

Adult Migrant English Program (AMEP) Impact Evaluation Project

Research paper B: Validating self-reported English proficiency levels in the Census (2022)

Executive Summary

- This Research Paper has validated the self-reported proficiency in spoken English measure from the 2011 Australian Census of Population and Housing (hereafter called the Census) against the results of standardised English tests undertaken by Adult Migrant English Program (AMEP) participants upon program entry, to document how closely they align.
- This was done so that the next Research Paper in the series; Determinants of AMEP Participation, could use validated Census English proficiency information to help identify groups of migrants recorded in the Census who would have been eligible to participate in AMEP, but did not.
- The Census-based self-reported measure of English proficiency asks only about spoken English, not about reading, writing, or comprehension which are measured in AMEP entry tests.
- In this paper, AMEP clients were matched to the 2011 Census to validate the self-reported measure of English proficiency from the 2011 Census against the results of standardised English tests used in the AMEP. Results showed the Census measure correlated very well to the measures used by the AMEP. These results suggest Census self-reported proficiency in spoken English is a good proxy for general English proficiency in the absence of a standardised English test.
- Around 24% of all persons on the 2011 Census were born overseas in a non-English Speaking Background (NESB) country. Around 1-in-5 of these NESB migrants self-reported that they spoke only English at home.
- Of those NESB migrants who did not speak only English at home, less than 1-in-5 reported that they spoke English “not well” or “not at all”.
- Proficiency in spoken English was related to age of arrival in Australia. Migrants who arrived in Australia at younger ages reported higher levels of proficiency in spoken English and this pattern was true for both males and females.
- Proficiency in spoken English was also related to length of stay in Australia. Migrants who had been resident in Australia for longer reported higher levels of proficiency in spoken English and this pattern was true for both males and females.
- Holders of Skilled Migration visas reported higher proficiency in spoken English than holders of Family visas or Humanitarian visas.
- A range of socio-economic outcomes were better for migrants who reported they spoke English “well” or “very well” on Census 2011. These include better labour market outcomes, higher income levels, lower rates of public housing tenancy, and being less likely to receive any form of income support payment.

Introduction

The Adult Migrant English Program (AMEP) Impact Evaluation Project helps the Australian Government to better understand the drivers of AMEP participation, and the broader impacts participation have on employment and welfare outcomes for migrants. It consists of topical papers that utilise the broad ranging government information held within the Australian Bureau of Statistics' (ABS) Multi-Agency Data Integration Project (MADIP).

The AMEP Impact Evaluation Project began as a collaborative research initiative between the Department of Education, Skills and Employment (DESE) and the Australian Research Council Centre of Excellence for Children and Families Over the Life Course (the Life Course Centre) in July 2019. Jurisdiction over the AMEP subsequently moved from DESE to the Department of Home Affairs (the Department), making the Department custodians of the AMEP data and the key stakeholder in the AMEP Impact Evaluation Project.

This research paper has been co-funded by the Australian Government in partnership with the Life Course Centre.

Aim of the paper

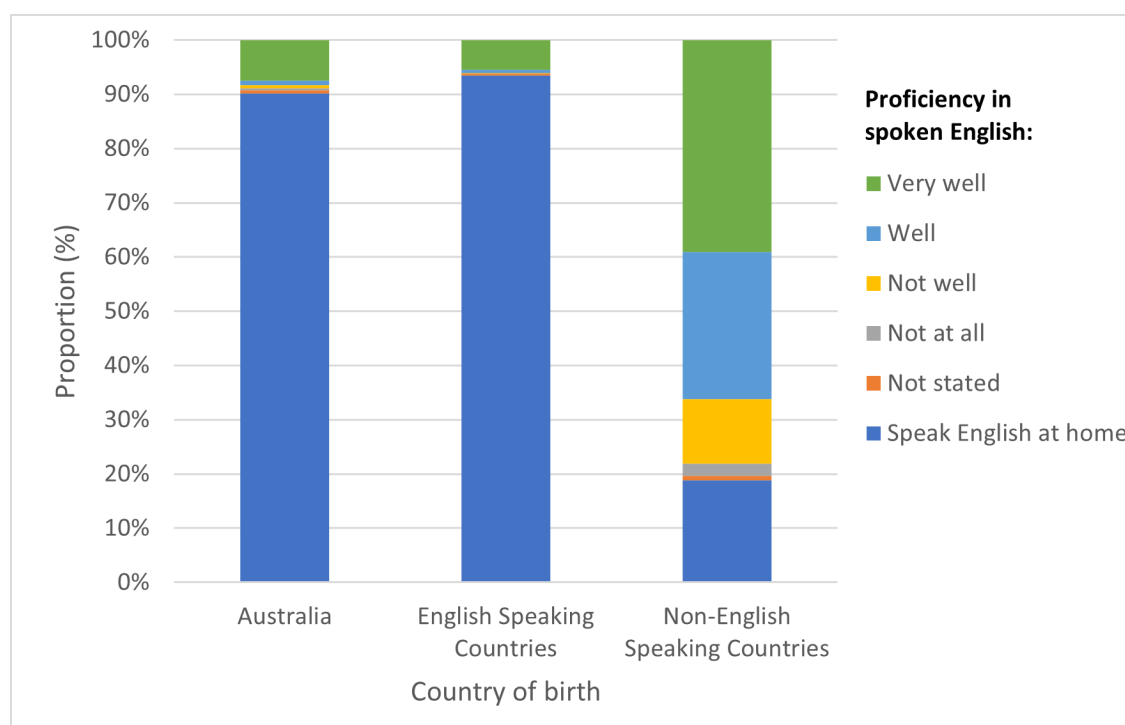
The aim of the research paper is to validate the English proficiency data published in the 2011 Census against AMEP English proficiency data in MADIP. This paper looks to validate the self-reported measure of English proficiency reported in the Census against the results of standardised English tests undertaken by AMEP clients to identify cohorts eligible for program entry.

Before now, it has not been possible to make accurate estimates from readily available Census data as to which migrants are, or would have been, eligible for AMEP, based on their self-reported English language proficiency. Knowing how well self-reported English language proficiency, as measured on the Census, maps to independent measures of English proficiency in AMEP client data provides the first important step toward better understanding the drivers of AMEP participation. It also offers a rare opportunity to assess the credibility of a self-reported measure of English proficiency in the Census that is widely used in research and policy planning across the private and public sectors. Newly linked Census-AMEP-Migration data in MADIP makes these validation exercises possible for the first time.

Results

We begin by describing the self-reported measure of English proficiency from the 2011 Census. In the Census, if the person speaks a language other than English at home, a question about English proficiency is asked: “How well does the person speak English?”, with four potential responses: “very well”, “well”, “not well”, and “not at all”. It is important to note that the Census asks only about spoken English, not about reading, writing, or comprehension.

Figure 1: Census 2011 – Proficiency in spoken English by country of birth for all persons



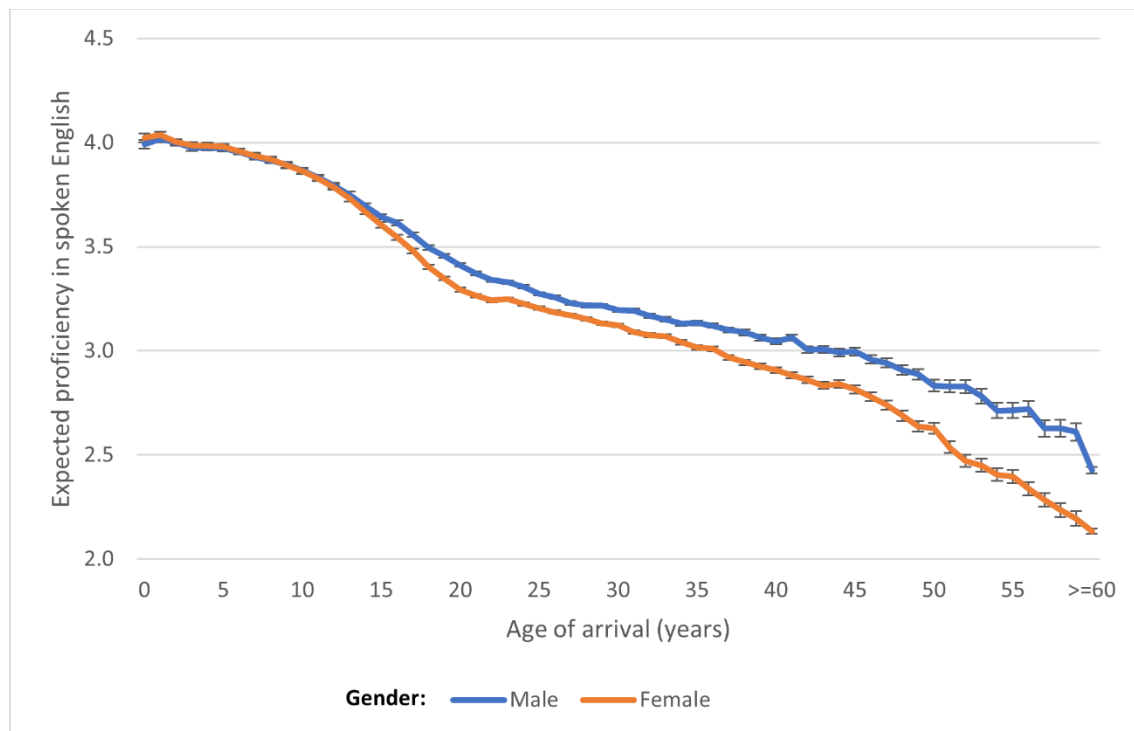
Notes: Sample of all 7.35 million persons available in the 2011 Census through MADIP, excluding overseas visitors and individuals with missing information on country of birth (about 1% of the total sample)

Around 24% of all persons on the 2011 Census were born overseas in a Non-English Speaking Background (NESB) country. Figure 1 shows that around 1-in-5 of these NESB migrants self-reported that they spoke English at home. Of those NESB migrants who did not speak English at home, 82% indicated that they spoke English “well” or “very well”. This means that around 18% of NESB migrants spoke English “not well” or “not at all” (includes around 1% “not-stated”). It should be noted that Figure 1 does not adjust for age of arrival, and many of these NESB migrants may have arrived as children. This is analysed in more detail below.

As expected, Figure 1 shows that 90% of Australian-born individuals, who represented 63% of all individuals in the 2011 Census data, spoke English at home and, among those who did not, 88% reported that they spoke English “well” or “very well”. Similarly, almost all (94%) migrants originating from English Speaking Background (ESB) countries (i.e., United Kingdom, Ireland, Canada, New Zealand, United States, and South Africa) spoke English at home, and 96% of those who did not speak English at home reported that they spoke English “well” or “very well”.

As detailed above, the Census question about proficiency in spoken English is not asked for individuals who spoke only English at home. As such, in what follows, we make the assumption that those who self-reported on the Census that they spoke only English at home, did so at the level of “very well”.

Figure 2: Census 2011 – Expected proficiency in spoken English by age of arrival in Australia for all migrants



Notes: Results (expected values (and 95% confidence intervals) of proficiency in spoken English) are obtained from an Ordinary Least Squares (OLS) regression of English proficiency on a constant, age (at the Census time, in categorical values), country of birth, gender, age of arrival (in categorical values) and interactions between age of arrival and gender. We assign 1 for “not at all”, 2 for “not well”, 3 for “well”, and 4 for “very well”. Sample of about 1.245 million migrants recorded in 2011 Census, excluding overseas visitors and individuals with invalid information. Age of arrival is the difference (in years) between year of arrival in Australia and year of birth. Year of arrival is identified from responses to Census question 13: “In what year did the person first arrive in Australia to live here for one year or more?”.

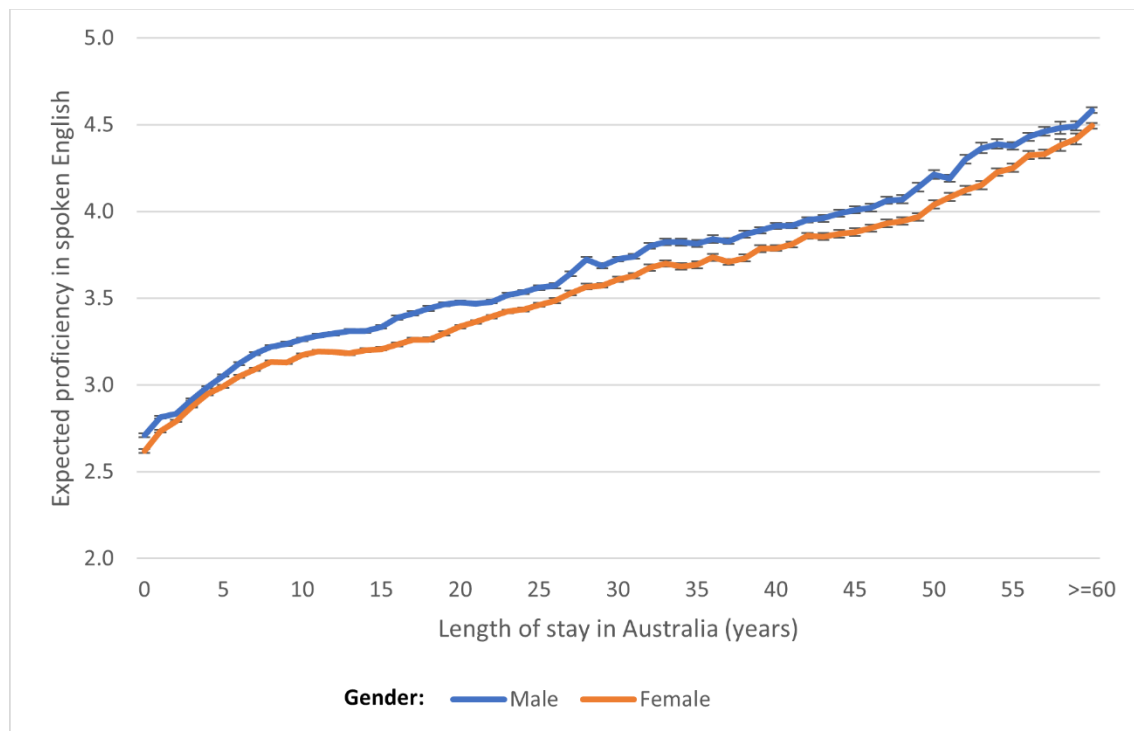
To validate the English proficiency information available in the 2011 Census, Figures 2–9 correlate self-reported proficiency in spoken English with some observable characteristics which have typically been found to be strongly associated with English skills.¹

Figure 2 presents expected levels of proficiency in spoken English by migrants’ age of arrival and suggests that migrants who arrived in Australia at earlier ages achieved higher levels of proficiency in spoken English and this pattern holds for both males and females. Figure 2 additionally suggests that while there is no noticeable gender difference in proficiency in spoken English at earlier ages, male migrants tend to report a higher level of proficiency in spoken English than female migrants from an age of arrival of 16 years old and upwards, and the gender gap widens as age of arrival rises.

Coincidentally, persons who arrive as adults (aged 18 years and above) are beyond standard school-age. We can see from Figure 2 that school-aged arrivals show both a higher proficiency in spoken English than is observed for later ages of arrival, and no obvious gender gap in spoken English proficiency.

¹ For a literature review, see: Ginsburgh, V., Weber, S., 2020. The Economics of Language. Journal Of Economic Literature 58, 348-404.

Figure 3: Census 2011 – Proficiency in spoken English by length of stay in Australia for all migrants



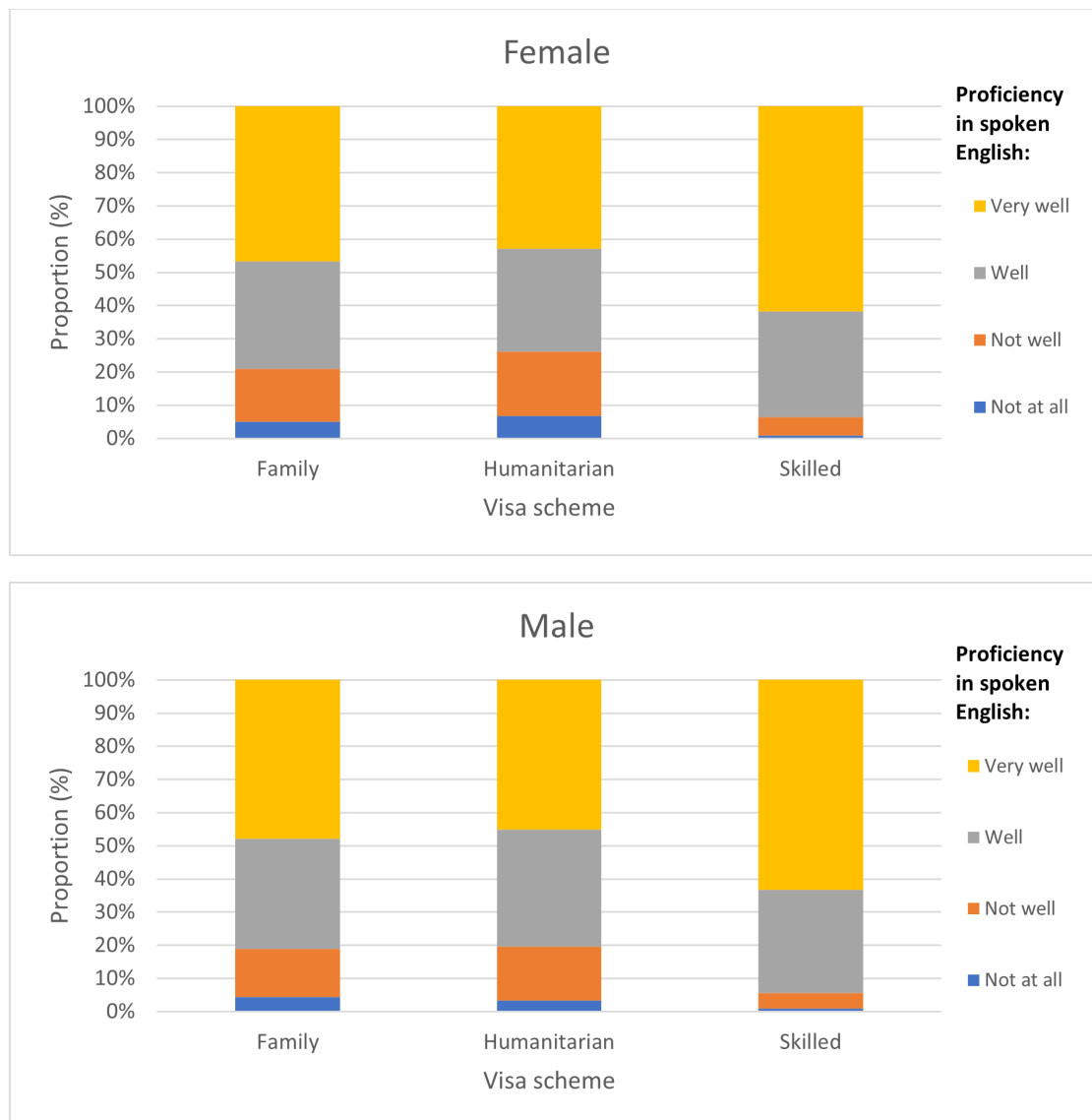
Notes: Results (expected values (and 95% confidence intervals) of proficiency in spoken English) are obtained from an Ordinary Least Squares (OLS) regression of English proficiency on an intercept, age (at the Census time, in categorical values), country of birth, gender, length of stay in Australia (in categorical values) and interactions between length of stay in Australia and gender. We assign 1 for “not at all”, 2 for “not well”, 3 for “well”, and 4 for “very well”. Sample of about 1.245 million migrants recorded in 2011 Census, excluding overseas visitors and individuals with invalid information. Length of stay is the difference (in years) between census year (i.e., 2011) and year of arrival in Australia.

Figure 3 shows that, on average, migrants who have lived in Australia for longer have higher proficiency in spoken English compared to new arrivals. While this is cross-sectional rather than longitudinal data, it does suggest that time spent in Australia is related to improvements in spoken English, and that the time trend for improvement might be almost linear. Another observation is the consistency of the gender gap across time, with proficiency in spoken English higher across all years for males compared with females.

Figure 3 also highlights the effect of spending time in a new country and being exposed to a new language. It must be noted that the sample contains both ESB and NESB migrants, and that the composition of ESB versus NESB migrants to Australia may have looked different over time.

Adult Migrant English Program (AMEP) Impact Evaluation Project

Figure 4: Linked Census – Migration. Proficiency in spoken English by visa program for permanent resident migrants

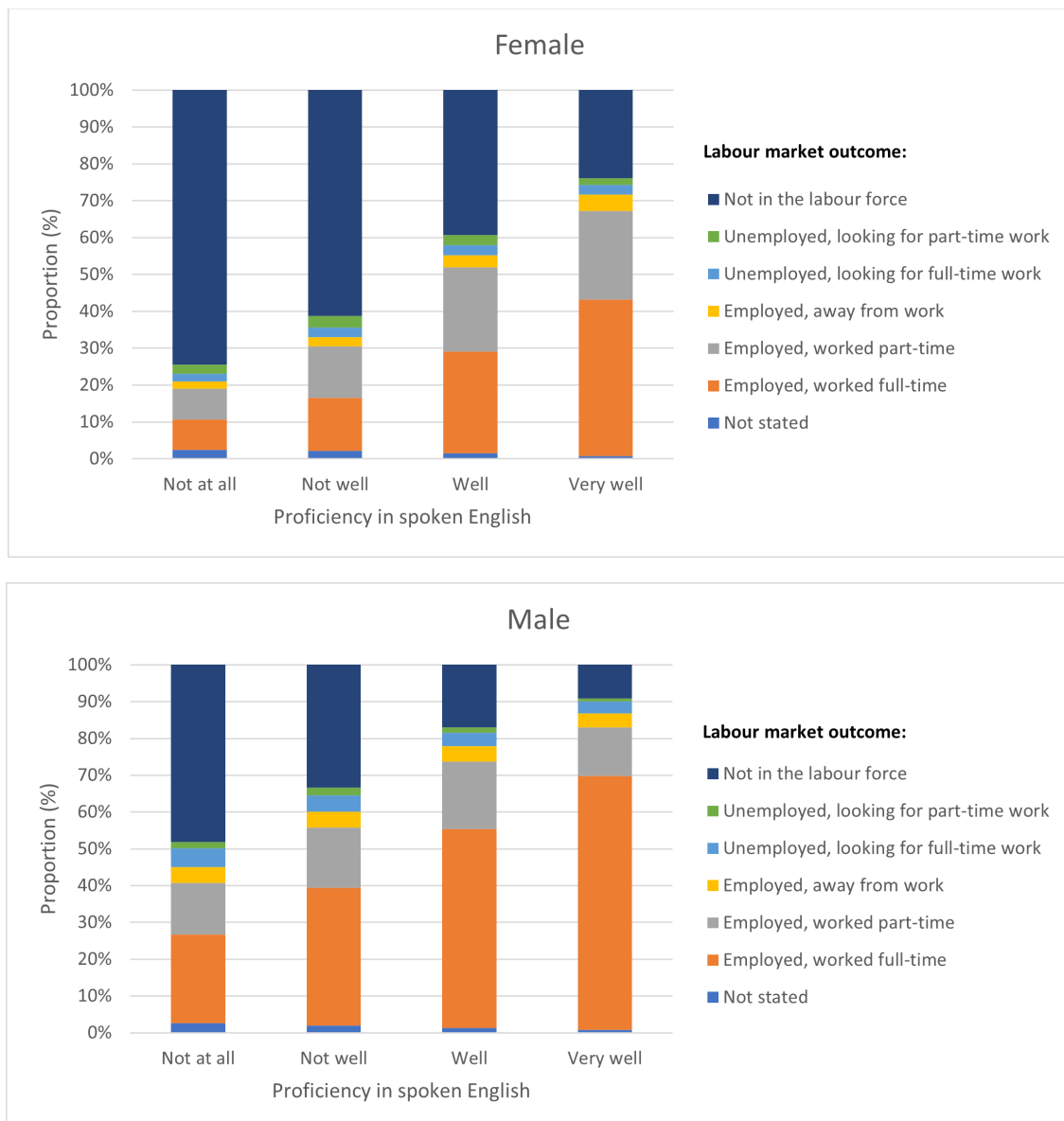


Notes: Sample of about 620,000 permanent resident migrants in merged Migration and 2011 Census database.

Figure 4 presents proficiency in spoken English by visa schemes, suggesting that, consistent with the design of Australian migration policies, Skilled visa holders are more proficient in spoken English than Family visa or Humanitarian visa holders. For Skilled visa holders of both genders, only a small proportion (under 5%) spoke English “not well” or “not at all”, compared with 20% or more for Family visa or Humanitarian visa holders.

Adult Migrant English Program (AMEP) Impact Evaluation Project

Figure 5: Census 2011 – Migrant labour market outcomes by proficiency in spoken English

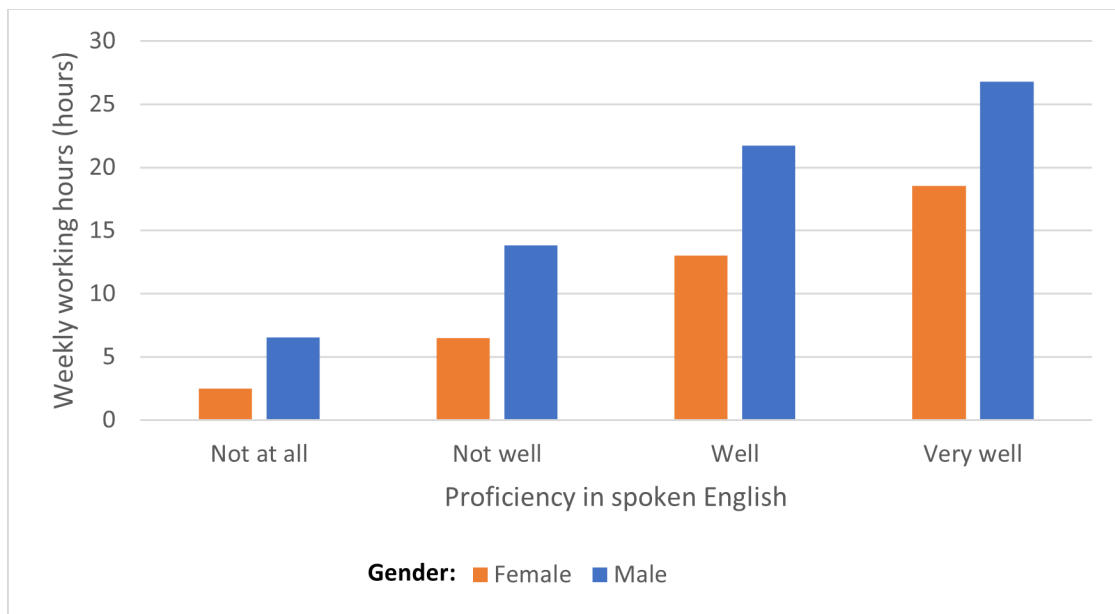


Notes: Sample of about 966,000 working-age (i.e., 24 to 64 (65) years old for female (male)) migrants recorded in 2011 Census data. Labour market outcomes are derived from responses to Census question 34 asking: “Last week, did the person have a full-time or part-time job of any kind?”

Figure 5 reports labour market outcomes by proficiency in spoken English. Working-age migrants with higher levels of English proficiency displayed higher rates of labour market participation compared with migrants with low proficiency in spoken English. In particular, 82% of male migrants who spoke English “very well” worked, either in part- or full-time employment. By contrast, only 38% of males who spoke English “not well” or “not at all” were working. The same pattern is observed for female migrants. Specifically, 66% of female migrants who spoke English “very well” worked while only 17% of individuals who spoke English “not well” or “not at all” worked.

Adult Migrant English Program (AMEP) Impact Evaluation Project

Figure 6: Census 2011 – Weekly working hours for migrants by proficiency in spoken English

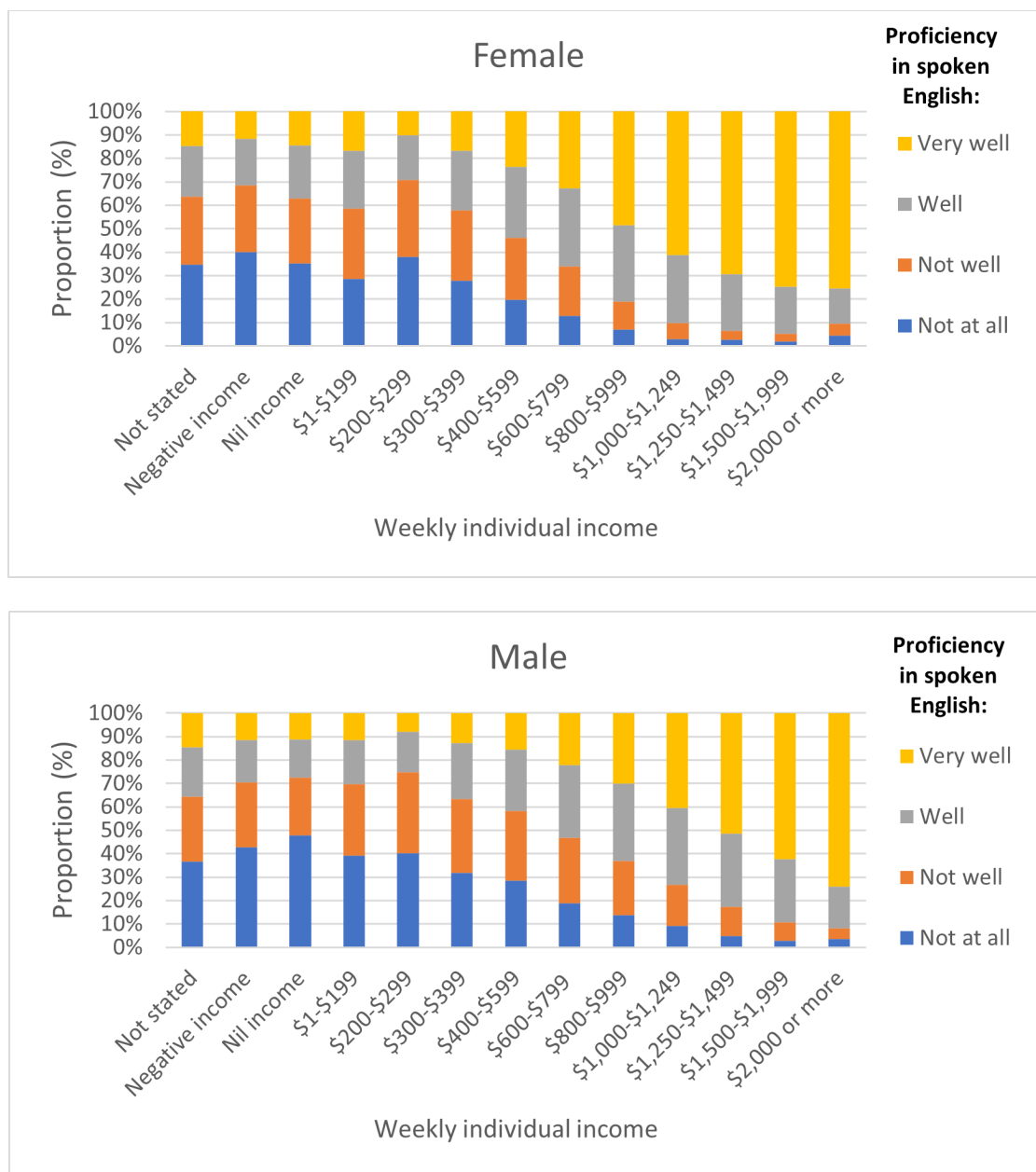


Notes: Sample of about 966,000 working-age migrants recorded in 2011 Census data. Working hour is constructed from responses to a question asking: "Last week, how many hours did the person work in all jobs?". Working hour is set at zero for individuals whose employment status is recorded as unemployed, not stated or out of the labour force.

We also observe from Figure 6 that working-age migrants with a higher proficiency in spoken English worked substantially longer hours per week. One interpretation could be that better English skills have facilitated greater employment opportunities.

Adult Migrant English Program (AMEP) Impact Evaluation Project

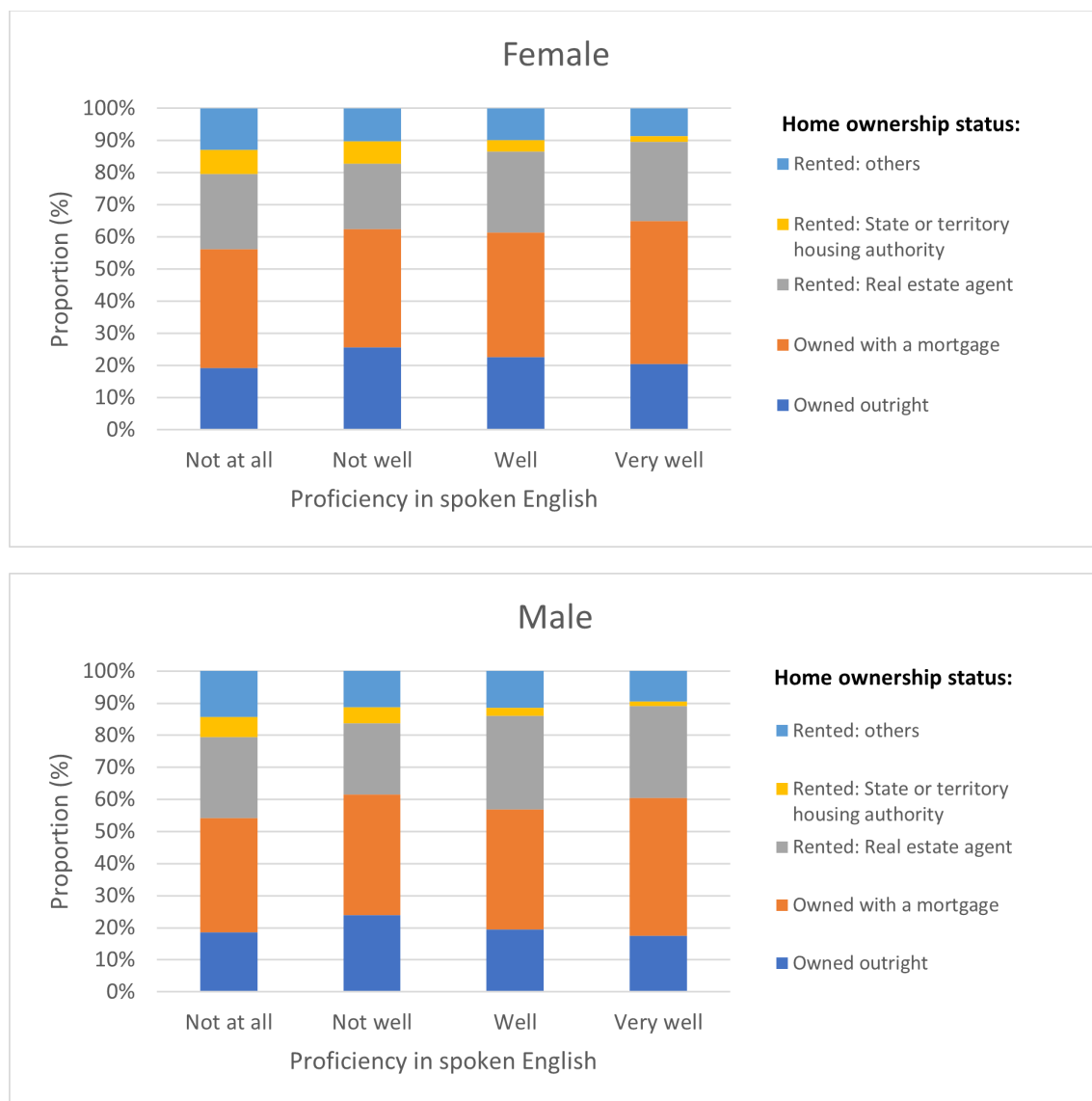
Figure 7: Census 2011 – Weekly individual income for migrants by proficiency in spoken English



Notes: Sample of about 966,000 working-age migrants recorded in 2011 Census data. Weekly income is derived from responses to a question: "What is the total of all wages/salaries, government benefits, pensions, allowances and other income the person usually receives?"

Consistent with these working patterns, working-age migrants with better English skills earned more and this is the case for both males and females (see Figure 7)

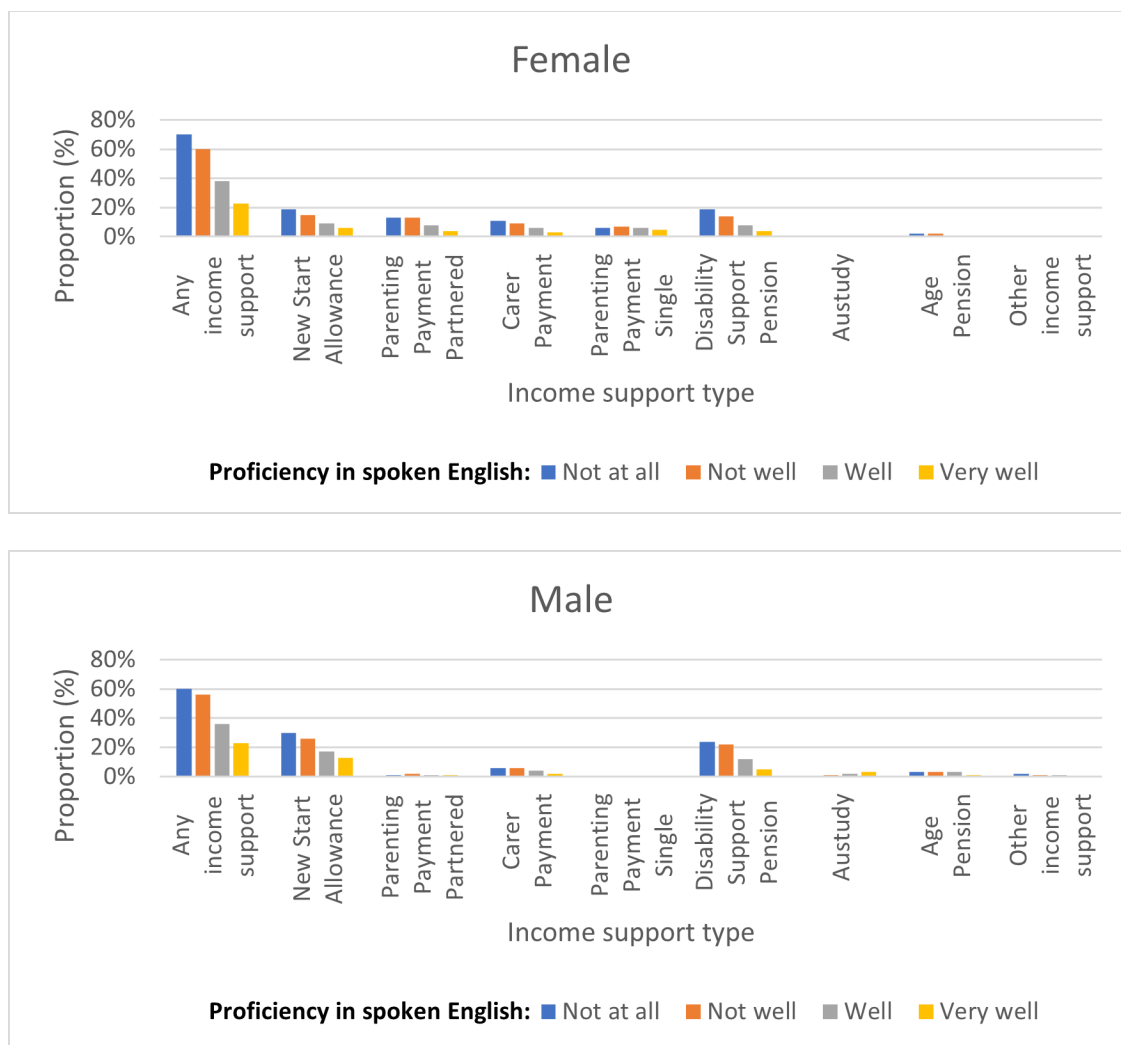
Figure 8: Census 2011 – Home ownership status for migrants by proficiency in spoken English



Notes: Sample of about 966,000 working-age migrants recorded in 2011 Census data.

Figure 8 suggests that, while there are only small differences in the rates of home ownership by proficiency in spoken English, working-age migrants with better English language skills are less likely to rely on public housing (Rented: State or territory housing authority) and this pattern holds for both male and female migrants. Particularly, working-age migrants who spoke English “very well” are about four times less likely to live in public housing authority accommodation than those who spoke English “not at all”. Of note, over 50% of migrants owned their homes outright, or with a mortgage, across all levels of English proficiency, including for those who spoke little to no English. In observing the consistency of these home ownership rates across levels of English proficiency, we must consider that thousands of these migrants have lived and worked in Australia for decades, many of whom could have purchased their homes long ago, and under different market conditions to those experienced in recent years.

Figure 9: Linked Census – DOMINO. Income support status for migrants by proficiency in spoken English



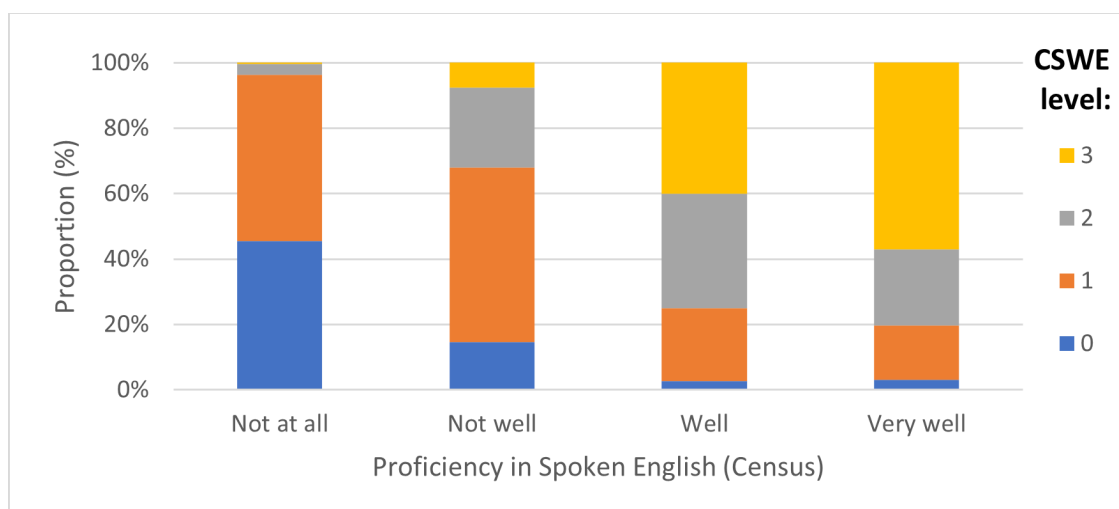
Notes: Sample of about 899,000 working-age migrants in matched 2011 Census – DOMINO datasets. Income support status is measured in 2011-12 financial year.

Similarly, Figure 9 shows that migrants with better English language skills are considerably less likely to receive any type of income support. Specifically, only 23% of working-age male migrants who spoke English “very well” received any type of income support while 60% of those who spoke English “not well” or “not at all” did. Figure 9 further indicates that the most common types of income support received by migrants are New Start Allowance, Disability Support Pension and Parenting Payment Partnered (for females only) and the relationship between proficiency in spoken English and income support status is highly apparent for these three income support modes.

The above analysis in Figures 1–9 reaffirms commonly hypothesised associations between English proficiency and socio-economic characteristics such as age, gender, and labour market outcomes. These associations thus provide indirect evidence supporting the credibility of the proficiency in spoken English measure available in the Census. To further validate this self-reported measure, we correlate it with some professionally assessed indicators of English proficiency available for a subset of linked Census - AMEP clients.

Adult Migrant English Program (AMEP) Impact Evaluation Project

Figure 10: Linked Census – AMEP. Proficiency in spoken English and CSWE level attained in AMEP entrance assessment



Notes: Sample of 16,650 matched AMEP-Census individuals who took the AMEP entrance test 12 months around the Census date.

The AMEP currently uses the Australian Core Skills Framework (ACSF) of four macro skills in listening, reading, speaking and writing to assess potentially eligible clients' English language skills. However, the International Second Language Proficiency Rating (ISLPR) scale is referenced in this report, as this scale was used in the AMEP during the period analysed in this project (July 2003 to June 2019). ISLPR is denoted in an ascending order from 0, 0+, 1-, 1, 1+, 2, 2+, 3, 3+, 4, 4+ up to 5. Each of the four macros skill is assessed and a score from the scale is applied. A potentially eligible AMEP client's English proficiency is assessed by qualified personnel and the results from these ISLPR tests are used to determine their eligibility for the program. If the client achieved a score that is less than 2 on the ISLPR scale, in any one or more of the four core skills, they are considered to have less than Functional English² and so become eligible to participate in the program.

We employ a sample of AMEP clients who appear in both Census and AMEP datasets. Furthermore, we focus on a matched sample of AMEP clients who took the AMEP entrance tests within 12 months around the Census date of 9 August 2011. This time window is chosen to make the timing of the two sets of English proficiency measures relatively close, allowing a meaningful comparison. Moreover, this approach leads to a reasonably large sample of around 16,650 AMEP clients to provide a statistically reliable analysis.

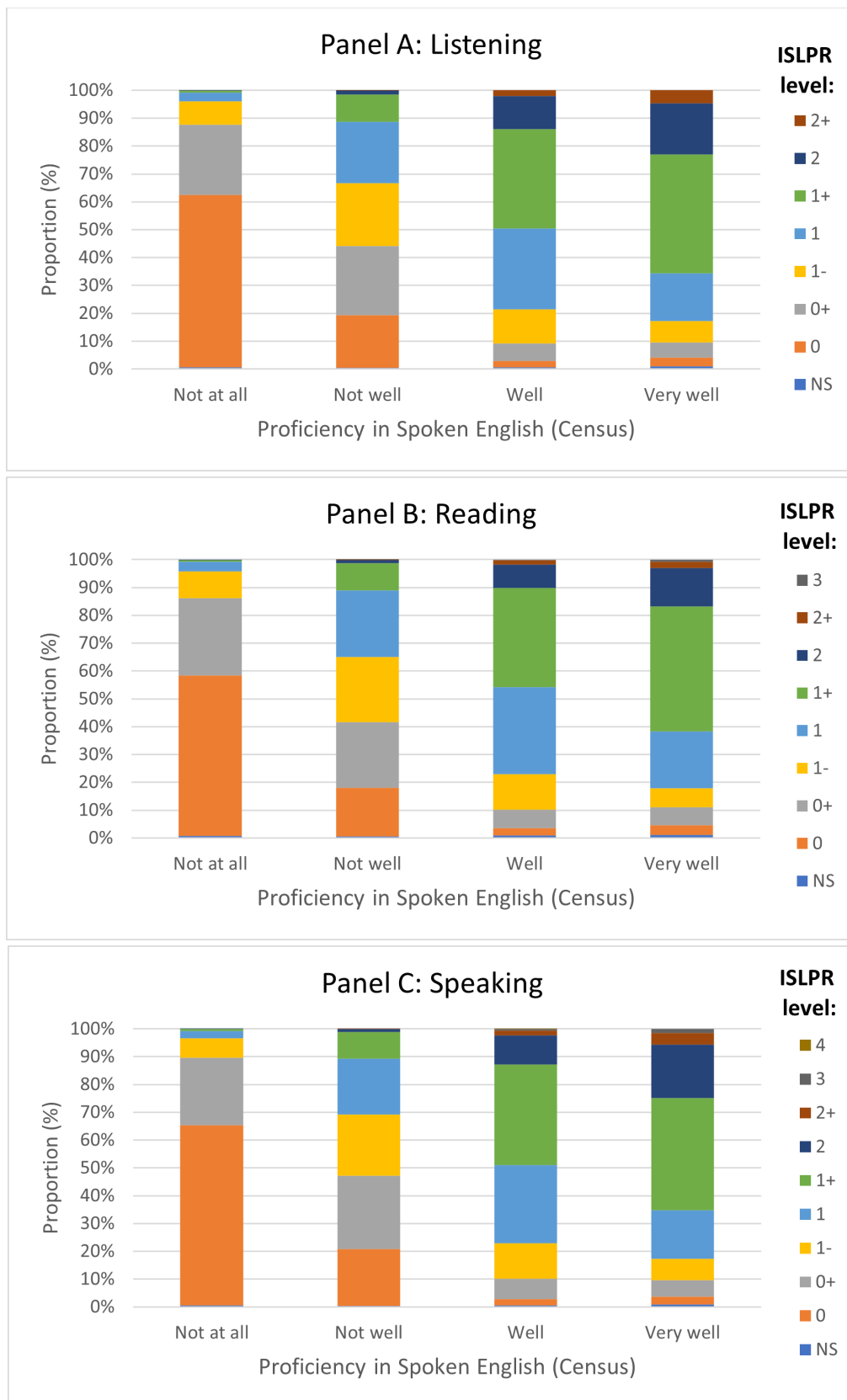
Figure 10 shows that self-reported proficiency in spoken English recorded in Census corresponds relatively well to Certificate in Spoken and Written English (CSWE stages attained in the AMEP entrance assessment). In particular, 80% of AMEP clients who reported speaking English "very well" on the Census received the top two CSWE stages (i.e., 2 and 3). Furthermore, 95% of AMEP clients reporting they spoke English "not well" or "not at all" on the Census were assigned the two lowest CSWE stages (i.e., 0 and 1).

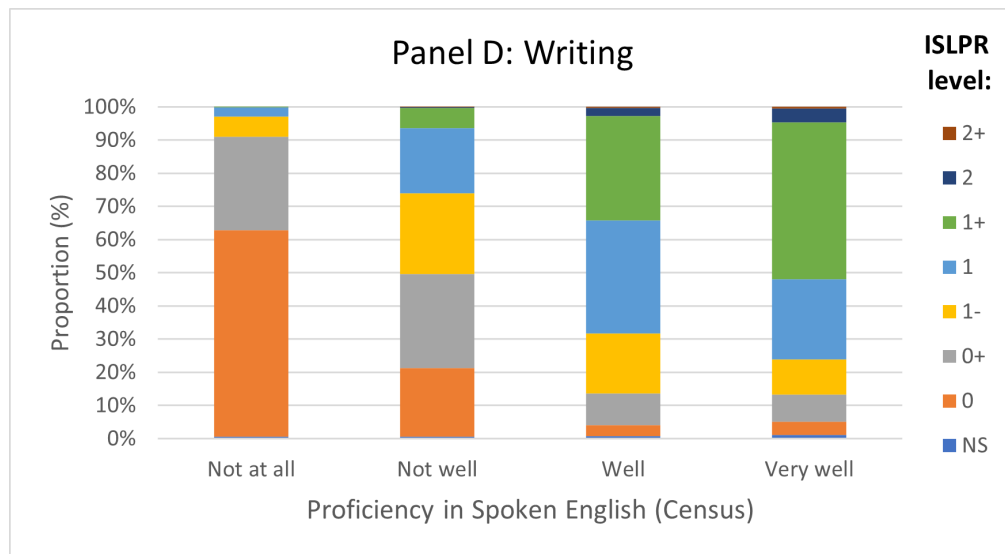
This relationship suggests Census reported proficiency in spoken English is a good proxy for general English proficiency in the absence of standardised assessments such as ACSF and ISLPR.

² In 2021, AMEP eligibility was expanded from Functional to Vocational English. For more information about 2021 AMEP reforms, see Data Note #1.

Adult Migrant English Program (AMEP) Impact Evaluation Project

Figure 11: Linked Census – AMEP. Proficiency in spoken English and initial ISLPR assessments





Notes: Sample of 16,650 matched AMEP-Census individuals who took the ISLPR assessment 12 months around the Census date.

Figure 11 also indicates a reasonably good agreement between the Census' self-reported measure of proficiency in spoken English with each of the four ACSF core language skills in listening, reading, speaking and writing.

Figure 11 further shows that, consistent with the eligibility condition set out by the AMEP, very few matched Census-AMEP clients achieved a score higher than 2 in the ISLPR scale for any of the four core language skills. Furthermore, all individuals who self-reported they spoke English "not well" or "not at all" on Census, had a score lower than 2 in the ISLPR scale in all four CSWE macro skills, suggesting that these two lowest levels of the Census's self-rated proficiency in spoken English can be used to determine if an individual has less than Functional English in the absence of a standardised English test.

Table 1: Correlation structure among various English proficiency measures

	Census English speaking	CSWE entry level	Initial ISLPR listening	Initial ISLPR reading	Initial ISLPR speaking	Initial ISLPR writing
Census English speaking	1.00					
CSWE entry level	0.59	1.00				
Initial ISLPR listening	0.63	0.84	1.00			
Initial ISLPR reading	0.61	0.84	0.90	1.00		
Initial ISLPR speaking	0.64	0.84	0.95	0.89	1.00	
Initial ISLPR writing	0.61	0.85	0.88	0.91	0.88	1.00

Notes: Sample of 16,650 matched AMEP-Census individuals who took the AMEP entrance test 12 months around the Census date. All responses are linearly transformed to get the correlations reported in this Table. In particular, for Census' English speaking, we assign 1 for "not at all", 2 for "not well", 3 for "well", and 4 for "very well". Similarly, for CSWE entry level, we set CSWE stage 0 to zero, CSWE stage 1 to 1, CSWE stage 2 to 2, and CSWE stage 3 to 3. Finally, we assign zero (0) for ISLPR score of 0, 1 for ISLPR score of 0+, 2 for ISLPR score of 1-, 3 for ISLPR score of 1, 4 for ISLPR score of 1+, 5 for ISLPR score of 2, 6 for ISLPR score of 2+, 7 for ISLPR score of 3, 8 for ISLPR score of 4, and 9 for ISLPR score of 5. All correlations are statistically significant at the 1% level.

Table 1 sets out strong correlations between the Census' self-reported measure of proficiency in spoken English and other professionally assessed indicators, with the magnitude of correlation ranging from 0.59 (for the CSWE entry level) to 0.64 (for initial ISLPR speaking) and all correlations are statistically significant at the 1% level. The fact that the highest correlation is observed for initial ISLPR speaking is consistent with the Census measure, which asks about proficiency in spoken English. Overall, the above analysis has provided both indirect and direct evidence supporting the credibility of the self-reported measure on proficiency in spoken English available on the Census.

Conclusion

This study matched AMEP clients to the 2011 Census to validate the self-reported measure of English proficiency against the results of standardised English tests used in the AMEP. The Census measure correlated very well to the measures used by the AMEP. Findings allowed us to identify persons from the 2011 Census who may have been eligible for enrolment in AMEP by virtue of their low proficiency in spoken English, but who did not participate in the program, as reported in Research Paper C: *Determinants of AMEP Participation*.

In the process of validating the self-reported proficiency in spoken English measure from the Census, we also showed that around one-quarter of all persons recorded on the 2011 Census were born overseas in a non-English Speaking Background country, and around 1-in-5 of these people spoke only English at home. Migrants who arrived in Australia at younger ages reported higher levels of proficiency in spoken English, as did those who had been resident in Australia for longer timeframes. Lastly, a range of socio-economic outcomes were better for migrants who reported they spoke English "well" or "very well" on the 2011 Census.

In summary, this paper has validated the self-reported proficiency in spoken English measure from the 2011 Census against a number of demographic variables and found it to be suitable as a proxy for general English proficiency in the absence of a standardised English test. This result supports more rigorous future evaluation of AMEP effectiveness, verifies the quality of the self-reported proficiency in spoken English measure from the Census for use in other research, and showcases the power of the MADIP investment to improve the quality of program evaluation more generally in Australia.

Data notes

1. In 2020, the Australian Government announced major reforms to the AMEP, including the extension of eligibility from functional English to vocational English, and the removal of time limits for eligible migrants who were in Australia on or before 1 October 2020.
2. This report uses the superseded Functional English as this was the eligibility requirement during the period analysed in this project (July 2003 to June 2019).

Adult Migrant English Program (AMEP) Impact Evaluation Project

3. The International Second Language Proficiency Rating (ISLPR) rating scale was used in previous tender periods - July 2003 to June 2017. ISLPR determines eligibility and entry point into the CSWE framework at a particular Certificate level. CSWE Certificates are mapped to ISLPR levels.
4. Certificate in Spoken and Written English (CSWE) was the sole AMEP curriculum up until June 2017. Curriculum flexibility was introduced from July 2017 onwards - AMEP service providers can currently choose the curriculum that best meets their clients' needs.
5. The Census-based self-reported measure on English proficiency asks only about spoken English. It does not ask about reading, writing or comprehension.
6. The current linked AMEP-MADIP dataset has incomplete tuition hours information for about 34,000 AMEP clients who enrolled in AMEP before 2011. This missing data issue means that some results may be biased, and care should be taken with interpretation.
7. Breaking down results into male and female is based on gender identity as recorded on the historical data sets used in the analysis presented here. None of these data sources provided gender-diverse identification options at the time of their collection, though most have now been updated to accommodate this for future collections. Therefore, representation of male and female in this paper may be skewed towards sex at birth, and not take into account a person's preferred gender identity.
8. Figures could also highlight the effect of spending time in a new country and being exposed to a new language.
9. The sample taken includes both English Speaking Background (ESB) and Non-English Speaking Backgrounds (NESB) or Culturally and Linguistically Diverse (CaLD) migrants.
10. All client information used in this study was managed in a secure data environment, de-identified and access restricted only to authorised researchers.

Acknowledgment

This Research Paper was analysed, compiled, and written by researchers from Telethon Kids Institute and The University of Western Australia as part of the Australian Research Council Centre of Excellence for Children and Families over the Life Course (The ARC Life Course Centre - CE200100025), with input from the Department of Home Affairs (Australian Government).