

# Teaching Math Effectively: Insights from LSAY and PISA on Primary Teacher Proficiency

Tashya Sathyajit (29672732)

## Table of contents

<b>1</b>	<b>Abstract</b>	<b>1</b>
<b>2</b>	<b>Background and motivation</b>	<b>2</b>
<b>3</b>	<b>Objectives and significance</b>	<b>3</b>
<b>4</b>	<b>Methodology</b>	<b>4</b>
4.1	How did we identify primary school teachers? . . . . .	4
4.2	What did these teachers study? . . . . .	4
4.3	How did we explore teacher performance relative to other professionals? . . . .	5
4.4	How did we complete the Mathematical Proficiency Analysis? . . . . .	7
4.5	Working with Sampling Weights and Plausible Values Accurately . . . . .	7
<b>5</b>	<b>Results</b>	<b>8</b>
5.0.1	How well did individuals who became primary school teachers perform in mathematics across the years? . . . . .	9
5.0.2	How do teachers' math scores compare to the rest of the sample (non primary school teachers)? . . . . .	9
5.0.3	How are they performing in comparison to other 'professionals'? . . . .	10
<b>6</b>	<b>Discussion, Limitations and Future Research</b>	<b>11</b>
<b>7</b>	<b>References</b>	<b>14</b>

## 1 Abstract

This report investigates the mathematics performance of individuals who became primary school teachers in Australia, using the Longitudinal Surveys of Australian Youth (LSAY) and

the Program For International Student Assessment (PISA) datasets. This analysis was done for the contribution to Grattan Institute’s report on teacher proficiency in mathematics. Results show that teachers consistently have a higher average math proficiency than non-teachers over time, with better performance across PISA testing years. However, they tend to score lower than professionals in fields like medicine, law, and engineering. Limitations include disproportionate teacher counts in some PISA years, potential attrition bias, confounding variables such as socioeconomic status, and the evolving nature of math proficiency over time.

## 2 Background and motivation

Over the past two decades, Australian students’ performance in the mathematics components of international assessments like PISA have shown [significant declines](#) as seen in Figure 1. These declines persist even after adjusting for changes in family academic resources, affecting students across all states, social groups, and school types (albeit with variations). This trend has raised concerns among educators, researchers and policymakers as mathematics proficiency is not only foundational for broader academic success but crucial for Australia’s future economic competitiveness.

Research indicates a positive correlation between teacher mathematics knowledge (as assessed through certification exams or subject-matter tests) and [student achievement](#). This underscores the importance of qualified primary school teachers in improving children’s math outcomes, supporting the notion that teachers with stronger mathematical knowledge are [likely to teach it more effectively](#).

While research consistently suggests that a teacher’s proficiency in mathematics is a crucial determinant of their effectiveness in the classroom, impacting student achievement directly (Morsy et al. 2018) the relationship between a teacher’s subject knowledge and student performance is complex, and extends beyond the teacher’s own content mastery. It involves the ability to translate that knowledge into effective pedagogical practices, adapt to curriculum needs, engage in productive professional conversations and address student misconceptions. When students are taught by teachers who have little mathematical content *and* pedagogical content knowledge, learning suffers (Baumert et al., 2010; Hill et al., 2005).

Despite these nuances, the positive correlation between teachers’ competence in mathematics and student achievement is clear (Hill et al., (2005)). While this correlation does not strictly imply causation, understanding the mathematics proficiency of primary school teachers remains a critical step in improving student outcomes.

This analysis seeks to deepen the understanding of the mathematics proficiency of individuals who eventually become primary school teachers in Australia by leveraging longitudinal data from PISA and LSAY. By examining trends over time and comparing the math performance of primary school teachers, the broader Australian population and other professionals, this study aims to provide insights into the readiness of teachers to deliver effective math instruction.

Understanding these trend is crucial, as they provide insights into the mathematics performance of individuals who later became primary school teachers, and may assist in guiding professional development initiatives focused on targeting and strengthening the mathematics proficiency of primary school teachers, ultimately leading to improved student outcomes across Australia.

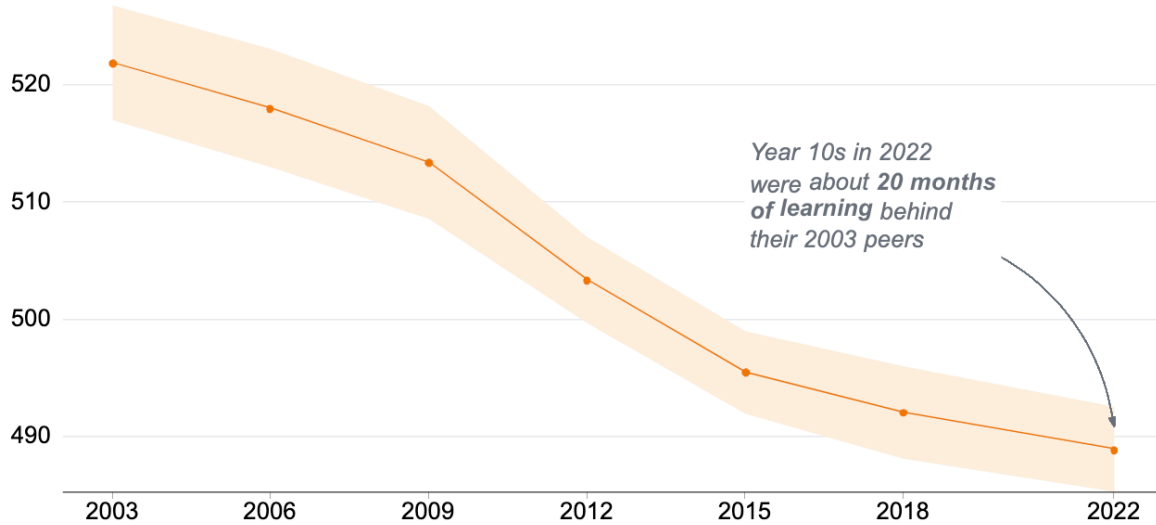


Figure 1: Decline in PISA Scores Across the Years

### 3 Objectives and significance

The primary objectives of this analysis was to utilise longitudinal survey data to understand the mathematics proficiency of students who eventually became primary school teachers in Australia. Using the LSAY and PISA datasets, the analysis seeks to address three main questions:

1. **Teacher performance over time:** How well did individuals who became primary school teachers perform in mathematics across the years?
2. **Comparison to other Professionals:** How are they performing in comparison to other ‘professionals’?
3. **Comparison to the General Population:** How do teachers’ math scores compare to the rest of the sample (non primary school teachers)?

This analysis is significant as it provides critical insights into the math skills of individuals who become primary school teachers, a factor that can directly impact student outcomes. By identifying trends and proficiency gaps, the analysis can allow for a deeper understanding of

Table 1: Variables for ‘What kind of work do you do in this job?’ Questions

Year	Variables
<b>2003</b>	LDD025, LED025, LFD025, LGD025, LHD029, LID029, LJD029, LKD029
<b>2006</b>	ANZSCO, LBD011, LCD024, LDD024, LED029, LFD029, LGD029, LHD029, LID029, LJD029, LKD029
<b>2009</b>	ANZSCO73, LBD014, LCD029, LDD029, LED029, LFD029, LGD029, LHD029, LID029, LJD029, LKD029
<b>2015</b>	LSAYID, LBEM004, LCD022, LDD022, LED022, LFD025, LGD022, LHD022

the performance of primary school teachers in Australia and inform potential policy recommendations.

## 4 Methodology

Data was sourced from multiple LSAY cohorts (2003, 2006, 2009 and 2015), each containing demographic data from student responses and their linked performance in PISA. Data was selected from these years because post 2003, LSAY participants were recruited from schools that also took part in PISA, allowing for the linkage of PISA scores to LSAY data.

### 4.1 How did we identify primary school teachers?

For each cohort, variables relating to kind of work were selected for each wave of LSAY data aforementioned. These were located using the LSAY data dictionary sourced [here](#).

To identify primary school teachers, we used the variables found in Table 1. These variables were linked to questions regarding the kind of work the respondent did in their job. These variables were selected and a flag was created separating individuals who responded with ANZSCO code 2412 which corresponds to the occupation of a primary school teacher. This flag was used to identify individuals who eventually became primary school teachers, and those that did not respond with this code were classified as non-teachers.

### 4.2 What did these teachers study?

For exploratory understanding, we also wanted to see the area of study of individuals who later became teachers. We did this by utilising the variables seen in Table 2, which related

Table 2: Variables for Main Area of Study Questions

Year	Variables
<b>2003</b>	LBCZ087, LCCZ087, LDCZ087, LEC087, LFC087, LGC087, LHC087, LIC087, LJC088, LKC088
<b>2006</b>	LBC087, LCC087, LDC087, LEC087, LFC087, LGC088, LHC088, LIC088, LJC088, LKC088
<b>2009</b>	LBC089, LCC088, LDC088, LEC088, LFC088, LGC088, LHC088, LIC088, LJC088, LKC088
<b>2015</b>	LCC088, LDC088, LEC088, LFC088, LGC088, LHC088

to demographic questions about what the respondent’s area of study was. We again filtered the student IDs of these individuals with those who responded with the [Australian Standard Classification of Education \(ASCED\)](#) code 70103 for primary school education.

After doing this, we also did a reverse in seeing what the individuals who became primary school teachers studied primary to entering that field of work. Using the student IDs of those who responded as primary school teachers, we looked again at the variables for area of study across the waves and explored what ASCED scores they provided. The area of studies were consistent with the pathway to becoming a primary school teacher as seen in

### 4.3 How did we explore teacher performance relative to other professionals?

This part of the analysis was driven more by curiosity than formal research aims, exploring how primary school teachers’ mathematics proficiency compares to that of other professionals. Due to this, the analysis focused only on the 2003 PISA cohort which had the largest sample of teachers (view Figure 6) and the final wave of that cohort. By examining the last wave of this cohort, the analysis captures participants’ final professional roles, though this approach results in a lower teacher count than the total count for the entire 2003 cohort as seen in Table 3.

To define professional categories, the analysis used the ANZSCO codes sourced from the Australian Bureau of Statistics [website](#). The respondent’s occupation code response to variable LKD029 corresponding to LSAY question relating to kind of work was used to identify whether they belonged to one of the professional groups of interest. A new variable was established which created a flag for individuals based on their occupation code (e.g. Teachers coded as ‘1’, while other professional groups were coded ‘2’ to ‘6’ in accordance to their ANZSCO code). After this the **Rrepest** package was used to estimate the weighted average proficiency scores and standard deviations for each group, using plausible values and their respective weights for mathematical scores.

Code	Field_of_Education
<b>70101</b>	Teacher Education: Early Childhood
<b>70103</b>	Teacher Education: Primary
<b>70109</b>	Teacher Education: Vocational Education and Training
<b>70199</b>	Teacher Education, n.e.c.
<b>70115</b>	English as a Second Language Teaching
<b>90503</b>	Children's Services
<b>91703</b>	Religious Studies
<b>92101</b>	Sport and Recreation Activities
<b>100101</b>	Music
<b>80101</b>	Accounting
<b>100705</b>	Written Communication
<b>100103</b>	Drama and Theatre Studies
<b>80399</b>	Business and Management, n.e.c.
<b>70105</b>	Teacher Education: Secondary
<b>91701</b>	Philosophy
<b>60901</b>	Optometry
<b>61399</b>	Public Health, n.e.c.
<b>60101</b>	General Medicine
<b>81101</b>	Banking and Finance
<b>91523</b>	Literature
<b>90513</b>	Counselling
<b>90701</b>	Psychology
<b>79999</b>	Education, n.e.c.
<b>60501</b>	Pharmacy
<b>61799</b>	Rehabilitation Therapies, n.e.c.

Table 3

Table 4 Professionals, Corresponding ANZSCO Codes, and Counts

Professionals	ANZSCO Code	Count
Primary School Teacher	2412	82
Medical Practitioners	253	48
Midwifery and Nursing Professionals	254	94
Legal Professionals	271	50
Accounts Auditors and Company Secretaries	221	92
Engineering Professionals	233	88

#### 4.4 How did we complete the Mathematical Proficiency Analysis?

The mathematics performance of primary school teachers was assessed using PISA scores. The analysis aimed to compare teachers' scores against the national proficiency standard which has been set at a proficiency level 3 or [482 PISA points](#). The analysis incorporated survey weights to ensure representativeness of the data and address differences in sample sizes across cohorts. It also had to account for the use of plausible values undertaken by PISA.

#### 4.5 Working with Sampling Weights and Plausible Values Accurately

To handle PISA's plausible values and sampling weights, the analysis utilised [intsvy](#) and [rrepest](#) R packages, which are designed to accurately estimate statistics under the complex survey design used by PISA.

These allowed for the calculation of mean scores, confidence intervals and standard errors while accounting for the inherent variability within the dataset.

##### But what are plausible values?

International surveys such as PISA report student performance using plausible values (PVs). Plausible values are a statistical technique that represent uncertainty in measurements by generating a range of possible values a student may have obtained as opposed to a single estimate. This approach assists in conveying the variability in student performance more accurately, considering that not all students answer the same questions (PISA is a large test) or have the same background information.

To derive PVs, PISA uses [Item Response Theory \(IRT\)](#), which predicts the likelihood of a student answering a question correctly by assessing the difficulty of each question and how well it distinguishes between students of different ability levels. For example, if the question is deemed a relatively conceptually 'easy' question, and the estimated ability of the student is high, the model predicts a high probability of a correct response. PISA can determine

if a question is ‘easy’ or ‘hard’ based on the statistical analysis of response patterns from a broad group of students, using IRT to estimate how challenging each question is relative to the separate proficiency levels.

Looking at Figure 2 retrieved from Wu (2022), the process starts with a prior distribution. This is an initial estimate of the student’s proficiency based on background variables like socioeconomic status or parental education. This prior estimation serves as a rough starting point that as the students begin to answer questions, get finetuned as PISA updates this estimate using student responses and resulting in a posterior distribution. This updating distribution combines that initial prior estimate with the observed test results.

After this, multiple imputed values are taken from this posterior distribution. Essentially, PISA generates plausible values that are a representation of the range of abilities a student may reasonably have.

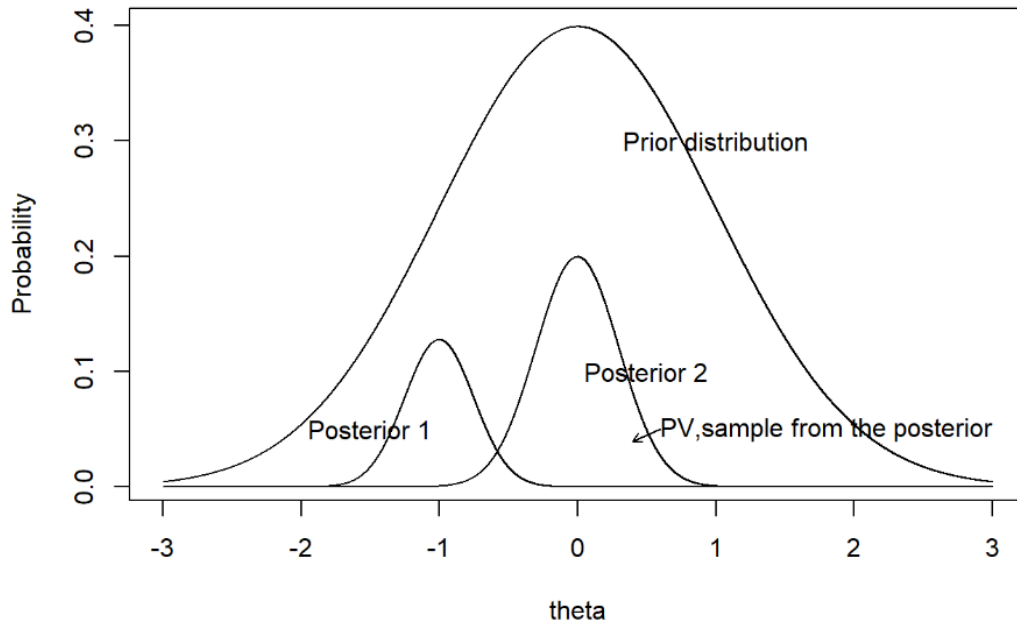


Figure 2: Prior and Posterior Distributions

## 5 Results

As earlier mentioned in the report, the project sought to address three key questions through the exploration of the mathematical performance of individuals who became primary school teachers.



Table 5: PISA Teacher vs Non-Teacher Math Proficiency

Year	Teacher	Mean	SE_Mean	Std	SE_Std
2003	False	530.3	2.4	92.9	2.4
2003	True	554.0	10.5	64.8	5.3
2006	False	518.6	3.0	87.9	1.2
2006	True	550.7	10.4	67.2	8.2
2009	False	513.7	2.7	95.1	2.5
2009	True	552.7	7.2	63.9	8.4
2015	False	494.0	1.8	92.5	0.9
2015	True	519.4	20.2	61.0	19.0

### 5.0.1 How well did individuals who became primary school teachers perform in mathematics across the years?

The analysis reveals that individuals who eventually became primary school teachers consistently performed better in mathematics compared to their non-teaching peers over the years. To quantify this, the weighted average proficiency of teachers and non-teachers was calculated taking into consideration the total number of teachers and non-teachers which served as the basis for weighting the average proficiency scores.

The results show as depicted in Figure 3, and illustrates that over the years considered in this analysis, just over 80% of teachers met or exceed the national proficiency standard. In comparison, just over 60% of the non-teachers met or exceeded the national proficiency standard of 482 PISA points. This may suggest that individuals who eventually become primary school teachers have a higher proficiency in mathematics compared to their non-teaching peers.

### 5.0.2 How do teachers' math scores compare to the rest of the sample (non primary school teachers)?

When trying to see how teachers performed in each year, we were able to find that teachers tended to perform better across each year of testing. As seen in Figure 4, teachers (red) tend to do consistently better than non-teachers (orange) across the years.

The boxplots shows that the median, interquartile range, and 10th and 90th percentiles. The boxplots for non-teachers have longer whiskers, indicating a broader range of scores and greater variability in math proficiency compared to teachers. The wider spread suggests that non-teachers include individuals with a diverse range of mathematical skills, spanning from very low to very high.

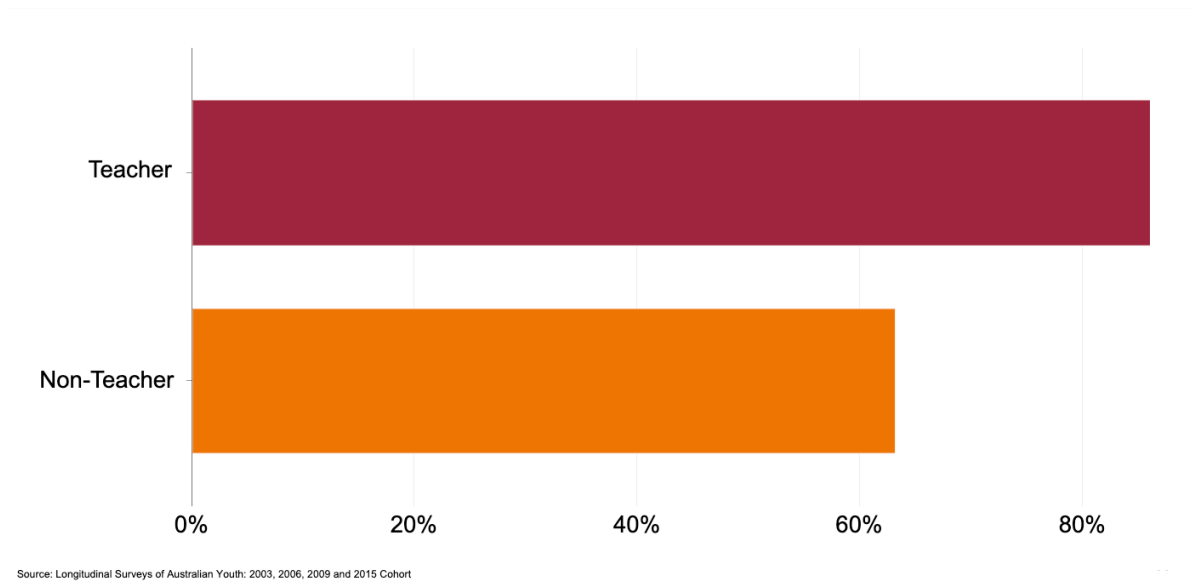


Figure 3: Weighted Average of Performance Across Years: Teachers vs Non-Teachers

For teachers, the whiskers are shorter, indicating more consistency in their math scores. However, the standard error (SE) for the teacher group is higher, this is likely due to their smaller sample size, which increases the uncertainty in estimating the average score.

Ultimately, while non-teachers exhibit more variability in individual scores, the higher standard error for teachers reflects the effect of smaller sample sizes on the precision of the average, not necessarily a greater variability in the scores itself.

### 5.0.3 How are they performing in comparison to other 'professionals'?

The project had a smaller focus on understanding how teachers' mathematical proficiency compared to other professionals, driven largely by curiosity rather than formal research aims. The selection of professional categories was based on existing reports, and the analysis was conducted on the 2003 PISA cohort, the year with the largest sample of teachers. However this still had its limitations due to the smaller sample size of the relative professional groups within a total sample of 10,370, limiting the overall robustness of the findings.

Despite this, as seen in Figure 5, teachers performed lower than most other professional groups in terms of average mathematical proficiency at age 15. Professions such as medical practitioners, legal professionals, engineers and auditors scored higher on average, indicating better mathematical capability amongst these groups. However, teachers scored higher than the 'Other' category, which represents the broader population to align with the specific professions of interest. This suggests that while teachers were not among the highest performing

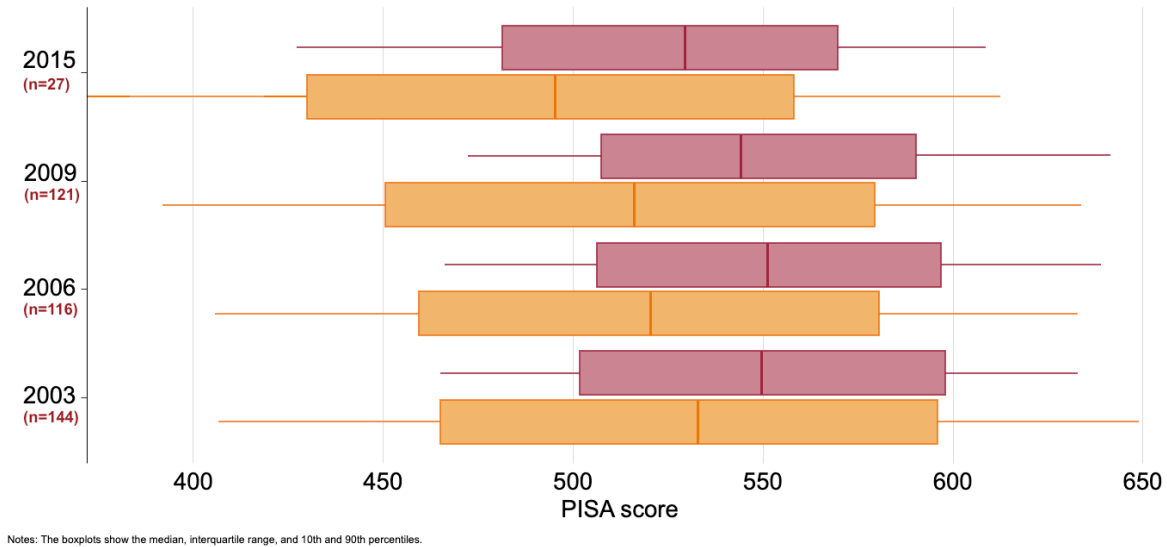


Figure 4: Proficiency of Teachers vs Non-Teachers Across Years

professional groups, they seemed to demonstrate better mathematical proficiency than the general cohort.

Despite these findings, it is important to consider the larger confidence intervals observed for certain professional groups, such as legal professionals, engineers, and auditors. These wider intervals indicate greater uncertainty in the estimates and may result from factors like smaller sample sizes and higher variability within these groups.

## 6 Discussion, Limitations and Future Research

The findings of this analysis provide insights into the mathematical proficiency of individuals who eventually become primary school teachers, indicating a generally higher proficiency than their non-teaching peers across multiple years. Teachers tend to meet or exceed national proficiency standards more frequently than non-teachers, suggesting a relatively stronger foundation in mathematical skills. This trend is evident over time, as teachers consistently achieve higher median scores with less variability compared to non-teachers, as seen in Figure 4.

The more consistent performance among teachers is suggested by the slightly shorter whiskers in the boxplots, indicating a relatively uniform level of mathematical competence. In contrast, the longer whiskers for non-teachers reflect greater variability, indicating a broader range of skills. However, it is important to note that the whiskers for both groups are relatively long, suggesting a wide distribution of scores overall, with non-teachers demonstrating only slightly more spread. The higher standard error for teachers, meanwhile, is primarily due to their

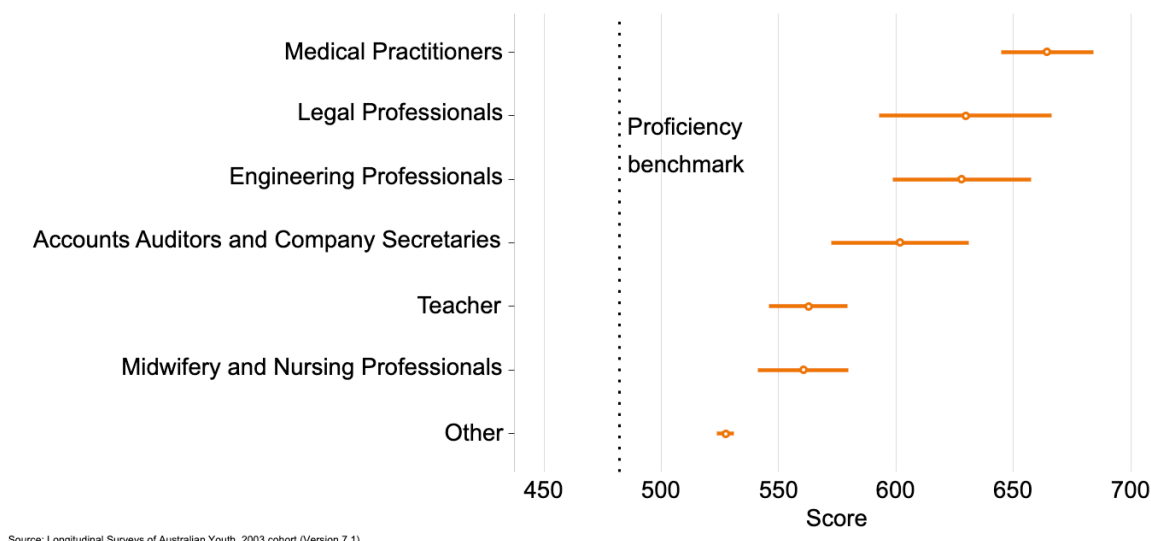


Figure 5: Teachers in Comparison to Other Professionals

significantly smaller sample size, increasing uncertainty in the estimated average scores. This however, reflects the inherent limitations of sample size than actual variability in proficiency.

When comparing teachers to other professional groups, as seen in Figure 5, the analysis reveals that while future teachers performed better than the general population, they performed lower than medical practitioners, legal professionals, engineers, and auditors and performed quite similarly to midwifery and nursing professionals. This aligned with a preconceived idea that STEM professions would attract students with more confidence or proficiency in mathematics. However, the large confidence intervals observed for some professional groups highlight greater uncertainty of where the true average lies due to smaller sample sizes and potential variability within these groups.

However, there are some considerations that are important to pay mind to upon the results of this analysis. Despite the relatively strong performance of teachers in mathematics, as aforementioned, effective teaching requires more than just content knowledge. Research frameworks emphasise that teachers must understand how to present content effectively, integrate it with the curriculum, and adapt it to diverse learning styles (Norton, 2018). Additionally, although we can see that individuals who became teachers in later life had a higher proficiency in mathematics when compared to the general population (save for the professions explored deeper), there are limitations to the analysis.

#### **Small sample size for teachers and professional groups:**

The sample size for teachers as seen in Figure 6 and other professionals as seen in Table 3 was relatively small, which limits the robustness of the findings. The smaller sample size is

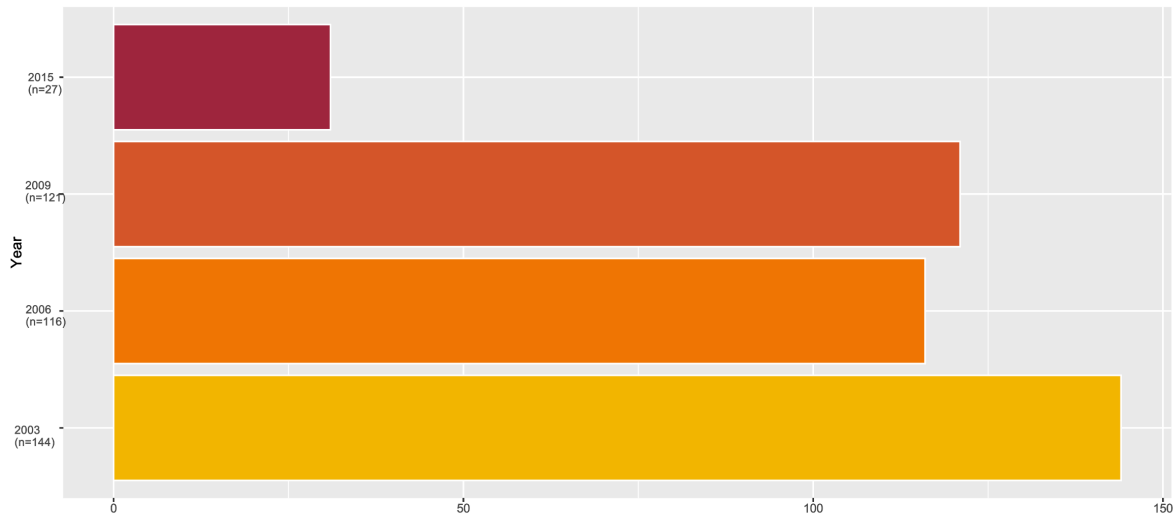


Figure 6: Count of Teachers per Year

also associated with larger standard errors, increasing uncertainty in the estimated averages of those groups. The lack of statistical power due to small sample sizes may have affected the precision of the estimates, potentially leading to biased results or challenges in generalising the results.

#### Attrition Rate in Longitudinal Data:

The analysis is based on longitudinal data, which is inherently prone to attrition over time. As participants drop out of the study, the sample size decreases which may introduce potential bias and affect the reliability of the results. Further, the later PISA cohorts haven't completed all the waves which contributes to the smaller sample size for teachers in the later years. This attrition bias may have influenced the results, particularly in the later years, and may have implications for the generalisability of the findings.

#### Scope of PISA:

As mentioned prior, while PISA tests part of the expected learning outcomes of for mathematics, it does not test how good an individual will be at conveying this knowledge to others or the entire range of mathematical knowledge required in teacher training or professional practice. This could suggest that although individuals who performed well in PISA who later became teachers may be good at mathematics, they may not necessarily be good at teaching it.

Additionally, PISA only tracks individuals from the age of 15 to 25, and this does not account for the many other ways individuals go into the pathway of being teachers. For example, mathematical proficiency can be enhanced through alternative pathways or lived work experience before the individual moved into teaching. This could also suggest that the analysis

may not capture the full range of individuals who become teachers and their mathematical proficiency.

**High Standard Error:**

## 7 References

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47(1), 133-180. <https://doi.org/10.3102/0002831209345157>

Gotsis, T. (2020). NSW school education: PISA 2018, socioeconomic background and proposals for reform. <https://www.parliament.nsw.gov.au/researchpapers/Pages/NSW-school-education-PISA-2018.aspx>

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. *American Educational Research Journal*, 42(2), 371-406. <https://doi.org/10.3102/00028312042002371>

Morsy, L., Khavenson, T., & Carnoy, M. (2018). How international tests fail to inform policy: The unsolved mystery of Australia's steady decline in PISA scores. *International Journal of Educational Development*, 60, 60-79. <https://doi.org/https://doi.org/10.1016/j.ijedudev.2017.10.018>

Norton, S. (2019). The relationship between mathematical content knowledge and mathematical pedagogical content knowledge of prospective primary teachers. *Journal of Mathematics Teacher Education*, 22(5), 489-514. <https://doi.org/10.1007/s10857-018-9401-y>

OECD. (2009). Plausible Values. <https://doi.org/doi:https://doi.org/10.1787/9789264056275-7-en>

Okubo, T. (2022). Theoretical considerations on scaling methodology in PISA. <https://doi.org/doi:https://doi.org/10.1787/9789264056275-7-en>

Wu, M. (2022, June 7) Chapter 14 Estimating Population Characteristics - Part II: Plausible Values. *A Course on Test and Item Analyses*. Retrieved October 28, 2024, from <https://www.edmeasurementsurveys.com/Ch14.html>