new format in the 2018 Census.

These test results provide some evidence of improved data quality compared with 2013 Census results. The new questionnaire design may be contributing to this, however it may also be related to greater use of online forms, the relatively low non-response associated with using online forms, and the voluntary nature of this test. These test results suggest that in 2018 it will be possible to produce higher quality data for occupied dwelling type than has been produced previously.

Note: The occupied dwelling type derivation itself needs to be checked as a separate exercise to make sure it is working correctly.

## Unoccupied dwelling type

Unoccupied dwellings in Wanganui, Census Test 2017 and 2013 Census

	Census Test 2017		2013 Census	
	Number	Percent	Number	Percent
Residents away		45.2		26.9
Empty dwelling		54.8		73.1
Total unoccupied dwellings				

Comments:

*Frequency of responses* – There was a low number of unoccupied dwellings in this test. The split between unoccupied, residents away and unoccupied, empty was different to that in 2013 for this area. In 2013, most unoccupied dwellings in this area were empty, but in this test there was a fairly even split between empty and residents away.

This data may not be comparable with the 2013 data because a major part of the process for determining whether a dwelling was occupied or not on census night was not tested in this test due to non-response follow-up being de-scoped. Only one area unit received a follow-up visit after census day.

It is difficult to tell how good the quality of this data is. In general it is expected to be harder to collect this information under the new collection model (because it reduces the amount of information on occupancy that can be collected during the enable phase compared with what could be collected under the previous model). This may tend to reduce data quality. Information from the field officer debrief indicates some difficulties with collecting this information. It was hard to determine dwelling occupancy when no contact had been made and the dwelling was not very visible, and difficult to distinguish between unoccupied, empty and unoccupied, residents away. It was noted that use of the same criteria for determining both these statuses could result in incorrect decisions.

Note: All primary sampling units in the Wanganui area were selected for the test except for 16 mesh blocks where the Household Labour Force Survey was in operation.

*Assessment* – As far as is possible to tell from this test, it appears that we will still be able to produce data on unoccupied, empty dwellings and unoccupied dwellings who had residents who were away. It may be helpful to review the criteria used by field staff.

Note: For 2018 dwelling occupancy status will be imputed if it has not been possible to determine the occupancy status.

## Rooms

Subject population: private dwellings, but this analysis includes all dwellings in the dwelling dataset. Generally these are private, but there may be a few that are really non-private. This is expected to make little difference to the results. It may slightly decrease data quality.

**10** For each type of room listed below, count the number in this dwelling.

Write '0' if this dwelling has no rooms of the type in the list.

<input type="text"/>	<input type="text"/>	kitchens
<input type="text"/>	<input type="text"/>	lounges, living rooms, or family rooms
<input type="text"/>	<input type="text"/>	dining rooms
<input type="text"/>	<input type="text"/>	bedrooms, including any sleepouts furnished as bedrooms
<input type="text"/>	<input type="text"/>	conservatories you can sit in
<input type="text"/>	<input type="text"/>	studies, studios, hobby rooms

Count any open-plan rooms as separate rooms.  
For example, a kitchen-dining room is two separate rooms

## Number of bedrooms

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
One bedroom		3.5		8.6		4.5
Two bedrooms		18.2		29.2		20.4
Three bedrooms		50.3		43.9		49.0
Four bedrooms		22.2		13.6		20.5
Five bedrooms		4.8		3.8		4.6
Six bedrooms		0.6		0.7		0.6
Seven bedrooms		0.2		0.1		0.2
Eight bedrooms		0.0		0.1		0.0
Nine bedrooms		0.0		0.0		0.0
Fourteen or more bedrooms		0.2		0.1		0.2
Total stated						
Not stated		3.7		12.2		5.5
Total						

Comments:

*Non-response* – Overall non-response is fairly acceptable. It is a little higher than for the 2013 Census, which is consistent with the previous trend.

Non-response for paper forms is higher than desirable. Breaking this down by the bilingual indicator shows that about half of this was on English forms and half on bilingual forms. The percentage using English paper

forms who did not respond was higher than acceptable (9.1 percent) and the percentage using bilingual paper forms was particularly high (20.5 percent). Viewing some images indicates that this high percentage for not stated was often due to true non-response (sometimes just to the bedrooms part of the question), but that responses in the wrong place, respondents marking boxes instead of writing a number, and responses not being captured has also contributed to it.

*Response unidentifiable* – No responses were coded to unidentifiable.

*Frequency of responses* – The overall distribution looks sensible and as expected, with three being the most common number of bedrooms.

Checking images for forms where the number of bedrooms was relatively high (six, seven, eight) or very high (eg 30, 35) showed no evidence of numeric misrecognition. All numerical responses viewed had been recognised correctly, including one with a scribbled out response beside the intended response. Some of these responses of a high number of bedrooms (six, seven or eight) seem to be correct and are just large houses. In other cases there was evidence that the dwelling was really non-private rather than private or possibly that a hotel/motel manager has mistakenly indicated all the bedrooms and other room types in the hotel or motel units rather than the number in their home.

There was a small number of responses of zero bedrooms, some of which were for mobile dwellings or dwellings described as “other”, and some of which appear to be due to respondent error.

*Assessment* – This new question style mostly appears to be working acceptably well and is suitable for inclusion in the 2018 Census in this type of format. However, there is some possibility that this new style of question is negatively affecting the quality of the bedrooms data.

The issue of respondents marking the bedroom box instead of giving the number suggests that some minor changes to the questionnaire design might be needed. It might be worth considering moving the bedrooms response option to the top to help respondents understand that a number is required or seeing if there is another way of making it very clear that a number is required (put “count” in bold?). Making sure all responses are captured and addressing the issue of respondents putting responses in the wrong place would also help ensure the quality of this data.

## Numbers of other room types

Limited analysis has been done for other room types – kitchens, living rooms, dining rooms, conservatories, and studies.

Room type	Percent not stated		
	Online	Paper	Total
Kitchens	2.5	13.0	4.7
Lounges, living rooms, or family rooms	2.6	13.3	4.9
Dining rooms	11.2	25.3	14.2
Conservatories you can sit in	49.1	45.9	48.4
Studies, studios, hobby rooms	42.0	45.2	42.7

Comments:

*Non-response* – Non-response for dining rooms was high on online forms, and especially high on English forms. Non-response for kitchens and living rooms was high on paper forms. Non-response for conservatories and studies was extremely high.

The overall level of non-response for kitchens and living rooms is acceptably low. Along with bedrooms, low non-response for these room types is important for being able to derive the total number of rooms. Non-response for these room types was 10 percent higher on bilingual forms than English forms.

Viewing images showed that non-response on bilingual forms was often due to responses being put in the wrong place. Some (apparent) non-response on English forms appears to be a scanning error (ie there was a response on the form in the correct place but it has not been captured during scanning). Many dwellings with non-response for these room types also had non-response for most, if not all, of the form.

As found in previous tests, it seems likely that many respondents are leaving boxes blank if they do not have any rooms of that type rather than writing '0' and that this is why non-response for dining rooms is high, and non-response for conservatories and studies is extremely high. (It is quite common for small units not to have a dining room/separate dining area.)

*Frequency of responses* – Viewing paper form images for dwellings with high numbers of certain room types showed several causes:

- misrecognition of the intended answer (an answer of '1' for kitchen with a scribbled out '1' beside it was recognised as 14)
- a scanning/data capture error (respondent clearly wrote '1' for kitchen, but it was scanned/captured as 12)
- the dwelling was really non-private not private
- facetious response. (The dwelling with 25 of each room type appears to be a facetious response as occupant names included "Government Spy" and "Mickey Mouse".)










These scanning errors occurred on bilingual and English forms.

It may be worthwhile investigating how responses of 10 or 20 could be accidentally entered on the online form.

As for bedrooms, a few respondents marked boxes instead of writing the number. A mark for a room type could indicate zero rooms of that type or one room, depending on whether the respondent used "0" or "1" for other room types. For this test, these responses were coded to "-", which is not a proper category. For census proper these responses will need to be coded to response unidentifiable. These respondents seem to be saying yes, I have a kitchen, yes I have a lounge, and it isn't until they get to bedrooms that they realise a number is required.

## Number of rooms

Number of rooms\* by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
One room		0.1		0.0		0.0
Two rooms		0.0		0.1		0.0
Three rooms		1.6		4.1		2.0

Four rooms		4.3	10.2	5.4
Five rooms		12.9	20.5	14.3
Six rooms		29.1	29.4	29.2
Seven rooms		24.1	17.5	22.8
Eight rooms		14.2	8.6	13.1
Nine rooms		7.4	5.1	6.9
Ten rooms		3.4	3.0	3.3
Eleven rooms		1.6	0.8	1.4
Twelve rooms		0.5	0.5	0.5
Thirteen rooms		0.2	0.0	0.2
Fourteen rooms		0.1	0.1	0.1
Fifteen rooms		0.1	0.0	0.1
Sixteen rooms		0.2	0.0	0.2
Seventeen rooms		0.0	0.0	0.0
Eighteen rooms		0.0	0.0	0.0
Nineteen rooms		0.0	0.1	0.0
Twenty or more rooms		0.3	0.2	0.3
Total stated				
Response unidentifiable		0.0	1.4	0.3
Not stated		3.9	17.5	6.8
Total				

\*using version A of the derivation

#### Comments:

*Non-response* – The overall not stated percentage (which means that insufficient information was provided to determine the number of rooms) is higher than desirable, and the percentage not stated for paper forms is much higher than desirable. Not stated for online forms was acceptably low.

Viewing some images for paper forms showed that some of this was true non-response, or insufficient response, but respondents marking boxes instead of giving a number, putting their response in the wrong place, or responses not being captured have also contributed to this.

*Response unidentifiable* – Very few responses were coded to response unidentifiable, which is good for data quality.

*Frequency of responses* – The distribution mostly looks sensible, with six as the most common number of rooms (eg three bedrooms, kitchen, lounge, dining), followed by seven (which could mean an additional bedroom, study, or conservatory), and almost none with only one or two rooms.

*Assessment* – The question used to derive this data mostly appears to be working acceptably well and suitable for inclusion in this format in the 2018 Census, but the issue of respondents marking boxes rather than giving a number needs to be addressed. This may require some changes to questionnaire design. The issues of responses not being captured or being put in the wrong place also need to be addressed.

Note: The results here use version A of the derivation. These will need to be compared with version B and a decision will need to be made about which version to use. The percentage of dwellings for which the number of rooms could not be determined may be different using the other version of this derivation.



For census proper it is recommended that edits or evaluation checks be used to identify dwellings treated as private that are really non-private and private dwellings where the responses for room types relate to a non-private dwelling. Criteria for such an edit could include a dwelling description response indicating a non-private dwelling (eg "hotel"), high number of bedrooms, a high number of kitchens (more than two), a high number of living rooms (more than three), or high number of dining rooms (more than two), and relationship to reference person responses indicating things such as "client", "service user".

## Main types of heating

**11** Mark as many spaces as you need to show which types of heating are used most often in this dwelling.

- heat pump
- electric heater (including bar, panel, oil-filled or fan)
- fixed gas heater
- portable gas heater
- wood burner
- pellet fire
- coal burner
- other. Print type of heating used:

or  don't use any form of heating

Subject population: private dwellings, but the following analysis has been done for all dwellings in the dwelling dataset. Generally these are private but a few previously unidentified non-private dwellings have been included in this dataset.

Main types of heating used by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
No heating used		1.3		2.3		1.5
Heat pump		30.1		21.9		28.3
Electric heater		43.2		41.7		42.9
Fixed gas heater		35.7		28.5		34.2
Portable gas heater		10.2		14.1		11.1
Wood burner		36.6		33.2		35.9
Pellet fire		0.4		0.3		0.4
Coal burner		0.4		0.5		0.4
Other types of heating		3.3		4.2		3.5
Total stated						

Residual categories for main types of heating used by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Response unidentifiable		0.1		0.3		0.1
Response outside scope		0.2		0.3		0.2
Not stated		2.3		4.8		2.8
Total						

Fuel types used to heat dwelling by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Electricity		62.6		57.5		61.5
Gas		44.0		39.8		43.1
Wood		36.9		33.3		36.1
Coal		0.4		0.5		0.4
Home heating oil		0.0		0.4		0.1
No fuels used in this dwelling		1.3		2.3		1.5
Other fuel(s)		1.4		3.0		1.7
Total stated						

Residual categories for fuel types used to heat dwelling by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Response unidentifiable		0.1		0.3		0.1
Response outside scope		0.2		0.3		0.2
Not stated		2.3		4.8		2.8
Total						

Number of heating types by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
No heating types used		1.3		2.3		1.5
One heating type used		48.6		59.1		50.8
Two heating types used		40.2		32.1		38.5
Three heating types used		8.8		5.9		8.2
Four heating types used		1.0		0.6		0.9
Five heating types used		0.1		0.0		0.0
Total stated						
Response unidentifiable		0.1		0.3		0.1
Not stated		2.3		4.8		2.9
Total						

Number of heating fuels used by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
No heating fuels used		1.3		2.3		1.5
One fuel		56.5		65.0		58.3
Two fuels		37.9		29.7		36.2
Three fuels		4.1		2.8		3.9
Four fuels		0.1		0.2		0.1
Total stated						
Response unidentifiable		0.1		0.3		0.1
Not stated		2.3		4.8		2.9
Total						

Comments:

*Non-response* – Overall non-response is low and well within the acceptable level. In contrast to some other variables, it was also within the acceptable level for those using paper forms. Possibly respondents tend to feel that this is a relevant and important question to answer.

*Response unidentifiable* – Inconsistent multiple responses (ie those indicating use of heating and no use of heating) are coded to response unidentifiable. There were very few (eleven) inconsistent multiple responses, which is good for data quality.

*Response outside scope* –there were very few responses coded to “response outside scope”. These are responses that are not heating appliances eg ventilation systems, blankets, clothes, hot water bottle.

*Frequency of responses* – The distribution for main types of heating used looks sensible. This data shows that the most popular type of heating appliance in Wanganui is an electric heater. Wood burners are second most popular; fixed gas heaters third most popular. This differs somewhat from what would be expected nationally, but probably just reflects the characteristics of the area (relatively cold, older housing stock, lower income levels) rather than a data quality problem. Nationally, use of heat pumps is expected to be very common.

This data shows greater use of fixed gas heaters than portable gas heaters. This is consistent with expectations. Previous census data has showed decreasing use of portable gas heaters. Publicity on the risks and issues associated with these heaters is likely to be contributing to this decrease.

Responses of pellet fire and coal burner were very rare, and less common than written responses relating to central heating systems or open fireplaces.

As for the 2013 Census, the fuels data for this test shows that electricity was the most common fuel type. There appear to be very few cases where the fuel type could not be determined.

The percent of those who indicated that they did not use heating (1.5 percent of those who responded) is lower than the national result in the 2013 Census. This is probably to be expected, and is likely related to several factors: the nature of the sample for this test, the relative degree of coldness of this region, a predominance of housing styles more likely to require heating (eg separate houses versus joined units or small apartments), and lower fuel poverty in this area. (In 2013 the percentage not using heating was highest in Auckland and Northland.)

The distributions for number of heating types and number of heating fuels used look sensible. As expected, there were many multiple responses, with 47.7 percent of those who stated an answer using more than one type of heating (appliance). Comparing the distributions for number of heating types (appliances) and number of heating fuels shows the expected pattern – using two types of heating appliances is a little more common than using two types of heating fuels (eg because using a heat pump and an electric heater counts as two heating appliances but only one fuel type).

The numbers using many different types of heating appliances or fuels were low. Possibly at least some of these respondents may have incorrectly listed all the different types of heating they use rather than just those used most often (the main ones). As the numbers are low, this does not appear to be negatively affecting data quality and is not of great concern.

*Written responses* – There were ■■■ written responses, which is 5.3 percent of those who responded. This is a reasonably low level of written response, but if desired (eg to make processing easier), it could be reduced further by adding additional response categories on the form or changing the response categories used.

The most common types of written responses were gas central heating and open fireplace. There were also many other written responses that indicated use of a central heating system. A few written responses were unnecessary (eg electric fan heater) or were not a heating appliance (eg blankets).

*Assessment* – This variable is suitable for inclusion in the 2018 Census. Respondents seem to be very receptive to it. The question used in this test is performing well both on the online and paper forms and is suitable for inclusion in this format in the 2018 Census.

It appears that it will still be possible to produce data on the heating fuels used, using this question. ‘As-you-type’ lists for the online form should specify the fuel used to help ensure that good quality data on fuels can be produced.

These test results suggest that it may be useful to consider reviewing the response categories used on the form and the classification categories for main types of heating. Regional variations will need to be taken into account.

## Housing quality

### 14 Is this dwelling damp?

A damp dwelling may feel or smell damp, or have damp patches on the walls, ceiling, floor or window frames.

- yes – always  
 yes – sometimes  
 no  
or  don't know

### 15 Can you see mould in any part of this dwelling that, in total, is larger than an A4 sheet of paper?

- Mould (mildew) may grow on the walls, ceiling, floor, doors, window frames, curtains or blinds.
- Mould can be black, white, green, brown, red, etc.
- An A4 sheet is the size of 1 page of this 4 page form.

- yes – always  
 yes – sometimes  
 no  
or  don't know

### 16 Which of these things are available here in this dwelling?

**Don't** count anything that is disconnected or broken.

- cooking facilities  
 tap water that is safe to drink  
 kitchen sink  
 fridge  
 bath or shower  
 toilet  
 electricity supply  
or  none of these are available in this dwelling

## Dwelling mould indicator

Subject population: private dwellings, but this analysis includes all dwellings in the dwelling dataset for the test. Generally these are private, but it may include a few that are really non-private and were not identified as such. This should make no or little difference to the results.

Dwelling mould indicator by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Mould over A4 size – always		2.5		3.4		2.7
Mould over A4 size – sometimes		9.4		10.3		9.6
No mould/mould smaller than A4 size		87.1		85.1		86.7
Don't know		0.9		1.2		1.0
Total stated						
Response unidentifiable						
Not stated		2.4		7.3		3.5
Total						

### Comments:

*Non-response* – Overall non-response is acceptably low at 3.5 percent, but the level of non-response for paper forms is of some concern. Checking a sample of images for paper forms shows that respondents putting responses in the wrong place has contributed to this, and that solving this problem would reduce the amount of data coded to not stated.

Cross-tabulating tenure of household by dwelling mould indicator shows that only 1.7 percent of those who rented did not answer the mould question. Low non-response for renters is good, as housing quality for renters is of particular interest. Non-response to the mould question was also acceptably low for all other household tenures, with the exception of “dwelling not owned and not held in a family trust, rental arrangements not further defined”. The high figure for this category is not of concern as there were very few households in it so it is based on very small numbers.

In most cases where the dwelling mould indicator was not stated, tenure of household was also not stated. There does not appear to be a particular issue with answering the mould question.

These are very good results. Results in census proper might not be quite as good as for this voluntary test, but based on these results, it looks like it will be possible to use census data to analyse mould amounts by household tenure.

Tenure of household by dwelling mould indicator, all modes

	Mould over A4 size – always	Mould over A4 size – sometimes	No mould/mould smaller than A4 size	Don't know	Not stated	Total	Percentage not stated for mould
Dwelling owned or partly owned, mortgage arrangements not further defined							3.0
Dwelling owned or partly owned, mortgage payments made							0.5
Dwelling owned or partly owned, mortgage payments not made							0.5
<b>Total owned</b>							<b>0.6</b>
Dwelling not owned and not held in a family trust, rental arrangements not further defined							17.6
Dwelling not owned and not held in a family trust, rent payments made							1.7
Dwelling not owned and not held in a family trust, rent payments not made							2.5
<b>Total not owned</b>							<b>1.9</b>
Dwelling held in a family trust, mortgage arrangements not further defined							2.5
Dwelling held in a family trust, mortgage payments made							0.6
Dwelling held in a family trust, mortgage payments not made							0.4
<b>Total in family trust</b>							<b>0.6</b>
Not stated							
<b>Total</b>							

*Frequency of responses* – The data looks sensible overall and shows the expected patterns.

Most respondents indicated that no mould was present or the amount present was smaller than A4 size. As expected, responses that it was present sometimes were more common than responses indicating that it was always present and the percentages for those using paper forms were a little higher than for those using online forms. The total percentage indicating that mould over A4 size was always or sometimes present (12.3 percent) fits with expectations based on results from the New Zealand General Social Survey that around 10 percent of dwellings would have a total mould amount over A4 size.

The low percentage of don't know responses is good from a data quality point of view.

Cross-tabulation with tenure of household shows that, as expected, those who rented their home were the most likely to have a total mould amount over A4 size (13.1 percent for online, 11.4 percent for paper). This was true for forms filled in online and those filled in on paper. Home owners who made mortgage payments were more likely to have a total mould amount over A4 size than those who did not make mortgage payments.

Tenure of household by dwelling mould indicator and percentage with mould over A4 size, online forms

	Online								
	Mould over A4 size – always	Mould over A4 size – sometime	Total mould over A4 size	No mould/mould smaller than A4 size	Don't know	Total stated	Percentage with mould over A4 size	Not stated	Total online
Dwelling owned or partly owned, mortgage arrangements not further									
Dwelling owned or partly owned, mortgage payments made									
Dwelling owned or partly owned, mortgage payments not made									
<b>Total owned</b>									
Dwelling not owned and not held in a family trust, rental arrangements not									
Dwelling not owned and not held in a family trust, rent payments made									
Dwelling not owned and not held in a family trust, rent payments not made									
<b>Total not owned</b>									
Dwelling held in a family trust, mortgage arrangements not further defined									
Dwelling held in a family trust, mortgage payments made									
Dwelling held in a family trust, mortgage payments not made									
<b>Total dwelling held in a family trust</b>									
Not stated									
<b>Total</b>									



Tenure of household by dwelling mould indicator and percentage with mould over A4 size, paper forms

	Paper								
	Mould over A4 size – always	Mould over A4 size – sometime	Total mould over A4 size	No mould/mould smaller than A4 size	Don't know	Total stated	Percentage with mould over A4 size	Not stated	Total paper
Dwelling owned or partly owned, mortgage arrangements not further									
Dwelling owned or partly owned, mortgage payments made									
Dwelling owned or partly owned, mortgage payments not made									
<b>Total owned</b>									
Dwelling not owned and not held in a family trust, rental arrangements not									
Dwelling not owned and not held in a family trust, rent payments made									
Dwelling not owned and not held in a family trust, rent payments not made									
<b>Total not owned</b>									
Dwelling held in a family trust, mortgage arrangements not further defined									
Dwelling held in a family trust, mortgage payments made									
Dwelling held in a family trust, mortgage payments not made									
<b>Total dwelling held in a family trust</b>									
Not stated									
Total									

Cross-tabulating dwelling mould indicator by sector of landlord does not show any issues with answering the mould question for certain landlord sectors. For those who indicated they rented their home, non-response to the mould question was acceptably low for those in private rentals and in social housing. (It looks very high for the 'iwi, hapu or Maori land trust' category, but this percentage is based on very small numbers and so not of concern.)

Dwelling mould indicator by sector of landlord, for those who rented their home

	Private person, trust, or business	Local authority or city council	Housing New Zealand Corporation	Iwi, hapū, or Māori land trust	Other community housing provider	Other state-owned corporation or state-owned enterprise, or government department or ministry	Don't know	Not stated	Total
Mould over A4 size – always									
Mould over A4 size – sometimes									
No mould/mould smaller than A4 size									
Don't know									
Not stated									
Total									
Percent not stated for mould									

*Assessment* – The test data indicates that this question is working acceptably well and is suitable for inclusion in this format in the 2018 Census. There do not appear to be any issues with respondents of particular household tenures or landlord types being unwilling to answer.

Putting additional help information in the question is worth considering. There has been some discussion that one aspect that may not be clear to respondents is how to answer if there are various patches of mould (each smaller than A4 size) in different places inside the dwelling. This could affect many respondents. Putting an instruction about this into the question notes (paper form) /static help (online form) may be useful.

## Dwelling dampness indicator

Subject population: private dwellings, but the analysis below includes all dwellings in the dwelling dataset for the test. Generally these are private dwellings, but it could include a few unidentified non-private dwellings.

Dwelling dampness indicator by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Always damp		1.3		2.2		1.5
Sometimes damp		15.6		16.1		15.7
Not damp		81.4		79.8		81.1
Don't know		1.7		1.9		1.7
Total stated						
Response unidentifiable		0		0		0
Not stated		2.4		6.8		3.3
Total						

### Comments:

*Non-response* – Overall non-response is acceptably low. As for mould, non-response on paper forms was a bit higher than desirable. Checking images for some of these forms showed that putting responses in the wrong place (usually on bilingual forms) has contributed to this non-response.

Breaking the non-response down by tenure of household does not indicate any areas of concern. Non-response was quite low for all tenure categories except “dwelling not owned, not in a family trust and rental arrangements not further defined” which has a very low number of dwellings, and so is not of concern.

Tenure of household by dwelling dampness indicator, all modes

	Always damp	Sometimes damp	Always or sometimes damp	Not damp	Don't know	Total stated	Percentage always or sometimes damp	Not stated	Total	Percentage not stated for dampness
Dwelling owned or partly owned, mortgage arrangements not further defined										2.4
Dwelling owned or partly owned, mortgage payments made										0.4
Dwelling owned or partly owned, mortgage payments not made										0.4
Total owned										0.5
Dwelling not owned and not held in a family trust, rental arrangements not further defined										5.9
Dwelling not owned and not held in a family trust, rent payments made										1.7
Dwelling not owned and not held in a family trust, rent payments not made										1.3
Total not owned and not in family trust										1.8
Dwelling held in a family trust, mortgage arrangements not further defined										2.5
Dwelling held in a family trust, mortgage payments made										0.3
Dwelling held in a family trust, mortgage payments not made										0.4
Total dwelling held in family trust										0.5
Not stated										
Total										

Cross-tabulating the dampness indicator by sector of landlord does not show any issues with answering the dampness question for certain landlord sectors. For those who indicated that they rented their home, non-response to the dampness question was acceptably low for those in private rentals and in social housing. (It looks very high for the 'iwi, hapu or Maori land trust' sector of landlord category, but this percentage is based on very small numbers and so not of concern.)

Dwelling dampness indicator by sector of landlord, for those who rented their home

	Private person, trust, or business	Local authority or city council	Housing New Zealand Corporation	Iwi, hapū, or Māori land trust	Other community housing provider	Other state-owned corporation or state-owned enterprise, or government department or ministry	Don't know	Not stated	Total
Always damp									75
Sometimes damp									382
Not damp									955
Don't know									55
Not stated									26
Total									1493
Percent not stated for dampness									

*Frequency of responses* – The data looks sensible overall, and generally as expected with the majority reporting no dampness and dampness some of the time being more common than dampness all the time. The total percentage indicating dampness (always or sometimes) is higher than expected (around 10 percent based on NZGSS results) at 17.2 percent. This may reflect characteristics of the sample for this test or housing quality in this area. Levels may vary by region.

The percentage indicating that they didn't know is very low, which is good for data quality.

Cross-tabulating the dwelling dampness indicator with tenure of household shows that, as expected, rented dwellings were the most likely to be damp, at 23.8 percent. This is lower than recent findings from the Building Research Association (Branz) that almost a third of rental houses felt damp. This may reflect differences in the measure/question used and in the samples, including the voluntary nature of this test. The results from this test for owned dwellings (13.2 percent) are fairly similar to those for this Branz research (11 percent).

As for mould, dampness was more common in owned dwellings for which mortgage payments were being made than those for which no mortgage payments were made. This is probably to be expected. Those who have paid off a mortgage are likely to have greater financial resources to fix housing quality problems (eg by installing ventilation systems, double glazing etc).

*Assessment* – The test data indicates that this variable is suitable for inclusion in the 2018 Census. The question used is working acceptably well and can be included in this format for the 2018 Census.

## Access to basic amenities

Subject population: private dwellings, but this analysis includes all dwellings in the dwelling dataset. This should make no, or very little, difference to the results.

Access to basic amenities by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Cooking facilities		99.6		98.9		99.4
Tap water that is safe to drink		98.0		97.0		97.8
Kitchen sink		99.4		99.5		99.4
Refrigerator		98.4		96.8		98.1
Bath or shower		99.3		99.5		99.3
Toilet		99.5		99.6		99.5
Electricity supply		99.4		99.0		99.3
None of these		0.1		0.3		0.2
Total stated						
Response unidentifiable		0		0		0
Not stated		2.2		6.7		3.2
Total dwellings						

Number of basic amenities by mode

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
No basic amenities available		0.1		0.3		0.2
One amenity		0.1		0.1		0.1
Two amenities		0.0		0.1		0.0
Three amenities		0.0		0.0		0.0
Four amenities		0.1		0.3		0.1
Five amenities		0.4		0.7		0.5
Six amenities		4.3		5.4		4.5
Seven amenities		95.0		93.2		94.7
Total stated						
Not stated		2.3		6.8		3.3
Total						

Comments:

*Non-response* – Overall non-response is acceptably low. As for many other variables, non-response for paper forms was a bit higher than desirable but has been balanced out by low non-response for online forms and greater use of online forms. Checking a small sample of images for paper forms showed that (apparent) non-response was often due to respondents putting responses in the wrong place (all those seen happened to be bilingual forms in this instance). Of the images viewed, there were a few cases of true non-response to this particular question, and a few cases of non-response to the entire dwelling form.

*Response unidentifiable* – No responses were coded to response unidentifiable for this test.

The amenities variable in the final ('cleanish') dataset indicates two cases of apparent inconsistent response (ie access to amenities and no access to amenities). Viewing the images showed that in both cases the intended response was clear and the data would not need to be coded to response unidentifiable. One respondent ticked every amenity and put a dash in the response oval for none. The other respondent marked every amenity and wrote '0' in the response oval for none. For census proper, it is expected that an edit would be in place to deal with these types of responses.

*Frequency of responses* – The data looks sensible and as expected. The vast majority of dwellings have all amenities listed in the question.

The amenities most likely to be missing were tap water that is safe to drink, followed by a refrigerator. As expected, the number of dwellings with none of the specified amenities was very, very low and the number with only one or a few was also very low. For dwellings that lacked amenities, the most common situation was to lack one of them – 4.5 percent of dwellings in this test lacked one of these basic amenities. Dwellings that lacked one or two amenities were most likely to be rented.

Viewing images of a sample of paper forms where the data indicates one or more missing amenities showed that this was the intended response in almost every case. (There was one possible exception involving a pen dot.) They appear to be genuinely lacking the amenity/amenities indicated. Further checking showed that respondents in dwellings that lacked amenities such as a fridge tended to be renting, on low incomes, living alone, and often also lacked access to other things eg a telephone, cellphone, Internet, a motor vehicle.

However results from the follow-up survey indicated that some cases of dwellings appearing to lack amenities was due to respondent error ie the respondent missed marking a box/ boxes which they should have marked. It is unknown whether the level of respondent error for this question would be any higher than for any other question.

Responses that the dwelling lacked tap water that was safe to drink may have been partly due to the heavy rain and flooding that occurred in the area around the time of the test. There is evidence from the follow-up survey that there could be a more general issue with tap water quality in this area. A respondent to this survey commented that the quality of the tap water isn't great and they don't drink it without boiling it first.

*Assessment* – This variable appears to be suitable for inclusion in the 2018 Census. The question used generally appears to be working well and suitable for inclusion in this format. Given the element of respondent error identified in the follow-up survey, it is recommended that careful checking of this data be done during evaluation to try to identify any respondent errors. Caution may also be needed when reporting this data and drawing conclusions/taking actions based on it.

## Access to telecommunication systems

**12** Mark as many spaces as you need to show which of these are available here in this dwelling.

Don't include:

- anything that is disconnected or broken
- anything that can be used only for work.

a cellphone / mobile phone

a telephone

Internet access

or  none of these

*Subject population* – Households in private occupied dwellings, but the analysis below is for all dwellings in the dwelling dataset as there was no variable available to restrict the data to households. Generally these are private dwellings. This should make little, or no, difference to the results.

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Access to a cellphone		92.6		77.3		89.4
Access to a telephone		78.3		77.5		78.1
Access to the Internet		93.0		50.7		84.1
No access to telecommunication systems		0.4		2.1		0.8
Total stated						
Not stated		2.4		4.9		2.9
Total						

*Comments:*

*Non-response* – Non-response was acceptably low overall, at 2.9 percent, and acceptably low for both the paper and (especially for) the online mode. This non-response level is lower than in the previous three censuses.

Respondents generally seem happy to answer this question.

*Response unidentifiable* – No responses were coded to response unidentifiable for this test. However there was a small number of apparently inconsistent multiple responses of having access to one or more telecommunication system and also having no access to any telecommunication systems – seven or nine depending on how the number was produced. Data quality-wise, it is good that there is only a very small number of these types of responses. These were not investigated further for this test, but would need to be for census proper (and coded to response unidentifiable). It is important to see whether this was the intended, possibly mistaken, response by the respondent, or was due to the processing system not being able to distinguish unintended answers (eg crossings out, scribbles, big ticks from another question etc) from intended answers.

*Frequency of responses* – The data looks sensible overall. It shows an increase in access to the Internet, an increase in access to a cellphone, and a decrease in access to a (landline) telephone, all of which is consistent with the national trends in previous census data.

As expected, the percentage with no access to any of these telecommunication systems was very low. This percentage includes a few cases of apparent inconsistent multiple response, which could be partly due to respondent error and would be coded to response unidentifiable for census proper, so the true percentage may be lower still.

*Assessment* – This question appears to be working well, and – as previously – appears to be suitable for inclusion in the 2018 Census in this format. Non-response may be higher in census proper than in this voluntary test, however greater use of online forms looks promising

Note: this variable was coded incorrectly in the test ie not as per the classification. This needs to be fixed for census proper.





Note: for those who usually travel to work, the main means of travel to work data comes from question 46 on the individual form. For those who usually work at home, this data (ie the worked at home category) comes from question 45 (workplace address) on the individual form.

## Main means of travel to work

Subject population: employed usual residents aged 15 years and over

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Worked at home		12.3		14.3		12.6
Drove a private car, truck or van		65.0		63.4		64.8
Drove a company car, truck or van		11.8		10.5		11.7
Passenger in a car, truck, van or company bus		2.8		1.8		2.6
Public bus		0.3		0.8		0.3
Train		0.1		0.4		0.1
Bicycle		2.6		2.3		2.6
Walked or jogged		3.7		5.2		3.9
Other		1.5		1.3		1.5
Total stated						
Response unidentifiable		0.0		1.7		0.2
Not stated		0.9		1.7		1.0
Total						

Comments:

*Non-response:* Non-response to this question appears to be exceptionally low, even on paper.

*Response unidentifiable:* There were very few unidentifiable responses, which is good for data quality. These were multiple responses.

*Frequency of response:* The relative frequencies of the different modes look sensible. The data is dominated by the driving a private car, truck, or van category. Being a passenger in a car, truck, or van appears to be a relatively rare mode of travel to work in this area, maybe because there is no particular pressure to car-pool in Wanganui and parking is easy.

As expected, walking or jogging was more common than using a bicycle.

Very few people used a public bus. This probably reflects the relatively limited provision of bus services in Wanganui and relative ease of using private transport in this area. Most of the few respondents who indicated that they travel by train either worked or lived in Wellington.

The percentage who worked at home was higher than the national figure in 2013 (8.8 percent). This does not seem unreasonably high. The 2013 data is not completely comparable as it is based on a different concept (census day travel) and relates to whether people worked at home on that particular day. This higher figure may also reflect regional differences.

The number of 'other' responses is relatively low, suggesting we have got the mix of response boxes for certain modes and 'other' for any remaining modes about right. As there was no write-in box, it is not possible to know what these responses were.

*Assessment:* The new (changed) question appears to be working very well and is suitable for inclusion in 2018 in this format.

One aspect that may need discussion is whether the motorbike category should be re-instated.

## Workplace address (analysis of non-response only)

Subject population: employed usual residents aged 15 years and over who mostly worked away from home.

Non-response to workplace address by mode

	Number	Percent
<b>Paper form</b>		
Workplace building name	■	35.6
Workplace street or shopping centre	■	26.0
Workplace suburb	■	39.6
Workplace city	■	18.7
<i>Paper complete non-response to question</i>	■	11.1
<b>Total paper</b>	■	
<b>Online form</b>		
Workplace address	■	4.4
<i>Online complete non-response to question</i>	■	4.4
<b>Total Online</b>	■	
<i>Overall complete non-response</i>	■	5.4
<b>Total Overall</b>	■	

Comments: Overall non-response was fairly acceptable at 5.4 percent and non-response for the online form was within the acceptable range. Having a dropdown list may have contributed to this – 24 percent of online respondents used this list.

For the paper form, complete non-response to this question was higher than acceptable and non-response to each part of the question was very high, particularly for building name and suburb. For those working in central business districts, suburb may not seem relevant and it may generally still be possible to code the geographic location if the other parts of the question have been answered.

Viewing a small sample of forms did not show any evidence of this being due to respondents writing in the wrong place. It appears that they don't know some details of their workplace address, have privacy concerns, or don't see the relevance of the question.

*Assessment* – This limited analysis does not raise any new issues for this question, which has not been changed from 2013. It remains suitable for inclusion. Online completion may be particularly useful for ensuring the quality of this data.

## Travel to education

**20** What is the **one main way** you usually travel to your place of education – that is, the one you use for the greatest distance?

If you don't have a usual method, select the method you used most recently.

- study at home → go to **22**
- drive a car, truck or van
- passenger in a car, truck or van
- bicycle
- walk or jog
- school bus
- public bus
- train
- ferry
- other, eg taxi, motorbike

**21** Where are you attending, studying or enrolled?

*Give all of the following, if possible:*

name of pre-school, school or other place of education


campus and/or suburb


city, town or district


## Main means of travel to education

Subject population: New Zealand residents who participated in study

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
Study at home		10.4		9.8		10.3
Drive a car, truck or van		14.3		19.8		15.2
Passenger in a car, truck or van		44.2		31.8		42.1
Bicycle		4.7		2.4		4.3
Walk or jog		17.4		25.5		18.8
School bus		5.0		4.7		4.9
Public bus		1.4		2.4		1.6
Train		0.0		0.0		0.0
Ferry		0.0		0.0		0.0
Other		2.6		3.7		2.8
Total stated						
Response unidentifiable		0.0		5.9		1.1
Not stated		0.2		0.9		0.3
Total						

Comments:

*Non-response* – There was almost no non-response, even for those using paper forms. Possible explanations for this are that respondents (those in this test at least) are generally keen to answer this question and regard it as important.

*Unidentifiable responses* – For the paper forms, 5.9 percent of responses were coded to unidentifiable. These are cases of multiple response. This indicates that some respondents had difficulty deciding how to answer, misinterpreted the question, or did not fully read the question and did not realise that a single response was required. As most forms were filled in online, and the online form allows only one response, the level of unidentifiable responses is still acceptably low overall and has not negatively affected data quality.

Potentially we could introduce an edit for multiple response on paper forms, with the corrective action being to determine which of the modes given would have been used for the longest distance. For example, a few of the multiple responses given were drove a private car and waked or jogged in which case drove a private car would almost certainly be the one used for the longest distance and we could code the response to this. Likewise, there were a few responses of public bus and walk or jog which could potentially be coded to public bus.

However, many of the other multiple responses given are odd combinations. They appear to be all the different modes the respondent uses and to be modes used on different days rather than the modes used to complete a full journey. Examples:

- drive private car and passenger in private car
- drive private car and bicycle
- bicycle, walk or jog, and public bus
- drive private car, passenger in private car, public bus
- passenger in private car and school bus
- drive private car and school bus

- study at home, drive private car
- school bus and public bus

It is not clear which of these responses would be their usual, main way of getting to their place of education. So it may be best not to have an edit on multiple response in the processing system, leave multiple responses coded to response unidentifiable, monitor the amount of this during evaluation, and then determine whether any corrective action is appropriate.

*Frequency of responses* – The data looks sensible and as generally expected, with a wider spread of travel modes than for travel to work, and passenger in a car, truck or van being the most common mode. The percentage using a bus was relatively low, with public bus being less common than school bus, as expected. There were no answers of train or ferry which is as expected for this area. The percentage who walked or jogged is maybe higher than might be expected, but maybe this mode is more common (easier) in smaller urban areas such as Wanganui.

The amount of ‘other’ responses is quite low at 2.8 percent which suggests that the choice of separate response categories for certain modes versus ‘other’ for anything else is working reasonably well.

#### Main means of travel to education by age group

	Under five years		Five to 17 years		18 Years and over		Total
	Number	Percent	Number	Percent	Number	Percent	
Study at home	1	0.5	3	2.6	100	38.7	104
Drive a car, truck or van	10	12.8	10	8.0	100	39.5	120
Passenger in a car, truck or van	100	74.0	100	49.8	10	6.3	210
Bicycle	1	0.5	10	5.4	10	2.0	21
Walk or jog	10	9.6	100	23.4	10	6.7	120
School bus	1	1.8	10	6.7	1	0.3	12
Public bus	0	0.0	1	0.9	10	4.3	11
Other	1	0.9	10	3.2	10	2.1	21
<b>Total stated</b>	100		100		100		300
Response unidentifiable	1		1		1		3
Not stated	1		1		1		3
<b>Total</b>	102		102		102		306

When broken down by age, the data looks mostly as expected but reveals some impossible responses – driving for children aged under five years, or five to 15 years. This appears to be respondent error. Checking images showed that drive was the response given. The suspected cause is that parents filling out forms for children and teenagers answered from their point of view rather than their child’s, and the real answer is passenger in a car, truck or van. We may need to consider having an edit to correct such responses. Probably the correct response in this situation is passenger in a car, truck or van.

‘Other’ modes were most common for those aged five to 17 years. Possibly this is responses such as skateboards and scooters.

*Assessment* – In general this question appears to be working well and suitable for inclusion in this format for the 2018 Census. Most respondents appear to be answering correctly and it appears to be working well for all age groups.

Edits for this question need to be developed. We may want to consider whether a motorbike category is needed.

## Educational institution location

Subject population: usual residents who participated in study and travelled to it (ie those who studied at home are excluded).

*Non-response* – Non-response for educational institution name was well within the acceptable level (2.5 percent) and non-response for city was fairly acceptable (5.8 percent), but non-response for suburb was very high. (Note: currently in the test data, there are no not stated or response unidentifiable categories for these variables. These figures have been calculated by applying filters in the excel spreadsheet.)

Breaking this down by age showed that non-response for name and city is acceptably low for those under 18 years, but higher than desirable for respondents aged 18 years or over. Non-response for suburb was very high for all age groups. This may indicate that they don't know the suburb or don't see its relevance. If the name and city has been provided, but the suburb has not been provided, usually it should still be possible to determine the geographic location of the educational institution so this high non-response for suburb should not have a major negative effect on data quality.

Non-response for educational institution location by age group

	Under five years		5-17 years		18 years and over		Total	
	Number	Percent	Number	Percent	Number	Percent	Number	Percent
Educational institution name	0	0.0	■	0.9	■	11.9	■	2.5
Educational institution suburb	54	24.3	■	32.7	■	29.9	■	31.6
Educational institution city	7	3.2	■	4.3	■	14.4	■	5.8
<b>Total</b>	<b>222</b>		■		■		■	

Based on these results, it is expected that data quality for educational institution location will be high for pre-school and school-aged students, but may be lower than desirable for tertiary students.

*Quality of responses given* – The quality of the name information – when provided – appears to be very good in most cases, with specific, full names of institutions being given. Checking what is provided at those given for children under five years did not show any data quality issues. The responses given seem to be places that provide early childhood education rather than only childcare, even if they have 'daycare' or 'childcare' in their name.

The quality of the suburb information, when provided, generally looks fine. Some respondents appear to have given the street name instead of the suburb in the boxes intended for the suburb. This is probably not a big issue as it provides more specific information than required and helps with determining the location, but it might complicate automatic coding processes.

The quality of the information given on the city, when provided, generally looks good. Occasionally respondents gave a more detailed response in the boxes intended for the city response (eg suburb and city) or a more vague response (eg region).

*Assessment* – This variable is suitable for inclusion in the 2018 Census in the current format. Although many respondents do not answer it fully, it is expected that the information they do provide will still be sufficient to produce good quality data. It appears that this information can be successfully collected for those under five years as well as for school and tertiary students. There were doubts earlier about whether good quality information on this could be produced for pre-schoolers, but these results suggest that data quality issues are more likely to affect the data for tertiary students.

For the online form, it is recommended that the questionnaire design or presentation of this question be reviewed to see if it can be enhanced to improve the quality of responding from tertiary students. To achieve the best data quality possible, automatic coding processes may need to allow for certain pieces of information being in a different place than intended.

## Number of motor vehicles

**13** How many motor vehicles (not counting motorbikes) do the people who live here have available for their use?

**Don't include:**

- vehicles that belong to visitors
- vehicles that this household borrows occasionally from another household
- vehicles that can be used only for work or only on the farm
- motorbikes or scooters.




























none

1

2

or  number of motor vehicles

Subject population: households in occupied private dwellings, but the analysis below is for all dwellings in the dwelling dataset because the variable needed to restrict the data to households only was not available. Generally the dwellings in this dataset are private. It should not make much difference to the results.

	Online		Paper		Total	
	Number	Percent	Number	Percent	Number	Percent
No motor vehicle		5.1		15.7		7.4
One motor vehicle		41.0		51.3		43.2
Two motor vehicles		39.4		24.3		36.2
Three motor vehicles		9.8		6.3		9.1
Four motor vehicles		3.2		1.6		2.9
Five or more motor vehicles		1.4		0.9		1.3
<b>Total Stated</b>						
Not stated		2.3		5.0		2.9
<b>Total</b>						

Comments:

*Non-response* – The overall level of non-response is acceptable, and better than in the 2013 Census. This may be partly due to the compliant nature of respondents in this voluntary test, but greater use of online forms could also be contributing to it.

*Frequency of responses* – The overall distribution looks sensible. The number with no motor vehicle shows a small decrease compared with national 2013 Census data, which is consistent with the previous trend for this category. The number with three or more motor vehicles is lower than for the national 2013 Census data. This may reflect the skewed nature of the sample for this test (older people, smaller households). As expected, those using the online form were more likely to have motor vehicle access than those using the paper form, and tended to have a higher number of motor vehicles than paper-form fillers.

Checking images for cases where the vehicle number was relatively or very high (five to eleven) did not show any evidence of facetious answers. In all cases checked, the response given had been recognised correctly ie there was no numeric misrecognition. In all but one case, these responses of high vehicle numbers seem believable – at least there was no evidence to suggest otherwise. The one case (11 vehicles) that did not seem believable was a retired person living alone in a rented council flat who indicated that they used the bus. In another case with a high vehicle number, it appears that the dwelling is really non-private, not private. (There was evidence that people with major disabilities were present and relationship to reference person responses of ‘clients’, so it is probably a non-private residential and community care dwelling). In census proper, evaluation checks on the occupied dwelling type data should see cases such as this correctly reclassified as non-private and thus removed from the data for variables that have households in occupied private dwellings as their subject population.

Checking some responses of three or four vehicles did not show any cases of numeric misrecognition or obvious respondent error. These responses were relatively common so errors in these categories could have a relatively significant effect on the data.

Whether people are including only those vehicles that should be included and not others (eg work vehicles) is unknown. Only cognitive testing could reveal this.

*Assessment* – This question appears to be working well and is suitable for inclusion in the 2018 Census in this format. The planned change to output more detailed information and make data on exact vehicle number available looks like it can go ahead. During evaluation it will be important to check cases where a high vehicle number has been given, and take corrective action where appropriate, to ensure that this data is of good quality.

## Work variables

### Work and Labour Force Status

Work and Labour Force Status (WKLFS) is derived from five questions on the individual form – job indicator, hours worked, seeking paid work, job search methods and availability for work.



**38** In the 7 days that ended Sunday 2 April, which of these did you do?

- I worked for pay, profit or income for an hour or more
- I worked in a family business or family farm **without pay** → go to **39**
- I work in a job, business or farm, but I was not working last week for some reason

or  none of these → go to **48**

**39** How many hours, to the nearest hour, do you usually work each week?

in your main job:    hours

in all other jobs for pay, profit or income:    hours

**48** Did you look for paid work in the last 4 weeks?

- yes
- no → go to **50**

**49** Mark as many spaces as you need to show all the ways you looked for paid work in the last 4 weeks.

- looked at job advertisements
- wrote, phoned or applied in person to an employer
- contacted Work and Income to look for a job
- contacted friends or relatives for help in finding a job
- contacted career advisers or vocational guidance officers
- other methods

**50** If a paid job had been available, would you have started last week?

- yes
- no

*Non-response categories* – not applicable as the data is imputed for the subject population of New Zealand adults (aged 15 and over).

*Frequency of responses* –

**Work and Labour Force Status, Whanganui TA residents aged 15 and over, 2013 and 2017 Test**

Work and Labour Force Status (% of total)	2013	2017
Employed full-time	40.5	47.4
Part- time	14.2	7.3
Unemployed	5.8	2.9
Not in the labour force	39.5	42.4
	100%	100%

**Work and Labour Force Status, Whanganui TA, by mode of completion, 2017 Census Test**

WKLFS	Online	%	Paper	%	Grand Total
Employed full-time	████	52.6	████	27.1	████
Employed part-time	████	6.1	████	12.0	████
Unemployed	████	3.0	████	2.4	████
Not in the labour force	████	38.2	████	58.6	████
Sum of stated	████	100	████	100	████
To be imputed	████	3.5	████	23.9	████
Grand Total	████		████		████

*Assessment* – 2017 test data is not comparable with the 2013 Census data. There was an issue with the capture of hours worked online, which may have impacted on the data. A work-around was used to populate the hours worked in “all other jobs”, namely that job hours 1 = job hours 2, which would have inflated full-time work status and vice versa for part-time. This must be fixed for the 2018 Census.

Analysis of the differences in online and paper responses reflects the higher proportions of older age people completing the test on paper, and a high level of imputed responses for paper forms due to high non-response on page 4 of the paper form. Overall, around 8% of WKLFS responses were to be imputed in the test, which was only slightly higher than in 2013 (Whanganui data).

**Job (Work) Indicator**

Q 38 on the Individual Form (see above)

*Non-response categories* – overall the non-response rate was acceptable but paper non-response was high. The work questions are all on the last page (page 4) of the questionnaire and paper non-response was relatively high for questions on this page. Unidentifiable responses were minimal.

*Frequency of responses* – Although there is no comparative data available for 2013 on Supercross, the data looks sensible with 51% of all respondents working for pay profit or income, and 45% not working at the time of the census. Differences between paper and online forms reflect the bias towards older respondents on paper.

**Job Indicator, Whanganui TA resident adults, 2017 Census Test**

Job Indicator	Online	Paper	Grand Total
Family business/farm w/out pay	1.9	3.3	2.2
Work, but not this week	1.6	1.5	1.6
None of the above	41.0	61.7	45.4
Stated	100%	100%	100%
Unidentifiable	1.8	1.7	1.8
Not Stated	0.7	16.3	4.5
Grand Total	████	████	████

Analysis by age data (below) also looks sensible with higher proportions of youth and retirement age people not in work. As noted above, paper respondents are biased towards the older age groups and those not in work. Not stated responses were highest among 65 years and over paper respondents, some of whom may have thought the question was not relevant to them.

**Job Indicator by Age Groups, Whanganui adult residents, 2017 Test**

<i>Online Responses</i>	15-24	25-39	40-64	65+	Total
Worked for pay, profit etc	41.6	73.1	76.7	20.3	55.6
Family business without pay	1.3	1.1	2.4	1.7	1.9
Usually work but not last week	1.6	1.7	2.2	0.6	1.6
None of the above	55.5	24.1	18.7	77.4	41.0
Unidentifiable	0.2	2.4	2.7	0.6	1.8
Not stated	1.0	0.5	0.4	1.1	0.7

<i>Paper Responses</i>	15-24	25-39	40-64	65+	Total
Worked for pay, profit etc	26.2	63.6	56.9	12.1	33.4
Family business without pay	2.2	3.8	4.0	3.0	3.3
Usually work but not last week	1.8	1.3	3.0	0.5	1.5
None of the above	69.8	31.4	36.2	84.4	61.7
Unidentifiable	1.3	3.8	3.4	0.3	1.7
Not stated	7.3	6.0	10.7	23.3	16.3

*Assessment* – in general this question appears to be working well and is suitable for inclusion in this format in the 2018 Census. The paper non-response rate is higher than acceptable.








## Hours Worked in Employment per Week

Q 39 on the Individual Form (see above)

*Issue affecting analysis* – there was an issue with the capture of hours worked online, which has impacted on the data. A work-around was used to populate the hours worked in “all other jobs”, namely that job hours 1 = job hours 2, which has probably inflated full-time work status and vice versa for part-time. This must be fixed for the 2018 Census.

*Non-response categories* – non-response to this question appears to be low, although the analysis was main job only due to data quality issues described above.

*Frequency of responses* – analysis of the hours worked in ‘main job’ correlate very well with the 2013 Census data for Whanganui adults who said they were working, or usually worked, as shown below.

<i>Hours Worked</i>	<i>2017 Census Test</i>		<i>2013 Census</i>
	Number	%	%
Less than 20 hours		14.9	15.6
20-40 hours		58.7	57.0
More than 40 hours a week		26.4	27.4
Total Stated		100	100
Unidentifiable/out of scope		0.6	
Not stated		1.5	
Grand Total			

## Status in Employment

- 41** In that job, which one of these were you?
- a paid employee
  - self-employed and **not** employing others
  - an employer of other person(s) in my own business
  - working in a family business or family farm **without pay**

*Issue affecting analysis* – no major issues found.

*Non-response categories* – non-response to this question was very low (1% of eligible respondents) and the routing appeared to work well on paper, with only a few non-working respondents answering the question.

*Frequency of responses* – the data looks sensible, with the majority of workers in paid employment as shown below.

### Employment Status – % of respondents who answered the work indicator question, 2017 Census Test

Employment Status	Online	Paper	Total
Paid employee	80.9	78.1	80.5
Employer	6.2	4.7	6.0
Self-employed	11.2	12.1	11.4
Family business or farm w/out pay	1.7	5.1	2.2

The data correlates well with the Whanganui TA data from the 2013 Census.

Employment Status	2013 Census	2017 Test
Paid Employee	81.7%	80.5%
Employer	6.3%	6.0%
Self-employed	9.8%	11.4%
Family business etc. without pay	2.2%	2.2%

*Assessment* –the question appears to be working well and is suitable for inclusion in this format in the 2018 Census.

## Occupation



**Detailed Occupation Codes (N=20 or over)**

**Managers**

Chief Exec or Managing Director	■
Retail Manager (General)	■
Corporate General Manager	■
Policy and Planning Manager	■
Sales and Marketing Manager	■
Mixed Crop and Livestock Farmer	■
Project Builder	■
Corporate Services Manager	■
Finance Manager	■
Production Manager (Forestry)	■
Child Care Centre Manager	■
Dairy Cattle Farmer	■
Sheep Farmer	■
Hotel or Motel Manager	■

**Professionals**

Secondary School Teacher	■
Primary School Teacher	■
Early Childhood Teacher	■
Registered Nurse (Aged Care)	■
Accountant (General)	■
Nurse Manager	■
Social Worker	■
Registered Nurse NEC	■
Solicitor	■
Management Consultant	■
Welfare Worker	■
General Practitioner	■
Mechanical Engineer	■
Financial Investment Advisor	■
Registered Nurse ( Medical Practice)	■
Minister of Religion	■

**Technician and Trades**

Electrician (General)	■
Chef	■
Motor Mechanic (General)	■
Painting Trades Worker	■
Plumber (General)	■
Hairdresser	■
Gardener (General)	■
Carpenter	■
Butcher/Smallgoods	■
Cook	■

**Community and Personal Service Workers**

Personal Care Assistant	■
Prison Officer	■
Community Worker	■



*Issue affecting analysis* – there was a large number of unidentifiable responses, but as this was a low priority variable, there was very little processing done for the 2017 Test due to time constraints.

*Residual categories* – as above, unidentifiable responses were high but this should not be the case for 2018.

*Frequency of responses* – it is difficult to analyse the data due to the high number of unidentifiable responses. The data below indicates that some categories may be easier to autocode than others – for example, education and training, whereas “manufacturing” counts were relatively low compared to the 2013 Census data for Whanganui.

**Industry, Whanganui TA respondents in employment, 2013 Census and 2017 Test**

Industry	2013 Census	2017 Census Test
Agriculture, Forestry, Fishing	6.5	5.2
Mining	0.1	0.0
Manufacturing	14.8	7.3
Electricity Gas Water and Waste	0.7	0.4
Construction	7.4	9.7
Wholesale Trade	2.8	1.9
Retail Trade	10.8	7.6
Accommodation and Food Services	5.1	4.4
Transport, Postal and Warehousing	2.9	3.0
Info Media and Telecommunications	1.0	0.9
Financial and Insurance Services	1.7	1.9
Retail, Hiring and Real Estate	1.8	3.2
Professional, Scientific and Technical	4.9	7.0
Admin and Support Services	2.5	4.1
Public Administration and Safety	6.7	7.5
Education and Training	9.2	13.9
Health Care and Social Assistance	14.8	15.0
Art and Recreation Services	1.5	2.1
Other Services	4.7	4.7
<b>Total Stated</b>	<b>100.0</b>	<b>100.0</b>
Other Residuals	0.2	n/a
Unidentifiable	0.1	35.6
Not Stated	2.4	3.2

**Workplace Address**

*Issue affecting analysis* – Physical workplace address was not analysed for this report.

*Residual categories* – within an acceptable range for this question.

*Frequency of responses* – 86% of respondents who were working at the time said that they worked away from home.



**45** In that job, did you mostly:

- work at home? → go to **51**
- work away from home?

*Print the full address of the place you mostly worked at. Include, if possible, all of these:*

Workplace Address- home or away	Online	Paper	Grand Total
Work at home	12.4	15.2	12.8
Work away from Home	87.6	82.8	86.9
Unidentifiable	0	1.8	0.3
Not Specified	0.9	8.3	2.0

## Job Search/Availability for Work

**48** Did you look for paid work in the last 4 weeks?

- yes
- no → go to **50**

**49** Mark as many spaces as you need to show all the ways you looked for paid work in the last 4 weeks.

- looked at job advertisements
- wrote, phoned or applied in person to an employer
- contacted Work and Income to look for a job
- contacted friends or relatives for help in finding a job
- contacted career advisers or vocational guidance officers
- other methods

**50** If a paid job had been available, would you have started last week?

- yes
- no

*Issue affecting analysis* – none known. ■■■ respondents answered the question on the paper form when they should have skipped it as they were in the workforce and ■■■ respondents skipped the job indicator question but answered Q48.

*Residual categories* – generally quite low, 3-6% across the 3 questions relating to job searching and availability for work.

*Frequency of responses* – Of those respondents who were eligible (not working in the past week) to answer the job search question, 91% were not looking for paid work. Of those respondents who were actively seeking work, the most common method used was looking at job advertisements.

Job Search (not in the workforce)

Looked for paid work	■	9.4
Did not look for paid work	■	90.6
Not stated	■	3.1
Grand Total	■	

Of those who looked for paid work in the past 4 weeks...

Method of Job Searching	Number	%
Looked at job ads	■	83.5
Wrote, phoned, applied in person	■	43.4
Contacted WINZ	■	29.2
Contacted friends, relatives for help	■	30.5
Contacted careers advisers	■	11.2
Other methods	■	24.7
Total looked for work last 4 weeks	■	100.0
No response	■	2.8

Of those who were not working in the last 7 days.....

Available for paid work last week

	Number	%
Yes	■	15.6
No	■	84.4
Not stated	■	5.6

Seeking work (in the past 4 weeks)

	Number	%
Yes	■	9.4
No	■	90.6
Not stated	■	3.1