

The Literacy and Numeracy Test for Initial Teacher Education Students

Administration Report 2019 Test Window 1

18 February – 3
March 2019



Confidential

Introduction

The first test window of 2019 for the Literacy and Numeracy Test for Initial Teacher Education Students was held from Monday 18 February to Sunday 3 March 2019. This window offered both test centre and remote proctoring test sessions. Test centre sessions were held in both the capital cities and regional locations.

A total of 11419 test session bookings were received during the registration period: Wednesday 16 January to Tuesday 29 January 2019. Both the test centre and remote proctoring test sessions were administered successfully and smoothly in general.

Pre-Test Activities and Candidate Communication

Website and Communication

The test website was the main information source for candidates to find out about the test, to register, and to locate important test administration procedures. The teacher test email and phone helpdesk were in operation during standard business hours to assist candidates with their enquiries. All email and phone enquiries were generally responded to within 24 business hours.

Email Reminders

Registered candidates were sent reminder emails leading up to the start of the test window.

For test centre candidates, reminders were sent to notify candidates that Admission Tickets had been released. Candidates were also encouraged to check their test registration session details via their candidate account carefully and to ensure that they had arranged appropriate photo identification prior to their scheduled test sitting(s). Subsequent reminders for test centre candidates focused on reminding candidates of important test day procedures, permitted test items, the misconduct policy and the general test procedures on the test day.

For remote proctoring candidates, the first email notified candidates that Admission Tickets had been released. Candidates were reminded to complete the technical readiness checks, to download and install the ACER online test application and to schedule their test sessions with ProctorU if they had not already done so.

Subsequent emails for remote proctoring candidates focused on test day procedures, such as the technical and test room requirements, permitted items and misconduct policy for the test. Candidates were given step-by-step instructions for connecting with ProctorU on their scheduled test day. Information on breaks, test security, and how to manage disruption and technical issues was also provided.

Practice Material

Sample and Practice Questions were available on the test website. An interactive online version of each was also accessible to all registered test candidates. Answers were provided for all questions.

Registrations and Refunds

Registrations for Test Window 1 opened on Wednesday 16 January and closed on Tuesday 29 January 2019. All candidates were required to register online at the ACER test website. Upon registration, candidates received an email confirming their registration details and payment. Candidates registered for the Literacy and Numeracy test components separately using their ACER candidate account.

A total of 5700 candidates registered for the literacy test, and a total of 5719 candidates registered to sit the numeracy test. This equates to 11419 test registrations.

The refund deadline was Tuesday 29 January 2019. 75 refund requests were received and processed by the deadline. A further 9 refunds were processed on compassionate grounds after the deadline for varied reasons such as the death of an immediate family member, major health issues

and illness, hospitalisation and surgery.

Test Centres

ACER managed and ran all regional test centres in Victoria and also set up a new test centre in Granville to replace the Parramatta test centre in New South Wales. In response to the substantial rising costs in regional test centres, ACER decided to use a different test centre in Armidale via Saxons instead of Cliftons for this test window. Experienced test supervisors in Armidale were recruited and trained by ACER. The Adelaide Saxons test centre that was run in November 2018 continued to run alongside the established Cliftons test centre to cater to the increasing candidate demand in Adelaide.

Other than a network issue experienced at the Adelaide Saxons test centre on 22 February 2019, all the test sittings generally went smoothly. Investigation undertaken by the Adelaide Saxons test centre revealed that the network issue was an isolated once-off incident related to traffic levels and congestion on the upstream network of the building. A network review and deep dive technical incident response have been implemented to evaluate the devices in the network, their settings, firmware and topology to ensure that no such issues will arise in future. The centre will also conduct enhanced monitoring to the network to identify issues more quickly and to make changes to the environment as necessary to ensure stability and operability for future test sittings.

Feedback on all other new test centres and the test supervisors was positive and ACER would continue to use the centres and supervisors for subsequent test windows.

Admission Tickets and Photo Identification

Admission tickets were released to candidates via their ACER account on Monday 11 February 2019. The ticket stated the personal details, test session and venue details for test centre candidates, and the usernames required to access the tests for remote proctoring candidates.

Candidates had to check their personal details and their higher education provider and declare that the information was correct before they could access their ticket. This was done to minimise the number of errors that had to be rectified during and after the test period.

Candidates were also reminded that they needed to have their admission ticket and appropriate photo identification to be permitted to sit the test. Candidates were permitted to show the test supervisors or their online proctor either an electronic or printed copy of their admission ticket.

Test Day

Attendance

A total of 5700 candidates registered for the literacy test, and a total of 5719 candidates registered to sit the numeracy test by the close of registrations. 9 candidates were then refunded after the close of registrations. At the end of the test window, 5486 candidates had sat the literacy test, and 5480 candidates had sat the numeracy test.

Table 1 provides a summary of attendance for 2019 Test Window 1.

HEPs	Literacy			Numeracy		
	Registered	Present	Attendance	Registered	Present	Attendance
Alphacrucis College	13	13	100.00%	13	13	100.00%
Australian Catholic University	596	579	97.15%	621	603	97.10%
Australian College of Physical Education	13	12	92.31%	10	9	90.00%
Avondale College	28	27	96.43%	29	27	93.10%
Central Queensland University	105	99	94.29%	100	92	92.00%

Charles Darwin University	57	51	89.47%	56	51	91.07%
Charles Sturt University	65	64	98.46%	65	64	98.46%
Christian Heritage College	4	4	100.00%	5	5	100.00%
Curtin University	157	150	95.54%	154	146	94.81%
Deakin University	242	237	97.93%	242	232	95.87%
Eastern College Australia	2	2	100.00%	2	2	100.00%
Edith Cowan University	49	46	93.88%	50	49	98.00%
Excelsia College	6	5	83.33%	6	5	83.33%
Federation University Australia	86	85	98.84%	89	87	97.75%
Flinders University	405	393	97.04%	410	397	96.83%
Griffith University	80	78	97.50%	92	87	94.57%
Holmesglen TAFE	7	7	100.00%	8	8	100.00%
James Cook University	41	39	95.12%	42	38	90.48%
La Trobe University	222	214	96.40%	237	222	93.67%
Macquarie University	132	127	96.21%	127	121	95.28%
Melbourne Polytechnic	4	4	100.00%	2	2	100.00%
Monash University	114	98	85.96%	88	77	87.50%
Montessori World Educational Institute	1	1	100.00%	1	1	100.00%
Murdoch University	28	28	100.00%	30	30	100.00%
Queensland University of Technology	479	458	95.62%	471	450	95.54%
RMIT University	89	86	96.63%	91	88	96.70%
Southern Cross University	150	144	96.00%	146	141	96.58%
Swinburne University of Technology	307	299	97.39%	299	285	95.32%
Tabor Adelaide	22	22	100.00%	21	21	100.00%
The University of Adelaide	69	66	95.65%	62	60	96.77%
The University of Melbourne	41	36	87.80%	46	42	91.30%
The University of New England	162	159	98.15%	168	163	97.02%
The University of New South Wales	100	97	97.00%	99	97	97.98%
The University of Newcastle	262	254	96.95%	270	256	94.81%
The University of Notre Dame Australia	89	86	96.63%	96	91	94.79%
The University of Queensland	55	53	96.36%	60	58	96.67%
The University of Sydney	85	84	98.82%	87	85	97.70%
The University of Western Australia	32	32	100.00%	31	31	100.00%
University of Canberra	51	47	92.16%	52	50	96.15%
University of South Australia	383	374	97.65%	373	361	96.78%
University of Southern Queensland	59	58	98.31%	50	50	100.00%
University of Tasmania	84	83	98.81%	89	88	98.88%

University of Technology, Sydney	39	33	84.62%	35	32	91.43%
University of the Sunshine Coast	79	77	97.47%	85	81	95.29%
University of Wollongong	154	152	98.70%	156	156	100.00%
Victoria University	235	219	93.19%	223	210	94.17%
Western Sydney University	211	204	96.68%	223	216	96.86%
Grand Total	5694	5486	96.35%	5712	5480	95.94%

Table 1: Summary of attendance by Higher Education Provider for 2019 Test Window 1

Test Dates

2019 Test Window 1 was held from 18 February to 3 March 2019. Remote proctoring sessions could be scheduled by candidates between 18 February and 28 February 2019. The remote proctoring test window was extended until 2 March 2019 for a small number of candidates who experienced significant technical issues and needed to reschedule their sessions.

Reporting and Test Start Times

There were four test sessions at the test centres per day. Candidates sitting the morning literacy test session were advised to report to their test centre at 7.30am. The morning numeracy test session had a reporting time of 10.30am. Candidates who registered to sit the afternoon literacy test session had to report at 1.30pm and those sitting the afternoon numeracy test session were told to report at 4.30pm.

Candidates who elected to sit the test by remote proctoring were able to book their test from the list of available session times in the ProctorU booking system. Remote proctoring candidates were able to reschedule their test session(s) during the test window at their convenience.

All candidates were reminded via email that they should allow at least 2½ hours for all testing procedures to be completed for each test component. Remote proctoring candidates were advised to check and ensure that they had allowed enough time between test sessions to complete the pre-test checks and registration procedures if they wished to book consecutive literacy and numeracy test sessions.

Reasonable Adjustments

74 applications for reasonable adjustments were received and approved by ACER for 2019 Test Window 1.

All applications were approved after supporting documentation, such as medical documentation and/or psychologist reports, was reviewed by the committee which was made up of 2 or more ACER staff.

The types of adjustments granted for 2019 Test Window 1 included:

- Rest breaks
- Additional test time
- Permission to bring medication and/or diabetes/glucose monitoring kit
- Permission to consume food/drink relating to medical condition
- Permission to bring a desk lamp if required
- Permission to wear Irlen Spectral Filters or tinted lenses
- Permission to wear ear plugs or hearing aids
- Permission to increase the font size
- Permission to use a text-to-speech software
- Permission to use more scratch paper
- Having written test instructions, test timing warnings and additional test supervisor support by the side when required
- Permission to have a chair without armrests

- Special seating arrangements for a corner of the test room or in a smaller room.

The adjustments were well managed by the test supervisors and online proctors.

Management of Incidents

ACER received daily summary reports and incident reports for individual events via the ACER online incident reporting system from the Chief Supervisors at the test centres. For remote proctoring, incidents for individual candidate test sessions were flagged and escalated in the ProctorU system and notifications were emailed to ACER. Candidates were also advised to contact ACER by email if they experienced technical problems or other incidents during their test session(s).

ACER reviewed all incident reports raised by test supervisors and ProctorU on a daily basis, and also followed up on incidents and complaints raised directly by candidates during the test window.

Test Centres

Incidents reported for 2019 Test Window 1 are summarised in Table 2 below.

Test Centre	Number of Incidents Reported	Incident Type(s)	Resolution/Action (if required)
Capital cities			
Adelaide (Saxons)	14	1 candidate requested permission to take anxiety medication should she feel the need to even though she had not applied for reasonable adjustments. 1 candidate was still doing the first section and would not log off at the end of the session. The candidate approached the Chief Supervisor and mentioned that she had epilepsy. She also stated that she had been nervous. 1 candidate requested to wear computer glasses. 1 diabetic candidate had a dog on a lead wearing a green vest when she reported to the test centre. On the dog coat were the words "Diabetic Alert Service Dog, please do not pat, dog is working". Dog was not permitted in the test room. 1 candidate's phone rang when the test started. The candidate stood up and went to her bag in the room to turn the mobile off, without seeking permission from the test supervisor.	Test supervisor checked the medication and permission was given. Test supervisor approached the candidate to log off. The supervisor had to shut down the candidate's computer. Candidate also confirmed that she knew the length of the test and heard the 30 minutes' warning to move to section 2. Candidate was advised to contact ACER. Test supervisor checked the glasses and permission given. No issues arose during her test session. Candidate was advised that an incident report would be filed and to contact ACER.

		1 candidate queried the test content. 19 candidates experienced technical issues including frozen screens. 26 candidates experienced issues on their computers due to the internet/technical disruptions. 1 candidate's screen went blank. On the top left corner of the screen, it showed "Host unreachable". 2 candidates presented photo ID that contained mismatches in DOB or names on the admission tickets.	Referred to ACER test developers for review. Test supervisor contacted IT technician and candidates were given extra time to cover lost time. The total occurrences were five. ACER would perform a psychometric analysis to investigate if there was any impact of the disruption on candidates' performance in the test. Test supervisor contacted ACER and extra time was provided to the candidates for each occurrence to cover time lost. There were 13 occurrences in total. ACER would perform a psychometric analysis to investigate if there was any impact of the disruption on candidates' performance in the test. Issues resolved by technician and given extra time to cover lost time. The candidate completed the test before the exam time elapsed. Candidates were admitted to the test after verification of their identity was made by the test supervisors.
Adelaide (Cliftons)	9	1 candidate presented ID that contained mismatches in DOB or names on the admission tickets. 2 candidates queried the test content. Sensor lights went off on one side of the room for no more than 5 seconds. Fire alarm sounded just prior to the pre-test registration. The building was evacuated with all students in attendance at that time for approximately 15 minutes. 1 candidate's mobile phone rang unanswered for less than 30 seconds.	Candidate was admitted to the test after verification of their identity was made by the test supervisor. Referred to ACER test developers for review. Room was still lit due to all the natural light coming in the windows. No complaints were received. No candidate made any reference to the evacuation affecting their concentration for the test. The candidate was given their phone and asked to turn it off under the supervision of the test

		1 candidate could not log in to the test using the username and password provided. 1 candidate, approved for reasonable adjustment, was under impression that she would be in a room on her own or with other students who also had extra time. Upon completion of the test, candidate complained that it was disruptive when all the remaining students left the testing room and the test supervisor announced to the candidates to leave the room quietly.	supervisor. The phone was placed back in the bag and the bag was placed out the front in the designated area. The candidate was moved to alternative PC and the candidate still could not log in. The test supervisor provided spare login details to avoid further delay. Test supervisor filed the incident report and asked her to contact ACER.
Ballarat	4	1 candidate arrived late.	Candidate was not admitted to the test.
Brisbane	7	1 candidate presented ID that contained mismatches in DOB or names on the admission tickets. 1 candidate presented ID that contained mismatches in name on the admission tickets. 1 candidate queried if a colon (:) needed to be included in the answer. 1 candidate presented a medical certificate stating that towels were needed to be placed under the left elbow due to a medical condition. 1 candidate raised her hand and asked for a bathroom break. She also requested access to her personal belongings for medication and other items because of sudden onset of medical condition. Female test supervisor checked the contents taken from the candidate's bag.	Candidate was admitted to the test after verification of their identity was made by the test supervisor. Candidate was admitted to the test after verification of the identity was made by the test supervisors. Test supervisor informed candidate to answer the question as best she could and an incident report would be filed. Test supervisor forwarded the medical certificate to the test venue manager and permission was given. Female test supervisor checked the contents taken from the candidate's bag before the candidate went to the bathroom and after the candidate had returned to the test room. The candidate took her medication under the supervision of the test supervisor. Additional 5 minutes were given to the candidate.

		1 candidate did not turn the phone off. There was a message tone from the bag.	Test supervisors were unable to identify the bag and there was no complaint from any candidate.
Canberra	0	NA	NA
Darwin	1	The screen showed a blank white page instead of displaying the login page.	Technical issues were resolved promptly and no candidate was impacted.
Hobart	1	1 candidate's computer turned off by itself.	Candidate was moved to alternative PC and candidate completed the test before the exam time elapsed and did not require extra time.
Melbourne	28	3 candidates experienced minor technical issues relating to navigation, onscreen calculator and screen brightness. 1 candidate approached a test supervisor to explain that she would like to go back and review her answers. However, the candidate had already clicked on 'finish the test' and 'return to test list' and was therefore unable to return to her answers. 1 candidate's mobile phone rang during the exam. Candidate had switched his phone off but the alarm function still went off. 1 candidate informed the Chief Supervisor that she was feeling unwell and requested extra time or a separate exam room. 1 candidate had a fainting spell in the registration area and was attended to by several staff members. She appeared visibly unwell and was having trouble keeping her balance. Candidate informed chief supervisor that she had a variety of serious health concerns, but expressed that she still wished to attempt the test.	Candidates were moved to alternative PCs and given extra time to cover lost time. Candidate was informed that she was unable to return to her answers as she had already clicked past the finish point. Candidate's bag was removed from the test room and returned to him at the end of the test. Candidate was informed that reasonable adjustments could only be granted by ACER and was advised to contact ACER for further advice. Candidate was placed in the smallest test room and was brought to the room a couple of minutes early to get settled. The following arrangements were made for the candidate (in consultation with ACER/Cliftons Exam Coordinator): -the candidate was seated in the smallest test room; -the candidate was physically assisted to the test room; and -the candidate was provided with a printed copy of all exam announcements (including the

		1 candidate suffered two coughing spells during the test session. 1 candidate seemed to have an incessant low bronchial cough and sniffing which appeared to be disturbing 2 candidates nearby. 2 candidates reported that they were disturbed by another candidate who had an incessant low bronchial cough and sniffing. 8 candidates queried the test content. 4 candidates did not bring appropriate photo ID to the test. 2 candidates presented for the wrong exam session (different date to that listed on their admission tickets) 1 candidate was granted a reasonable adjustment with 30 minutes extra time and a 5 minute rest break. All test supervisors were informed of this. However one supervisor initially asked the candidate to stop working at the conclusion of the test for all other candidates. 1 candidate presented at the registration area and requested to be provided with a paper test, as she was under	30 minute/10 minute warning announcements). Candidate was removed from the test room to recover and to minimise disruption to other candidates. The candidate eventually recovered and was able to return to the room and complete the test. No other candidates raised complaints about the brief disturbance. Candidate was offered tissues (accepted) and a mint lolly to soothe his throat (declined). He completed his test early and left the test room; test supervisors offered him a seat in the far back corner of the room for his next test session to minimise any potential disturbance to other candidates. Candidates were offered the option to move seats – one candidate accepted and was granted an additional 5 minutes to compensate for the disturbance. The other candidate declined the offer but told supervisors he would inform them if he wanted to be moved. Referred to ACER test developers for review. Candidates were not admitted to the test. Candidates were not admitted to the test. Other supervisors in the room quickly corrected the situation with minimal disruption to the candidate. Candidate completed the test with extra time with no issues or complaints. Candidate was informed that she was unable to submit the test on paper/unable to
--	--	--	--

		the impression that the test could be submitted on paper. She then asked if the test (computer-based) could be done from home and was informed in the negative. 1 candidate arrived early for her test session and attempted to enter one of the test rooms while the previous test session was still in progress.	complete the test from home as she was already registered to sit at a test centre. Candidate was asked if she would like to attempt the available computer-based test – was registered and issued with her login details. Candidate completed the test and was advised an incident report would be filed and submitted to ACER. Candidate was stopped by a test supervisor and prevented from entering the test room. Test supervisor explained to candidate that she would need to return to the registration area at the time listed on her admission ticket. Chief supervisor was notified and candidate was directed back to the registration area, where she was advised to return at the time showed on her admission ticket.
Newcastle	1	1 candidate presented ID that contained mismatches in DOB or names on the admission tickets.	Candidate was admitted to the test after verification of their identity was made by the test supervisors.
Perth	0	NA	NA
Sydney	22	5 candidates experienced minor technical issues including frozen screens and blank screens. 1 candidate complained at end of test that the font size on the screen was difficult to read. 1 candidate presented looking visibly unwell but remained at the venue to complete the test. 1 candidate claimed he thought there was a break	Technical issues were resolved promptly. In cases which the issues could not be resolved, candidates were moved to alternative PCs and given extra time to cover lost time. Candidate was advised that she should have alerted invigilator at the beginning of the test so that she could enlarge the font size by using her mouse. Candidate's font size was the same as other candidates. Chief supervisor was called as test supervisors were becoming increasingly concerned about the candidate's wellbeing. Candidate was removed from the room and supervised until she could return to the test room. Extra time was granted to compensate for the time spent outside the test room. Candidate was informed that all test instructions were read out

		between the calculator and non-calculator questions. He did not attempt the non-calculator questions. He claimed he was given the wrong information and requested to re-attempt the test without any extra charge. 4 candidates queried the test content. 2 candidates presented without an admission ticket. 5 candidates presented ID that contained mismatches in DOB or names on the admission tickets. 1 candidate was late for the test. 1 candidate presented for numeracy after missing his literacy exam. 1 candidate experienced issues with logging into the test.	as per the ACER test administration manual and the instructions were also clearly displayed in the tutorial of the test. Referred to ACER test developers for review. Candidates were not admitted to the test. Candidates were admitted to the test after verification of their identity was made by the test supervisors. Candidate was not admitted to the test. Candidate was advised to contact ACER for further advice. Candidate was provided with spare log-ins and completed the test.
Townsville	6	1 candidate reported that she did not like the seat she had been assigned – one at the very front of the room in the centre (she felt that everyone could see her screen). 1 candidate reported that the pen at her station had stopped working. The pen was immediately replaced but also caused trouble and was also replaced. 2 candidates had difficulty locating the test centre. Test supervisor reported that she felt offended by the verbal behaviour of 1 candidate, who was annoyed at the difficulty	Candidate was asked if she would prefer to sit somewhere else. Supervisor offered to move her to one of the spare locations but she chose to stay in the centre seat. Supervisor immediately replaced pens at candidate's workstation. A new batch of pens was brought into the test room and the whole box of spare original pens was removed and no longer used for this test window. All original pens were removed at the end of this test session and only the replacements were provided. Candidates were informed that the feedback would be provided to ACER so they can be addressed for future candidates. Supervisor spoke with candidate privately, informing her that her behaviour had offended another and that an incident report would be submitted to ACER.

		she experienced in locating the test centre. 1 candidate reported, at the conclusion of the test, that she had been distracted by the loud behaviour of another candidate who seemed quite frustrated.	Candidate understood and apologised for her behaviour. Candidate was advised that an incident report would be submitted to ACER.
Wollongong	2	1 candidate presented herself on the wrong test day. 1 candidate experienced issues with logging into the test.	Candidate was not admitted to the test. She was advised to contact ACER for further advice and instruction. Candidate was provided with spare log-ins and completed the tests.
Total	95		

Table 2: 2019 Test Window 1 Test Centre Incident Summary

Remote Proctoring

Table 3 below provides a breakdown of the total number of remote proctoring incidents.

Total sessions	No. of investigated incidents	Percentage	Cases with significant investigations	Percentage
2358	214	9%	12	0.5%

Table 3: 2019 Test Window 1 Remote Proctoring Incident Summary

ACER investigated all reported incidents by reviewing the ProctorU test session records, the ACER test delivery system records and server logs, and information from the candidates. There were 12 incidents which required significant investigation involving a detailed study of chat logs, test and time logs to determine the problem and the loss of test time where applicable.

The majority of incidents were mainly due to use of unauthorised programs, disconnection, test timing and unauthorised breaks. More minor incidents included the ACER test application not being installed properly, candidates not following test instructions or having prohibited items on their desks when they commenced the test sessions.

Where it was established that candidates lost test time due to technical issues, additional time in lieu was granted and candidates were able to reschedule the remaining proportion of their test time with ProctorU and complete their test during the test window.

Post-Test Activities

Review of Test Content

Following each test window, questions and feedback from candidates about the test content were summarised in a spreadsheet and referred to the lead test developers for their review and consideration. Most of these questions and feedback required no further action. From the 11419 test sessions conducted in 2019 Test Window 1, 18 questions were raised by candidates. Of these, 16 required no action. Action was taken for 1 numeracy question (involving a slight rephrasing of a sentence or the question) and 1 literacy question (involving the correction of the paragraph number to be referred to).

Table 4 below provides a detailed summary of the actions taken.

Test component	Count of Action
Literacy	
Make correction to the question itself	1
No action required	8
Literacy Total	9
Numeracy	
To rephrase the question itself	1
No action required	8
Numeracy Total	9
Grand Total	18

Table 4: Summary of actions taken based on candidate feedback in 2019 Test Window 1

It was noted by the test developers that it can be difficult to assess candidate queries where the feedback simply mentions 'ambiguity' or a general inability to understand the question. In the literacy test, common concerns raised by candidates include the intent of the questions being 'unclear', 'subjective' and the language used being 'confusing'. In the numeracy test, there were some concerns raised about possible missing images or answers not being found in the options provided. These were all found to be baseless. However, relevant actions were taken to rephrase a sentence contained in the question to provide greater clarity for a few queries that were raised by candidates.

It is evident that in most cases, candidates did not understand the questions, which is actually what the test is assessing. ACER reviews the instructions given to test administrators on an ongoing basis regarding test content queries to ensure that there is a mechanism for all candidates to provide feedback, whilst also maintaining the security and integrity of the test at all times.

Results

Test Window 1 results were released to DET, VIT and HEPs via the ACER portal on 25 March 2019. The portal allows DET to access de-identified results, and HEP and VIT nominated contacts are able to access, search and verify results for their students.

Results would be released to candidates via their ACER candidate account on 1 April 2019.

Candidates would receive an email notification from ACER about the release of results, with instructions on how to access the results and a document on how to interpret their report. The test website would also be updated to notify candidates that results had been released.

Social Media

ACER continues to monitor four restricted access Facebook groups set up for the test.

- LANTITE 2019: There are currently 2049 members and the group describes itself as 'A support and study group for students sitting the LANTITE teacher education tests';
- Numeracy and Literacy Test Support Group: There are currently 3066 members and they describe themselves as 'A support group for those who wish to develop their Literacy and Numeracy skills for the ACER exam';
- Lantite Test Help: There are currently 373 members and they describe themselves as 'LANTITE test help to help you in achieving your Lantite success dreams. We have written study guides and are here to answer any questions that you might have regarding your LANTITE tests'.
- Place to vent about LANTITE: There are currently 259 members and the group describes itself as 'This is a place for people to have an open discussion about their experiences with the LANTITE. This is a safe place where you won't be judged on your successes or failures.'

Once again, the majority of comments and posts from group members provide general exam information, general exam tips, free study links including specific literacy and numeracy support sites, NAPLAN references, links to the equivalent UK teacher test and so forth. There are also links to tutorial support, commercial providers selling a variety of literacy and numeracy practice test materials.

ACER has not found an instance where the content of the test or test security has been breached and will continue to monitor social media to ensure that candidates do not compromise the test integrity or engage in misconduct by publishing secure test information online.

Conclusion

After extensive discussions with DET and the ACER legal team in 2018, "special testing conditions" were renamed "reasonable adjustments" on both the test website and in the registration system to reflect a more accurate and better understanding of what special accommodations mean. ACER has not received any feedback on this change in terminology from the candidates or the HEPs.

ACER managed all the Victorian test centres in this test window and also set up a new test centre in Granville to replace the Parramatta test centre in New South Wales successfully. Following the success, ACER plans to set up and manage test centres in Geelong and Wagga Wagga in Test Window 2, and to use Saxons to provide most of the regional test centres in Queensland and in Bathurst. Under the current arrangement with Saxons, ACER would recruit and train test supervisors for the centres provided by Saxons.

It is hoped that this three-pronged approach in extending the test centre network will reduce the risk of the reliance on using only one contractor for the provision of all test centres and to deliver the test in a more cost-effective manner without affecting the quality standard.



LITERACY AND NUMERACY TEST FOR INITIAL TEACHER EDUCATION STUDENTS

2016 Technical Report

FOR THE DEPARTMENT OF EDUCATION AND
TRAINING, MAY 2017

CONTENTS

1	OVERVIEW	1
	1.1 Administration.....	1
	1.2 Candidate results	2
	1.3 Test design and in-test trialling for replenishment of item pool	3
	1.4 Validation of the standards.....	3
2	TEST ADMINISTRATION WINDOWS 1–4 IN 2016	4
	2.1 Demographic Characteristics of Candidates	4
	2.2 Demographic Characteristics of Candidates by Test Windows	12
	2.3 Accessibility and Accommodations	16
3	CANDIDATE PERFORMANCE	17
	3.1 Scale Score Distributions	17
	3.2 Candidate Scale Scores by Subscales and Subgroups.....	18
	3.3 Performance of Resit Candidates	25
4	CANDIDATE PERFORMANCE BY TEST WINDOWS.....	26
	4.1 Distributions of Candidate Scale Scores by Subscales and Subgroups.....	26
	4.2 Performance of Resit Candidates by Test Windows	28
5	PHASE 2 ITEM TRIAL ANALYSIS	30
	5.1 In-test trialling.....	30
	5.2 Item trial analysis	30
6	PIAAC ITEM ANALYSIS AND EQUATING.....	33
	6.1 Background.....	33
	6.2 Selection of PIAAC items.....	33
	6.3 Equating Methodology	33
	6.4 Outcome of Equating	34
7	PHASE 3 TEST DEVELOPMENT.....	36
8	CONCLUSION	36
9	REFERENCES	36
10	APPENDICES	37
	Appendix 1: Score Frequency Distribution	37
	Appendix 2: Performance by Demographic Characteristics and Test Windows..	40
	Appendix 3: Performance by Test Centres and Remote Proctoring by Test Window	47
	Appendix 4: Trial Item Analysis Tables.....	51
	Appendix 5: Analysis of Trial Item Differential Functioning (DIF)	59

List of Figures

Figure 1: Candidate pathways and outcomes in 2016	3
Figure 2: Distribution of candidate scale scores for literacy	17
Figure 3: Distribution of candidate scale scores for numeracy	18
Figure 4: Score distribution by gender and age	23
Figure 5: Score distribution by program type and year level	23
Figure 6: Score distribution by course category	24
Figure 7: Score distribution by test centres and remote proctoring	24
Figure 8: Score distribution by gender and test window	44
Figure 9: Score distribution by age group and test window	45
Figure 10: Score distribution by program type, year level and test window, literacy	45
Figure 11: Score distribution by program type, year level and test window, numeracy	46
Figure 12: Score distribution by course category and test window	46
Figure 13: Gender DIF plots	59
Figure 14: Age group DIF plots	61
Figure 15: Program type DIF plots	63
Figure 16: Course category DIF plots	65

List of Tables

Table 1: Number of sittings by location of test centre by testing window, including resits	1
Table 2: Summary sittings by testing window, including resits	1
Table 3: Number of sittings and pass rates by attempt and by testing window.....	2
Table 4: Demographic characteristics of candidates who sat the test in 2016	5
Table 5: Demographic characteristics of candidates who sat the test twice (first resit).....	6
Table 6: Demographic characteristics of candidates who sat the test three times (two resits).....	8
Table 7: Demographic characteristics of candidates who achieved no standard in 2016	9
Table 8: Number and proportion of candidates who participated by test centre	10
Table 9: Number and proportion of candidates who participated at test centres and by remote proctoring	11
Table 10: Number and proportion of candidates who achieved no standard by test centre	11
Table 11: Number and proportion of candidates who achieved no standard by test centre area and by remote proctoring	12
Table 12: Demographic characteristics of candidates by test windows – literacy	12
Table 13: Demographic characteristics of candidates by test windows - numeracy	13
Table 14: Number of candidates in test centres by test windows – literacy	14
Table 15: Number of candidates in test centres by test windows – numeracy	15
Table 16: Number and proportion of candidates in test centres and by remote proctoring by test window – literacy	15
Table 17: Number and proportion of candidates in test centres and by remote proctoring by test window – numeracy	16
Table 18: Candidate performance overall and by subscale for first-attempt candidates	19
Table 19: Performance by demographic characteristics for first-attempt candidates.....	19
Table 20: Performance by test centres and remote proctoring	22
Table 21: Number and proportion of first-attempt candidates who did not achieve the standard.....	24
Table 22: Resit candidate performance overall and by subscale.....	25
Table 23: Candidate performance by test window, overall and by subscale.....	26
Table 24: Number and proportion of candidates who did not achieve the standard by test window	27
Table 25: Resit candidate performance by test window, overall and by subscale	28
Table 26: Distribution of the trial literacy items across the scale	32
Table 27: Distribution of the trial numeracy items across the scale.....	32
Table 28: Test scale scores equivalent to the 70th percentile of the Australian adult population, and existing scale score standards.....	34
Table 29: Equivalent Australian adult population percentiles for existing test standards.....	34
Table 30: Comparison of outcomes between the existing and the higher standards for first-attempt candidates	35
Table 31: Comparison of outcomes between the existing and the higher standards for first to third attempt candidates	35
Table 32: Literacy score frequency distribution of first-attempt candidates.....	37
Table 33: Numeracy score frequency distribution of first-attempt candidates.....	38
Table 34: Performance by demographic characteristics in test window 1	40
Table 35: Performance by demographic characteristics in test window 2 (including resits)	41
Table 36: Performance by demographic characteristics in test window 3 (including resits)	42
Table 37: Performance by demographic characteristics in test window 4 (including resits)	43
Table 38: Performance by test centres and remote proctoring in test window 1.....	47
Table 39: Performance by test centres and remote proctoring in test window 2.....	48
Table 40: Performance by test centres and remote proctoring in test window 3.....	49
Table 41: Performance by test centres and remote proctoring in test window 4.....	50
Table 42: Summary of trial item analysis for literacy	51
Table 43: Summary of trial item analysis for numeracy	54
Table 44: List of potential gender DIF items	60
Table 45: List of potential age group DIF items	62

Table 46: List of potential program type DIF items.....	64
Table 47: List of potential course category DIF items	66

1 OVERVIEW

1.1 Administration

The Literacy and Numeracy Test for Initial Teacher Education Students (“the test”) was conducted across Australia for the first time, in four testing windows from May 2016 to December 2016. A total of 13,345¹ candidates sat at least one component of the test in this period. Of these, 748 candidates resat either the literacy test or numeracy test or both tests after their first attempt. Over 75% of the candidates were female, with slightly more candidates from primary than from secondary courses (41% and 38% respectively).

Students from 47 institutions sat the test in 2016. It was offered at test centres in all states and territories, or via remote proctoring under prescribed conditions. On first sitting, the majority of candidates sat the test at one of 21 test centres, while 18% of candidates took the test via remote proctoring. Candidates taking the test for a second or third time were more likely to do so via remote proctoring, rising to 43% for those requiring a third sitting.

For testing window 1, most test centres were located in capital cities, and for testing windows 2, 3 and 4, more test centres became available in regional cities. For testing windows 3 and 4, remote proctoring was made available to all candidates, regardless of their location. Table 1 shows the number of sittings by location for each testing window.

Table 1: Number of sittings by location of test centre by testing window, including resits²

Location of test centre	TW1		TW2		TW3		TW4	
	Literacy	Numeracy	Literacy	Numeracy	Literacy	Numeracy	Literacy	Numeracy
Capital cities	2787	2809	2367	2376	2223	2255	1658	1659
Regional cities	265	263	619	618	610	613	435	434
Remote proctoring	779	772	506	494	329	342	871	897
Total candidates	3831	3844	3492	3488	3162	3210	2964	2990

A small number of candidates chose to attempt only one of the component tests at each testing window as shown by Table 2.

Table 2: Summary sittings by testing window, including resits

	TW1	TW2	TW3	TW4	Total
Both literacy and numeracy	3765	3327	2980	2749	12821
Literacy only	66	165	182	215	628
Numeracy only	79	161	230	241	711
Total sittings	3910	3653	3392	3205	14160

Testing conditions were modified to accommodate 75 candidates with special needs in 2016. Accessible versions of the test were also available for candidates who required supportive technology such as a screen reader; however, the accessible versions were not required in 2016.

¹ In 2016, one candidate’s results were excluded from the reported totals due to a cancelled result from test misconduct

² Tables 1 and 2 include resit candidates in testing windows 2, 3 and 4

The administration of the test is described in Sections 1 and 2 of this report.

1.2 Candidate results

In 2016, 93.3% of the 13,083 candidates who attempted the literacy test achieved the standard at their first sitting. For numeracy, 92.5% of the 13,084 candidates achieved the standard at their first sitting. Of the candidates who sat *both* literacy and numeracy, 2.7% failed to achieve either standard at their first sitting.

Under the standard resit allowance, candidates who do not achieve the standard on their first attempt are permitted up to two additional sittings. On second and third sittings (resit 1 and resit 2) for literacy, of the 341 and 25 candidates, 67.7% and 76.0%, respectively, achieved the standard. On second and third sittings for numeracy, of the 406 and 40 candidates, 55.4% and 37.5%, respectively, achieved the standard. The two candidates who were permitted a fourth attempt at numeracy in 2016 both achieved the standard on their fourth attempt.

Table 3: Number of sittings and pass rates by attempt and by testing window

		TW1	Pass rate	TW2	Pass rate	TW3	Pass rate	TW4	Pass rate	Total	Pass rate
Literacy	First sitting	3831	94.5	3397	93.3	3017	91.9	2838	93.4	13083	93.3
	Second sitting	0	NA	95	67.4	137	67.2	109	68.8	341	67.7
	Third sitting	0	NA	0	NA	8	87.5	17	70.6	25	76.0
	Total sittings	3831		3492		3162		2964		13449	
Numeracy	First sitting	3844	93.1	3376	92.6	3041	92.2	2823	91.7	13084	92.5
	Second sitting	0	NA	112	54.5	154	53.2	140	58.6	406	55.4
	Third sitting	0	NA	0	NA	15	33.3	25	40.0	40	37.5
	Fourth sitting	0	NA	0	NA	0	NA	2	100.0	2	100.0
	Totals sittings	3844		3488		3210		2990		13532	

After one or more sittings, 95.2% of the 13,083 candidates who attempted the literacy component of the test had achieved the literacy standard. Similarly, after one or more sittings, 94.2% of the 13,084 candidates who attempted the numeracy component of the test had achieved the numeracy standard.

By the end of 2016, 13,345 candidates had attempted one or more components of the test. Most candidates (12,823) attempted both components. After resits, 11,828 candidates had achieved *both* the literacy and numeracy standards on the test. Of the remaining candidates, 634 had achieved the literacy standard only, 500 had achieved the numeracy standard only and 383 achieved no standard.³

Figure 1 summaries the different pathways of the 13,345 candidates.

³ 383 candidates achieved no standard because they attempted only one component of the test (one or more times in 2016) and did not achieve the standard on that component or because they attempted both components of the test (one or more times in 2016) and did not meet the standard on both components.

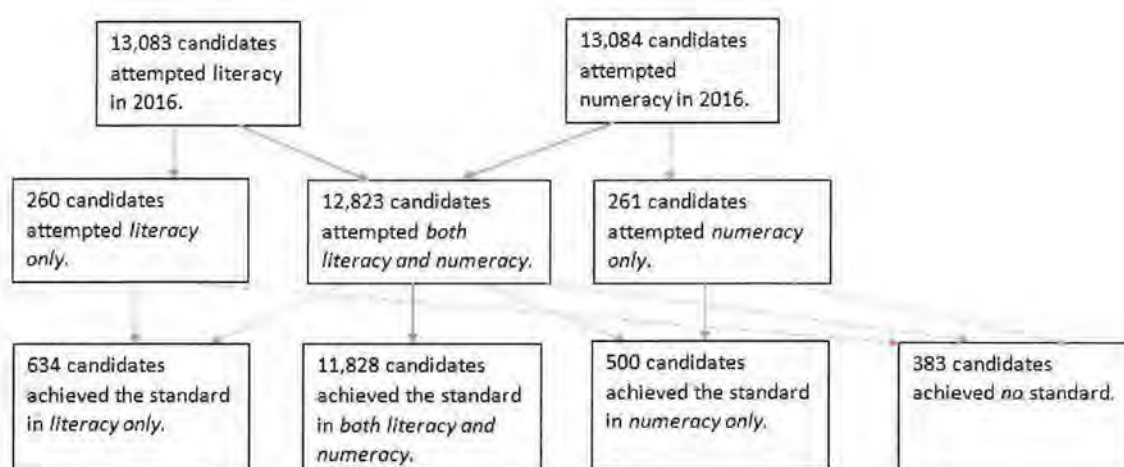


Figure 1: Candidate pathways and outcomes in 2016

Candidates' results for 2016 are described in Section 3 of this report.

1.3 Test design and in-test trialling for replenishment of item pool

Throughout 2016 there were 12 equivalent test forms for literacy and 14 equivalent test forms for numeracy. For literacy, each test form was five 12-item clusters totalling 60 items. For numeracy, the test was divided into two sections as follows: section 1 (calculator available) comprising four 12-item clusters (48 items), and section 2 (calculator not available) comprising two 6-item clusters (12 items) totalling 60 items.

In order to augment and replenish the pool of items available for the test in future administrations, items were trial-tested within the live instruments. These "Phase 2" items were administered in small clusters (one to five items) at the end of each test, and did not contribute to the candidates' scores. A total of 160 items for each of literacy and numeracy were administered in the in-test trial clusters. A minimum of 250 candidates was administered each of these items.

An example of one literacy test and one numeracy test with in-test trial clusters are shown below.

Literacy test 1	C1	C2	C3	C4	C5	C6	Trial C1
-----------------	----	----	----	----	----	----	----------

	Section 1					Section 2		
Numeracy test 1	CA1	CA2	CA3	CA4	Trial CA1	CN1	CN2	Trial CN1

The in-test trialling and its outcomes are described in Section 4 of this report.

1.4 Validation of the standards

In testing windows 3 and 4, selected items from the Programme for the International Assessment of Adult Competencies (PIAAC) were administered in similar in-test clusters. The purpose of this was to validate the provisional benchmarks and standards set by expert groups against empirical data on the literacy and numeracy of the whole Australian adult population. Fifteen (15) PIAAC literacy items and 16 PIAAC numeracy items were included in clusters in the test for this purpose. Like the in-test trial items, the PIAAC items did not contribute to the candidates' scores.

The equating study and its outcomes are described in Section 5 of this report.

2 TEST ADMINISTRATION WINDOWS 1–4 IN 2016

This section covers the demographic characteristics of candidates who sat the Literacy and Numeracy Test for Initial Teacher Education Students (the test) in 2016. For details on test centres, remote proctoring and other administrative matters, see each of the four 2016 administration reports.

2.1 Demographic Characteristics of Candidates

A total of 13,345 candidates sat the Literacy and/or Numeracy Test for Initial Teacher Education Students in 2016. The candidates were students of the following 47 institutions:

- Alphacrucis College
- Australian Catholic University
- Australian College of Physical Education
- Avondale College
- Central Queensland University
- Charles Darwin University
- Charles Sturt University
- Christian Heritage College
- Curtin University of Technology
- Deakin University
- Eastern College Victoria
- Edith Cowan University
- Excelsia College
- Federation University Australia
- Flinders University
- Griffith University
- Holmesglen TAFE
- James Cook University
- La Trobe University
- Macquarie University
- Melbourne Polytechnic
- Monash University
- Montessori World Educational Institute
- Morling College
- Murdoch University
- Queensland University of Technology
- RMIT University
- Southern Cross University
- Swinburne University of Technology
- The University of Adelaide
- The University of Melbourne
- The University of New England
- The University of New South Wales
- The University of Newcastle
- The University of Notre Dame Australia
- The University of Queensland
- The University of Sydney
- The University of Western Australia
- University of Canberra
- University of South Australia
- University of Southern Queensland
- University of Tasmania
- University of Technology Sydney
- University of the Sunshine Coast
- University of Western Sydney
- University of Wollongong
- Victoria University

The following demographic analysis separates the 2016 candidates into three groups for each component of the test: first-attempt candidates (all candidates who attempted the test), second-attempt candidates (first resit) and third-attempt candidates (second resit).

Table 4 shows the demographic characteristics of the first-attempt candidates for each component of the test in 2016.⁴ It shows that the majority of candidates (76%)⁵ were female, and most candidates (65%) were in the age group 17–25. The majority of candidates (62%) were enrolled in an undergraduate course. Of the total number of candidates who participated, the largest group of undergraduate candidates was those in their fourth year (42%), and the largest group of postgraduate candidates was those in their second year (16%). In terms of course category, candidates were mainly enrolled in Teacher Education: primary (41%), followed by Teacher Education: secondary (38%), Teacher Education: other (16%), Teacher Education: early childhood (5%) and Teacher Education: special education (1%).

Table 4: Demographic characteristics of candidates who sat the test in 2016

Characteristic	Category	Literacy		Numeracy	
		N	%	N	%
Gender	Female	9970	76.2	9991	76.4
	Male	3110	23.8	3090	23.6
	Indeterminate/intersex	3	0.0	3	0.0
Age	17–25	8544	65.3	8558	65.4
	26–30	1910	14.6	1903	14.5
	31–35	889	6.8	885	6.8
	36–40	683	5.2	682	5.2
	41–45	554	4.2	553	4.2
	46+	500	3.8	500	3.8
	Not disclosed	3	0.0	3	0.0
International Students	No	12647	96.7	12658	96.7
	Yes	436	3.3	426	3.3
English as a First Language	Yes	11737	89.7	11758	89.9
	No	1346	10.3	1326	10.1
Indigenous	No	12569	96.1	12576	96.1
	Yes	203	1.6	198	1.5
	Not disclosed	311	2.4	310	2.4
Residential Area	Metropolitan areas	10718	81.9	10731	82.0
	Provincial areas	2271	17.4	2261	17.3
	Remote areas	47	0.4	47	0.4
	International	32	0.2	30	0.2
	Invalid or Missing	15	0.1	15	0.1
Program Type	Undergraduate	8078	61.7	8078	61.7
	Postgraduate	5005	38.3	5006	38.3
Program Type by Year Level	Undergraduate first year	104	0.8	103	0.8
	Undergraduate second year	244	1.9	234	1.8
	Undergraduate third year	1942	14.8	1950	14.9
	Undergraduate fourth year	5438	41.6	5438	41.6

⁴ This cohort is referred to as the first-attempt candidates. Subsequent tables show demographic characteristics for candidates who resat the test

⁵ In the descriptive text accompanying the tables throughout the report, most percentages are rounded to the nearest whole per cent.

	Undergraduate fifth year or above	350	2.7	353	2.7
	Postgraduate first year	1843	14.1	1836	14.0
	Postgraduate second year	2048	15.7	2046	15.6
	Postgraduate third year	219	1.7	217	1.7
	Postgraduate fourth year	582	4.5	588	4.5
	Postgraduate fifth year or above	313	2.4	319	2.4
Course Category	Teacher education: early childhood	612	4.7	616	4.7
	Teacher education: primary	5358	41.0	5409	41.3
	Teacher education: secondary	4964	37.9	4935	37.7
	Teacher education: special education	87	0.7	86	0.7
	Teacher education: other	2062	15.8	2038	15.6

In 2016, a total of 366 candidates resat the literacy test and a total of 448 candidates resat the numeracy test after they did not achieve the standard on their first attempt. Among these resit candidates, 25 candidates sat the literacy test three times and 40 candidates sat the numeracy test three times. Two candidates were authorised to sit the numeracy test for a fourth time⁶.

Table 5 shows the demographic characteristics of the candidates who sat the test twice (first resit) during 2016. It shows that the majority of these candidates (82% for literacy, 92% for numeracy) were female, and candidates were mostly in the age group 17–25 (74% for literacy, 72% for numeracy). The majority of candidates (69% for literacy, 70% for numeracy) were enrolled in an undergraduate course. This pattern matches that for the first-attempt candidates, but each of the categories listed above (female; age 17–25; undergraduate) was more highly represented in the first-resit cohort than in the first-attempt cohort.

Table 5 also shows that the proportion of candidates for whom English was not their first language was 20% of the number of first-resit candidates for literacy and 13% for numeracy whereas the proportion was only 10% of the number of first-attempt candidates described in Table 4.

Table 5: Demographic characteristics of candidates who sat the test twice (first resit)

Characteristic	Category	Literacy		Numeracy	
		N	%	N	%
Gender	Female	280	82.1	373	91.9
	Male	61	17.9	33	8.1
	Indeterminate/intersex	0	0.0	0	0.0
Age	17–25	251	73.6	294	72.4
	26–30	60	17.6	56	13.8
	31–35	14	4.1	11	2.7
	36–40	6	1.8	12	3.0
	41–45	4	1.2	16	3.9
	46+	6	1.8	17	4.2
	Not disclosed	0	0.0	0	0.0

⁶ The demographic information for the two candidates who attempted the test four times was not reported because of small sample size.

International Students	No	324	95.0	399	98.3
	Yes	17	5.0	7	1.7
English as a First Language	Yes	272	79.8	355	87.4
	No	69	20.2	51	12.6
Indigenous	No	327	95.9	382	94.1
	Yes	5	1.5	8	2.0
	Not disclosed	9	2.6	16	3.9
Residential Area	Metropolitan areas	286	83.9	338	83.3
	Provincial areas	55	16.1	68	16.8
	Remote areas	0	0.0	0	0.0
	International	0	0.0	0	0.0
	Invalid or Missing	0	0.0	0	0.0
Program Type	Undergraduate	235	68.9	284	70.0
	Postgraduate	106	31.1	122	30.1
Program Type by Year Level	Undergraduate first year	0	0.0	0	0.0
	Undergraduate second year	0	0.0	3	0.7
	Undergraduate third year	30	8.8	22	5.4
	Undergraduate fourth year	190	55.7	238	58.6
	Undergraduate fifth year or above	15	4.4	21	5.2
	Postgraduate first year	21	6.2	21	5.2
	Postgraduate second year	40	11.7	48	11.8
	Postgraduate third year	7	2.1	6	1.5
	Postgraduate fourth year	25	7.3	26	6.4
	Postgraduate fifth year or above	13	3.8	21	5.2
Course Category	Teacher education: early childhood	26	7.6	32	7.9
	Teacher education: primary	143	41.9	185	45.6
	Teacher education: secondary	85	24.9	88	21.7
	Teacher education: special education	3	0.9	5	1.2
	Teacher education: other	84	24.6	96	23.7

Table 6 shows the demographic characteristics of candidates who participated in the test three times (two resits) during 2016. It shows that almost all these candidates were female (92% for literacy, 93% for numeracy), and again mostly in the age group 17–25 (76% for literacy, 75% for numeracy). The majority of candidates (72% for literacy, 55% for numeracy) were enrolled in an undergraduate course. This pattern also matches that for the first-attempt candidates, but each of the categories listed above (female; age 17–25; undergraduate) was more highly represented in the two resit cohort than in the first-attempt cohort.

Table 6 also shows that the proportion of candidates for whom English was not their first language was 32% of twice resitting candidates for literacy and 25% for numeracy whereas the proportion was only 10% of the ‘one sit only’ candidates described in Table 4.

Table 6: Demographic characteristics of candidates who sat the test three times (two resits)

Characteristic	Category	Literacy		Numeracy	
		N	%	N	%
Gender	Female	23	92.0	37	92.5
	Male	2	8.0	3	7.5
	Indeterminate/intersex	0	0.0	0	0.0
Age	17–25	19	76.0	30	75.0
	26–30	5	20.0	7	17.5
	31–35	1	4.0	0	0.0
	36–40	0	0.0	0	0.0
	41–45	0	0.0	2	5.0
	46+	0	0.0	1	2.5
	Not disclosed	0	0.0	0	0.0
International Students	No	25	100.0	39	97.5
	Yes	0	0.0	1	2.5
English as a First Language	Yes	17	68.0	30	75.0
	No	8	32.0	10	25.0
Indigenous	No	24	96.0	35	87.5
	Yes	0	0.0	2	5.0
	Not disclosed	1	4.0	3	7.5
Residential Area	Metropolitan areas	23	92.0	34	85.0
	Provincial areas	2	8.0	6	15.0
	Remote areas	0	0.0	0	0.0
	International	0	0.0	0	0.0
	Invalid or Missing	0	0.0	0	0.0
Program Type	Undergraduate	19	76.0	26	65.0
	Postgraduate	6	24.0	14	35.0
Program Type by Year Level	Undergraduate first year	0	0.0	0	0.0
	Undergraduate second year	0	0.0	0	0.0
	Undergraduate third year	1	4.0	1	2.5
	Undergraduate fourth year	18	72.0	22	55.0
	Undergraduate fifth year or above	0	0.0	3	7.5
	Postgraduate first year	0	0.0	2	5.0
	Postgraduate second year	3	12.0	6	15.0
	Postgraduate third year	0	0.0	0	0.0
	Postgraduate fourth year	3	12.0	5	12.5
	Postgraduate fifth year or above	0	0.0	1	2.5
Course Category	Teacher education: early childhood	0	0.0	2	5.0
	Teacher education: primary	13	52.0	23	57.5
	Teacher education: secondary	5	20.0	4	10.0
	Teacher education: special education	0	0.0	0	0.0
	Teacher education: other	7	28.0	11	27.5

Table 7 shows the demographic characteristics for the 383 candidates who achieved no standard at the end of 2016. These candidates may have had one, two or three attempts at the test. Thirteen (13) attempted literacy only and 38 attempted numeracy only and did not meet the test standard. The majority (274) attempted both literacy and numeracy and did not meet the test standard. It is expected that some of these candidates will resit the test again in 2017.

The demographics of this group is similar to those described in Table 6 (the candidates who had three test attempts in 2016). The proportions of female candidates, undergraduate candidates and candidates from early childhood courses were higher in this group of candidates than in those who achieved the standard in one or both components.

Table 7: Demographic characteristics of candidates who achieved no standard in 2016

Characteristic	Category	Literacy		Numeracy	
		N	%	N	%
Gender	Female	287	90.3	312	92.0
	Male	30	9.4	27	8.0
	Indeterminate/intersex	1	0.3		
Age	17–25	214	67.3	231	68.1
	26–30	48	15.1	45	13.3
	31–35	12	3.8	17	5.0
	36–40	14	4.4	13	3.8
	41–45	19	6.0	22	6.5
	46+	11	3.5	11	3.2
	Not disclosed	0	0.0	0	0.0
International Students	No	303	95.3	327	96.5
	Yes	15	4.7	12	3.5
English as a First Language	Yes	238	74.8	268	79.1
	No	80	25.2	71	20.9
Indigenous	No	298	93.7	321	94.7
	Yes	12	3.8	13	3.8
	Not disclosed	8	2.5	5	1.5
Residential Area	Metropolitan areas	261	82.1	284	83.8
	Provincial areas	56	17.6	54	15.9
	Remote areas	1	0.3	1	0.3
	International	0	0.0	0	0.0
	Invalid	0	0.0	0	0.0
Program Type	Undergraduate	214	67.3	230	67.8
	Postgraduate	104	32.7	109	32.2
Program Type by Year Level	Undergraduate first year	10	3.1	9	2.7
	Undergraduate second year	8	2.5	5	1.5
	Undergraduate third year	59	18.6	69	20.4
	Undergraduate fourth year	126	39.6	135	39.8
	Undergraduate fifth year or above	11	3.5	12	3.5
	Postgraduate first year	27	8.5	29	8.6
	Postgraduate second year	30	9.4	30	8.8
	Postgraduate third year	8	2.5	9	2.7

	Postgraduate fourth year	25	7.9	27	8.0
	Postgraduate fifth year or above	14	4.4	14	4.1
Course Category	Teacher education: early childhood	41	12.9	42	12.4
	Teacher education: primary	134	42.1	159	46.9
	Teacher education: secondary	71	22.3	67	19.8
	Teacher education: special education	4	1.3	2	0.6
	Teacher education: other	68	21.4	69	20.4

Table 8 shows the number and proportion of candidates participating in each test centre. At their first attempt, about 76% of candidates sat the test at three test centres: Melbourne CBD, Sydney CBD and Brisbane CBD.

Table 8: Number and proportion of candidates who participated by test centre⁷

Test Centre	First Attempt				Second Attempt				Third Attempt			
	Literacy		Numeracy		Literacy		Numeracy		Literacy		Numeracy	
	N	%	N	%	N	%	N	%	N	%	N	%
Adelaide CBD	131	1.2	131	1.2	1	0.4	2	0.7	0	0.0	0	0.0
Albury	150	1.4	142	1.3	1	0.4	4	1.5	1	6.7	1	5.0
Armidale	99	0.9	98	0.9	0	0.0	1	0.4	0	0.0	0	0.0
Ballarat	173	1.6	174	1.6	4	1.7	3	1.1	0	0.0	0	0.0
Bathurst	64	0.6	65	0.6	1	0.4	1	0.4	0	0.0	0	0.0
Brisbane CBD	1327	12.4	1326	12.4	18	7.5	22	8.0	1	6.7	1	5.0
Bundaberg	37	0.4	38	0.4	0	0.0	0	0.0	0	0.0	0	0.0
Canberra CBD	317	3.0	316	2.9	0	0.0	4	1.5	0	0.0	1	5.0
Darwin	35	0.3	35	0.3	0	0.0	0	0.0	0	0.0	0	0.0
Hobart	50	0.5	52	0.5	0	0.0	0	0.0	0	0.0	0	0.0
Melbourne CBD	3728	34.8	3747	34.9	105	43.8	110	40.2	8	53.3	10	50.0
Mildura	23	0.2	23	0.2	3	1.3	4	1.5	0	0.0	0	0.0
Newcastle	569	5.3	573	5.3	17	7.1	14	5.1	0	0.0	0	0.0
Parramatta	362	3.4	356	3.3	5	2.1	11	4.0	0	0.0	0	0.0
Perth CBD	124	1.2	123	1.2	2	0.8	1	0.4	0	0.0	1	5.0
Sunshine Coast/Maroochydore	127	1.2	122	1.1	0	0.0	0	0.0	0	0.0	0	0.0
Sydney CBD	3107	29.0	3125	29.1	76	31.7	85	31.0	5	33.3	6	30.0
Townsville	10	0.1	10	0.1	0	0.0	0	0.0	0	0.0	0	0.0
Wagga Wagga	86	0.8	87	0.8	4	1.7	5	1.8	0	0.0	0	0.0
Warrnambool	46	0.4	50	0.5	1	0.4	1	0.4	0	0.0	0	0.0
Wollongong	144	1.3	139	1.3	2	0.8	6	2.2	0	0.0	0	0.0
Total	10,709	100.0	10,732	100.0	240	100.0	274	100.0	15	100.0	20	100.0

Table 9 shows the number and proportion of candidates participating in test centres and by remote proctoring. Of the total number of first-attempt candidates, 18% sat the test via remote proctoring. Access to remote proctoring for all candidates was made available from testing window three in 2016. Remote proctoring was increasingly used among resit candidates. For example, in the literacy test, about 30% of second-attempt candidates and 40% of third-attempt candidates chose to use remote proctoring.

⁷ Some test centres were not offered in particular test windows. For example in windows 3 and 4, candidates would not have been able to resit the test at the Hobart test centre as this test centre was not offered.

Table 9: Number and proportion of candidates who participated at test centres and by remote proctoring

Location of testing	First Attempt				Second Attempt				Third Attempt			
	Literacy		Numeracy		Literacy		Numeracy		Literacy		Numeracy	
	N	%	N	%	N	%	N	%	N	%	N	%
Test Centre	10709	81.9	10732	82.0	240	70.4	274	67.5	15	60.0	20	50.0
– Capital Cities	8819	67.4	8855	67.7	202	59.2	224	55.2	14	56.0	19	47.5
– Regional Cities	1890	14.5	1877	14.4	38	11.1	50	12.3	1	4.0	1	2.5
Remote Proctoring	2374	18.2	2352	18.0	101	29.6	132	32.5	10	40.0	20	50.0
Total	13083	100.0	13084	100.0	341	100.0	406	100.0	25	100.0	40	100.0

Table 10 shows the location by test centre of the 383 candidates who achieved no standard in 2016. The distribution of these candidates by test centre location is very similar to other candidates.

Table 10: Number and proportion of candidates who achieved no standard by test centre

Test Centre	Literacy		Numeracy	
	N	%	N	%
Adelaide CBD	4	1.5	3	1.1
Albury	1	0.4	2	0.7
Armidale	1	0.4	1	0.4
Ballarat	5	1.9	4	1.4
Bathurst	2	0.8	3	1.1
Brisbane CBD	30	11.5	29	10.2
Bundaberg	3	1.1	3	1.1
Canberra CBD	10	3.8	10	3.5
Darwin	1	0.4	1	0.4
Hobart	0	0.0	0	0.0
Melbourne CBD	101	38.7	113	39.8
Mildura	0	0.0	0	0.0
Newcastle	8	3.1	9	3.2
Parramatta	17	6.5	20	7.0
Perth CBD	3	1.1	3	1.1
Sunshine Coast/Maroochydore	2	0.8	2	0.7
Sydney CBD	68	26.1	76	26.8
Townsville	0	0.0	0	0.0
Wagga Wagga	1	0.4	1	0.4
Warrnambool	1	0.4	1	0.4
Wollongong	3	1.1	3	1.1
Subtotal (test centre)	261	100.0	284	100.0

Table 11 shows the number and proportion of the 383 candidates who achieved no standard in 2016 by test centre area and remote proctoring. The distribution of these candidates by location and remote proctoring is very similar to other candidates.

Table 11: Number and proportion of candidates who achieved no standard by test centre area and by remote proctoring

Location of testing	Literacy		Numeracy	
	N	%	N	%
– Capital Cities	217	68.2	235	69.3
– Regional Cities	44	13.8	49	14.5
Remote Proctoring	57	17.9	55	16.2

2.2 Demographic Characteristics of Candidates by Test Windows

For both literacy and numeracy, the number of candidates in 2016 was greatest in test window 1 at almost 4000 and then declined steadily to approximately 3000 in test window 4. The actual numbers for test windows 1–4 were: 3831, 3492, 3162 and 2964 for literacy and 3844, 3488, 3210, and 2990 for numeracy. Table 12 and Table 13 show the number and proportion of candidates who participated in the test by demographic characteristics for each testing window.

The distribution of candidates by demographic characteristics for literacy was similar to the distribution for numeracy. For example, Table 12 shows that for literacy, across four test windows, the majority of candidates were female (in the range of 72% to 80%), and candidates were mostly in the age group 17–25 (in the range of 60% to 69%). The majority of candidates sitting the literacy test in each testing window were enrolled in an undergraduate course. The proportion of candidates in each course category remained relatively constant across the testing windows. These demographic characteristics of candidates in each test window were consistent to the demographic characteristics of candidates overall.

Over testing windows 1–4, the proportion of fourth-year undergraduates sitting the literacy test decreased from 65% to 24% while the proportion of third-year undergraduates increased from 1% to 29%. A similar pattern of change was observed in postgraduate candidates taking the literacy test. The proportion of second-year postgraduate candidates dropped from test window 1 (22%) to test window 4 (9%), while the proportion of first-year postgraduate candidates increased from test window 1 (5%) to test window 4 (22%). It is noted that test window 1 was an early priority window offered for final year undergraduate and postgraduate students ahead of national implementation commencing 1 July 2016.

Table 12: Demographic characteristics of candidates by test windows – literacy

Characteristic	Category	TW1		TW2		TW3		TW4	
		N	%	N	%	N	%	N	%
Gender	Female	3052	79.7	2712	77.7	2278	72.0	2231	75.3
	Male	778	20.3	778	22.3	884	28.0	733	24.7
	Indeterminate/intersex	1	0.0	2	0.1	0	0.0	0	0.0
Age	17–25	2624	68.5	2299	65.8	2111	66.8	1780	60.1
	26–30	493	12.9	564	16.2	454	14.4	464	15.7
	31–35	247	6.5	216	6.2	208	6.6	233	7.9
	36–40	192	5.0	180	5.2	155	4.9	162	5.5
	41–45	152	4.0	123	3.5	117	3.7	166	5.6
	46+	123	3.2	109	3.1	117	3.7	157	5.3
	Not disclosed	0	0.0	1	0.0	0	0.0	2	0.1
International Students	No	3722	97.2	3360	96.2	3054	96.6	2860	96.5
	Yes	109	2.9	132	3.8	108	3.4	104	3.5
English as a First Language	Yes	3539	92.4	3058	87.6	2805	88.7	2624	88.5
	No	292	7.6	434	12.4	357	11.3	340	11.5
Indigenous	No	3676	96.0	3371	96.5	3032	95.9	2841	95.9
	Yes	69	1.8	43	1.2	49	1.6	47	1.6
	Not disclosed	86	2.2	78	2.2	81	2.6	76	2.6
Residential Area	Metropolitan areas	3132	81.8	2962	84.8	2603	82.3	2330	78.6
	Provincial areas	672	17.5	519	14.9	544	17.2	593	20.0
	Remote areas	15	0.4	5	0.1	7	0.2	20	0.7

	International	10	0.3	2	0.1	6	0.2	14	0.5
	Invalid or Missing	2	0.1	4	0.1	2	0.1	7	0.2
Program Type	Undergraduate	2509	65.5	2079	59.5	1992	63.0	1752	59.1
	Postgraduate	1322	34.5	1413	40.5	1170	37.0	1212	40.9
Program Type by Year Level	Undergraduate first year	0	0.0	39	1.1	36	1.1	29	1.0
	Undergraduate second year	12	0.3	74	2.1	78	2.5	80	2.7
	Undergraduate third year	22	0.6	450	12.9	651	20.6	850	28.7
	Undergraduate fourth year	2475	64.6	1335	38.2	1119	35.4	717	24.2
	Undergraduate fifth year or above	0	0.0	181	5.2	108	3.4	76	2.6
	Postgraduate first year	200	5.2	533	15.3	487	15.4	644	21.7
	Postgraduate second year	858	22.4	607	17.4	357	11.3	269	9.1
	Postgraduate third year	71	1.9	39	1.1	61	1.9	55	1.9
	Postgraduate fourth year	193	5.0	122	3.5	147	4.7	148	5.0
	Postgraduate fifth year or above	0	0.0	112	3.2	118	3.7	96	3.2
Course Category	Teacher education: early childhood	164	4.3	199	5.7	124	3.9	151	5.1
	Teacher education: primary	1849	48.3	1337	38.3	1190	37.6	1138	38.4
	Teacher education: secondary	1296	33.8	1347	38.6	1187	37.5	1224	41.3
	Teacher education: special education	35	0.9	19	0.5	19	0.6	17	0.6
	Teacher education: other	487	12.7	590	16.9	642	20.3	434	14.6

The observations and patterns described above for literacy candidates across the four testing windows are also pertinent for numeracy as shown in Table 13.

Table 13: Demographic characteristics of candidates by test windows - numeracy

Characteristic	Category	TW1		TW2		TW3		TW4	
		N	%	N	%	N	%	N	%
Gender	Female	3073	79.9	2721	78.0	2333	72.7	2276	76.1
	Male	770	20.0	766	22.0	877	27.3	713	23.9
	Indeterminate/intersex	1	0.0	1	0.0	0	0.0	1	0.0
Age	17-25	2635	68.6	2299	65.9	2148	66.9	1801	60.2
	26-30	492	12.8	554	15.9	454	14.1	467	15.6
	31-35	248	6.5	220	6.3	209	6.5	219	7.3
	36-40	191	5.0	180	5.2	158	4.9	165	5.5
	41-45	155	4.0	124	3.6	122	3.8	170	5.7
	46+	123	3.2	110	3.2	119	3.7	166	5.6
	Not disclosed	0	0.0	1	0.0	0	0.0	2	0.1
International Students	No	3736	97.2	3366	96.5	3106	96.8	2890	96.7
	Yes	108							
English as a First Language	Yes	3555	92.5	3074	88.1	2863	89.2	2653	88.7
	No	289	7.5	414	11.9	347	10.8	337	11.3
Indigenous	No	3686	95.9	3375	96.8	3071	95.7	2863	95.8
	Yes	69	1.8	39	1.1	52	1.6	48	1.6
	Not disclosed	89	2.3	74	2.1	87	2.7	79	2.6
Residential Area	Metropolitan areas	3151	82.0	2969	85.1	2640	82.2	2345	78.4
	Provincial areas	666	17.3	509	14.6	554	17.3	606	20.3

	Remote areas	15	0.4	5	0.1	8	0.3	19	0.6
	International	10	0.3	1	0.0	6	0.2	13	0.4
	Invalid or Missing	2	0.1	4	0.1	2	0.1	7	0.2
Program Type	Undergraduate	2520	65.6	2070	59.4	2033	63.3	1766	59.1
	Postgraduate	1324	34.4	1418	40.7	1177	36.7	1224	40.9
Program Type by Year Level	Undergraduate first year	0	0.0	38	1.1	36	1.1	29	1.0
	Undergraduate second year	12	0.3	69	2.0	76	2.4	80	2.7
	Undergraduate third year	27	0.7	444	12.7	665	20.7	837	28.0
	Undergraduate fourth year	2481	64.5	1338	38.4	1141	35.6	738	24.7
	Undergraduate fifth year or above	0	0.0	181	5.2	115	3.6	82	2.7
	Postgraduate first year	200	5.2	533	15.3	488	15.2	638	21.3
	Postgraduate second year	854	22.2	614	17.6	368	11.5	265	8.9
	Postgraduate third year	70	1.8	37	1.1	59	1.8	57	1.9
	Postgraduate fourth year	200	5.2	120	3.4	143	4.5	156	5.2
	Postgraduate fifth year or above	0	0.0	114	3.3	119	3.7	108	3.6
Course Category	Teacher education: early childhood	165	4.3	199	5.7	125	3.9	161	5.4
	Teacher education: primary	1869	48.6	1352	38.8	1231	38.4	1166	39.0
	Teacher education: secondary	1292	33.6	1330	38.1	1189	37.0	1216	40.7
	Teacher education: special education	35	0.9	19	0.5	20	0.6	17	0.6
	Teacher education: other	483	12.6	588	16.9	645	20.1	430	14.4

Table 14 and Table 15 show the number and proportion of candidates participating in each test centre in test windows 1–4 for literacy and for numeracy respectively. It can be seen that some test centres in some locations were not available for all test windows.

Table 14: Number of candidates in test centres by test windows – literacy

Test Centre	TW1		TW2		TW3		TW4	
	N	%	N	%	N	%	N	%
Adelaide CBD	26	0.9	26	0.9	50	1.8	30	1.4
Albury	57	1.9	36	1.2	40	1.4	19	0.9
Armidale	NA ⁸	NA	36	1.2	25	0.9	38	1.8
Ballarat	NA	NA	40	1.3	80	2.8	57	2.7
Bathurst	NA	NA	18	0.6	28	1.0	19	0.9
Brisbane CBD	407	13.3	236	7.9	329	11.6	374	17.9
Bundaberg	NA	NA	9	0.3	8	0.3	20	1.0
Canberra CBD	191	6.3	58	1.9	39	1.4	29	1.4
Darwin	7	0.2	5	0.2	13	0.5	10	0.5
Hobart	44	1.4	6	0.2	NA	NA	NA	NA
Melbourne CBD	998	32.7	1118	37.4	1029	36.3	696	33.3
Mildura	NA	NA	19	0.6	7	0.3	NA	NA
Newcastle	198	6.5	153	5.1	155	5.5	80	3.8

⁸ NA indicates that the test centre was not available for that particular testing window

Parramatta	NA	NA	137	4.6	137	4.8	93	4.4
Perth CBD	45	1.5	32	1.1	17	0.6	32	1.5
Sunshine Coast/Maroochydore	NA	NA	46	1.5	21	0.7	60	2.9
Sydney CBD	1069	35.0	886	29.7	746	26.3	487	23.3
Townsville	10	0.3	NA	NA	NA	NA	NA	NA
Wagga Wagga	NA	NA	40	1.3	31	1.1	19	0.9
Warrnambool	NA	NA	28	0.9	19	0.7	NA	NA
Wollongong	NA	NA	57	1.9	59	2.1	30	1.4
Total	3052	100.0	2986	100.0	2833	100.0	2093	100.0

Table 15: Number of candidates in test centres by test windows – numeracy

Test Centre	TW1		TW2		TW3		TW4	
	N	%	N	%	N	%	N	%
Adelaide CBD	28	0.9	25	0.8	51	1.8	29	1.4
Albury	54	1.8	35	1.2	40	1.4	18	0.9
Armidale	NA ⁹	NA	36	1.2	25	0.9	38	1.8
Ballarat	NA	NA	40	1.3	80	2.8	57	2.7
Bathurst	NA	NA	18	0.6	29	1.0	19	0.9
Brisbane CBD	407	13.3	236	7.9	334	11.7	372	17.8
Bundaberg	NA	NA	9	0.3	9	0.3	20	1.0
Canberra CBD	191	6.2	58	1.9	43	1.5	30	1.4
Darwin	7	0.2	5	0.2	13	0.5	10	0.5
Hobart	44	1.4	8	0.3	NA	NA	NA	NA
Melbourne CBD	1008	32.8	1124	37.5	1036	36.1	699	33.4
Mildura	NA	NA	19	0.6	8	0.3	NA	NA
Newcastle	199	6.5	153	5.1	156	5.4	79	3.8
Parramatta	NA	NA	136	4.5	136	4.7	95	4.5
Perth CBD	45	1.5	32	1.1	16	0.6	32	1.5
Sunshine Coast/Maroochydore	NA	NA	45	1.5	18	0.6	59	2.8
Sydney CBD	1079	35.1	888	29.7	762	26.6	487	23.3
Townsville	10	0.3	0	NA	NA	NA	NA	NA
Wagga Wagga	NA	NA	40	1.3	33	1.2	19	0.9
Warrnambool	NA	NA	30	1.0	21	0.7	NA	NA
Wollongong	NA	NA	57	1.9	58	2.0	30	1.4
Total	3072	100.0	2994	100.0	2868	100.0	2093	100.0

Table 16 and Table 17 show the number and proportion of candidates participating in test centres and by remote proctoring in each test window for literacy and numeracy respectively. It can be seen that the proportion of candidates sitting in test centres in capital cities ranged between 68% and 73% for test windows 1 to 3 but dropped to 56% for test window 4. For test centres in regional cities, the proportion rose from only 7% for test window 1 to a peak of 19% for test window 3. For remote proctoring, the proportion fluctuated between 10% (test window 3) and 30% (test window 4).

Table 16: Number and proportion of candidates in test centres and by remote proctoring by test window – literacy

Location of testing	TW1		TW2		TW3		TW4	
	N	%	N	%	N	%	N	%
Test Centre	3052	79.7	2986	85.5	2833	89.6	2093	70.6
– Capital Cities	2787	72.8	2367	67.8	2223	70.3	1658	55.9
– Regional Cities	265	6.9	619	17.7	610	19.3	435	14.7
Remote Proctoring	779	20.3	506	14.5	329	10.4	871	29.4
Total	3831	100.0	3492	100.0	3162	100.0	2964	100.0

⁹ NA indicates that the test centre was not available for that particular testing window

Table 17: Number and proportion of candidates in test centres and by remote proctoring by test window – numeracy

Location of testing	TW1		TW2		TW3		TW4	
	N	%	N	%	N	%	N	%
Test Centre	3072	79.9	2994	85.8	2868	89.4	2093	70.0
– Capital Cities	2809	73.1	2376	68.1	2255	70.3	1659	55.5
– Regional Cities	263	6.8	618	17.7	613	19.1	434	14.5
Remote Proctoring	772	20.1	494	14.2	342	10.7	897	30.0
Total	3844	100.0	3488	100.0	3210	100.0	2990	100.0

2.3 Accessibility and Accommodations

In 2016, seventy-five candidates (0.5%) required special testing conditions as follows:

Types of conditions:

- Dyslexia, Dyscalculia, ADD
- Anxiety, Depression, Panic Attacks
- PTSD
- Rheumatoid Arthritis
- Recovery from Stroke
- Visual Migraine
- Temporary physical conditions – e.g. broken leg, wrist, back injury
- Diabetes
- Glaucoma suffering from visual fatigue
- Partial Bilateral hearing loss
- Charcot-Marie-Tooth Syndrome and Muscular Dystrophy

Types of accommodations granted:

- Extra time (20 minutes per test component)
- Rest breaks (5 mins per test component)
- Seated near bathroom
- Seated at the front of the test room (for hearing loss)
- Seated in a quiet room
- Permission to bring blood-insulin monitor, food and drink
- Permission to take medication
- Permission to stand and stretch
- Emergency Action Plan (for Epilepsy – seizures)

Online test versions (one for each of literacy and numeracy) that are accessible to vision impaired candidates using screen readers were created using a selection of the most accessible items. Some items were modified in order to be more amenable to the most common screen readers. For example, graphs were replaced with tables of the data, and additional HTML (alt-text visible only to screen readers) was added. An accessible set of online sample items was also created for each component of the test.

An accessibility consultant was employed in 2016 to undertake and report on the online accessible tests and sample items using a screen reader. Based upon their report, further modifications were made including changes both to the delivery system and to specific test items. The consultant returned in 2017 to evaluate the effectiveness of the modifications. An ACER IT staff member worked closely with the consultant on this second visit. A second report was received and further modifications were made. The accessible versions of the test were not required by any candidate in 2016. It is anticipated that both online and paper versions of the accessible test forms will be utilised by candidates during 2017.

3 CANDIDATE PERFORMANCE

This section describes the performance of candidates who participated in the test in 2016. The analysis breaks the cohort of candidates into two groups: first-attempt candidates (the majority) and those who did not achieve the standard at their first attempt and resat the test. It presents the distributions of candidate performance overall, by subscale and by their collected demographic information: gender, age group, program type, program type by year level, course category, and locations of the test centres.

3.1 Scale Score Distributions

Figure 2 and Figure 3 present the distributions of first-attempt candidate performance on the literacy test and numeracy test respectively. The vertical line in each figure represents the standard for that component of the test.

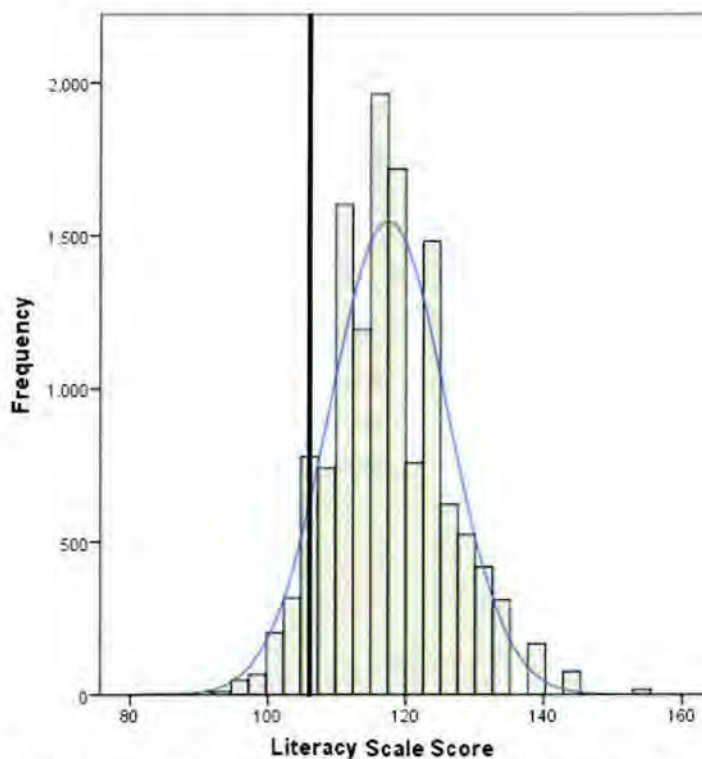


Figure 2: Distribution of candidate scale scores for literacy¹⁰

¹⁰ In 2016, the literacy standard scale score was 106

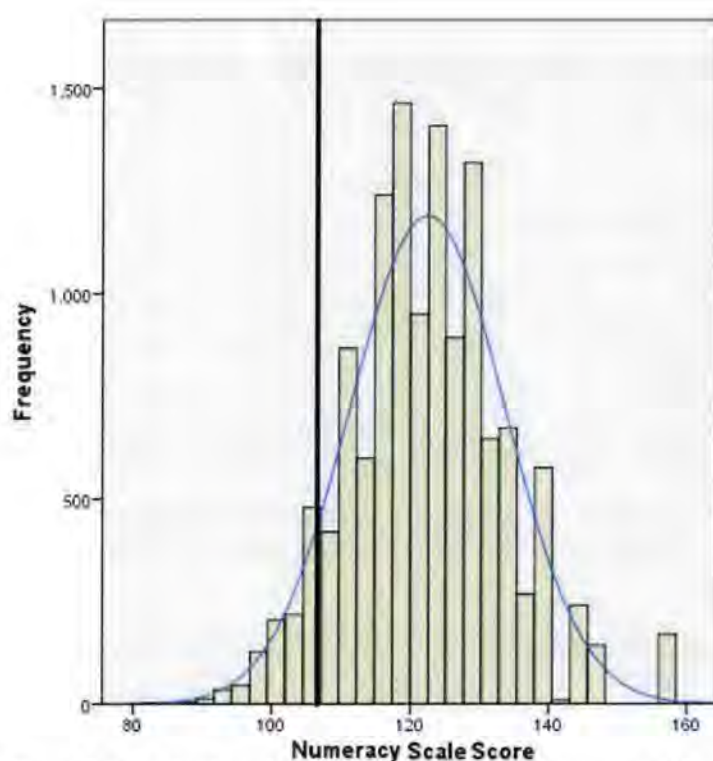


Figure 3: Distribution of candidate scale scores for numeracy¹¹

Figure 2 and Figure 3 show that scores in both tests are approximately normally distributed and that the tests spread candidates acceptably across the score scales. For both literacy and numeracy, the majority of candidates achieved scale scores above the standard at their first attempt, and it can be seen that a proportion of candidates achieved scale scores below the standard at their first attempt.

3.2 Candidate Scale Scores by Subscales and Subgroups

Table 18 shows the performance of first-attempt candidates in 2016. It shows the number of candidates, mean scale scores and standard deviation of the scale scores, overall and by subscale. The pass rates for the literacy and numeracy tests for this cohort of candidates are also shown in this table. The overall mean scale score for literacy was 117.5 with a pass rate of 93.3%; and the overall mean scale score for numeracy was 122.4 with a pass rate of 92.4%. This table also shows the performance of candidates on each subscale. The performance of candidates by subscale was similar to the performance of candidates on the whole test, except that the average performance of the numeracy subscale 'calculator not available' was lower (3.2 scale score points) than the average performance of the numeracy subscale 'calculator available'. Scale score frequency distributions for the candidates who participated in the tests are shown in Appendix 1.

¹¹ In 2016, the numeracy standard scale score was 107

Table 18: Candidate performance overall and by subscale for first-attempt candidates

Component	Whole test and subscale	Valid N	Mean	S.D.	Pass Rate
Literacy	Overall	13,083	117.5	8.5	93.3
	Reading		117.4	9.1	
	Technical skills of writing		117.5	10.2	
Numeracy	Overall	13,084	122.4	11.2	92.4
	Number & algebra		121.8	12.1	
	Measurement & geometry		121.5	11.5	
	Statistics & probability		122.7	11.9	
	Calculator available		122.5	10.9	
	Calculator not available		119.3	12.2	

Table 19 shows the number of candidates, mean scale score and standard deviation of the scale scores, and pass rate by demographic characteristics. Performance of any subgroup with a sample size of less than 10 was not reported.

Table 19: Performance by demographic characteristics for first-attempt candidates

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	9970	117.1	8.6	92.8	9991	121.0	10.9	90.9
	Male	3110	118.6	8.3	95.0	3090	127.1	10.9	97.3
	Indeterminate/intersex	3	-	-	-	3	-	-	-
Age	17–25	8544	116.2	7.8	92.6	8558	121.5	10.8	91.9
	26–30	1910	118.8	8.9	93.1	1903	123.7	11.7	93.2
	31–35	889	119.7	8.9	96.1	885	124.5	11.4	94.8
	36–40	683	120.7	9.2	96.3	682	125.0	11.4	94.7
	41–45	554	121.1	9.7	95.3	553	124.5	12.9	92.0
	46+	500	121.4	9.6	95.2	500	124.3	12.1	92.4
International Students	No	12,647	117.6	8.5	93.6	12,658	122.4	11.2	92.4
	Yes	436	113.7	8.3	86.2	426	124.2	11.3	93.4
English as a First Language	Yes	11,737	117.9	8.4	94.3	11,758	122.5	11.1	92.9
	No	1346	113.5	8.5	84.8	1326	121.5	12.1	88.2
Indigenous	No	12,569	117.5	8.5	93.4	12,576	122.5	11.2	92.5
	Yes	203	114.8	8.8	88.2	198	117.9	10.7	85.4
	Not disclosed	311	118.4	8.5	93.2	310	123.8	11.3	93.9
Residential Area	Metropolitan areas	10,718	117.5	8.6	93.4	10,731	122.6	11.3	92.2
	Provincial areas	2271	117.1	8.2	93.0	2261	121.7	10.6	93.1
	Remote areas	47	119.8	9.5	95.7	47	123.3	10.7	95.7
	International	32	122.3	7.6	100.0	30	128.4	9.8	100.0
	Invalid or Missing	15	118.7	6.9	100.0	15	125.2	6.9	100.0
Program Type	Undergraduate	8078	116.1	7.8	92.5	8078	120.8	10.5	91.3
	Postgraduate	5005	119.7	9.1	94.7	5006	125.1	11.7	94.3
Program Type by Year Level	Undergraduate first year	104	114.9	9.3	82.7	103	121.9	12.4	88.3
	Undergraduate second year	244	116.7	8.6	91.0	234	122.6	11.1	93.6
	Undergraduate third year	1942	115.7	7.6	92.2	1950	120.5	10.4	91.4

	Undergraduate fourth year	5438	116.1	7.7	92.9	5438	120.7	10.5	91.2
	Undergraduate fifth year or above	350	117.5	8.6	91.7	353	122.6	10.8	91.8
	Postgraduate first year	1843	120.1	9.1	95.9	1836	126.5	11.3	95.8
	Postgraduate second year	2048	120.4	9.0	95.6	2046	125.4	11.6	95.4
	Postgraduate third year	219	119.2	8.6	93.2	217	123.9	11.6	91.7
	Postgraduate fourth year	582	116.9	9.6	89.3	588	121.3	12.3	88.3
	Postgraduate fifth year or above	313	118.6	9.1	92.3	319	123.6	12.0	91.2
Course Category	Teacher education: early childhood	612	113.8	8.0	86.9	616	116.9	11.0	81.8
	Teacher education: primary	5358	116.9	8.3	93.1	5409	121.1	10.7	91.6
	Teacher education: secondary	4964	119.3	8.6	95.4	4935	125.4	11.4	95.4
	Teacher education: special education	87	114.3	8.6	88.5	86	117.8	8.7	87.2
	Teacher education: other	2062	115.7	8.2	91.0	2038	120.7	10.6	90.9

The t-test and Cohen's *d* for effect size were used to determine if group mean scale scores were significantly different for first-attempt candidates. Only differences where $p \leq 0.05$ and $d > 0.2$ are reported here as significant.

There were significant differences in mean scale scores for the demographic characteristics as follows:

- Candidate age, cultural background (indigenous or non-indigenous), program type and course category for both literacy and numeracy
- Language background, country of origin (international or domestic) and residential area for literacy but not numeracy.
- Gender for numeracy but less so for literacy.

Table 19 shows that female candidates performed similarly to the male candidates in literacy, but male candidates outperformed female candidates in numeracy. The numeracy mean scale score of male candidates (127.1) was significantly higher than the mean scale score of female candidates (121.0), but the same was not true for literacy. In literacy, the pass rate of female candidates (92.8%) was marginally lower than that of male candidates (95.0%). In numeracy, the pass rate of male candidates (97.3%) was higher than the pass rate of female candidates (90.9%).

For literacy and numeracy, achievement on the test tended to increase with the age of the candidates, with the pattern being stronger for literacy. For literacy, the youngest group of candidates aged 17–25 (mean scale score 116.2) achieved lower than candidates aged over 25 (119.8). For numeracy, candidates aged 17–25 (121.5) achieved lower than candidates aged over 25 (124.2). In both cases the differences were moderately significant (effect size 0.4). The age groups above 40 had a slightly higher performance than other age groups in literacy, and the age group 36–40 had a slightly higher performance in numeracy than other age groups.

While the mean scale score of international candidates (113.7) for literacy was significantly lower (effect size = 0.5) than the mean scale score of other candidates (117.6), the same was not true for numeracy. For numeracy, while the mean scale score of international candidates (124.2) was higher than the mean scale score of other candidates (122.4), the difference was not significant (effect size < 0.2).

While the mean scale score for literacy of candidates for whom English was a first language (117.9) was significantly higher than the mean scale score for literacy of other candidates (113.5), the same was not true for numeracy where the difference (122.5 and 121.5) was not significant.

For numeracy, the mean scale scores of candidates who identified as Aboriginal or Torres Strait Islanders (117.9) were significantly lower (effect size = 0.4) than the mean scale scores of other candidates (122.5). The same was true for literacy (114.8 and 117.5) but the difference was less significant (effect size = 0.3).

Residential postcode data were used to place candidates into three main categories: metropolitan, provincial and remote. Where postcodes could not be matched to either indicator they were indicated as international, missing or invalid. For both literacy and numeracy, there was little difference in achievement by residential area. For literacy only, the mean scale score of candidates with residential postcodes in remote areas (119.8) was higher than the mean scale score of candidates with residential postcodes in provincial areas (117.1). For numeracy, the differences were smaller. However, the number of candidates in remote areas is small (< 50).

For both literacy and numeracy, the mean scale scores of postgraduate candidates were significantly higher than for undergraduate candidates. The difference was greater for numeracy (125.1 and 120.8) than for literacy (119.7 and 116.1).

Unlike for age, there was no tendency for mean scale scores to increase with year of course. This was true for both literacy and numeracy and for both undergraduate and postgraduate courses. Perhaps the reason for this was that there is a mixture of ages in the year levels, or that candidates who participated in the test were not representative of all students by year of course.

The performance of undergraduate candidates was similar across year levels. The mean literacy scale scores were in the range of 114.9 to 117.5 for different year levels; and the mean numeracy scale scores were in the range of 120.5 to 122.7. Slightly larger variation in the performance of postgraduate candidates was observed across year levels. Performance of postgraduate fourth year candidates was the lowest for all year levels in both literacy (116.9) and numeracy (121.3). In literacy, the mean scale scores were in the range of 116.9 to 120.4, and in numeracy, the mean scale scores were in the range of 121.3 to 126.5.

Findings for the characteristic program type by year level were mixed and of moderate significance. For literacy, the mean scale score of undergraduate candidates in their fifth year or above (117.5) was greater than the mean scale score of undergraduate candidates in their first year (114.9). This was not the case for numeracy, however. For numeracy, the mean scale score of postgraduate candidates in their first year (126.5) was greater than the mean scale score of postgraduate candidates in their fifth year or above (123.6). For both literacy and numeracy, the mean scale scores of postgraduate candidates in their third year were greater than the mean scale scores of postgraduate candidates in their fourth year. The difference for literacy (119.2 and 116.9) was quite similar to the difference for numeracy (123.9 and 121.3).

There was some variation in the performance of candidates in the four course categories: early childhood school education, primary school education, secondary school education, and special education. Candidates in the early childhood school education course category had the lowest performance with a mean literacy scale score of 113.8 and a mean numeracy scale score of 116.9. Candidates in secondary school education had the highest performance with a mean literacy scale score of 119.3 and a mean numeracy scale score of 125.4. The difference was larger in numeracy. The ordering of mean scale scores by course was also the same for both literacy and numeracy: secondary was highest, followed by primary, then by other, then by special education, and then by early childhood. For both literacy and numeracy, the mean scale scores of candidates in secondary courses (119.3 and 125.4) were higher than the mean scale scores of candidates in primary courses, significantly so for numeracy (116.9 and 121.1). The mean

scale score of candidates in primary courses were significantly higher than the mean scale scores of candidates in early childhood courses, again significantly so for numeracy (113.8 and 116.9).

Table 20 shows the performance by test centres and remote proctoring. Insignificant differences were found in the performance of first-attempt candidates who sat the test at different test centres. The candidates at the Hobart test centre had the highest mean score (123.2) in literacy. The candidates at the Perth CBD test centre had the highest mean score (127.8) in numeracy. However the numbers at these centres are proportionally low (50 and 52, and 124 and 123 respectively) and it is possible that they represent the more confident or able candidates. The performance of candidates using remote proctoring was similar to the performance of candidates who took tests in capital city test centres and regional city test centres. The mean score of candidates using remote proctoring was 118.5 in literacy and 123.4 in numeracy.

The last three rows in Table 20 compare the performance of those who sat the test by remote proctoring with the performance of those who sat the test at test centres in the state/territory capital cities and in regional centres. While for both literacy and numeracy the mean scale scores of candidates who sat the test using remote proctoring were one scale score point higher than those who sat in test centres, the differences were not significant.

Table 20: Performance by test centres and remote proctoring

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Test Centre	Adelaide CBD	131	119.6	8.9	94.7	131	123.8	11.2	96.2
	Albury	150	116.8	7.8	96.0	142	121.5	9.6	90.8
	Armidale	99	118.0	7.0	98.0	98	122.8	11.6	93.9
	Ballarat	173	115.2	7.7	91.9	174	119.7	10.6	93.7
	Bathurst	64	118.8	7.9	95.3	65	124.0	12.7	93.8
	Brisbane CBD	1327	117.7	8.2	94.4	1326	122.4	11.1	92.2
	Bundaberg	37	117.5	8.1	89.2	38	120.0	10.1	92.1
	Canberra CBD	317	117.3	8.0	92.7	316	122.0	11.2	92.7
	Darwin	35	115.5	8.3	94.3	35	119.1	8.9	94.3
	Hobart	50	123.2	9.2	98.0	52	125.0	11.2	96.2
	Melbourne CBD	3728	116.8	8.5	91.8	3747	121.9	11.4	91.4
	Mildura	23	113.1	7.6	82.6	23	119.1	8.8	87.0
	Newcastle	569	117.7	7.9	94.6	573	122.4	10.9	94.1
	Parramatta	362	115.4	8.2	89.8	356	120.6	11.9	87.9
	Perth CBD	124	122.6	9.9	96.8	123	127.8	12.9	94.3
	Sunshine C./Maroochy.	127	118.3	7.6	96.1	122	122.6	9.4	94.3
	Sydney CBD	3107	117.5	8.6	93.6	3125	122.6	11.2	92.9
	Townsville	10	117.1	8.2	100.0	10	123.1	10.6	100.0
	Wagga Wagga	86	116.8	7.4	93.0	87	122.7	8.8	96.6
	Warrnambool	46	116.8	7.2	95.7	50	122.0	13.1	86.0
	Wollongong	144	117.2	7.7	94.4	139	122.8	11.0	94.2
	Remote Proctoring	2374	118.5	8.8	94.3	2352	123.4	11.1	93.2
	Capital Cities	8819	117.3	8.5	93.0	8855	122.3	11.3	92.2
	Regional Cities	1890	116.9	7.9	93.5	1877	121.8	11.0	92.4

Figure 4 to Figure 7 display scale score distributions of demographic characteristics for first-attempt candidates in 2016. The left panel of each figure shows literacy scale score distributions and the right panel of each figure shows numeracy scale score distributions. The horizontal lines in each figure represent the standard scale score for each component of the test.

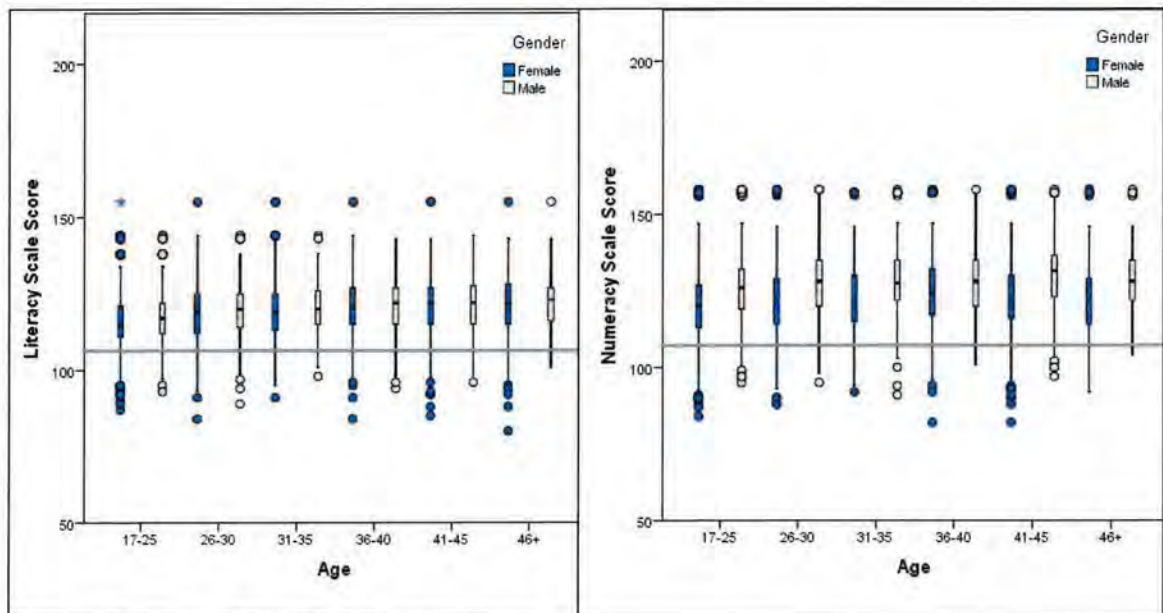


Figure 4: Score distribution by gender and age

Figure 4 shows that, for all age groups, the difference in achievement between male candidates and female candidates is more pronounced for numeracy than for literacy. However, in each age category and for each gender, there are candidates who achieve well above the standard.

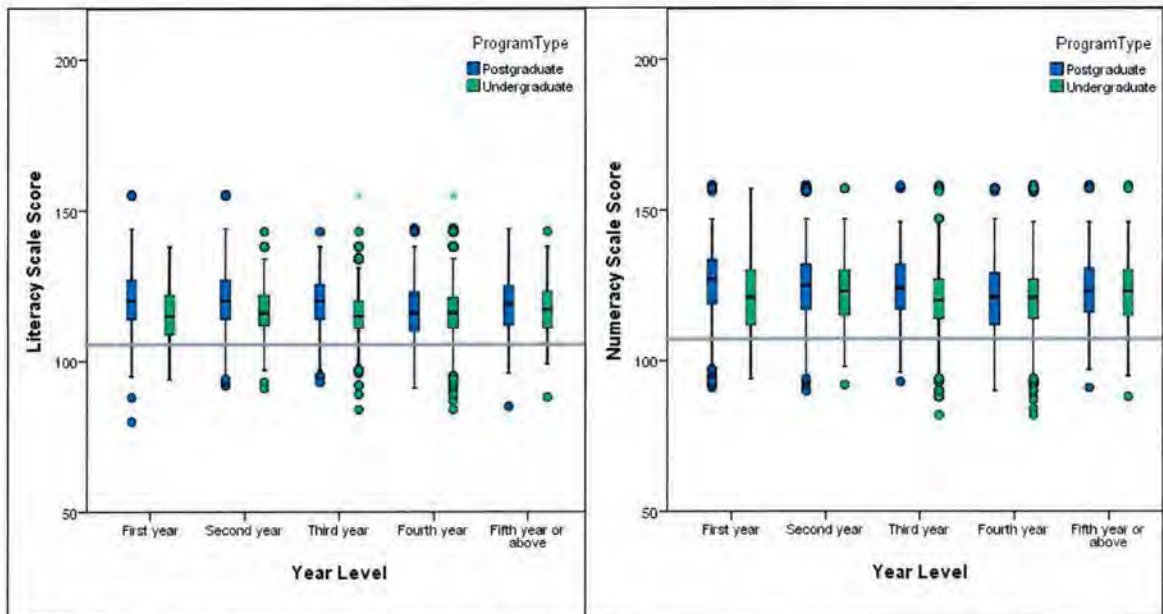


Figure 5: Score distribution by program type and year level

Figure 5 shows that, for all year levels, the difference in achievement between postgraduate candidates and undergraduate candidates is similar for literacy and numeracy. As for gender, at each year level and in each program type, there are candidates who achieve well above the standard.

3.3 Performance of Resit Candidates

Table 22 shows the performance of candidates who resat the test, at their second attempt and third attempt, overall and by subscale. As would be expected, the performance of resit candidates was significantly lower than the performance of the majority of candidates who achieved the standard at their first attempt. For example, for the 341 candidates who were making their second attempt for literacy, their overall mean scale score for literacy was 107.5 with a pass rate of 67.7% (compared to 117.5 and 93.3% for the first-attempt candidates). For the 406 candidates who were making their second attempt for numeracy, their overall mean scale score for numeracy was 107.5 with a pass rate of 55.4% (compared to 122.4 and 92.4% for the first-attempt candidates).

The mean pass rates of resit candidates in numeracy were much lower than the mean pass rates in literacy. The performance of candidates in their third attempt had a higher pass rate for literacy (76.0%), but a lower pass rate for numeracy (37.5%), than the second-attempt candidates. However, the number of candidates (25 and 40) in each of these categories was small. In terms of subscale scores, the mean score of resit candidates was similar for reading and technical skills of writing. However, the mean score of resit candidates on the 'calculator not available' subscale was significantly lower (nearly 5 scale score points) than the mean score of resit candidates on the 'calculator available' subscale. It might be expected that individual candidates would perform better with each resit, as a result of greater familiarity with the format of the test or of additional study that they might have done following a previous unsuccessful attempt. If so, this was more evident in literacy than in numeracy.

Table 22: Resit candidate performance overall and by subscale

Component	Whole test and subscale	Second Attempt (Resit 1)				Third Attempt (Resit 2)			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Literacy	Overall	341	107.5	5.0	67.7	25	107.7	5.0	76.0
	Reading		107.2	5.5			107.3	5.5	
	Technical skills of writing		107.8	7.8			108.5	6.5	
Numeracy	Overall	406	107.5	6.0	55.4	40	105.4	5.6	37.5
	Number & algebra		106.1	7.7			103.9	7.2	
	Measurement & geometry		108.0	8.2			106.8	8.7	
	Statistics & probability		109.6	7.4			107.0	6.1	
	Calculator available		108.5	6.0			106.2	6.0	
	Calculator not available		103.7	12.0			102.3	7.8	

4 CANDIDATE PERFORMANCE BY TEST WINDOWS

This section presents candidate performance by test window. The performance of resit candidates is also described by test windows.

4.1 Distributions of Candidate Scale Scores by Subscales and Subgroups

Table 23 shows the performance of all candidates (first-attempt and resits) who sat the test for each test window in 2016. The overall mean scale scores for literacy were similar across the four test windows (in the range of 116.4 to 117.6) with a pass rate in the range of 90.8% to 94.5%; and the overall mean scale score for numeracy was in the range of 121.3 to 122.5 with a pass rate in the range of 89.7% to 93.1%. The performance of candidates by subscale was similar to the average performance of candidates on the whole test, except that the average performance on the 'calculator not available' subscale was consistently significantly lower than the overall numeracy score.

Table 23: Candidate performance by test window, overall and by subscale

Component	Test Window	Whole test and subscale	Valid N	Mean	S.D.	Pass Rate
Literacy	TW1	Overall	3831	117.6	8.4	94.5
		Reading		117.5	8.9	
		Technical skills of writing		117.5	10.0	
	TW2	Overall	3492	117.3	8.6	92.6
		Reading		117.3	9.3	
		Technical skills of writing		117.5	10.1	
	TW3	Overall	3162	116.4	8.6	90.8
		Reading		116.5	9.3	
		Technical skills of writing		116.5	10.0	
	TW4	Overall	2964	117.4	8.8	92.4
		Reading		117.2	9.2	
		Technical skills of writing		117.5	10.7	
Numeracy	TW1	Overall	3844	122.5	11.0	93.1
		Number & algebra		122.1	11.9	
		Measurement & geometry		121.5	11.4	
		Statistics & probability		122.5	11.5	
		Calculator available		122.5	10.8	
		Calculator not available		119.5	11.9	
	TW2	Overall	3488	122.2	11.4	91.4
		Number & algebra		121.3	12.2	
		Measurement & geometry		121.6	11.6	
		Statistics & probability		122.6	12.1	
		Calculator available		122.3	11.0	
		Calculator not available		119.0	12.5	
	TW3	Overall	3210	121.5	11.7	90.0
		Number & algebra		120.8	12.5	

		Measurement & geometry		120.7	11.7	
		Statistics & probability		122.0	12.3	
		Calculator available		121.7	11.3	
		Calculator not available		118.2	12.7	
	TW4	Overall	2990	121.3	11.5	89.7
		Number & algebra		120.7	12.5	
		Measurement & geometry		120.4	11.8	
		Statistics & probability		121.7	12.0	
		Calculator available		121.5	11.1	
		Calculator not available		118.3	13.0	

Appendix 2 provides a detailed analysis of candidate performance by demographic characteristics and test windows. Candidate performance of any subgroup with a sample size less than 10 was not reported.

Appendix 3 provides a detailed analysis of performance by test centres and remote proctoring by test window. The distributions of candidate performance within a subgroup were similar across test windows.

Table 24 shows that the proportion of candidates (first-attempt and resits) who did not achieve the standard on the literacy test was the highest in test window 3 (9.2%) and the lowest in test window 1 (5.5%). The proportion of candidates who did not achieve the standard on the numeracy test was the highest in test window 4 (10.3%) and the lowest in test window 1 (6.9%). The proportion of candidates who did not achieve the standard on both tests in test windows 3 and 4 (3.6% and 3.7%) was slightly higher than that in test windows 1 and 2 (2.4% and 2.6%).

It might be expected that in the early years of implementation of the test, variation across test windows will be more marked than in later years, as candidates might initially respond with varying degrees of urgency to the requirement to sit for the test. As sitting for the test becomes a more routine procedure, patterns of participation are likely to stabilise.

Table 24: Number and proportion of candidates who did not achieve the standard by test window

Test Window	Test(s)	Number	% of Candidates
TW1	literacy	212	5.5%
	numeracy	267	6.9%
	both tests	89	2.4%
TW2	literacy	259	7.4%
	numeracy	301	8.6%
	both tests	88	2.6%
TW3	literacy	290	9.2%
	numeracy	321	10.0%
	both tests	110	3.7%
TW4	literacy	226	7.6%
	numeracy	307	10.3%
	both tests	99	3.6%

4.2 Performance of Resit Candidates by Test Windows

Table 25 shows the performance of candidates who resat the test, overall and by subscale, and by test attempt, for each test window. There were no resit candidates in test window 1.

In each test window, the mean scores of resit candidates who had a second attempt (resit 1) were close to or marginally above the standard in both literacy and numeracy.

For literacy, the candidates who sat the test for the third time (resit 2) had very similar mean scores to those candidates who sat the test for the second time. For numeracy, however, the mean scores of candidates who sat the numeracy test for the third time was more than 2 scale score points lower than those candidates who sat the test for the second time for both testing windows. The mean score of these candidates was also below the numeracy standard. It can be seen that the difference in mean scale scores for the 'calculator available' and the 'calculator not available' subscales was greater (4 to 5 scale score points) for the resit cohort than for the first-attempt candidates (3 scale score points).

Furthermore, while between 71% and 88% of third-attempt candidates achieved the literacy standard, only 33% to 40% of third-attempt candidates achieved the numeracy standard. However, because of the relatively small sizes of these groups, percentage comparisons should be taken as indicative only.

Although resitting a test might be expected to bring individual candidates greater familiarity with the format, and greater opportunity to prepare through study (with the potential for an improved score), it is not surprising that the mean score should in fact decrease for the third sitting, as the more able candidates would have been removed from the resitting pool by achieving the test standard at the second sitting.

Table 25: Resit candidate performance by test window, overall and by subscale

Component	TW	Whole test and subscale	Second Attempt (Resit 1)				Third Attempt (Resit 2)			
			N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Literacy	TW2	Overall	95	107.3	5.0	67.4				
		Reading		107.1	5.4					
		Technical skills of writing		107.6	7.4					
	TW3	Overall	137	107.5	5.1	67.2	8	107.8	3.5	87.5
		Reading		106.9	5.5			107.0	4.3	
		Technical skills of writing		108.2	7.9			109.4	5.3	
	TW4	Overall	109	107.6	5.0	68.8	17	107.6	5.6	70.6
		Reading		107.5	5.5			107.4	6.1	
		Technical skills of writing		107.5	7.9			108.1	7.1	
Numeracy	TW2	Overall	112	106.9	6.0	54.5				
		Number & algebra		105.3	7.4					
		Measurement & geometry		107.0	7.4					
		Statistics & probability		109.4	7.3					
		Calculator available		107.9	5.9					
		Calculator not available		102.6	11.2					
	TW3	Overall	154	107.5	5.9	53.2	15	105.0	4.5	33.3
		Number & algebra		106.0	8.2			102.1	6.0	
		Measurement & geometry		108.9	8.4			108.7	7.8	

		Statistics & probability		109.1	7.7			106.8	6.0	
		Calculator available		108.5	6.0			105.9	5.0	
		Calculator not available		103.8	11.6			101.9	8.1	
	TW4	Overall	140	108.0	6.0	58.6	25	105.6	6.3	40.0
		Number & algebra		106.8	7.4			104.9	7.7	
		Measurement & geometry		107.8	8.4			105.7	9.2	
		Statistics & probability		110.3	7.1			107.1	6.3	
		Calculator available		109.0	5.9			106.3	6.6	
		Calculator not available		104.4	12.9			102.6	7.8	

5 PHASE 2 ITEM TRIAL ANALYSIS

5.1 In-test trialling

For each of the literacy and numeracy test components, 160 of the 190 Phase 2 items were trialled within the live tests. The Phase 2 items were placed in small clusters (5-item clusters for literacy, and 4-item calculator available clusters and a 1-item calculator not available cluster for numeracy) at the end of each test. Candidates were unaware of the location of these trial items. The trial items did not contribute to a candidate's score. The items were trialled in multiple testing windows until sufficient candidates had attempted them. In this way, robust trial item estimates were obtained to enable selection of new, balanced clusters for refreshment of the tests in testing window 3 and testing window 4 in 2017. Selected items from the Program for the International Assessment of Adult Competencies (PIAAC) were also trialled in testing windows 3 and 4 in 2016. See Section 6 for more detail on the PIAAC item trial and resulting analysis.

5.2 Item trial analysis

For Rasch scaling analysis, the software ACER ConQuest was used to fit the trial item response data to the Rasch measurement model (Rasch, 1980). In the first round of item analysis, 8 literacy items and 10 numeracy items were removed because of their poor fit or inadequate discrimination. Item statistics for the remaining 152 literacy trial items and the remaining 150 numeracy trial items are provided in Appendix 4, including item difficulty (logit, scale score and facility), item discrimination (item rest corr.), item fit (infit) and the number of candidates attempting each item (cases). The item facility and discrimination statistics are part of classical test analysis; the other statistics are part of item response theory analysis. In these tables, items are ordered by item difficulty from hard to easy.

The *item facility* statistic expresses the percentage of individuals who are successful in answering each question or item on a test.

The *item discrimination* expresses the correlation between the individual's score and the aggregate score on the set of items in the same test. The item discrimination index used throughout this report is item-rest correlation, in which the aggregate score excludes the score of the item under examination.

One of the *item difficulty* statistics is expressed in terms of 'logit' – a metric used to standardise the test results across different test instruments. The logit transformation puts individuals (cases) and items from different linked test forms onto one scale, allowing fair judgements to be made about the relative difficulty of the items. Importantly, it also makes it possible to judge the relative proficiency of candidates, in spite of the fact that they have been administered different tests that may have had different levels of difficulty. The item estimates were calibrated based on large sample sizes. For item analysis, the mean difficulty (in logits) of the items for each construct was set at 0.

The item difficulty in scale score is a transformation of item difficulty from logit on to a scale with a mean of 100 and standard deviation of 10. The transformation also adjusts for an equating shift enabling comparison to the provisional recommended benchmarks that were defined on the scale of the field trial analysis.

The *item fit* statistics are a measure of the extent to which an item is contributing to the measurement of the characteristic of interest. In the case of the item infit mean square, values near one are desirable. An infit mean square greater than one is often associated with a low discrimination index, and an infit mean square less than one is often associated with a high discrimination index. The mean of the infit mean squares for a scale is 1.0.

The last column in Table 42 and Table 43 in Appendix 5 lists the number of candidates who saw the question. The minimum numbers of observations were 252 for literacy items and 271 for numeracy items.

The mean item facilities for literacy items and numeracy items were 76.4 and 78.2 respectively. These item facilities were found to be slightly lower than the mean item facilities (77.5 for literacy items and 82.0 for numeracy items) in the existing item pool of live tests.

Differential item functioning (DIF) analysis was performed for subgroups on gender, age group, program type, and course category. The details are available in Appendix 5. For trial items, DIF analysis was performed for subgroups on gender, age group, program type, and course category. The findings of the DIF analysis are used when selecting items for test forms to ensure that no subgroup of candidates is disadvantaged.

For the purpose of the DIF analysis, age group categories have been combined into two main categories: 17–25 and 26+. Due to small sample sizes in two course categories (Teacher education: special education and Teacher education: others), only three course categories were used for DIF analysis, and two course categories (Teacher education in early childhood and Teacher education in primary) have been combined as a single category.

Items were trialled in two batches. The data for the first batch of trial items were collected during test window 1 of the Aug/Sept 2015 and 2016 sittings, and the data for the second batch of trial items were collected during test windows 2–4 in 2016. DIF analysis was performed separately for each batch of trial items.

The interaction of item and subgroups was modelled using ConQuest software. The mean item difficulty in each of the two independent sets of item difficulties was centred at zero to adjust for group difference in ability.

Figure 13 in Appendix 5 shows the DIF plot for gender. The left panels of the figure show the literacy items DIF plots for each batch of trial items, and the right panels of the figure show the numeracy items DIF plots. The DIF plots in the top panels are for the first batch of items; the plots in the bottom panels are for the second batch of items.

On each DIF plot, an item is represented by one point on the plot. A diagonal line serves as the reference line, with confidence interval limits indicated by the curved lines on either side of the reference line. If the relative item difficulty for an item is not different between the two groups after taking their overall performance on the test into account, the point representing the item should lie on or close to the diagonal. The distance of a point from the diagonal reflects the size of any potential DIF. Any item that falls outside the two lines representing the confidence limits may warrant investigation for potential DIF.

Most items fall within the confidence interval limits or close to the confidence interval. A few items are relatively distant from the confidence lines.

Table 44 in Appendix 5 lists potential gender DIF items with a standardised difference of 2.5 and above. The table shows that for literacy, six items significantly favoured females and twelve items significantly favoured males. Of the fourteen numeracy items showing gender DIF, eight items significantly favoured females and six items significantly favoured males. It is normal for some items to show gender DIF and it is acceptable to include these items in the live test provided other item characteristics are favourable. What is important is the overall balance of the test; that is, that the test does not have more items favouring one gender over another.

In addition to refreshing the test, the trial items are also used to contribute to the described proficiency scales and the pool of released sample items. For these reasons, a good spread of the trial items across the three bands of the scale is critical. Table 26 and Table 27 show the distribution of the trial items across sets 1, 2 and 3 for literacy and numeracy respectively. Items in set 1 are used to describe what candidates below the standard (Band 1) can do, items in set 2 are used to describe what candidates at and

above the standard (Band 2) can do, and items in set 3 are used to describe what candidates well above the standard (Band 3) can do.

Table 26: Distribution of the trial literacy items across the scale

	Reading	Technical Skills of Writing	Total
Above Set 3	3	2	5
Set 3	10	9	19
Set 2	43	15	58
Set 1	15	20	35
Below set 1	28	7	35
Total	99	53	152

Table 27: Distribution of the trial numeracy items across the scale

	Number & algebra	Measurement & geometry	Statistics & probability	Calculator available	Calculator not available	Total
Above Set 3	0	0	1	1	0	1
Set 3	4	6	7	16	1	17
Set 2	25	23	15	51	12	63
Set 1	27	11	15	39	14	53
Below set 1	5	8	3	13	3	16
Total	61	48	41	120	30	150

It can be seen that for both literacy and numeracy, the trial item pool is well-targeted, particularly for sets 1 and 2, the critical bands describing what candidates below the standard (Band 1) can do, and what candidates at or above the standard (Band 2) can do.

6 PIAAC ITEM ANALYSIS AND EQUATING

6.1 Background

In order to validate the provisional benchmarks and standards set by expert groups, a separate equating study was conducted to provide information on how the literacy and numeracy skills of candidates achieving the test standards compare to those of Australian adults as measured by the OECD Programme for International Assessment of Adult Competencies (PIAAC). PIAAC was administered to a representative sample of Australian adults (aged 16–74) in 2011–12.

6.2 Selection of PIAAC items

Permission to use the PIAAC items for the equating study was kindly granted by the OECD. For the purpose of the study, a set of PIAAC items (15 literacy items and 16 numeracy items) were selected as common items for scaling. These items were grouped into clusters: four numeracy clusters (four items in each cluster) and three literacy clusters (five items in each cluster). Each PIAAC item cluster was included at the end of a test, replacing the in-test trial items. The tests with PIAAC items were randomly assigned to the candidates. The PIAAC items were not used for estimating candidate performance.

The constructs of PIAAC literacy and numeracy assessments are similar to those of the test, and the PIAAC assessment framework was a significant reference for the test assessment framework. The emphasis in PIAAC on using literacy and numeracy in real-life contexts is broadly similar to that of the test's reading and numeracy constructs. It should be noted, however, that in PIAAC there is no equivalent to the test's assessment of the technical skills of writing.

Like the test, PIAAC was delivered online. Item types were selected response (simple multiple choice, true/false, hotspot) or short constructed response. The PIAAC items were re-formatted slightly so that they were not readily distinguishable from the test items and so that their psychometric properties were retained. Scoring rules of PIAAC items were unchanged.

For numeracy, the PIAAC items were selected so that the proportions of items in each content subscale (Number and algebra; Measurement and geometry; and Statistics and probability) were close to those in the test. All PIAAC numeracy items are calculator allowed and therefore could only be included in the calculator-available sections of the test's numeracy tests.

6.3 Equating Methodology

The PIAAC tests were calibrated with a two-parameter IRT model (2PL) (see PIAAC, 2016). Under the 2PL model, each item has two item parameter estimates: item difficulty and item discrimination. On the other hand, the test was calibrated with the Rasch model. Under the Rasch model, each item has only one parameter (i.e. item difficulty). Using data collected for the equating study, it is possible to obtain item difficulty estimates of the common PIAAC items using the Rasch model. However, it is not recommended to equate two sets of item difficulty estimates of the common PIAAC items directly, because item discrimination estimates from the 2PL model would be ignored in such equating.

The equating method used in this study takes into account both item difficulty and item discrimination of the PIAAC items, and it also addresses the IRT model effect explicitly. The equating method consists of two components: common PIAAC item equating, and investigation of the IRT model effect.

Prior to performing common item equating, a joint calibration of the core items (the items used for determining candidate scores in the test) and common PIAAC items (for linking) was performed with a 2PL IRT model, excluding any trial items. This is a free calibration and no anchoring was used. After the calibration, two sets of 2PL item estimates were available for each PIAAC item: one on the free-calibrated scale, and the other on the PIAAC scale. The equating data set was of sufficient size for reliable

analysis and provided a minimum of 860 data points (candidates) for each PIAAC literacy item, and a minimum of 709 for each PIAAC numeracy item.

Next, the Stocking-Lord method (Stocking & Lord, 1983) was used to estimate scaling constants of a linear transformation relating the two sets of 2PL item estimates. The Stocking-Lord method is one of the most robust equating methods for when both item difficulty and item discrimination are to be considered for common item equating. The Stocking-Lord method calculates linear scaling constants by minimising the difference between the two test characteristic curves. To satisfy the symmetric property of linear transformation, the difference of test characteristic curves is minimised on each of the scales.

The second component of this equating method investigated the IRT model effect. The data set used in this analysis combines candidate response data sets from test windows 2–4 in 2016. This provided literacy and numeracy ability estimates for about 9600 candidates. IRT model effect was investigated by comparing two sets of candidate proficiency estimates obtained from this data set: candidate ability estimates, which were anchored on Rasch-based core items on the scale of test, and candidate ability estimates, which were anchored on free-calibrated 2PL-based core items.

By utilising the scaling constants obtained from these two components, the Australian adult population's proficiency percentile scores on the PIAAC scale can be transformed onto the scale of the test for each component. This allows comparison of the 70th percentile score (i.e. the minimum score of the top 30 per cent) of Australian adults' literacy and numeracy in relation to benchmarks and standards on the test scale.

6.4 Outcome of Equating

The projected test scale score equivalent to the 70th percentile of Australian adults is 107 for literacy, and 110 for numeracy (Table 28). For literacy, the projected scale score is similar to the scale score of the existing standard (106). However, for numeracy, the projected scale score is marginally higher than the scale score of the existing standard (107).

Table 28: Test scale scores equivalent to the 70th percentile of the Australian adult population, and existing scale score standards

Component	Test scale score equivalent to the 70th percentile of the Australian adult population	Test scale score of existing standard
Literacy	107	106
Numeracy	110	107

The existing test standards are equivalent to the 68th percentile of the Australian adult population for literacy, and the 62nd percentile for numeracy (Table 29).

Table 29: Equivalent Australian adult population percentiles for existing test standards

Component	Equivalent percentile of the Australian adult population for existing standard
Literacy	68th percentile
Numeracy	62nd percentile

The equating shows that the professional judgements of the expert groups when determining the provisional standards were commensurate with the desired empirical outcomes. For literacy, the experts arrived at a standard range of 106 to 110 scale score points. The scale score equivalent to the 70th percentile is 107, which is within the original range.

For numeracy, the experts arrived at a standard range of 107 to 109 scale score points. The scale score equivalent to the 70th numeracy percentile is 110, which is marginally outside and above the original range.

Based on the 2016 data¹², the estimated percentage of candidates achieving the test scale score equivalent to the 70th percentile of Australian adults would be 91.3 per cent for literacy and 88.0 per cent for numeracy. The difference in the percentage of candidates who achieved the existing standard and the estimated percentage who would achieve the higher standard is 2.1 per cent fewer for literacy and 4.4 per cent fewer for numeracy (Table 30).

Table 30: Comparison of outcomes between the existing and the higher standards for first-attempt candidates

Component	Total candidates	Candidates who achieved the existing standard		Candidates who would have achieved the higher standard		Difference	
		N	%	N	%	N	%
Literacy	13,083	12,212	93.3	11,939	91.3	-273	-2.1
Numeracy	13,084	12,094	92.4	11,517	88.0	-577	-4.4

Table 31 shows the estimated pass rates for candidates who sat the test more than once. Using the 2016 cohort, it is estimated that 21 or 6% of second-attempt literacy candidates would not have achieved the higher literacy standard and 73 or 18% of second-attempt numeracy candidates would not have achieved the higher numeracy standard. For third-attempt candidates, the figures are 4 or 16% and 7 or 18% respectively. Both fourth-attempt candidates would have achieved the higher numeracy standard.

Table 31: Comparison of outcomes between the existing and the higher standards for multi-attempt candidates

Component	Attempt	Total candidates	Candidates who achieved the existing standard		Candidates who would have achieved the higher standard		Difference	
			N	%	N	%	N	%
Literacy	2	341	231	67.7	210	61.6	-21	-6.2
	3	25	19	76.0	15	60.0	-4	-16.0
Numeracy	2	406	225	55.4	152	37.4	-73	-18.0
	3	40	15	37.5	8	20.0	-7	-17.5
	4	2	2	100.0	2	100.0	0	0.0

¹² The data consists of candidates in test windows 1–4 in 2016. For candidates who resat the test, only the result from their first attempt was included in the data.

7 PHASE 3 TEST DEVELOPMENT

During 2016 and 2017, 190 Phase 3 literacy items and 190 Phase 3 numeracy items mapped against the Assessment Framework were written by ACER and reviewed by the national expert groups. Approximately half of the items will be revised based upon expert group feedback and a small proportion will be retired. A selection of 160 literacy items and 160 numeracy items will be in-test trialled in testing windows 3 and 4 of 2017 and testing windows 1 and 2 of 2018.

8 CONCLUSION

The test was successfully administered in four testing windows in all Australian states and territories in metropolitan and regional testing centres and by remote proctoring to over 13,000 candidates. At the end of 2016, after second and third attempts were accounted for, 95.2% of candidates had achieved the literacy standard and 94.2% of candidates had achieved the numeracy standard. Equating of the test standards against the Australian population revealed that, in order to select candidates whose literacy and numeracy proficiency is equivalent to the top 30% of adult Australians, some adjustment of the standards in 2017 is advised.

In 2017, a refreshed version of the test will become available for testing windows 3 and 4. To provide candidates with additional information about their results in relation to the standard, a described proficiency scale will be released before testing window 3 with three described bands (two above the standard and one below the standard). Before the end of 2017, additional sample and practice items will be released and an additional secure emergency test will be created.

9 REFERENCES

- Adams, R.J, Wu, M.L, and Wilson, M.R. (2015). *ConQuest 4*. [computer program] Camberwell: Australian Council for Educational Research.
- Monseur, C. and Berezner, A. (2007). The Computation of Equating Errors in International Surveys in Education. *Journal of Applied Measurement*, 8(3), 323–335.
- PIAAC (2016). *Technical Report of the Survey of Adult Skills (PIAAC) (2nd Edition), Chapter 17: Scaling PIAAC Cognitive Data*. OECD.
- Rasch, G (1980). *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago: MESA Press (original work published 1960).
- Stocking, M. L. & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7(2), 201–210.

10 APPENDICES

Appendix 1: Score Frequency Distribution

Table 32: Literacy score frequency distribution of first-attempt candidates

Scale Score	Frequency	Percentile
80	1	0.0
84	2	0.0
85	1	0.0
87	1	0.0
88	2	0.1
89	2	0.1
90	1	0.1
91	3	0.1
92	6	0.1
93	6	0.2
94	5	0.2
95	9	0.3
96	12	0.4
97	28	0.6
98	23	0.8
99	45	1.1
100	50	1.5
101	52	1.9
102	101	2.7
103	150	3.8
104	168	5.1
105	203	6.7
106	273	8.7
107	303	11.1
108	373	13.9
109	371	16.7
110	504	20.6
111	533	24.7
112	576	29.1
113	519	33.0
114	681	38.2
115	636	43.1
116	710	48.5
117	624	53.3
118	334	55.9
119	626	60.6
120	763	66.5
121	193	68.0
122	566	72.3
123	774	78.2
124	124	79.1
125	588	83.6
126	144	84.7
127	483	88.4
128	79	89.0
129	445	92.4
131	417	95.6
134	311	98.0
138	169	99.3
143	54	99.7
144	22	99.9
155	17	100.0

Pass Level

Table 33: Numeracy score frequency distribution of first-attempt candidates

Scale Score	Frequency	Percentile
82	2	0.0
84	1	0.0
87	1	0.0
88	4	0.1
90	10	0.1
91	3	0.2
92	10	0.2
93	13	0.3
94	13	0.4
95	16	0.6
96	30	0.8
97	40	1.1
98	43	1.4
99	45	1.8
100	60	2.2
101	70	2.8
102	77	3.3
103	103	4.1
104	120	5.1
105	138	6.1
106	191	7.6
107	154	8.7
108	174	10.1
109	249	12.0
110	258	13.9
111	295	16.2
112	322	18.7
113	226	20.4
114	376	23.3
115	412	26.4
116	453	29.9
117	383	32.8
118	359	35.5
119	582	40.0
120	529	44.0
121	287	46.2
122	666	51.3
123	383	54.2
124	485	58.0
125	548	62.1
126	360	64.9
127	536	69.0
128	409	72.1
129	358	74.9
130	560	79.1
131	50	79.5
132	598	84.1
133	195	85.6
134	32	85.8
135	446	89.2
136	251	91.1
137	16	91.3
138	11	91.4
139	315	93.8
140	251	95.7
141	10	95.8
144	24	95.9
145	217	97.6
146	133	98.6

Pass Level

147	11	98.7
156	23	98.9
157	115	99.8
158	32	100.0

Appendix 2: Performance by Demographic Characteristics and Test Windows

Table 34 to Table 37 show performance by demographic characteristics for each test window. It can be seen that differences between the achievement of each group are consistent across test windows. In general, the overall findings in Section 3: Candidate Performance are also true for each test window and there are no major differences in the achievement of candidates in each group between test windows. One exception was test window 3 where, for literacy, the mean scale score of female candidates was more than two scale score points lower than male candidates, and the pass rate of female candidates was nearly 5% lower than that of male candidates. In numeracy, the pass rates of male candidates were more than 5% higher than the pass rates of female candidates in all test windows and 10% higher in test window 3.

Table 34: Performance by demographic characteristics in test window 1

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	3052	117.3	8.5	94.0	3073	121.2	10.8	92.0
	Male	778	118.8	7.9	96.4	770	127.6	10.5	97.3
	Indeterminate/intersex	1	-	-	-	1	-	-	-
Age	17–25	2624	116.6	7.7	94.6	2635	121.7	10.4	93.0
	26–30	493	118.3	9.6	90.3	492	122.8	11.8	92.1
	31–35	247	120.4	8.7	98.0	248	124.9	11.5	94.8
	36–40	192	121.7	8.9	97.4	191	126.1	10.8	96.9
	41–45	152	121.9	9.5	96.7	155	125.4	13.8	92.3
	46+	123	120.5	9.2	95.1	123	124.1	13.1	90.2
International Students	No	3722	117.8	8.3	94.7	3736	122.4	11.1	93.1
	Yes	109	112.7	8.9	85.3	108	124.1	10.5	92.6
English as a First Language	Yes	3539	118.0	8.2	95.4	3555	122.6	10.9	93.5
	No	292	113.0	8.9	83.2	289	121.5	12.5	87.5
Indigenous	No	3676	117.7	8.4	94.5	3686	122.6	11.0	93.2
	Yes	69	115.3	8.8	91.3	69	117.5	10.8	87.0
	Not disclosed	86	118.1	7.5	96.5	89	122.7	11.7	93.3
Residential Area	Metropolitan areas	3132	117.8	8.4	94.6	3151	122.8	11.2	93.3
	Provincial areas	672	116.9	8.0	93.8	666	120.8	10.1	91.6
	Remote areas	15	120.5	9.6	100.0	15	125.4	8.8	100.0
	International	10	124.9	9.4	100.0	10	128.3	9.3	100.0
	Invalid or Missing	2	-	-	-	2	-	-	-
Program Type	Undergraduate	2509	116.5	7.6	94.3	2520	121.0	10.3	92.3
	Postgraduate	1322	119.8	9.3	94.9	1324	125.3	11.9	94.4
Program Type by Year Level	Undergraduate first year	0				0			
	Undergraduate second year	12	122.5	8.2	100.0	12	127.3	11.1	100.0
	Undergraduate third year	22	117.0	6.9	95.5	27	120.5	10.7	88.9
	Undergraduate fourth year	2475	116.5	7.6	94.2	2481	121.0	10.3	92.3
	Undergraduate fifth year or above	0				0			
	Postgraduate first year	200	119.2	9.8	95.0	200	126.0	11.3	93.5
	Postgraduate second year	858	120.8	8.9	96.5	854	126.3	11.7	96.0
	Postgraduate third year	71	117.8	9.3	88.7	70	122.2	11.1	91.4
	Postgraduate fourth year	193	117.0	9.5	89.6	200	121.6	12.5	89.5

	Postgraduate fifth year or above	0				0			
Course Category	Teacher education: early childhood	164	114.1	8.1	85.4	165	116.6	11.0	83.0
	Teacher education: primary	1849	117.4	8.4	94.6	1869	121.5	10.6	93.0
	Teacher education: secondary	1296	119.1	8.3	96.0	1292	125.3	11.3	95.3
	Teacher education: special education	35	114.8	9.8	88.6	35	118.5	8.1	88.6
	Teacher education: other	487	116.1	7.8	93.4	483	121.1	10.3	91.1

Table 35: Performance by demographic characteristics in test window 2 (including resits)

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	2712	116.9	8.6	92.0	2721	120.8	11.1	89.9
	Male	778	118.6	8.6	94.6	766	127.2	11.1	96.5
	Indeterminate/intersex	2	-	-	-	1	-	-	-
Age	17–25	2299	116.2	8.1	91.7	2299	121.5	11.3	90.6
	26–30	564	118.5	8.9	92.9	554	123.6	11.2	92.8
	31–35	216	119.5	8.6	94.4	220	124.4	11.1	95.0
	36–40	180	120.4	9.7	95.6	180	124.0	12.0	93.3
	41–45	123	120.3	10.1	95.9	124	122.5	12.6	90.3
	46+	109	120.6	10.2	96.3	110	122.6	12.1	90.0
International Students	No	3360	117.4	8.6	92.8	3366	122.2	11.4	91.3
	Yes	132	114.2	8.1	87.1	122	124.0	11.4	93.4
English as a First Language	Yes	3058	117.8	8.5	93.6	3074	122.3	11.3	91.8
	No	434	113.7	8.6	85.3	414	121.5	12.0	88.2
Indigenous	No	3371	117.3	8.6	92.6	3375	122.2	11.4	91.3
	Yes	43	115.6	8.5	90.7	39	118.8	10.1	89.7
	Not disclosed	78	119.6	9.7	92.3	74	124.8	11.6	94.6
Residential Area	Metropolitan areas	2962	117.4	8.7	92.5	2969	122.3	11.6	91.0
	Provincial areas	519	116.8	7.9	93.1	509	122.0	10.2	93.5
	Remote areas	5	-	-	-	5	-	-	-
	International	2	-	-	-	1	-	-	-
	Invalid or Missing	4	-	-	-	4	-	-	-
Program Type	Undergraduate	2079	115.7	8.0	91.2	2070	120.6	10.8	90.0
	Postgraduate	1413	119.5	9.1	94.6	1418	124.6	11.8	93.3
Program Type by Year Level	Undergraduate first year	39	114.5	7.8	87.2	38	121.8	13.3	84.2
	Undergraduate second year	74	115.0	8.2	90.5	69	121.8	10.1	91.3
	Undergraduate third year	450	115.3	7.6	92.0	444	120.8	10.7	92.1
	Undergraduate fourth year	1335	115.6	7.9	90.9	1338	120.0	10.7	89.5
	Undergraduate fifth year or above	181	118.3	8.8	93.4	181	123.4	11.5	90.1
	Postgraduate first year	533	121.1	8.8	97.0	533	126.9	11.2	95.9
	Postgraduate second year	607	119.5	9.2	94.2	614	124.1	11.8	93.3
	Postgraduate third year	39	119.7	7.8	97.4	37	123.4	10.2	91.9
	Postgraduate fourth year	122	114.8	8.8	86.1	120	118.9	12.4	84.2

	Postgraduate fifth year or above	112	117.9	8.5	92.9	114	123.1	11.8	91.2
Course Category	Teacher education: early childhood	199	114.4	7.5	92.5	199	118.1	11.0	84.9
	Teacher education: primary	1337	116.4	8.2	92.4	1352	120.7	11.0	89.3
	Teacher education: secondary	1347	119.6	8.9	95.1	1330	125.7	11.4	95.4
	Teacher education: special education	19	112.6	6.1	84.2	19	115.5	7.4	94.7
	Teacher education: other	590	115.0	8.2	87.6	588	119.4	10.3	89.1

Table 36: Performance by demographic characteristics in test window 3 (including resits)

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	2278	115.8	8.7	89.5	2333	119.5	11.2	87.3
	Male	884	118.0	8.2	94.3	877	126.8	11.2	97.3
	Indeterminate/intersex	0				0			
Age	17–25	2111	115.0	7.8	89.4	2148	120.4	11.2	89.0
	26–30	454	118.0	9.1	91.6	454	122.9	12.5	90.5
	31–35	208	119.5	9.5	96.2	209	124.8	11.1	95.2
	36–40	155	120.4	9.0	95.5	158	125.0	11.9	94.3
	41–45	117	121.1	9.9	95.7	122	124.5	12.8	91.0
	46+	117	120.0	10.4	92.3	119	123.9	13.1	89.9
International Students	No	3054	116.6	8.6	91.2	3106	121.5	11.6	89.9
	Yes	108	112.2	7.4	80.6	104	123.3	12.8	92.3
English as a First Language	Yes	2805	116.9	8.5	92.1	2863	121.6	11.6	90.5
	No	357	112.6	8.1	81.0	347	120.9	12.5	85.6
Indigenous	No	3032	116.5	8.6	91.1	3071	121.6	11.7	90.2
	Yes	49	113.2	8.5	81.6	52	116.8	9.5	78.8
	Not disclosed	81	117.1	8.7	87.7	87	121.3	11.4	88.5
Residential Area	Metropolitan areas	2603	116.4	8.7	90.7	2640	121.6	11.6	89.8
	Provincial areas	544	116.4	8.3	91.4	554	121.3	11.7	90.6
	Remote areas	7	-	-	-	8	-	-	-
	International	6	-	-	-	6	-	-	-
	Invalid or Missing	2	-	-	-	2	-	-	-
Program Type	Undergraduate	1992	115.0	7.8	89.5	2033	119.6	10.9	88.5
	Postgraduate	1170	118.9	9.3	93.2	1177	124.8	12.2	92.5
Program Type by Year Level	Undergraduate first year	36	113.1	9.4	75.0	36	119.9	13.2	88.9
	Undergraduate second year	78	117.1	9.5	85.9	76	122.3	11.7	93.4
	Undergraduate third year	651	115.3	7.5	89.9	665	120.1	10.3	91.6
	Undergraduate fourth year	1119	114.6	7.7	90.0	1141	118.9	11.1	86.1
	Undergraduate fifth year or above	108	116.6	7.9	88.9	115	122.0	10.3	92.2
	Postgraduate first year	487	119.4	9.5	95.3	488	126.5	11.5	95.9
	Postgraduate second year	357	119.3	8.9	93.6	368	124.2	11.7	93.7
	Postgraduate third year	61	118.3	8.8	90.2	59	123.7	12.7	86.4

	Postgraduate fourth year	147	117.2	10.0	87.8	143	121.7	13.6	83.9
	Postgraduate fifth year or above	118	118.6	9.4	91.5	119	123.7	13.6	88.2
Course Category	Teacher education: early childhood	124	111.7	8.6	78.2	125	113.5	10.8	72.0
	Teacher education: primary	1190	115.5	8.3	89.4	1231	119.6	10.9	87.9
	Teacher education: secondary	1187	118.6	8.8	93.9	1189	125.1	11.7	94.8
	Teacher education: special education	19	115.2	6.7	100.0	20	118.4	10.4	85.0
	Teacher education: other	642	115.0	7.9	89.9	645	120.2	11.2	88.8

Table 37: Performance by demographic characteristics in test window 4 (including resits)

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	2231	117.2	8.8	92.0	2276	119.9	11.2	87.7
	Male	733	118.1	8.7	93.6	713	126.0	11.2	96.4
	Indeterminate/intersex	0				1	-	-	-
Age	17–25	1780	115.8	8.0	91.4	1801	120.0	10.9	88.3
	26–30	464	118.9	8.8	93.1	467	123.3	12.0	92.5
	31–35	233	118.6	9.1	94.4	219	123.0	12.3	91.3
	36–40	162	119.9	9.5	95.1	165	123.3	11.8	91.5
	41–45	166	120.4	9.8	92.2	170	123.3	12.8	90.0
	46+	157	123.2	9.0	95.5	166	124.1	11.3	93.4
International Students	No	2860	117.5	8.7	92.5	2890	121.3	11.5	89.6
	Yes	104	114.3	8.4	89.4	100	124.1	11.0	93.0
English as a First Language	Yes	2624	118.0	8.6	93.3	2653	121.6	11.4	90.7
	No	340	112.9	8.3	85.6	337	119.5	12.2	82.5
Indigenous	No	2841	117.5	8.7	92.6	2863	121.4	11.5	89.9
	Yes	47	113.4	10.0	80.9	48	116.4	12.4	77.1
	Not disclosed	76	117.3	8.9	90.8	79	123.1	12.0	91.1
Residential Area	Metropolitan areas	2330	117.3	8.7	92.6	2345	121.3	11.6	89.0
	Provincial areas	593	117.3	8.9	91.1	606	121.2	11.2	92.1
	Remote areas	20	122.0	9.2	95.0	19	125.3	13.4	89.5
	International	14	121.9	7.0	100.0	13	126.3	6.2	100.0
	Invalid or Missing	7	-	-	-	7	-	-	-
Program Type	Undergraduate	1752	116.0	8.1	91.3	1766	119.7	10.9	88.1
	Postgraduate	1212	119.5	9.3	94.0	1224	123.8	12.0	92.1
Program Type by Year Level	Undergraduate first year	29	117.7	10.7	86.2	29	124.2	10.1	93.1
	Undergraduate second year	80	117.1	8.0	95.0	80	122.1	11.5	92.5
	Undergraduate third year	850	115.9	7.9	92.8	837	120.2	10.4	90.1
	Undergraduate fourth year	717	115.9	8.1	89.7	738	118.9	11.2	85.9
	Undergraduate fifth year or above	76	115.2	9.0	86.8	82	117.0	10.7	81.7
	Postgraduate first year	644	119.8	8.9	95.2	638	125.5	11.6	95.0
	Postgraduate second year	269	120.5	9.5	95.5	265	123.4	11.8	92.8

Course Category	Postgraduate third year	55	120.3	8.7	94.5	57	124.9	12.8	93.0
	Postgraduate fourth year	148	116.5	10.0	87.8	156	118.9	11.7	82.1
	Postgraduate fifth year or above	96	118.0	9.7	90.6	108	120.6	11.8	87.0
	Teacher education: early childhood	151	112.9	8.1	83.4	161	116.2	11.1	75.2
	Teacher education: primary	1138	116.9	8.4	92.3	1166	120.1	10.9	88.5
	Teacher education: secondary	1224	119.1	8.6	94.7	1216	124.0	11.6	93.7
	Teacher education: special education	17	112.6	10.6	76.5	17	116.1	9.0	82.4
	Teacher education: other	434	115.7	9.0	89.9	430	119.5	11.3	87.7

Figure 8 to Figure 12 show achievement distributions in literacy and numeracy by demographic characteristics for each test window.

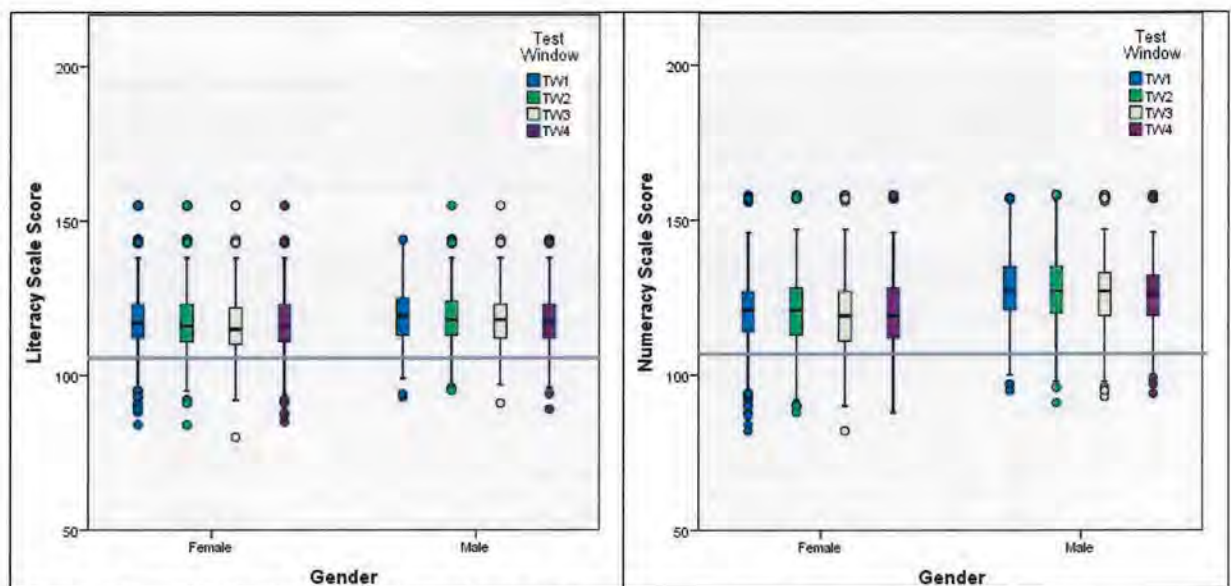


Figure 8: Score distribution by gender and test window

Figure 8 shows that, for literacy, the distributions of the scale scores of female candidates and male candidates are very similar. For numeracy, while the distributions of the scale scores of male candidates were higher up the scale than those of female candidates, it can be seen that in each testing window many female candidates achieved well above the numeracy standard.

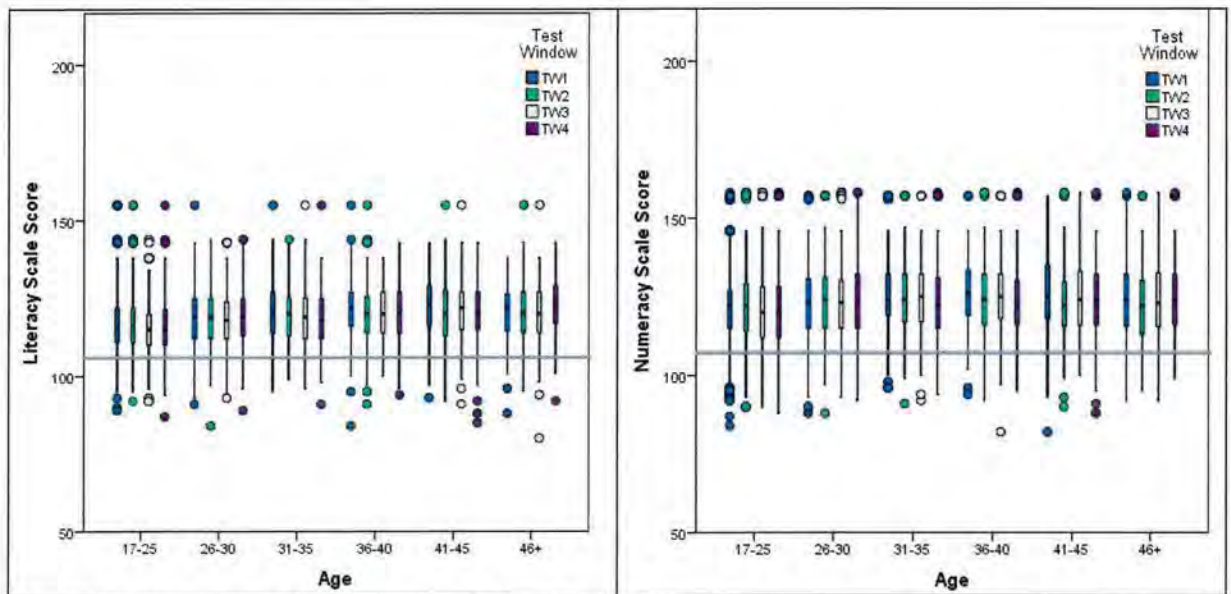


Figure 9: Score distribution by age group and test window

Figure 9 shows that, for both literacy and numeracy, the distributions of the scale scores of female candidates and male candidates are very similar. For numeracy, while the distributions of the scale scores of male candidates were higher up the scale than those of female candidates, it can be seen that in each testing window many female candidates achieved well above the numeracy standard.

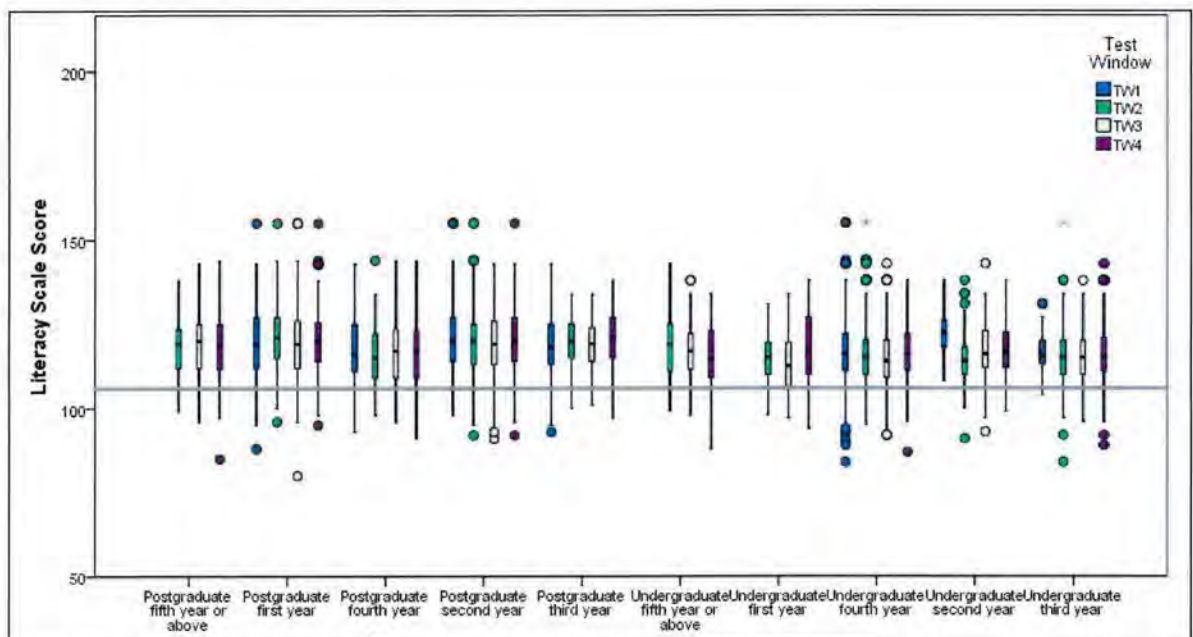


Figure 10: Score distribution by program type, year level and test window, literacy

Figure 10 shows that there are no observable patterns in the scale score distributions of the program types by year level between test windows for literacy.

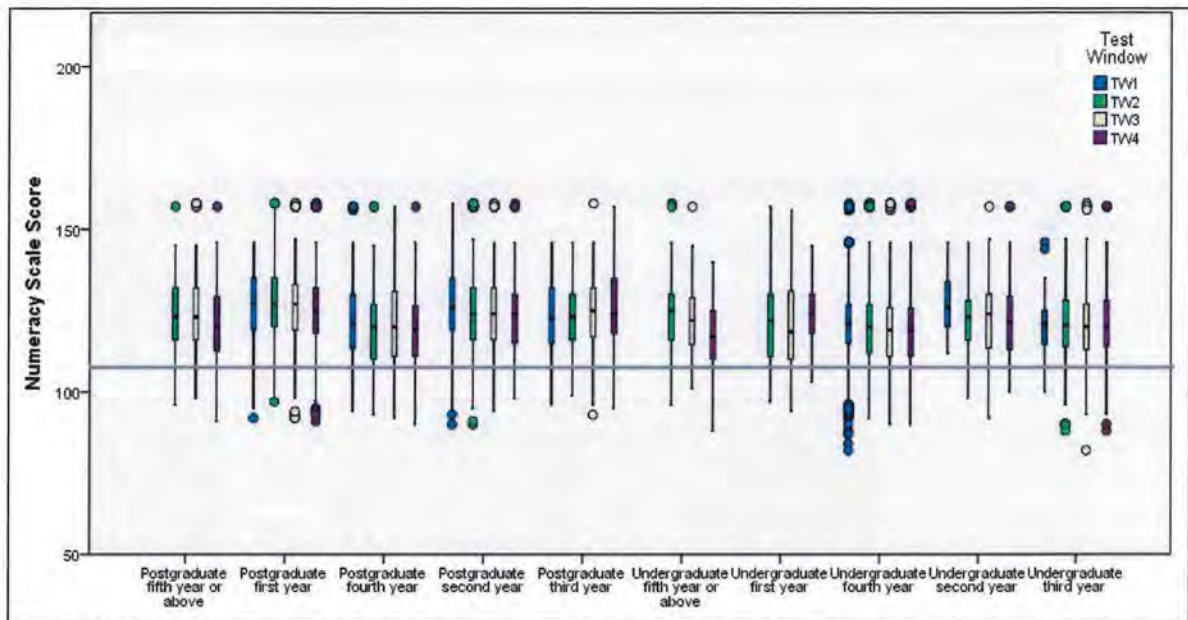


Figure 11: Score distribution by program type, year level and test window, numeracy

Similarly, Figure 11 shows that there are no observable patterns between test windows in the scale score distributions of the program types by year level for numeracy.

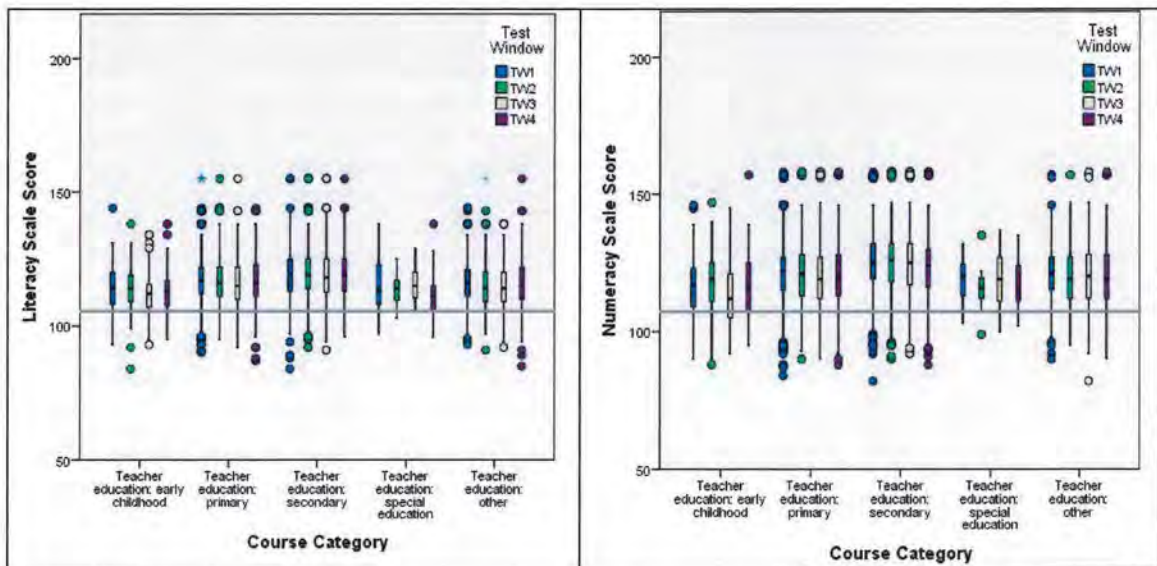


Figure 12: Score distribution by course category and test window

Figure 12 shows that there are no observable patterns between test windows in the scale score distributions of the course categories for both literacy and numeracy.

Appendix 3: Performance by Test Centres and Remote Proctoring by Test Window

Table 38 to Table 41 show performance by test centres and remote proctoring in test windows 1 to 4, respectively.

Table 38: Performance by test centres and remote proctoring in test window 1

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Test Centre	Adelaide CBD	26	117.7	9.2	88.5	28	122.7	10.7	100.0
	Albury	57	115.7	6.7	96.5	54	121.2	9.0	90.7
	Armidale	NA				NA			
	Ballarat	NA				NA			
	Bathurst	NA				NA			
	Brisbane CBD	407	117.7	7.7	96.6	407	123.1	10.6	94.3
	Bundaberg	NA				NA			
	Canberra CBD	191	116.6	8.0	92.7	191	121.0	11.1	91.6
	Darwin	7	-	-	-	7	-	-	-
	Hobart	44	122.9	9.5	97.7	44	124.5	11.1	95.5
	Melbourne CBD	998	116.7	8.0	94.0	1008	121.5	11.3	91.3
	Mildura	NA				NA			
	Newcastle	198	118.3	7.6	96.5	199	122.6	10.3	95.0
	Parramatta	NA				NA			
	Perth CBD	45	123.2	10.4	95.6	45	128.6	13.5	93.3
	Sunshine C./Maroochy.	NA				NA			
	Sydney CBD	1069	117.7	8.6	93.9	1079	122.5	11.1	93.0
	Townsville	10	117.1	8.2	100.0	10	123.1	10.6	100.0
	Wagga Wagga	NA				NA			
	Warrnambool	NA				NA			
	Wollongong	NA				NA			
	Remote Proctoring	779	118.3	8.9	94.4	772	123.4	10.8	94.3
	Capital Cities	2787	117.4	8.3	94.3	2809	122.3	11.2	92.6
	Regional Cities	265	117.7	7.5	96.6	263	122.3	10.1	94.3

Table 39: Performance by test centres and remote proctoring in test window 2

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Test Centre	Adelaide CBD	26	117.0	10.4	84.6	25	119.6	10.9	88.0
	Albury	36	117.4	8.9	91.7	35	121.4	9.6	91.4
	Armidale	36	116.9	6.1	100.0	36	123.0	12.9	91.7
	Ballarat	40	114.9	8.5	87.5	40	119.1	9.5	90.0
	Bathurst	18	118.2	6.7	100.0	18	123.1	11.9	94.4
	Brisbane CBD	236	116.5	8.0	92.8	236	120.1	11.1	88.1
	Bundaberg	9	-	-	-	9	-	-	-
	Canberra CBD	58	117.3	7.4	93.1	58	122.4	10.9	93.1
	Darwin	5	-	-	-	5	-	-	-
	Hobart	6	-	-	-	8	-	-	-
	Melbourne CBD	1118	117.2	8.8	90.9	1124	121.8	11.4	91.2
	Mildura	19	112.5	7.4	78.9	19	118.8	9.4	84.2
	Newcastle	153	116.3	7.5	92.2	153	122.0	11.7	92.2
	Parramatta	137	116.1	8.7	90.5	136	121.3	12.6	89.0
	Perth CBD	32	123.3	8.9	96.9	32	124.6	10.2	93.8
	Sunshine C./Maroochy.	46	117.5	6.8	97.8	45	123.1	8.2	100.0
	Sydney CBD	886	117.5	8.8	94.0	888	123.0	11.6	91.8
	Townsville	NA				NA			
	Wagga Wagga	40	114.7	7.7	87.5	40	120.8	9.1	92.5
	Warrnambool	28	117.2	6.8	96.4	30	126.8	11.5	96.7
	Wollongong	57	115.9	8.1	93.0	57	121.5	12.9	89.5
	Remote Proctoring	506	118.4	8.9	94.3	494	123.1	11.1	91.7
	Capital Cities	2367	117.3	8.7	92.3	2376	122.1	11.5	91.2
	Regional Cities	619	116.2	7.8	92.2	618	121.8	11.4	91.7

Table 40: Performance by test centres and remote proctoring in test window 3

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Test Centre	Adelaide CBD	50	119.6	7.7	100.0	51	127.3	11.5	100.0
	Albury	40	117.1	8.6	95.0	40	121.6	10.7	90.0
	Armidale	25	118.8	7.2	96.0	25	124.0	13.2	92.0
	Ballarat	80	114.3	6.3	93.8	80	120.3	10.6	93.7
	Bathurst	28	118.2	9.3	89.3	29	125.7	14.8	89.7
	Brisbane CBD	329	117.1	8.6	92.4	334	122.2	11.4	90.1
	Bundaberg	8	-	-	-	9	-	-	-
	Canberra CBD	39	118.8	6.7	97.4	43	124.6	11.3	93.0
	Darwin	13	118.7	8.3	100.0	13	120.5	9.0	100.0
	Hobart	NA				NA			
	Melbourne CBD	1029	115.5	8.4	89.4	1036	121.1	11.5	89.8
	Mildura	7	-	-	-	8	-	-	-
	Newcastle	155	116.9	8.3	92.3	156	120.8	10.8	90.4
	Parramatta	137	115.2	8.6	87.6	136	120.2	11.6	86.8
	Perth CBD	17	122.1	10.8	94.1	16	133.4	14.6	100.0
	Sunshine C./Maroochy.	21	119.1	8.5	90.5	18	121.4	9.7	94.4
	Sydney CBD	746	116.6	8.9	90.3	762	121.8	11.5	91.6
	Townsville	NA				NA			
	Wagga Wagga	31	116.9	7.8	90.3	33	122.2	9.9	93.9
	Warrnambool	19	115.6	7.9	94.7	21	114.8	11.9	71.4
	Wollongong	59	117.6	8.2	91.5	58	122.1	10.4	93.1
	Remote Proctoring	329	117.2	9.2	91.2	342	120.9	12.8	84.8
	Capital Cities	2223	116.3	8.6	90.6	2255	121.8	11.6	90.9
	Regional Cities	610	116.4	8.2	91.3	613	120.9	11.2	89.7

Table 41: Performance by test centres and remote proctoring in test window 4

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Test Centre	Adelaide CBD	30	123.1	8.8	96.7	29	121.2	10.7	89.7
	Albury	19	117.1	8.2	100.0	18	118.8	11.3	83.3
	Armidale	38	118.5	7.7	97.4	38	121.4	9.6	94.7
	Ballarat	57	116.3	8.7	91.2	57	118.5	11.5	93.0
	Bathurst	19	119.5	7.1	100.0	19	121.3	10.4	94.7
	Brisbane CBD	374	118.4	8.4	94.1	372	122.1	11.4	91.9
	Bundaberg	20	118.4	8.9	90.0	20	119.9	12.0	90.0
	Canberra CBD	29	119.3	10.8	86.2	30	121.2	13.3	86.7
	Darwin	10	112.6	9.1	90.0	10	115.2	9.0	80.0
	Hobart	NA				NA			
	Melbourne CBD	696	116.8	9.2	90.2	699	121.2	11.8	88.7
	Mildura	0				0			
	Newcastle	80	117.9	9.4	91.3	79	122.8	11.6	96.2
	Parramatta	93	114.2	6.7	90.3	95	118.2	12.4	82.1
	Perth CBD	32	120.0	10.7	93.8	32	125.5	13.8	87.5
	Sunshine C./Maroochy.	60	118.7	7.9	96.7	59	122.6	10.3	89.8
	Sydney CBD	487	116.6	8.1	93.4	487	120.5	10.7	89.7
	Townsville	NA				NA			
	Wagga Wagga	19	118.3	7.2	94.7	19	123.3	9.3	100.0
	Warrnambool	NA				NA			
	Wollongong	30	117.8	7.0	100.0	30	123.4	10.6	93.3
	Remote Proctoring	871	117.8	8.9	92.1	897	121.8	11.7	89.4
	Capital Cities	1658	117.3	8.8	92.2	1659	121.3	11.5	89.6
	Regional Cities	435	117.1	8.1	93.8	434	120.8	11.3	90.8

Comparing the mean scale scores for remote proctoring and test centres in capital cities and regional cities at the bottom of each test window table, it can be seen that while mean scale scores tended to be higher in test window 1 and lower in test windows 3 and 4, the differences were in the order of one scale score point only and not significant.

Appendix 4: Trial Item Analysis Tables

Table 42: Summary of trial item analysis for literacy

Item Label	Content Subscale	Item Difficulty Estimate (logit)	Item Difficulty Estimate (scale score)	Facility (%)	Item-Rest Corr.	Infit	Number of Cases
L007403	Tech. Skills of Writ.	3.641	136.4	13.61	0.29	0.97	316
L003509	Reading	3.375	133.8	18.00	0.10	1.06	550
L008201	Tech. Skills of Writ.	3.370	133.7	19.41	0.12	1.06	273
L003601	Reading	3.116	131.2	23.10	0.32	0.98	567
L002604	Reading	2.955	129.6	25.36	0.19	1.05	556
L003402	Reading	2.709	127.1	29.78	0.25	1.01	272
L002608	Reading	2.485	124.9	34.17	0.14	1.09	556
L007504	Tech. Skills of Writ.	2.385	123.9	34.43	0.39	0.92	305
L007301	Tech. Skills of Writ.	2.359	123.6	34.45	0.06	1.14	656
L002603	Reading	2.176	121.8	37.70	0.26	1.00	793
L003404	Reading	2.149	121.5	41.18	0.32	0.98	272
L002605	Reading	2.105	121.1	39.22	0.21	1.04	793
L007401	Tech. Skills of Writ.	1.968	119.7	41.77	0.27	1.02	316
L006801	Tech. Skills of Writ.	1.949	119.5	43.71	0.43	0.91	286
L002602	Reading	1.845	118.5	47.84	0.10	1.12	556
L002609	Reading	1.819	118.2	45.52	0.38	0.94	793
L008603	Tech. Skills of Writ.	1.805	118.1	47.28	0.36	0.95	294
L003001	Reading	1.719	117.2	50.00	0.29	1.00	252
L007603	Tech. Skills of Writ.	1.657	116.6	49.74	0.27	1.01	756
L007402	Tech. Skills of Writ.	1.651	116.5	48.73	0.29	1.02	316
L007605	Tech. Skills of Writ.	1.633	116.3	50.26	0.13	1.11	756
L003610	Reading	1.614	116.1	53.09	0.38	0.95	680
L003904	Reading	1.569	115.7	52.01	0.36	0.96	298
L006802	Tech. Skills of Writ.	1.404	114.0	55.59	0.23	1.06	286
L003603	Reading	1.354	113.5	58.32	0.41	0.94	571
L003302	Reading	1.304	113.0	58.21	0.46	0.89	402
L004105	Reading	1.247	112.5	60.27	0.30	0.99	297
L002601	Reading	1.216	112.2	59.02	0.26	1.01	793
L007604	Tech. Skills of Writ.	1.196	112.0	59.92	0.27	1.01	756
L003611	Reading	1.161	111.6	62.79	0.45	0.90	680
L003111	Reading	1.063	110.6	62.41	0.13	1.13	290
L008803	Tech. Skills of Writ.	1.050	110.5	63.16	0.46	0.90	285
L008703	Tech. Skills of Writ.	1.044	110.4	64.44	0.39	0.96	284
L003605	Reading	1.024	110.2	65.15	0.37	0.96	571
L003803	Reading	1.016	110.2	64.08	0.18	1.06	348
L003405	Reading	1.014	110.1	65.81	0.26	1.02	272
L003802	Reading	1.002	110.0	64.37	0.43	0.91	348
L003503	Reading	0.894	108.9	64.13	0.35	0.97	276
L004102	Reading	0.863	108.6	65.47	0.26	1.04	278
L002303	Reading	0.829	108.3	68.80	0.09	1.12	1237
L005901	Tech. Skills of Writ.	0.805	108.1	68.37	0.14	1.11	294
L004107	Reading	0.772	107.7	67.27	0.19	1.08	278
L007902	Tech. Skills of Writ.	0.757	107.6	70.07	0.25	1.02	284
L003503	Reading	0.734	107.3	70.55	0.29	0.99	550
L002401	Reading	0.717	107.2	71.51	0.37	0.95	1176
L003910	Reading	0.703	107.0	70.03	0.47	0.88	297
L003906	Reading	0.676	106.8	70.47	0.28	1.01	298

L002406	Reading	0.665	106.7	72.49	0.26	1.01	1174
L002607	Reading	0.674	106.7	70.24	0.34	0.96	793
L003310	Reading	0.607	106.1	72.14	0.37	0.95	402
L008805	Tech. Skills of Writ.	0.599	106.0	71.93	0.37	0.94	285
L003303	Reading	0.597	106.0	72.38	0.25	1.00	688
L003602	Reading	0.577	105.8	73.97	0.26	1.02	680
L007302	Tech. Skills of Writ.	0.538	105.4	75.09	0.11	1.12	273
L003609	Reading	0.530	105.3	74.39	0.38	0.95	570
L003002	Reading	0.498	105.0	74.60	0.34	0.98	252
L002304	Reading	0.471	104.7	75.26	0.29	0.99	1237
L007901	Tech. Skills of Writ.	0.462	104.6	75.35	0.36	0.95	284
L003507	Reading	0.451	104.5	75.64	0.24	1.01	550
L003507	Reading	0.437	104.4	72.92	0.45	0.90	277
L007501	Tech. Skills of Writ.	0.429	104.3	75.08	0.28	1.00	305
L003607	Reading	0.414	104.1	76.32	0.35	0.96	570
L003006	Reading	0.353	103.5	76.98	0.45	0.90	252
L002402	Reading	0.352	103.5	77.79	0.27	1.00	1175
L003608	Reading	0.345	103.5	77.79	0.36	0.95	680
L003005	Reading	0.328	103.3	77.38	0.33	0.97	252
L003502	Reading	0.330	103.3	77.64	0.12	1.06	550
L003604	Reading	0.259	102.6	79.12	0.40	0.91	680
L003502	Reading	0.227	102.3	76.53	0.19	1.04	277
L008501	Tech. Skills of Writ.	0.227	102.3	78.18	0.26	0.99	573
L004101	Reading	0.184	101.8	79.80	0.25	1.00	297
L005902	Tech. Skills of Writ.	0.171	101.7	79.25	0.26	1.01	294
L008502	Tech. Skills of Writ.	0.112	101.1	79.11	0.15	1.08	316
L002805	Reading	0.102	101.0	80.27	0.40	0.91	669
L003806	Reading	0.079	100.8	80.75	0.17	1.05	348
L002307	Reading	0.078	100.8	81.34	0.08	1.10	1238
L003908	Reading	0.068	100.7	80.54	0.38	0.93	298
L005301	Tech. Skills of Writ.	0.071	100.7	80.37	0.30	0.98	657
L007601	Tech. Skills of Writ.	0.068	100.7	80.69	0.21	1.02	756
L008901	Tech. Skills of Writ.	0.052	100.5	80.98	0.45	0.88	305
L002904	Reading	0.023	100.2	80.73	0.13	1.06	685
L003314	Reading	0.004	100.0	81.80	0.28	1.00	401
L007602	Tech. Skills of Writ.	-0.007	99.9	81.75	0.33	0.95	756
L002611	Reading	-0.053	99.5	83.60	0.27	0.98	555
L002807	Reading	-0.047	99.5	82.36	0.26	0.98	669
L003003	Reading	-0.058	99.4	82.94	0.50	0.86	252
L003809	Reading	-0.088	99.1	83.05	0.28	0.98	348
L008404	Tech. Skills of Writ.	-0.088	99.1	84.13	0.15	1.05	252
L008905	Tech. Skills of Writ.	-0.116	98.8	83.28	0.44	0.88	305
L007404	Tech. Skills of Writ.	-0.182	98.2	83.23	0.11	1.08	316
L008902	Tech. Skills of Writ.	-0.221	97.8	84.59	0.32	0.95	305
L007804	Tech. Skills of Writ.	-0.218	97.8	85.27	0.12	1.06	550
L008202	Tech. Skills of Writ.	-0.246	97.5	86.08	0.32	0.95	273
L002907	Reading	-0.251	97.5	84.38	0.14	1.04	685
L005904	Tech. Skills of Writ.	-0.263	97.4	85.03	0.37	0.94	294
L002405	Reading	-0.289	97.1	86.38	0.12	1.05	1175
L007001	Tech. Skills of Writ.	-0.449	95.5	87.09	0.27	0.98	573
L008203	Tech. Skills of Writ.	-0.457	95.4	88.28	0.27	0.97	273
L007803	Tech. Skills of Writ.	-0.505	95.0	88.36	0.19	1.01	550
L008801	Tech. Skills of Writ.	-0.520	94.8	87.72	0.34	0.95	285
L002813	Reading	-0.629	93.7	87.97	0.27	0.97	291
L003801	Reading	-0.726	92.7	89.94	0.27	0.96	348

L008403	Tech. Skills of Writ.	-0.752	92.5	90.87	0.33	0.93	252
L003106	Reading	-0.826	91.7	90.34	0.36	0.91	290
L002305	Reading	-0.855	91.5	91.28	0.13	1.02	1238
L008701	Tech. Skills of Writ.	-0.946	90.5	91.90	0.29	0.96	284
L007802	Tech. Skills of Writ.	-0.992	90.1	92.36	0.13	1.02	550
L003403	Reading	-1.009	89.9	92.65	0.15	1.02	272
L008405	Tech. Skills of Writ.	-1.029	89.7	92.86	0.20	0.99	252
L007005	Tech. Skills of Writ.	-1.080	89.2	92.50	0.06	1.04	573
L003312	Reading	-1.107	88.9	92.88	0.16	1.00	688
L005303	Tech. Skills of Writ.	-1.212	87.9	93.15	0.20	0.99	657
L008602	Tech. Skills of Writ.	-1.257	87.4	93.54	0.15	1.03	294
L004106	Reading	-1.267	87.3	93.94	0.27	0.95	297
L008402	Tech. Skills of Writ.	-1.305	87.0	94.44	0.25	0.96	252
L003309	Reading	-1.331	86.7	94.19	0.19	0.98	688
L003103	Reading	-1.399	86.0	94.46	0.22	0.99	307
L002811	Reading	-1.459	85.4	94.16	0.14	1.00	291
L007004	Tech. Skills of Writ.	-1.513	84.9	94.94	0.26	0.96	573
L003308	Reading	-1.528	84.7	95.02	0.13	1.01	402
L003109	Reading	-1.536	84.6	95.11	0.27	0.96	307
L002802	Reading	-1.596	84.0	94.85	0.14	0.99	291
L002806	Reading	-1.596	84.0	94.85	0.26	0.95	291
L003104	Reading	-1.682	83.2	95.52	0.26	0.94	290
L003108	Reading	-1.690	83.1	95.77	0.22	0.97	307
L006603	Tech. Skills of Writ.	-1.697	83.0	96.00	0.23	0.97	550
L002404	Reading	-1.702	83.0	96.09	0.27	0.96	1175
L004104	Reading	-1.707	82.9	95.96	0.16	1.00	297
L003307	Reading	-1.754	82.5	96.08	0.31	0.93	688
L003902	Reading	-1.779	82.2	95.97	0.11	1.01	298
L005302	Tech. Skills of Writ.	-1.814	81.9	96.04	0.31	0.96	657
L006601	Tech. Skills of Writ.	-1.902	81.0	96.83	0.18	0.98	252
L003501	Reading	-1.996	80.0	96.39	0.31	0.95	277
L003105	Reading	-2.081	79.2	97.07	0.25	0.96	307
L004108	Reading	-2.086	79.1	96.76	0.21	0.97	278
L003301	Reading	-2.124	78.8	97.24	0.24	0.95	688
L008503	Tech. Skills of Writ.	-2.143	78.6	97.21	0.19	0.97	573
L005304	Tech. Skills of Writ.	-2.148	78.5	97.11	0.21	0.99	657
L007801	Tech. Skills of Writ.	-2.173	78.3	97.45	0.12	1.00	550
L003102	Reading	-2.197	78.0	97.24	0.10	0.99	290
L003304	Reading	-2.495	75.1	98.01	0.14	0.99	402
L002302	Reading	-2.512	74.9	98.14	0.12	0.98	1237
L003501	Reading	-2.750	72.5	98.55	0.17	0.98	550
L002812	Reading	-3.017	69.8	98.81	0.12	0.98	670
L002906	Reading	-3.221	67.8	98.98	0.18	0.99	685
L002810	Reading	-3.308	66.9	99.10	0.23	0.96	669
L004103	Reading	-3.557	64.4	99.33	0.15	0.99	297
L004110	Reading	-3.633	63.7	99.28	0.14	0.98	278
L002809	Reading	-3.686	63.1	99.31	0.12	0.99	291
L002801	Reading	-3.718	62.8	99.40	0.12	0.99	669
L002902	Reading	-3.788	62.1	99.42	0.08	1.01	685
L002901	Reading	-4.078	59.2	99.56	0.11	1.00	685

Table 43: Summary of trial item analysis for numeracy

Item Label	Content Subscale	Calculator Subscale	Item Difficulty Estimate (logit)	Item Difficulty Estimate (scale score)	Facility (%)	Item-Rest Corr.	Infit	Number of cases
N029102	Statistics	Calculator available	3.886	138.9	18.63	0.18	1.06	612
N028501	Number	Calculator available	3.047	130.5	34.67	0.34	0.96	300
N061202	Statistics	Calculator available	3.017	130.2	34.55	0.41	0.93	550
N027101	Statistics	Calculator available	2.453	124.5	42.91	0.43	0.96	501
N050203	Number	Calculator available	2.416	124.2	41.92	0.63	0.75	291
N061103	Algebra	Calculator available	2.218	122.2	51.46	0.50	0.87	995
N061901	Statistics	Calculator available	2.157	121.6	49.06	0.54	0.84	477
N053202	Measure.	Calculator not available	2.146	121.5	47.08	0.50	0.93	291
N002401	Measure.	Calculator available	2.002	120.0	54.13	0.49	0.90	303
N060501	Measure.	Calculator available	1.989	119.9	54.91	0.22	1.13	499
N026402	Measure.	Calculator available	1.991	119.9	55.46	0.50	0.90	669
N060602	Statistics	Calculator available	1.885	118.9	53.55	0.44	0.95	310
N027602	Number	Calculator available	1.854	118.5	54.19	0.60	0.80	310
N002402	Measure.	Calculator available	1.800	118.0	58.09	0.32	1.08	303
N002002	Measure.	Calculator available	1.746	117.5	60.16	0.49	0.90	487
N029101	Statistics	Calculator available	1.729	117.3	59.17	0.30	1.04	600
N026002	Statistics	Calculator available	1.727	117.3	58.56	0.51	0.90	292
N027002	Statistics	Calculator available	1.686	116.9	58.07	0.49	0.94	322
N061201	Number	Calculator available	1.610	116.1	62.18	0.52	0.89	550
N053902	Number	Calculator available	1.598	116.0	61.20	0.41	1.00	299
N051402	Measure.	Calculator available	1.566	115.7	58.39	0.34	1.07	310
N029001	Statistics	Calculator available	1.574	115.7	62.95	0.53	0.88	413
N061001	Number	Calculator available	1.560	115.6	63.02	0.51	0.89	622
N051102	Probability	Calculator available	1.536	115.4	64.33	0.47	0.91	485
N053701	Measure.	Calculator available	1.523	115.2	61.18	0.49	0.94	322
N061002	Number	Calculator available	1.507	115.1	63.46	0.59	0.81	635
N060701	Statistics	Calculator available	1.489	114.9	65.06	0.20	1.16	498
N051101	Probability	Calculator available	1.476	114.8	65.37	0.39	0.99	488
N026001	Statistics	Calculator available	1.440	114.4	64.26	0.59	0.81	291

N053802	Measure.	Calculator not available	1.415	114.2	61.29	0.48	0.94	310
N050502	Measure.	Calculator available	1.386	113.9	65.29	0.48	0.92	291
N025802	Measure.	Calculator available	1.393	113.9	67.33	0.34	1.07	300
N027401	Number	Calculator available	1.341	113.4	64.34	0.52	0.88	673
N041102	Number	Calculator available	1.315	113.2	66.78	0.58	0.82	307
N061601	Number	Calculator available	1.299	113.0	67.05	0.53	0.86	346
N026901	Measure.	Calculator not available	1.266	112.7	69.06	0.44	0.93	501
N028302	Statistics	Calculator available	1.239	112.4	67.89	0.36	1.01	299
N026102	Measure.	Calculator available	1.197	112.0	70.42	0.26	1.15	311
N028401	Number	Calculator not available	1.184	111.8	67.66	0.52	0.89	501
N027001	Number	Calculator available	1.149	111.5	68.01	0.44	0.98	322
N027201	Measure.	Calculator available	1.135	111.4	70.77	0.52	0.88	284
N050602	Number	Calculator available	1.144	111.4	70.49	0.41	1.00	288
N028601	Number	Calculator available	1.125	111.3	72.30	0.27	1.10	964
N053201	Measure.	Calculator not available	1.100	111.0	69.30	0.42	1.01	316
N054301	Measure.	Calculator not available	1.084	110.8	72.33	0.22	1.16	300
N052301	Measure.	Calculator available	1.019	110.2	72.57	0.43	0.96	288
N028202	Number	Calculator available	1.007	110.1	70.91	0.52	0.89	471
N061502	Statistics	Calculator available	1.001	110.0	72.89	0.39	1.00	712
N061701	Number	Calculator not available	0.973	109.7	74.85	0.41	0.96	1002
N026301	Measure.	Calculator available	0.967	109.7	71.12	0.54	0.87	322
N025902	Measure.	Calculator available	0.961	109.6	73.64	0.46	0.93	550
N060301	Statistics	Calculator available	0.902	109.0	75.30	0.45	0.95	579
N026401	Measure.	Calculator available	0.882	108.8	75.83	0.32	1.05	666
N029202	Statistics	Calculator available	0.862	108.6	75.54	0.42	0.94	552
N053901	Number	Calculator available	0.864	108.6	74.25	0.40	0.98	299
N053801	Number	Calculator not available	0.850	108.5	74.91	0.32	1.04	271
N051602	Statistics	Calculator available	0.795	108.0	76.85	0.39	1.01	311
N027801	Statistics	Calculator available	0.775	107.8	74.29	0.17	1.20	673
N061902	Statistics	Calculator available	0.773	107.7	74.42	0.30	1.10	477
N028101	Number	Calculator not available	0.714	107.1	78.35	0.47	0.91	485
N026101	Measure.	Calculator available	0.707	107.1	78.14	0.30	1.08	311

N051301	Number	Calculator available	0.673	106.7	77.31	0.35	1.03	639
N061401	Geometry	Calculator available	0.663	106.6	77.98	0.35	0.98	613
N043401	Measure.	Calculator available	0.660	106.6	79.66	0.41	0.95	998
N052601	Number	Calculator not available	0.657	106.6	76.21	0.41	0.99	475
N027402	Number	Calculator available	0.568	105.7	77.53	0.41	0.97	672
N050001	Number	Calculator available	0.522	105.2	80.15	0.30	1.07	413
N052901	Number	Calculator not available	0.506	105.1	78.51	0.33	1.03	670
N025901	Measure.	Calculator available	0.473	104.7	80.76	0.38	0.99	551
N061102	Algebra	Calculator available	0.431	104.3	82.73	0.30	1.03	996
N062002	Number	Calculator available	0.425	104.3	82.00	0.25	1.05	550
N052802	Statistics	Calculator available	0.415	104.2	81.14	0.46	0.91	631
N053501	Measure.	Calculator not available	0.417	104.2	79.80	0.39	0.97	307
N052001	Measure.	Calculator available	0.410	104.1	78.01	0.57	0.82	291
N061402	Geometry	Calculator available	0.398	104.0	81.73	0.26	1.03	613
N020802	Number	Calculator available	0.401	104.0	81.60	0.20	1.16	288
N060402	Statistics	Calculator available	0.403	104.0	81.11	0.34	1.05	307
N051202	Number	Calculator available	0.382	103.8	82.54	0.39	0.98	670
N054202	Measure.	Calculator not available	0.370	103.7	81.79	0.52	0.88	302
N029201	Statistics	Calculator available	0.310	103.1	83.33	0.33	0.99	552
N050101	Measure.	Calculator available	0.297	103.0	82.98	0.42	0.94	711
N025801	Measure.	Calculator available	0.288	102.9	83.67	0.18	1.17	300
N027701	Measure.	Calculator available	0.265	102.7	84.30	0.40	0.95	535
N050901	Algebra	Calculator available	0.222	102.2	83.57	0.43	0.91	347
N022901	Statistics	Calculator available	0.183	101.8	85.64	0.41	0.93	954
N041101	Number	Calculator available	0.135	101.4	84.36	0.48	0.90	307
N051801	Number	Calculator not available	0.111	101.1	84.72	0.33	1.00	635
N054401	Number	Calculator available	0.086	100.9	85.21	0.42	0.95	284
N061302	Number	Calculator available	0.080	100.8	85.15	0.33	1.01	303
N050202	Statistics	Calculator available	0.075	100.8	82.47	0.31	1.03	291
N052501	Number	Calculator available	0.026	100.3	86.81	0.30	1.02	576
N053401	Statistics	Calculator available	0.010	100.1	86.67	0.40	0.94	300
N025501	Number	Calculator not available	-0.022	99.8	87.52	0.37	0.96	1507

N050301	Number	Calculator available	-0.050	99.5	87.63	0.40	0.93	485
N060601	Statistics	Calculator available	-0.055	99.5	85.48	0.42	0.92	310
N027302	Number	Calculator available	-0.089	99.1	85.86	0.36	0.98	502
N025502	Number	Calculator not available	-0.088	99.1	87.43	0.41	0.92	1344
N053301	Number	Calculator available	-0.094	99.1	87.67	0.43	0.91	300
N060502	Number	Calculator available	-0.100	99.0	88.02	0.24	1.04	501
N025601	Measure.	Calculator available	-0.120	98.8	86.29	0.35	0.98	671
N053001	Number	Calculator not available	-0.126	98.7	87.64	0.31	1.00	712
N050501	Measure.	Calculator available	-0.153	98.5	87.67	0.17	1.08	292
N060702	Statistics	Calculator available	-0.164	98.4	88.62	0.17	1.09	501
N028001	Number	Calculator available	-0.157	98.4	86.84	0.32	1.00	494
N054102	Number	Calculator available	-0.162	98.4	87.82	0.41	0.91	271
N025702	Number	Calculator not available	-0.190	98.1	88.09	0.46	0.89	1016
N050802	Measure.	Calculator not available	-0.221	97.8	87.99	0.27	1.06	308
N061101	Algebra	Calculator available	-0.269	97.3	89.82	0.31	0.99	1002
N050601	Number	Calculator available	-0.275	97.3	88.89	0.36	0.96	288
N060401	Statistics	Calculator available	-0.322	96.8	88.93	0.38	0.94	307
N028801	Number	Calculator available	-0.332	96.7	90.09	0.41	0.91	535
N028901	Number	Calculator not available	-0.348	96.5	89.44	0.33	1.00	284
N052502	Algebra	Calculator available	-0.361	96.4	90.12	0.40	0.92	577
N028701	Number	Calculator not available	-0.374	96.3	90.55	0.37	0.94	529
N053601	Number	Calculator not available	-0.403	96.0	90.60	0.27	0.99	553
N052202	Statistics	Calculator available	-0.406	95.9	90.37	0.37	0.95	301
N065001	Measure.	Calculator available	-0.441	95.6	90.20	0.42	0.87	347
N041602	Measure.	Calculator available	-0.439	95.6	90.31	0.28	1.02	413
N061801	Number	Calculator not available	-0.499	95.0	91.62	0.34	0.95	955
N027901	Measure.	Calculator not available	-0.509	94.9	90.78	0.25	1.00	347
N060901	Number	Calculator not available	-0.525	94.8	91.36	0.26	1.02	579
N051702	Statistics	Calculator available	-0.539	94.6	90.85	0.23	1.07	284
N052201	Statistics	Calculator available	-0.538	94.6	91.36	0.25	1.03	301
N026502	Statistics	Calculator available	-0.611	93.9	91.24	0.24	1.03	639
N027301	Measure.	Calculator available	-0.637	93.6	90.93	0.31	0.98	496

N050201	Statistics	Calculator available	-0.759	92.4	90.72	0.30	0.98	291
N054201	Measure.	Calculator not available	-0.790	92.1	92.93	0.21	1.03	283
N052101	Number	Calculator not available	-0.800	92.0	92.59	0.39	0.92	634
N028301	Statistics	Calculator available	-0.824	91.8	92.64	0.28	0.99	299
N060802	Measure.	Calculator available	-0.978	90.2	94.40	0.22	1.01	964
N050401	Statistics	Calculator available	-0.976	90.2	94.04	0.39	0.91	302
N054001	Statistics	Calculator available	-1.031	89.7	94.57	0.26	1.01	534
N052801	Statistics	Calculator available	-1.052	89.5	94.01	0.25	0.99	634
N062001	Number	Calculator available	-1.092	89.1	94.76	0.26	0.96	553
N025701	Measure.	Calculator not available	-1.192	88.1	94.56	0.25	0.98	1286
N051001	Measure.	Calculator available	-1.322	86.8	95.68	0.26	0.97	579
N061301	Number	Calculator available	-1.337	86.6	95.38	0.29	0.96	303
N062101	Number	Calculator available	-1.389	86.1	95.69	0.25	0.96	348
N027601	Measure.	Calculator available	-1.458	85.4	95.48	0.13	1.05	310
N028201	Number	Calculator available	-1.624	83.8	96.00	0.30	0.96	475
N051401	Number	Calculator available	-1.624	83.8	95.81	0.20	0.98	310
N051201	Statistics	Calculator available	-1.630	83.7	96.71	0.21	0.99	668
N026501	Statistics	Calculator available	-1.779	82.2	96.87	0.31	0.94	639
N051901	Measure.	Calculator available	-1.806	81.9	96.45	0.25	0.96	310
N050801	Number	Calculator not available	-1.859	81.4	97.25	0.14	1.00	291
N026801	Statistics	Calculator available	-1.928	80.7	97.42	0.21	0.96	271
N001701	Geometry	Calculator available	-2.097	79.0	97.94	0.18	0.99	535
N060801	Measure.	Calculator available	-2.150	78.5	98.13	0.18	0.98	960
N051902	Measure.	Calculator available	-2.448	75.5	98.06	0.17	0.99	310
N026201	Measure.	Calculator not available	-2.865	71.4	98.96	0.19	0.99	675

Appendix 5: Analysis of Trial Item Differential Functioning (DIF)

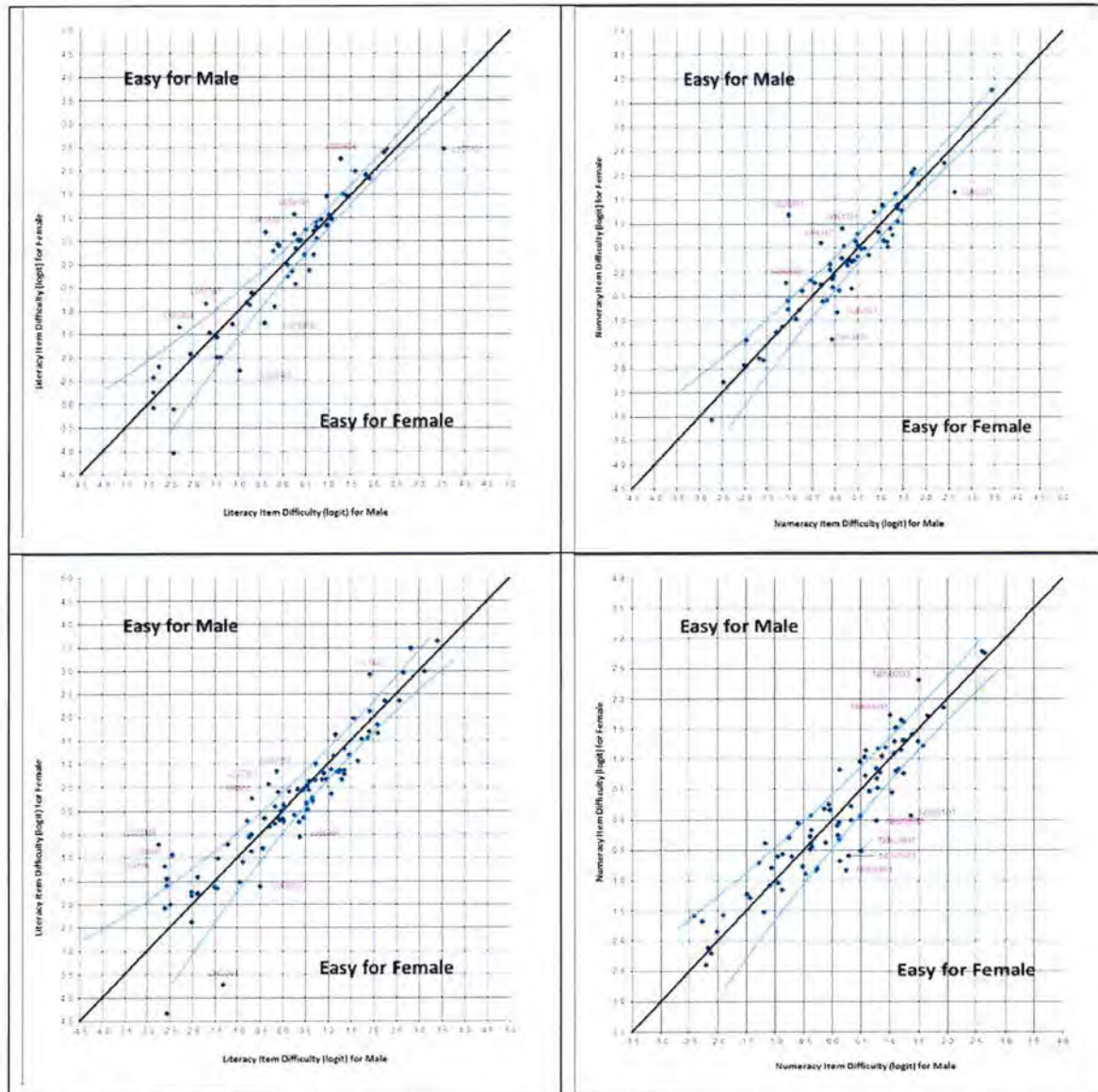


Figure 13: Gender DIF plots

Most items fall within the confidence interval limits or close to the confidence interval. A few items are relatively distant from the confidence lines. Table 44 lists potential gender DIF items with standardised difference of 2.5 and above. The table shows that for the 108 literacy items, six items significantly favoured females and twelve items significantly favoured males. Of the fourteen out of 108 numeracy items showing gender DIF, eight items significantly favoured females and six items significantly favoured males. It can be concluded that there is no significant bias towards one gender or the other in the test.

Table 44: List of potential gender DIF items

Item Label	Difference in Item Difficulties (logit) (Male–Female)	Standardised Difference in Item Difficulties (logit) (Male–Female)	Chi-square	Probability	Favours Female or Male?
L007301	1.08	8.64	74.65	0.00	Female
L002609	-0.98	-9.16	83.94	0.00	Male
L002401	-0.83	-8.55	73.18	0.00	Male
L003608	-1.08	-7.83	61.26	0.00	Male
L005303	-0.86	-3.85	14.79	0.00	Male
L003309	0.86	3.64	13.27	0.00	Female
L002404	-0.94	-4.37	19.06	0.00	Male
L008503	1.32	3.60	12.99	0.00	Female
L003402	-1.00	-4.95	24.46	0.00	Male
L003910	-0.98	-5.06	25.59	0.00	Male
L007901	-0.88	-4.20	17.68	0.00	Male
L005902	-0.94	-4.34	18.88	0.00	Male
L005904	0.93	3.87	14.96	0.00	Female
L003106	-2.03	-6.93	48.09	0.00	Male
L008405	-1.51	-4.27	18.24	0.00	Male
L008602	1.12	3.26	10.62	0.00	Female
L004106	-1.43	-4.04	16.36	0.00	Male
L003304	2.40	4.64	21.50	0.00	Female
N060501	0.97	7.00	48.97	0.00	Female
N028401	-2.21	-14.34	205.57	0.00	Male
N061701	-0.76	-6.72	45.14	0.00	Male
N043401	-0.91	-7.48	56.00	0.00	Male
N060702	-0.86	-4.08	16.66	0.00	Male
N060901	0.89	4.06	16.48	0.00	Female
N052801	1.32	5.36	28.77	0.00	Female
N050203	-0.80	-4.12	16.95	0.00	Male
N060602	-0.72	-4.00	15.98	0.00	Male
N026101	1.30	6.11	37.35	0.00	Female
N020802	0.78	3.37	11.35	0.00	Female
N053401	1.02	3.97	15.77	0.00	Female
N050501	0.89	3.35	11.24	0.00	Female
N060401	1.07	3.93	15.43	0.00	Female

Figure 14 shows the DIF plots for age groups (17–25 year olds vs 26+ year olds). The plots show that most items fall within the confidence interval limits.

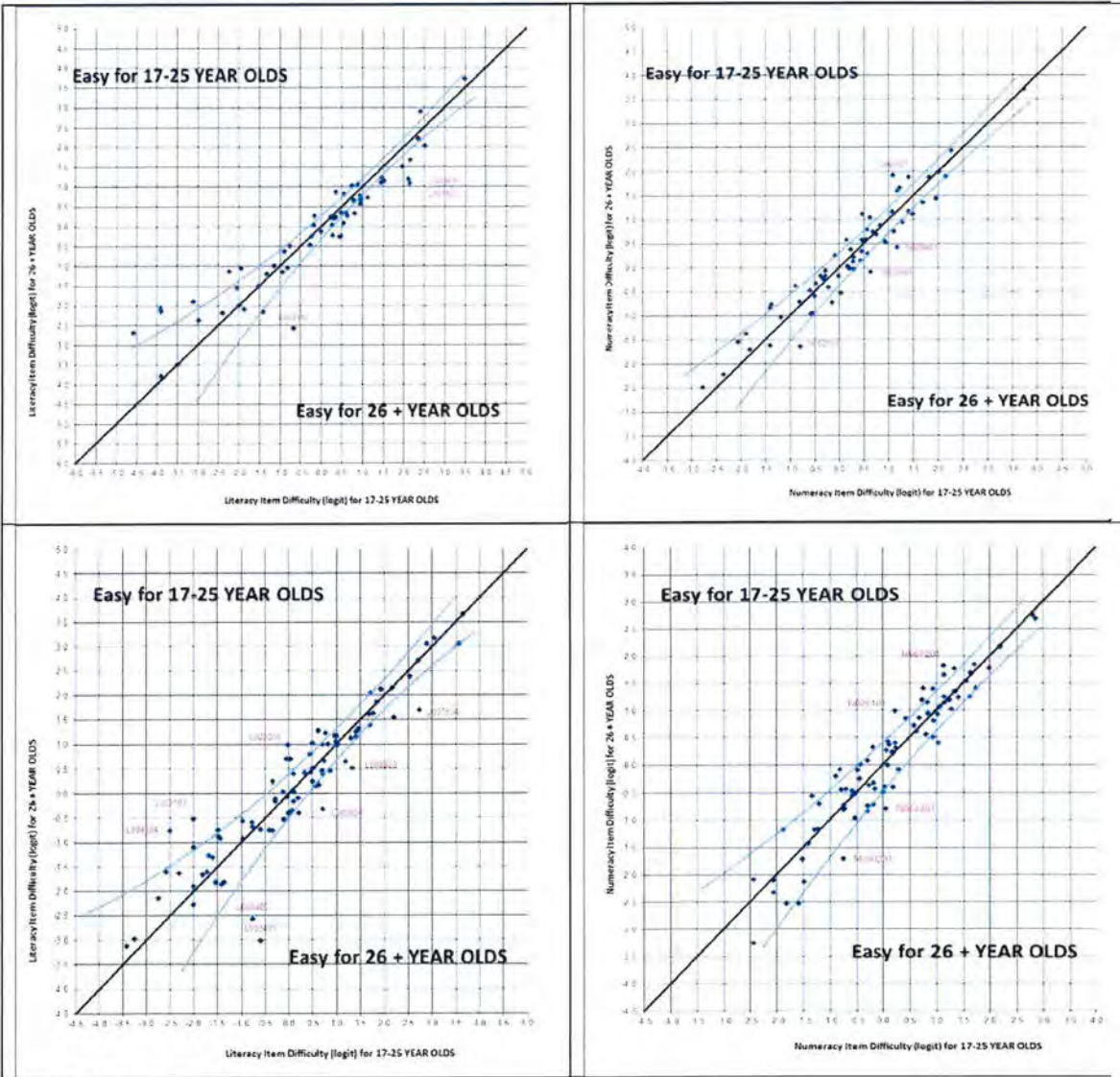


Figure 14: Age group DIF plots

Table 45 lists the items with potential age group DIF. Of the eleven literacy items, three favoured candidates aged 17–25 and eight favoured older candidates. Of the eight numeracy items, three favoured candidates aged 17–25 and five favoured older candidates. While there were more items favouring older candidates, for both literacy and numeracy, because each candidate attempted only 60 out of 108 items, the effect was lessened at the candidate level.

Table 45: List of potential age group DIF items

Item Label	Difference in Item Difficulties (logit) (17–25 Year Olds vs 26+ Year Olds)	Standardised Difference in Item Difficulties (logit) (17–25 Year Olds vs 26+ Year Olds)	Chi-square	Probability	Favours 26+ Year Olds or 17–25 Year Olds?
L007603	0.92	8.34	69.53	0.00	26+ Year Olds
L003610	1.06	8.92	79.59	0.00	26+ Year Olds
L003312	1.91	8.88	78.94	0.00	26+ Year Olds
L007504	1.05	5.60	31.36	0.00	26+ Year Olds
L008803	0.82	4.34	18.79	0.00	26+ Year Olds
L003507	1.03	5.00	25.00	0.00	26+ Year Olds
L003005	-1.00	-4.44	19.71	0.00	17–25 Year Olds
L003403	2.42	7.09	50.22	0.00	26+ Year Olds
L008405	1.82	5.09	25.95	0.00	26+ Year Olds
L003103	-1.48	-4.07	16.56	0.00	17–25 Year Olds
L004104	-1.73	-4.02	16.17	0.00	17–25 Year Olds
N060701	-0.84	-5.76	33.18	0.00	17–25 Year Olds
N028601	0.74	6.63	43.98	0.00	26+ Year Olds
N052601	0.72	4.25	18.05	0.00	26+ Year Olds
N052101	0.85	3.78	14.32	0.00	26+ Year Olds
N061201	-0.71	-5.10	25.99	0.00	17–25 Year Olds
N026101	-0.76	-3.59	12.89	0.00	17–25 Year Olds
N054401	0.84	3.26	10.62	0.00	26+ Year Olds
N050201	0.96	3.18	10.13	0.00	26+ Year Olds

Figure 15 shows the DIF plots for program type (undergraduate vs postgraduate).

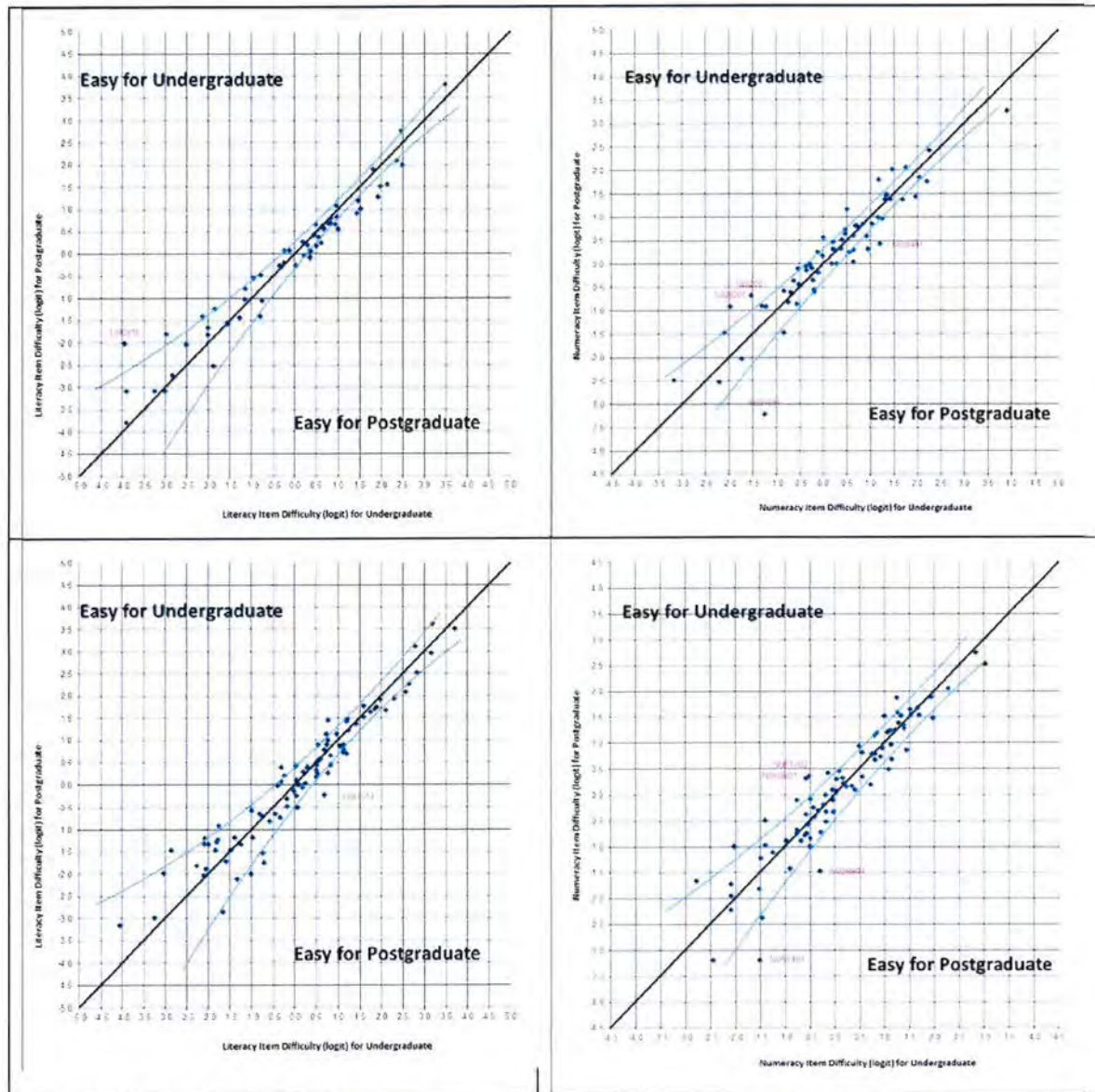


Figure 15: Program type DIF plots

Table 46 lists the items with potential program type DIF. Of these ten items, five favoured postgraduate students and five favoured undergraduate students.

Table 46: List of potential program type DIF items

Item Label	Difference in Item Difficulties (logit) (Undergraduate–Postgraduate)	Standardised Difference in Item Difficulties (logit) (Undergraduate–Postgraduate)	Chi-square	Probability	Favours Postgraduate or Undergraduate?
L002810	-1.93	-3.28	10.78	0.00	Undergraduate
L007501	0.90	4.52	20.46	0.00	Postgraduate
N028401	0.78	5.11	26.10	0.00	Postgraduate
N062001	-0.85	-3.03	9.21	0.00	Undergraduate
N051001	1.98	6.57	43.22	0.00	Postgraduate
N028201	-1.07	-3.14	9.85	0.00	Undergraduate
N061302	-0.88	-3.60	12.93	0.00	Undergraduate
N060601	-0.89	-3.68	13.53	0.00	Undergraduate
N028901	1.19	4.01	16.08	0.00	Postgraduate
N051401	1.69	4.08	16.65	0.00	Postgraduate

Figure 16 shows the DIF plots for course category (early childhood & primary vs secondary).

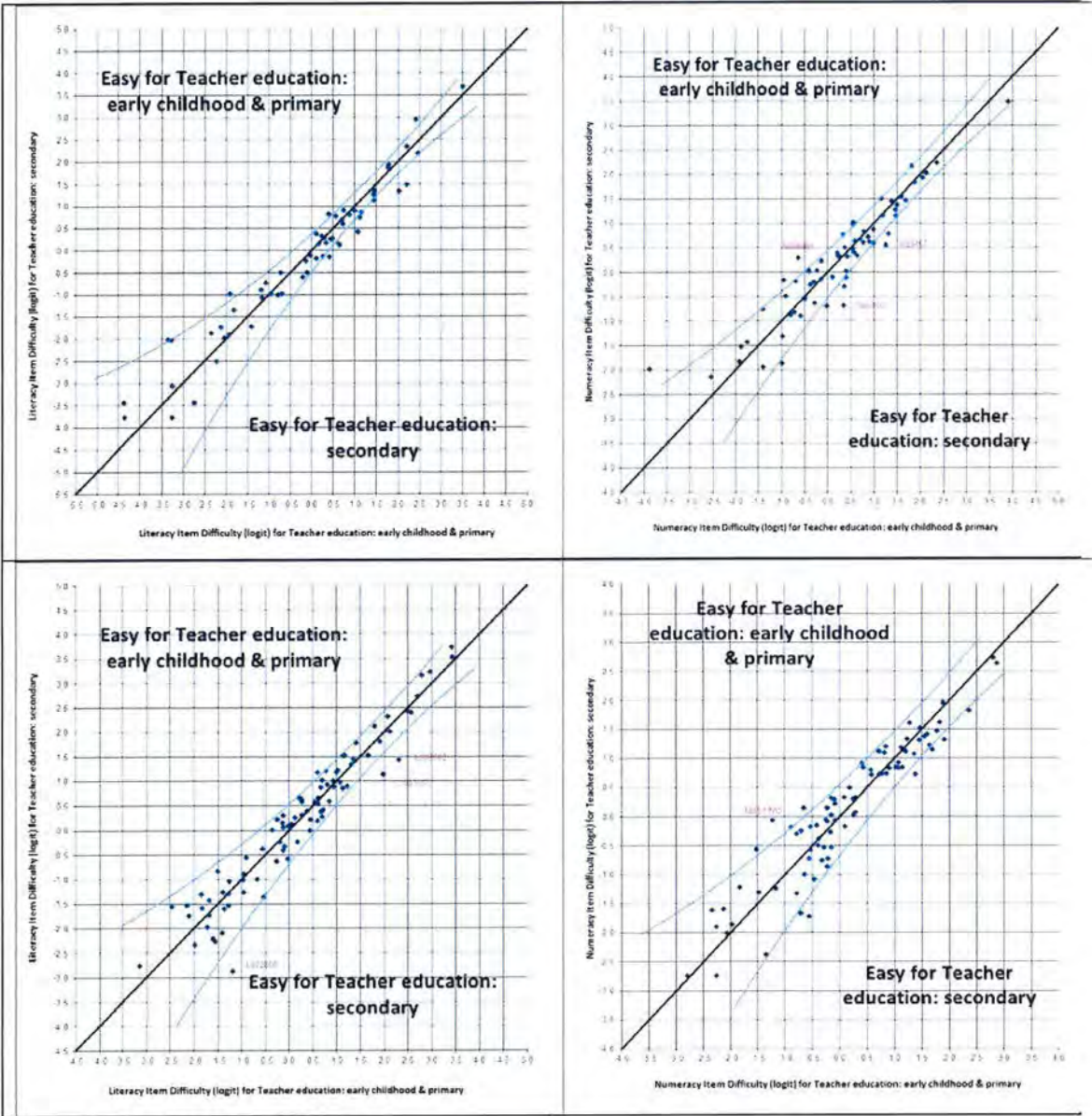


Figure 16: Course category DIF plots

Table 47 lists the items with potential course category DIF. Of the four numeracy items, two favoured students taking early childhood & primary courses and two favoured students taking secondary courses. All three literacy items favoured students taking secondary courses.

Table 47: List of potential course category DIF items

Item Label	Difference in Item Difficulties (logit) (Early childhood & Primary – Secondary)	Standardised Difference in Item Difficulties (logit) (Early childhood & Primary – Secondary)	Chi-square	Probability	Favours Secondary or Early childhood & Primary?
L008603	0.87	3.75	14.04	0.00	Teacher education: secondary
L003001	0.85	3.34	11.17	0.00	Teacher education: secondary
L002806	1.68	2.86	8.18	0.00	Teacher education: secondary
N028401	0.70	3.37	11.36	0.00	Teacher education: secondary
N050101	1.02	4.33	18.77	0.00	Teacher education: secondary
N028001	-0.93	-3.56	12.65	0.00	Teacher education: early childhood & primary
N051702	-1.15	-2.86	8.17	0.00	Teacher education: early childhood & primary



LITERACY AND NUMERACY TEST FOR INITIAL TEACHER EDUCATION STUDENTS

2017 Technical Report

FOR THE DEPARTMENT OF EDUCATION
AND TRAINING, JULY 2018

CONTENTS

1.	OVERVIEW	1
1.1	Administration	1
1.2	Candidate results	2
1.3	Test design and in-test trialling for replenishment of item pool	4
1.4	Differential test functioning.....	4
1.5	Revision of the standards	5
1.6	Comparison of 2016 and 2017 results.....	6
2.	TEST ADMINISTRATION WINDOWS 1–4 IN 2017	7
2.1	Demographic Characteristics of Candidates	7
2.2	Demographic Characteristics of Candidates by Test Windows	15
2.3	Accessibility and Accommodations.....	18
3.	CANDIDATE PERFORMANCE	21
3.1	Scale Score Distributions.....	21
3.2	Candidate Scale Scores by Subscales and Subgroups.....	23
3.3	Performance of Resit Candidates	29
4.	CANDIDATE PERFORMANCE BY TEST WINDOWS.....	32
4.1	Distributions of Candidate Scale Scores by Subscales and Subgroups	32
4.2	Performance of Resit Candidates by Test Windows.....	33
5.	TRIAL ITEM ANALYSIS	37
5.1	In-test trialling.....	37
5.2	Trial item analysis.....	37
5.3	Differential item functioning	37
6.	DIFFERENTIAL TEST FUNCTIONING	38
7.	REVISION OF THE STANDARDS	39
6.1	Summary of revision of the standards	39
6.2	Impact of revised standards on subgroups in 2017.....	39
8.	RELEASED ITEMS.....	43
9.	PHASE 4 TEST DEVELOPMENT.....	43
10.	CONCLUSION	43
11.	REFERENCES	43
12.	APPENDICES	45
	Appendix 1: Proportion of candidates by test centre and by attempt	45
	Appendix 2: Proportion of candidates by test centre and by test window.....	46
	Appendix 3: Score Frequency Distribution	48
	Appendix 4: Performance by Demographic Characteristics and Test Windows.....	51
	Appendix 5: Performance by Test Centres and Remote Proctoring by Test Window ...	58
	Appendix 6: Analysis of Differential Item Functioning (DIF).....	59
	Appendix 7: Impact of differential item functioning at the test level.....	74
	Appendix 8: Effect of revised standards on third- and fourth-attempt candidates	82

List of Figures

Figure 1: Distribution of candidate scale scores for literacy	21
Figure 2: Distribution of candidate scale scores for numeracy	22
Figure 3: Score distribution by gender and age.....	27
Figure 4: Score distribution by program type and year level	27
Figure 5: Score distribution by course category.....	28
Figure 6: Score distribution by location of testing	28
Figure 7: Score distribution by gender and test window	55
Figure 8: Score distribution by age group and test window	55
Figure 9: Score distribution by program type, year level and test window – literacy.....	56
Figure 10: Score distribution by program type, year level and test window – numeracy	56
Figure 11: Score distribution by course category and test window.....	57
Figure 12: Gender DIF plots	59
Figure 13: Age group DIF plots	61
Figure 14: Program type DIF plots.....	63
Figure 15: Course category DIF plots	64
Figure 16: Indigenous DIF plots	65
Figure 17: English as first language DIF plots.....	68
Figure 18: Location DIF plots.....	72
Figure 19: Comparison of Test Characteristic Curves by English as a first language – literacy test L47	75
Figure 20: Comparison of Test Characteristic Curves by English as a first language – literacy test L52	75
Figure 21: Difference in expected raw scores by English as a first language	76
Figure 22: Difference in expected raw scores by gender	78
Figure 23: Difference in expected raw scores by age.....	79
Figure 24: Difference in expected raw scores by course category	80
Figure 25: Difference in expected raw scores by Indigeneity	80
Figure 26: Difference in expected raw scores by program type	81
Figure 27: Difference in expected raw scores by residential area.....	81

List of Tables

Table 1: Number of sittings by location and by test window, including resits	1
Table 2: Summary of sittings by test window, including resits.....	2
Table 3: Summary of candidate results	2
Table 4: Summary of resit rates by year of registration and overall	3
Table 5: Number of sittings and pass rates by attempt and by test window in 2017.....	3
Table 8: Comparison of pass rates using provisional and revised standards in 2017 test windows 3 and 4	5
Table 9: Comparison of performance between 2016 and 2017 by subscale and mean scale score.....	6
Table 10: Demographic characteristics of unique candidates who sat the test in 2017 (including 2016 resitters).....	8
Table 11: Demographic characteristics of first-attempt candidates who sat the test in 2017.....	9
Table 12: Demographic characteristics of second-attempt candidates.....	10
Table 13: Demographic characteristics of third-attempt candidates	11
Table 14: Demographic characteristics of fourth-attempt candidates	12
Table 15: Demographic characteristics of candidates who achieved no standard by the end of 2017..	13
Table 16: Number of attempts by candidates who achieved no standard by the end of 2017.....	14
Table 17: Number and proportion of first-attempt candidates who participated at test centres and by remote proctoring	14
Table 18: Number and proportion of resit candidates who participated at test centres and by remote proctoring	15
Table 19: Demographic characteristics of candidates by test windows – literacy	16
Table 20: Demographic characteristics of candidates by test windows – numeracy	17
Table 21: Number and proportion of candidates in test centres and by remote proctoring by test window – literacy	18
Table 22: Number and proportion of candidates in test centres and by remote proctoring by test window – numeracy	18
Table 23: Largest accommodation groups	18
Table 24: Candidate performance overall and by subscale for first-attempt candidates.....	23
Table 25: Performance by demographic characteristics for first-attempt candidates.....	23
Table 26: Performance by test centres and remote proctoring.....	26
Table 27: Number and proportion of first-attempt candidates who did not achieve the standard.....	28
Table 28: Resit candidate performance overall and by subscale.....	29
Table 29: Change in scale score between first and second attempts by time	30
Table 30: Change in scale score between first and second attempts by performance (Band).....	30
Table 31: Candidate performance by number of test sittings in 2016 and 2017	30
Table 32: Candidate performance by test window, overall and by subscale.....	32
Table 33: Number and proportion of candidates who did not achieve the standard by test window	33
Table 34: Resit candidate performance by test window, overall and by subscale	35
Table 35: Distribution of Phase 3 trial items by Band	37
Table 36: Differential item functioning.....	38
Table 37: Comparison of outcomes between the old and the revised standards for first attempt candidates (test windows 3 and 4 only)	40
Table 38: Comparison of outcomes between the old and the revised standards for second attempt candidates	41
Table 39: Number and proportion of candidates who participated by test centre	45
Table 40: Number of candidates in test centres by test windows – literacy	46
Table 41: Number of candidates in test centres by test windows – numeracy	47
Table 42: Literacy score frequency distribution of first-attempt candidates.....	48
Table 43: Numeracy score frequency distribution of first-attempt candidates.....	49
Table 44: Candidates attempting the test in 2017 by Band achievement.....	50
Table 45: Performance by demographic characteristics in test window 1	51
Table 46: Performance by demographic characteristics in test window 2 (including resits)	52

Table 47: Performance by demographic characteristics in test window 3 (including resits)	53
Table 48: Performance by demographic characteristics in test window 4 (including resits)	54
Table 49: Performance by test centres and remote proctoring in test windows 1 to 4	58
Table 50: List of potential gender DIF items	60
Table 51: List of potential age group DIF items	62
Table 52: List of potential program type DIF items	64
Table 53: List of potential course category DIF items	65
Table 54: List of potential Indigenous DIF items	67
Table 55: List of potential English as first language DIF items	69
Table 56: List of potential region DIF items	73
Table 57: Number of items showing DIF on test L47 for the English as a first language variable	76
Table 58: Number of items showing DIF on test L52 for the English as a first language variable	78
Table 59: Comparison of outcomes between the old and the revised standards for third-attempt candidates	82
Table 60: Comparison of outcomes between the old and the revised standards for fourth-attempt candidates	84

1. OVERVIEW

1.1 Administration

The Literacy and Numeracy Test for Initial Teacher Education students ('the test') was conducted across Australia for the second year, in four test windows, from February 2017 to October 2017. In this period, 24 601 unique candidates attempted one or both components of the test (literacy and numeracy), of which 887 had initially registered for the test in 2016.

In 2017, 23 387 candidates sat the literacy component and 23 466 candidates sat the numeracy component for the first time. During 2017, there were 1819 resits of the literacy component for a second, third or fourth time, including the resits of 408 candidates who first registered in 2016. For numeracy, there were 1753 resits of the numeracy component in 2017 for a second, third or fourth time, including the resits of 493 candidates who first registered in 2016. Over 75% of the candidates were female, with slightly more candidates from primary than from secondary courses (41% and 38% respectively).

Students from 48 higher education providers sat the test in 2017, one more provider than in 2016. The test was offered at 23 test centres in all states and territories, or via remote proctoring under prescribed conditions.

On first sitting, the majority of candidates sat the test at a test centre. Candidates taking the test for a second, third or fourth time were more likely to do so via remote proctoring. For example, 56% of third-attempt candidates and 62% of fourth-attempt candidates sat the test remotely.

Table 1 shows the number of sittings by location for each test window (TW). Test window 1 was available by remote proctoring only.

Table 1: Number of sittings by location and by test window, including resits¹

Location of testing	TW1		TW2		TW3		TW4	
	Literacy	Numeracy	Literacy	Numeracy	Literacy	Numeracy	Literacy	Numeracy
Capital cities			4204	4229	4168	4174	4215	4209
Regional cities			673	674	790	793	746	746
Remote proctoring	4264	4287	2631	2637	1710	1695	1805	1775
Total sittings	4264	4287	7508	7540	6668	6662	6766	6730

A small number of candidates chose to attempt only one of the component tests at each test window as shown by Table 2.

¹ Tables 1 and 2 include resit candidates in all test windows

Table 2: Summary of sittings by test window, including resits

Test	TW1	TW2	TW3	TW4	Total
Both literacy and numeracy	4019	7055	6010	5944	23028
Literacy only	245	453	658	822	2178
Numeracy only	268	485	652	786	2191
Total sittings	4532	7993	7320	7552	27397

Testing conditions were modified to accommodate 179 candidates with special needs in 2017. Accessible versions of the test were also available for candidates who required supportive technology, such as a screen reader. The online accessible versions of the test were used twice for literacy and once for numeracy in 2017. A paper accessible version of the test was used once for literacy and once for numeracy in 2017.

The administration of the test is further described in Section 2 of this report.

1.2 Candidate results

Table 3 shows the number of candidates attempting each component and both components, and their pass rates at the end of 2017. The table shows how the pass rates increase as candidates resit and achieve the standard over time. For example, of those candidates who initially registered for the literacy component in 2016, the pass rate increased from 95.2% at the end of 2016 to 97.3% at the end of 2017. In 2017, some candidates resat the test up to three times. The revised standards were applied in 2017 in test windows 3 and 4.

In the two year period from 2016 to 2017, the number of unique candidates participating in one or more components of the test was 37 058. Of these, 36 470 sat the literacy component and 36 550 sat the numeracy component. Almost all candidates (35 962) attempted both components of the test while 509 attempted literacy only and 589 attempted numeracy only. At the end of 2017, of the 35 962 candidates who had attempted both components, 32 665 candidates had achieved both standards, an overall pass rate of 90.8%.

Table 3: Summary of candidate results

Component	Year of Registration	At End of Year	Number of Unique Candidates	Standard Achieved	Standard Not Achieved	Cancelled due to Misconduct	Pass Rate
Literacy	2016	2016	13083	12460	622	1	95.2
	2016	2017	13083	12732	350	1	97.3
	2017	2017	23387	21521	1866	0	92.0
	2016–17	2017	36470	34253	2216	1	93.9
Numeracy	2016	2016	13084	12324	760	0	94.2
	2016	2017	13084	12623	461	0	96.5
	2017	2017	23466	21655	1810	1	92.3
	2016–17	2017	36550	34278	2271	1	93.8
Both	2016–17	2017	35962	32665	3295	2	90.8

Table 4 shows the percentage of candidates by the number of attempts they had at each component of the test as at the end of 2017. It can be seen that in both years and for both components, approximately 95% of candidates attempt the test once, approximately 4% of candidates attempt the test twice and

approximately 1% of candidates attempt the test three or more times. For literacy, by the end of 2017, approximately 5% of the candidates who registered in 2016 had resat the test, which is the same for candidates who registered in 2017. For numeracy, by the end of 2017, approximately 6% of the candidates who registered in 2016 had resat the test, but only 4% of candidates who registered in 2017 had resat the test.

Table 4: Summary of resit rates by year of registration and overall

Domain	Year of Registration	Number of Unique Candidates	Unique candidates who had 1 attempt only (%)	Unique candidates who had 2 attempts only (%)	Unique candidates who had 3 attempts only (%)	Unique candidates who had 4 attempts only (%)
Literacy	2016	13083	94.7	3.9	1.1	0.2
	2017	23387	95.1	4.3	0.6	0.0
	2016–17	36470	95.0	4.1	0.8	0.1
Numeracy	2016	13084	93.9	4.2	1.5	0.4
	2017	23466	95.9	3.6	0.6	0.0
	2016–17	36550	95.2	3.8	0.9	0.1

Table 5 shows that of the 23 387 candidates who attempted the literacy component for the first time in 2017, 89.2% achieved the standard at their first sitting, compared to 93.3% in 2016. For numeracy in 2017, 90.0% of the 23 466 candidates achieved the standard at their first sitting, compared to 92.5% in 2016. The lower pass rates in test windows 3 and 4 reflect the revised standards and other factors addressed later in this report.

Under the standard resit allowance, candidates who do not achieve the standard on their first attempt are permitted up to two additional sittings. In a small number of cases, more than two resits may be granted in exceptional circumstances. On second and third sittings (resit 1 and resit 2) for literacy, of the 1488 and 296 candidates for these sittings, 53.6% and 39.9%, respectively, achieved the standard. Of the 35 candidates who were granted a fourth sitting, 40.0% were successful. On second and third sittings for numeracy, of the 1365 and 339 candidates for these sittings, 50.0% and 41.3%, respectively, achieved the standard. Of the 49 candidates who were granted a fourth sitting, 34.7% were successful.

Table 5: Number of sittings and pass rates by attempt and by test window in 2017

		TW1	Pass rate	TW2	Pass rate	TW3	Pass rate	TW4	Pass rate	Total	Pass rate
Literacy	First sitting	4052	92.6	7169	91.4	6164	86.8	6002	86.8	23387	89.2
	Second sitting	184	58.2	274	62.4	434	45.9	596	53.7	1488	53.6
	Third sitting	28	57.1	55	61.8	63	28.6	150	33.3	296	39.9
	Fourth sitting	0	NA	10	50.0	7	28.6	18	38.9	35	40.0
	Total sittings	4264		7508		6668		6766		25206	
Numeracy	First sitting	4056	92.5	7158	92.8	6232	88.3	6020	86.7	23466	90.0
	Second sitting	200	50.5	285	58.9	352	43.2	528	49.6	1365	50.0
	Third sitting	30	63.3	87	52.9	61	21.3	161	38.5	339	41.3

	Fourth sitting	1	100.0	10	60.0	17	35.3	21	19.0	49	34.7
	Totals sittings	4287		7540		6662		6730		25219	

Candidates' results for 2017 are described in more detail in Section 3 of this report.

1.3 Test design and in-test trialling for replenishment of item pool

In the first half of 2017, in test windows 1 and 2, there were 12 equivalent test forms for literacy and 14 equivalent test forms for numeracy. In the second half of 2017, in test windows 3 and 4, a proportion of the test forms were refreshed using items that were trialled in 2016. A further set of items was released as described in section 7. In these test windows, there were 10 equivalent test forms for literacy and 8 equivalent test forms for numeracy.

For literacy, each test form comprised five 12-item clusters (C1 to C5) totalling 60 items. For numeracy, the test was divided into two sections as follows: section 1 ('calculator available' – CA) comprising four 12-item clusters (48 items), and section 2 ('calculator not available' – CN) comprising two 6-item clusters (12 items), totalling 60 items.

In order to augment and replenish the pool of items available for the test in future administrations, items were trial-tested within the live instruments. These items were administered in small clusters (one to five items) at the end of each test, and did not contribute to the candidates' scores. One hundred and sixty items for each of literacy and numeracy were administered in the in-test trial clusters. At least 250 candidates were administered each of these items.

Examples of one literacy component and one numeracy component with in-test trial clusters are shown below.

Literacy component 1	C1	C2	C3	C4	C5	Trial C
-----------------------------	----	----	----	----	----	---------

	Section 1					Section 2		
Numeracy component 1	CA1	CA2	CA3	CA4	Trial CA	CN1	CN2	Trial CN

In the second half of 2017 and the first half of 2018, 160 Phase 3 literacy items and 165 Phase 3 numeracy items were in-test trialled. Of these, 21 literacy items and 8 numeracy items were judged to have unsatisfactory psychometric properties and were deleted from the pool. The remaining items were well-targeted for difficulty across the three reporting Bands as required by the test construct and assessment framework, thereby ensuring adequate test replenishment. The Phase 3 trial items revealed some differential item functioning (DIF); however, for both literacy and numeracy, the DIF was 'well-balanced' for most variables. For example, for the Age variable, while 8 numeracy items favoured candidates aged 17–25 years, 9 favoured candidates aged 26+ years. For more detail see Section 5.

1.4 Differential test functioning

Where multiple test forms are used in an assessment, it is important that each test is psychometrically equivalent and that no test significantly favours one subgroup over another, particularly in high-stakes testing. With over two years of data it was possible to reliably test for differential test functioning (DTF) for all background variables; in particular, the smaller subgroups (namely Indigenous candidates and international candidates). No significant DTF was found for any of the seven subgroups. The largest differences occurred on the literacy tests in regard to the English as first language variable, but these differences were still not significant and not uni-directional. A closer analysis of the items on two literacy papers revealed useful patterns to partly explain the differences and inform future test construction. For more detail see Section 6.

1.5 Revision of the standards

The provisional benchmarks and standards set by expert groups were validated in a separate equating study conducted in 2017. The equating study provided information on how the literacy and numeracy skills of candidates achieving the test standards compared to those of Australian adults as measured by the OECD Programme for International Assessment of Adult Competencies (PIAAC). PIAAC was administered to a representative sample of Australian adults (aged 16–74) in 2011–12.

The equating study showed that the professional judgements of the expert groups when determining the provisional standards were very close to, but slightly below, the desired empirical outcomes. The provisional literacy standard of 106 scale score points was equivalent to the 68th percentile (top 32%) of the Australian adult population. The provisional numeracy standard of 107 was equivalent to the 62nd percentile (top 38%) of the Australian adult population. The equating exercise revealed that the test scale score equivalent to the 70th percentile (top 30%) of Australian adults is 107 for literacy, and 110 for numeracy.

Considerable discussion about when to apply the revised standards took place between ACER and DET and between DET and the Governance Committee. The main concerns were fairness to students and transparency. Given that there would be resitting candidates in each test window, it was agreed that no test window would suit all candidates. It was agreed that the new standard would apply from and include test window 3 in 2017. It was also agreed to release the Described Proficiency Scale with sample items located in bands reflecting the new standards before 2017 test window 3 to facilitate student preparation. A letter announcing the new standards and explaining the reason for their revision was jointly developed by DET and ACER and forwarded to all HEPs before 2017 test window 3.

Table 8 compares the pass rates of candidates using both the provisional and revised standards in both literacy and numeracy in 2017 test windows 3 and 4 by the number of attempts. Earlier modelling predicted that the difference in the percentage of first-attempt candidates who achieved at the level of the old standards and who achieved the new standards would be –2.1% for literacy and –4.4% for numeracy. The actual difference was –2.7% for literacy and –4.5% for numeracy, as shown in Table 8.

Table 6: Comparison of pass rates using provisional and revised standards in 2017 test windows 3 and 4

Component	Attempt Number	Total Candidates	Pass rate using provisional standard (%)	Pass rate using revised standard (%)	Difference in pass rates (%)
Literacy	1st	12166	89.5	86.8	–2.7
	2nd	1030	58.3	50.4	–7.9
	3rd	213	43.2	31.9	–11.3
	4th	25	60.0	36.0	–24.0
		13434	86.3	83.1	–3.2
Numeracy	1st	12252	92.0	87.5	–4.5
	2nd	880	61.9	47.0	–14.9
	3rd	222	56.8	33.8	–23.0
	4th	38	39.5	26.3	–13.2
		13392	89.3	83.8	–5.5

The difference in pass rates due to the revised standards was greatest for resitting candidates: these were the candidates with scale scores closest to the standards.

1.6 Comparison of 2016 and 2017 results

Table 9 shows that the literacy mean scale scores for second-attempt and third-attempt candidates were 1 to 2 scale score points lower in 2017 than in 2016, indicating lower achievement as a cohort and partly contributing to the lower pass rates. The reverse was true for numeracy where numeracy mean scale scores for second-attempt and third-attempt candidates were 1 to 2 scale score points higher in 2017 than in 2016.

Table 7: Comparison of performance between 2016 and 2017 by subscale and mean scale score

Component	Attempt Number	Whole test and subscale	2016			2017		
			Number of sittings	Mean	Pass Rate	Number of sittings ²	Mean	Pass Rate
Literacy	1st	Overall	13083	117.5	93.3	23387	117.0	89.2
		Reading		117.4			117.1	
		Technical skills of writing		117.5			116.9	
	2nd	Overall	341	107.5	67.7	1488	106.5	53.6
		Reading		107.2			106.6	
		Technical skills of writing		107.8			106.1	
	3rd	Overall	25	107.7	76.0	296	105.3	39.9
		Reading		107.3			105.5	
		Technical skills of writing		108.5			105.1	
	4th	Overall	0	NA	NA	35	106.1	40.0
		Reading		NA	NA		105.5	
		Technical skills of writing		NA	NA		107.2	
Numeracy	1st	Overall	13084	122.4	92.4	2346-6	123.0	90.0
		Number & algebra		121.8			122.3	
		Measurement & geometry		121.5			122.8	
		Statistics & probability		122.7			123.0	
		Calculator available		122.5			123.2	
		Calculator not available		119.3			119.7	
	2nd	Overall	406	107.5	55.4	1365	108.1	50.0
		Number & algebra		106.1			106.4	
		Measurement & geometry		108.0			109.1	
		Statistics & probability		109.6			110.0	
		Calculator available		108.5			109.4	
		Calculator not available		103.7			103.4	
	3rd	Overall	40	105.4	37.5	339	107.2	41.3
		Number & algebra		103.9			105.7	
		Measurement & geometry		106.8			108.2	
		Statistics & probability		107.0			108.8	
		Calculator available		106.2			108.2	
		Calculator not available		102.3			103.5	
	4th	Overall	2	NA ³	NA	49	105.9	34.7
		Number & algebra		NA	NA		104.0	
		Measurement & geometry		NA	NA		107.0	
		Statistics & probability		NA	NA		108.2	
		Calculator available		NA	NA		106.7	
		Calculator not available		NA	NA		103.5	

² Sittings for attempt numbers beyond the first attempt include candidates who first registered in 2016

³ Not reported due to small (n = 2) group size

2. TEST ADMINISTRATION WINDOWS 1–4 IN 2017

This section covers the demographic characteristics of candidates who sat the test in 2017. Details on test centres, remote proctoring and other administrative matters can be found in each of the four 2017 test window administration reports.

2.1 Demographic Characteristics of Candidates

Over 24 000 candidates sat the test in 2017. In addition to the 47 institutions who participated in 2016, one additional institution, Tabor Adelaide, participated in 2017 as listed below.

Alphacrucis College	The University of Western Australia
Australian Catholic University	University of Canberra
Australian College of Physical Education	University of South Australia
Avondale College	University of Southern Queensland
Central Queensland University	University of Tasmania
Charles Darwin University	University of Technology Sydney
Charles Sturt University	University of the Sunshine Coast
Christian Heritage College	University of Wollongong
Curtin University	Victoria University
Deakin University	Western Sydney University
Eastern College Australia	
Edith Cowan University	
Excelsia College	
Federation University Australia	
Flinders University	
Griffith University	
Holmesglen TAFE	
James Cook University	
La Trobe University	
Macquarie University	
Melbourne Polytechnic	
Monash University	
Montessori World Educational Institute	
Morling College	
Murdoch University	
Queensland University of Technology	
RMIT University	
Southern Cross University	
Swinburne University of Technology	
Tabor Adelaide	
The University of Adelaide	
The University of Melbourne	
The University of New England	
The University of New South Wales	
The University of Newcastle	
The University of Notre Dame Australia	
The University of Queensland	
The University of Sydney	

Table 10 shows the demographic characteristics of all candidates who sat the test in 2017. This includes the candidates who first registered for the test in 2017 plus those who registered in 2016 and resat the test in 2017. It shows that the majority of candidates (75%)⁴ were female, and most candidates (62%) were in the age group 17–25. The majority of candidates (60%) were enrolled in an undergraduate course. Of the total number of candidates who participated, the largest group of undergraduate candidates was those in their fourth year (25%), down from 42% in 2016 as an increasing number of undergraduate candidates attempted the test in the first and second year of their course. The largest group of postgraduate candidates was those in their first year (20%, up from 9% in 2016). In terms of course category, candidates were mainly enrolled in primary teacher education courses (39–40%), followed by secondary (38%), other teacher education courses (15%), early childhood (7%) and special education (1%). These percentages were very similar to those in 2016.

The percentage of international candidates attempting the test doubled from 3% in 2016 to 6% in 2017. The percentage of Indigenous candidates stayed steady at 1.5%. The percentage of candidates from provincial areas declined slightly from 17% in 2016 to 15% in 2017.

Table 8: Demographic characteristics of unique candidates who sat the test in 2017 (including 2016 resitters)

Characteristic	Category	Literacy		Numeracy	
		N	%	N	%
Gender	Female	17828	74.9	17996	75.1
	Male	5956	25.0	5953	24.8
	Indeterminate/intersex	11	0.0	10	0.0
Age	17–25	14811	62.2	14931	62.3
	26–30	3648	15.3	3652	15.2
	31–35	1883	7.9	1880	7.8
	36–40	1499	6.3	1519	6.3
	41–45	1088	4.6	1099	4.6
	46+	866	3.6	878	3.7
International Students	No	22483	94.5	22645	94.5
	Yes	1312	5.5	1314	5.5
English as a First Language	Yes	20772	87.3	20914	87.3
	No	3023	12.7	3045	12.7
Indigenous	No	22944	96.4	23109	96.5
	Yes	368	1.5	369	1.5
	Not disclosed	483	2.0	481	2.0
Residential Area	Metropolitan areas	20116	84.5	20269	84.6
	Provincial areas	3482	14.6	3491	14.6
	Remote areas	121	0.5	120	0.5
	International	56	0.2	57	0.2
	Invalid or Missing	20	0.1	22	0.1
Program Type	Undergraduate	14306	60.1	14409	60.1
	Postgraduate	9489	39.9	9550	39.9
Program Type by Year Level	Undergraduate first year	1702	7.2	1738	7.3
	Undergraduate second year	1443	6.1	1459	6.1
	Undergraduate third year	3764	15.8	3756	15.7
	Undergraduate fourth year	5871	24.7	5881	24.5
	Undergraduate fifth year or above	625	2.6	649	2.7
	Undergraduate graduated	901	3.8	926	3.9
	Postgraduate first year	4852	20.4	4856	20.3
	Postgraduate second year	2410	10.1	2424	10.1
	Postgraduate third year	298	1.3	301	1.3
	Postgraduate fourth year	475	2.0	477	2.0
	Postgraduate fifth year or above	461	1.9	471	2.0
	Postgraduate graduated	993	4.2	1021	4.3

⁴ In the descriptive text accompanying the tables throughout the report, most percentages are rounded to the nearest whole per cent.

Course Category	Teacher education: early childhood	1613	6.8	1616	6.7
	Teacher education: primary	9380	39.4	9484	39.6
	Teacher education: secondary	9137	38.4	9152	38.2
	Teacher education: special education	151	0.6	147	0.6
	Teacher education: other	3514	14.8	3560	14.9

The following demographic analysis separates the 2017 candidates into five groups for each component of the test: first-attempt candidates, second-attempt candidates (first resit), third-attempt candidates (second resit), fourth-attempt candidates (third resit) and candidates who achieved no standard.

Table 11 shows the demographic characteristics of the first-attempt candidates for each component of the test in 2017.⁵ The demographic characteristics of this cohort are very similar to those described in Table 10 above.

Table 9: Demographic characteristics of first-attempt candidates who sat the test in 2017

Characteristic	Category	Literacy		Numeracy	
		N	%	N	%
Gender	Female	17484	74.8	17538	74.7
	Male	5892	25.2	5918	25.2
	Indeterminate/intersex	11	0.0	10	0.0
Age	17–25	14541	62.2	14617	62.3
	26–30	3569	15.3	3562	15.2
	31–35	1868	8.0	1850	7.9
	36–40	1484	6.3	1499	6.4
	41–45	1073	4.6	1079	4.6
	46+	852	3.6	859	3.7
International Students	No	22093	94.5	22161	94.4
	Yes	1294	5.5	1305	5.6
English as a First Language	Yes	20475	87.5	20533	87.5
	No	2912	12.5	2933	12.5
Indigenous	No	22562	96.5	22637	96.5
	Yes	356	1.5	356	1.5
	Not disclosed	469	2.0	473	2.0
Residential Area	Metropolitan areas	19792	84.6	19857	84.6
	Provincial areas	3400	14.5	3414	14.5
	Remote areas	120	0.5	119	0.5
	International	55	0.2	56	0.2
	Invalid or Missing	20	0.1	20	0.1
Program Type	Undergraduate	14030	60.0	14087	60.0
	Postgraduate	9357	40.0	9379	40.0
Program Type by Year Level	Undergraduate first year	1709	7.3	1739	7.4
	Undergraduate second year	1437	6.1	1459	6.2
	Undergraduate third year	3748	16.0	3747	16.0
	Undergraduate fourth year	5753	24.6	5758	24.5
	Undergraduate fifth year or above	606	2.6	613	2.6
	Undergraduate graduated	777	3.3	771	3.3
	Postgraduate first year	4867	20.8	4868	20.7
	Postgraduate second year	2387	10.2	2396	10.2
	Postgraduate third year	292	1.2	295	1.3
	Postgraduate fourth year	469	2.0	469	2.0
	Postgraduate fifth year or above	449	1.9	451	1.9
	Postgraduate graduated	893	3.8	900	3.8

⁵ This cohort is referred to as the first-attempt candidates. Subsequent tables show demographic characteristics for candidates who resat the test

Course Category	Teacher education: early childhood	1584	6.8	1575	6.7
	Teacher education: primary	9189	39.3	9244	39.4
	Teacher education: secondary	9047	38.7	9053	38.6
	Teacher education: special education	144	0.6	144	0.6
	Teacher education: other ⁶	3423	14.6	3450	14.7

In 2017, there were 1819 resits of the literacy component (14%, up from 2% in 2016) and a total of 1753 resits of the numeracy component (14%, up from 2% in 2016) after the standard was not achieved at previous sittings. For literacy, there were 1488 second attempts, 296 third attempts and 35 fourth attempts. For numeracy, the numbers were 1365, 339 and 49 respectively. These resit numbers included those of the candidates who did not achieve one or more standard in 2016.

Table 12 shows the demographic characteristics of the candidates who sat the test twice (first resit) during 2017. It shows that the overwhelming majority of these resit candidates were female (83% for literacy, 91% for numeracy), mostly in the age group 17–25 (71% for literacy, 63% for numeracy). The majority of candidates (68% for literacy, 67% for numeracy) were enrolled in an undergraduate course. This pattern matches that for the first-attempt candidates, but each of the categories listed above (female; age 17–25; undergraduate) was more highly represented in the second-attempt cohort than in the first-attempt cohort.

Table 12 also shows that the proportion of candidates for whom English was not their first language was 32% of the number of first-resit candidates for literacy (up from 20% in 2016) and 17% for numeracy (up from 13% in 2016), whereas for first-attempt candidates the proportion was only 13% as described in Table 11.

Table 10: Demographic characteristics of second-attempt candidates⁷

Characteristic	Category	Literacy			Numeracy		
		N	%	% of Total Sittings	N	%	% of Total Sittings
Gender	Female	1238	83.2	5.3	1244	91.1	5.3
	Male	248	16.7	1.1	120	8.8	0.5
	Indeterminate/intersex	2	0.1	0.0	1	0.1	0.0
Age	17–25	1049	70.5	4.5	853	62.5	3.6
	26–30	196	13.2	0.8	230	16.8	1.0
	31–35	80	5.4	0.3	92	6.7	0.4
	36–40	74	5.0	0.3	67	4.9	0.3
	41–45	46	3.1	0.2	66	4.8	0.3
	46+	43	2.9	0.2	57	4.2	0.2
International Students	No	1262	84.8	5.4	1331	97.5	5.7
	Yes	226	15.2	1.0	34	2.5	0.1
English as a First Language	Yes	1006	67.6	4.3	1137	83.3	4.8
	No	482	32.4	2.1	228	16.7	1.0
Indigenous	No	1399	94.0	6.0	1298	95.1	5.5
	Yes	41	2.8	0.2	34	2.5	0.1
	Not disclosed	48	3.2	0.2	33	2.4	0.1
Residential Area	Metropolitan areas	1262	84.8	5.4	1154	84.5	4.9
	Provincial areas	222	14.9	0.9	199	14.6	0.8
	Remote areas	3	0.2	0.0	8	0.6	0.0
	International	1	0.1	0.0	2	0.1	0.0
	Invalid or Missing	0	0.0	0.0	2	0.1	0.0
Program Type	Undergraduate	1010	67.9	4.3	916	67.1	3.9

⁶ The Other course categories includes curriculum studies, education studies, teacher education, and teacher education vocational.

⁷ Candidates who registered in both 2016 and 2017

	Postgraduate	478	32.1	2.0	449	32.9	1.9
Program Type by Year Level	Undergrad. 1st year	138	9.3	0.6	53	3.9	0.2
	Undergrad. 2nd year	65	4.4	0.3	41	3.0	0.2
	Undergrad. 3rd year	152	10.2	0.6	114	8.4	0.5
	Undergrad. 4th year	438	29.4	1.9	485	35.5	2.1
	Undergrad. 5th year+	53	3.6	0.2	72	5.3	0.3
	Undergrad. graduated	164	11.0	0.7	151	11.1	0.6
	Postgrad. 1st year	162	10.9	0.7	98	7.2	0.4
	Postgrad. 2nd year	98	6.6	0.4	110	8.1	0.5
	Postgrad. 3rd year	16	1.1	0.1	19	1.4	0.1
	Postgrad. 4th year	38	2.6	0.2	42	3.1	0.2
	Postgrad. 5th year+	27	1.8	0.1	41	3.0	0.2
	Postgrad. graduated	137	9.2	0.6	139	10.2	0.6
Course Category	Early childhood	148	9.9	0.6	156	11.4	0.7
	Primary	573	38.5	2.5	610	44.7	2.6
	Secondary	433	29.1	1.9	318	23.3	1.4
	Special education	8	0.5	0.0	7	0.5	0.0
	Other	326	21.9	1.4	274	20.1	1.2

Table 13 shows the demographic characteristics of candidates who participated in the test three times (two resits) during 2017. As for first-resit candidates, this cohort tended to be mostly female undergraduates aged 17–25, particularly for numeracy. These categories are more highly represented in the third-attempt cohort than in the first-attempt cohort, but are similar to the second-attempt cohort.

Table 13 also shows that for literacy, the proportion of candidates for whom English was not their first language was 34% of the two-resit candidates, whereas the proportion was only 13% of the first-attempt candidates described in Table 11. For numeracy, it was 18% (down from 25% in 2016) of two-resit candidates compared to 13% of the first-attempt candidates.

Table 11: Demographic characteristics of third-attempt candidates

Characteristic	Category	Literacy			Numeracy		
		N	%	% of Total Sitzings	N	%	% of Total Sitzings
Gender	Female	251	84.8	1.1	318	93.8	1.4
	Male	44	14.9	0.2	21	6.2	0.1
	Indeterminate/intersex	1	0.3	0.0	0	0.0	0.0
Age	17–25	204	68.9	0.9	226	66.7	1.0
	26–30	53	17.9	0.2	54	15.9	0.2
	31–35	12	4.1	0.1	18	5.3	0.1
	36–40	16	5.4	0.1	10	2.9	0.0
	41–45	6	2.0	0.0	15	4.4	0.1
	46+	5	1.7	0.0	16	4.7	0.1
International Students	No	251	84.8	1.1	334	98.5	1.4
	Yes	45	15.2	0.2	5	1.5	0.0
English as a First Language	Yes	196	66.2	0.8	279	82.3	1.2
	No	100	33.8	0.4	60	17.7	0.3
Indigenous	No	279	94.3	1.2	320	94.4	1.4
	Yes	10	3.4	0.0	13	3.8	0.1
	Not disclosed	7	2.4	0.0	6	1.8	0.0
Residential Area	Metropolitan areas	244	82.4	1.0	288	85.0	1.2
	Provincial areas	50	16.9	0.2	47	13.9	0.2
	Remote areas	0	0.0	0.0	3	0.9	0.0
	International	0	0.0	0.0	0	0.0	0.0
	Invalid or Missing	2	0.7	0.0	1	0.3	0.0
Program Type	Undergraduate	192	64.9	0.8	232	68.4	1.0

	Postgraduate	104	35.1	0.4	107	31.6	0.5
Program Type by Year Level	Undergrad. 1st year	21	7.1	0.1	4	1.2	0.0
	Undergrad. 2nd year	2	0.7	0.0	3	0.9	0.0
	Undergrad. 3rd year	12	4.1	0.1	8	2.4	0.0
	Undergrad. 4th year	96	32.4	0.4	120	35.4	0.5
	Undergrad. 5th year+	10	3.4	0.0	23	6.8	0.1
	Undergrad. graduated	51	17.2	0.2	74	21.8	0.3
	Postgrad. 1st year	22	7.4	0.1	7	2.1	0.0
	Postgrad. 2nd year	19	6.4	0.1	25	7.4	0.1
	Postgrad. 3rd year	4	1.4	0.0	5	1.5	0.0
	Postgrad. 4th year	4	1.4	0.0	11	3.2	0.0
	Postgrad. 5th year+	9	3.0	0.0	8	2.4	0.0
	Postgrad. graduated	46	15.5	0.2	51	15.0	0.2
Course Category	Early childhood	27	9.1	0.1	35	10.3	0.1
	Primary	119	40.2	0.5	157	46.3	0.7
	Secondary	75	25.3	0.3	76	22.4	0.3
	Special education	4	1.4	0.0	0	0.0	0.0
	Other	71	24.0	0.3	71	20.9	0.3

Table 14 shows demographic characteristics of the small number of candidates who were authorised to sit the test four times in 2017. This cohort had the highest proportion of female candidates (89% for literacy and 96% for numeracy), the highest proportion of graduates (69% for literacy and 47% for numeracy) and the highest proportion of candidates from primary courses (49% for both literacy and numeracy).

Table 12: Demographic characteristics of fourth-attempt candidates

Characteristic	Category	Literacy			Numeracy		
		N	%	% of Total Sitzings	N	%	% of Total Sitzings
Gender	Female	31	88.6	0.1	47	95.9	0.2
	Male	4	11.4	0.0	2	4.1	0.0
Age	17–25	26	74.3	0.1	35	71.4	0.1
	26–30	5	14.3	0.0	10	20.4	0.0
	31–35	0	0.0	0.0	2	4.1	0.0
	36–40	3	8.6	0.0	1	2.0	0.0
	41–45	1	2.9	0.0	0	0.0	0.0
	46+	0	0.0	0.0	1	2.0	0.0
International Students	No	34	97.1	0.1	48	98.0	0.2
	Yes	1	2.9	0.0	1	2.0	0.0
English as a First Language	Yes	25	71.4	0.1	40	81.6	0.2
	No	10	28.6	0.0	9	18.4	0.0
Indigenous	No	35	100.0	0.1	48	98.0	0.2
	Yes	0	0.0	0.0	1	2.0	0.0
Residential Area	Metropolitan areas	29	82.9	0.1	42	85.7	0.2
	Provincial areas	6	17.1	0.0	7	14.3	0.0
	Remote areas	0	0.0	0.0	0	0.0	0.0
Program Type	Undergraduate	18	51.4	0.1	36	73.5	0.2
	Postgraduate	17	48.6	0.1	13	26.5	0.1
Program Type by Year Level	Undergrad. 1st year	0	0.0	0.0	0	0.0	0.0
	Undergrad. 2nd year	0	0.0	0.0	0	0.0	0.0
	Undergrad. 3rd year	1	2.9	0.0	0	0.0	0.0
	Undergrad. 4th year	5	14.3	0.0	15	30.6	0.1
	Undergrad. 5th year+	0	0.0	0.0	5	10.2	0.0
	Undergrad. graduated	12	34.3	0.1	16	32.7	0.1
	Postgrad. 1st year	0	0.0	0.0	1	2.0	0.0
	Postgrad. 2nd year	1	2.9	0.0	2	4.1	0.0

Course Category	Postgrad. 3rd year	2	5.7	0.0	1	2.0	0.0
	Postgrad. 4th year	1	2.9	0.0	2	4.1	0.0
	Postgrad. 5th year+	1	2.9	0.0	0	0.0	0.0
	Postgrad. graduated	12	34.3	0.1	7	14.3	0.0
	Early childhood	3	8.6	0.0	5	10.2	0.0
	Primary	17	48.6	0.1	24	49.0	0.1
	Secondary	6	17.1	0.0	10	20.4	0.0
	Special education	1	2.9	0.0	0	0.0	0.0
	Other	8	22.9	0.0	10	20.4	0.0

Table 15 shows the demographic characteristics for the candidates who had achieved no standard at the end of 2017. The demographics of this group are similar to those of the previously described resit cohorts except that this group has the highest proportion of candidates from early childhood courses. While 7% of the first-attempt candidates were from early childhood courses, this proportion rose to 16% of those achieving no standard in 2017 for both literacy and numeracy.

Table 13: Demographic characteristics of candidates who achieved no standard by the end of 2017

Characteristic	Category	Literacy			Numeracy		
		N	%	% of Total Sitzings	N	%	% of Total Sitzings
Gender	Female	1669	83.4	7.1	1790	89.3	7.6
	Male	333	16.6	1.4	214	10.7	0.9
Age	17–25	1321	66.0	5.6	1230	61.4	5.2
	26–30	271	13.5	1.2	296	14.8	1.3
	31–35	134	6.7	0.6	128	6.4	0.5
	36–40	105	5.2	0.4	118	5.9	0.5
	41–45	82	4.1	0.4	107	5.3	0.5
	46+	89	4.4	0.4	125	6.2	0.5
International Students	No	1744	87.1	7.5	1949	97.3	8.3
	Yes	258	12.9	1.1	55	2.7	0.2
English as a First Language	Yes	1345	67.2	5.8	1650	82.3	7.0
	No	657	32.8	2.8	354	17.7	1.5
Indigenous	No	1904	95.1	8.1	1910	95.3	8.1
	Yes	55	2.7	0.2	55	2.7	0.2
	Not disclosed	43	2.1	0.2	39	1.9	0.2
Residential Area	Metropolitan areas	1736	86.7	7.4	1696	84.6	7.2
	Provincial areas	253	12.6	1.1	285	14.2	1.2
	Remote areas	12	0.6	0.1	17	0.8	0.1
	International	0	0.0	0.0	4	0.2	0.0
	Invalid or Missing	1	0.0	0.0	2	0.1	0.0
Program Type	Undergraduate	1527	76.3	6.5	1475	73.6	6.3
	Postgraduate	475	23.7	2.0	529	26.4	2.3
Program Type by Year Level	Undergrad. 1st year	357	17.8	1.5	256	12.8	1.1
	Undergrad. 2nd year	200	10.0	0.9	185	9.2	0.8
	Undergrad. 3rd year	402	20.1	1.7	359	17.9	1.5
	Undergrad. 4th year	372	18.6	1.6	456	22.8	1.9
	Undergrad. 5th year+	39	1.9	0.2	59	2.9	0.3
	Undergrad. graduated	157	7.8	0.7	160	8.0	0.7
	Postgrad. 1st year	188	9.4	0.8	179	8.9	0.8
	Postgrad. 2nd year	94	4.7	0.4	117	5.8	0.5
	Postgrad. 3rd year	14	0.7	0.1	25	1.2	0.1
	Postgrad. 4th year	27	1.3	0.1	39	1.9	0.2
	Postgrad. 5th year+	21	1.0	0.1	37	1.8	0.2
	Postgrad. graduated	131	6.5	0.6	132	6.6	0.6
Course Category	Early childhood	314	15.7	1.3	318	15.9	1.4
	Primary	780	39.0	3.3	827	41.3	3.5
	Secondary	416	20.8	1.8	411	20.5	1.8

	Special education	12	0.6	0.1	14	0.7	0.1
	Other	480	24.0	2.1	434	21.7	1.8

These candidates who had achieved no standard by the end of 2017 had up to four attempts at the test, as shown by Table 16. It is expected that some of these candidates will resit the test again in 2018.

Table 14: Number of attempts by candidates who achieved no standard by the end of 2017

Component	Year of registration	At end of	Number of 1-attempt candidates	Number of 2-attempt candidates	Number of 3-attempt candidates	Number of 4-attempt candidates
Literacy	2016	2017	182	90	60	19
	2017	2017	1385	390	89	2
Numeracy	2016	2017	204	132	96	29
	2017	2017	1379	352	77	3

Table 17 groups the location of testing into capital cities, regional cities and remote proctoring. It shows that while in 2016 only 18% of first-attempt candidates sat the test by remote proctoring, in 2017 this rose to 40%. The main reason for the larger percentage of first-attempt candidates sitting the test by remote proctoring was that remote proctoring was the only medium available in test window 1 in 2017. A more detailed breakdown by individual test centre may be found in Appendix 1.

Table 15: Number and proportion of first-attempt candidates who participated at test centres and by remote proctoring

Location of Testing	First Attempt			
	Literacy		Numeracy	
	N	%	N	%
Test Centres	13901	59.4	14004	59.7
– Capital Cities	11846	50.7	11941	50.9
– Regional Cities	2055	8.8	2063	8.8
Remote Proctoring	9486	40.6	9462	40.3
Total	23387	100.0	23466	100.0

Table 18 shows how resit candidates were distributed by location of testing in 2017. As for 2016, remote proctoring was increasingly used among resit candidates in 2017. For example, in the literacy component, while 41% of first-attempt candidates chose to use remote proctoring, this rose to 50% of second-attempt candidates and 56% of third-attempt candidates. Nearly two-thirds of fourth-attempt candidates sat the test by remote proctoring in 2017. Table 18 also shows that the percentage of candidates who sat the test at test centres in regional cities in 2017 was reasonably consistent by attempt, averaging 9% for first-, second- and third-attempt candidates.

Table 16: Number and proportion of resit candidates who participated at test centres and by remote proctoring

Location of Testing	Second Attempt				Third Attempt				Fourth Attempt			
	Literacy		Numeracy		Literacy		Numeracy		Literacy		Numeracy	
	N	%	N	%	N	%	N	%	N	%	N	%
Test Centres	751	50.5	659	48.3	131	44.3	143	42.2	13	37.1	19	38.8
– Capital Cities	628	42.2	538	39.4	102	34.5	117	34.5	11	31.4	16	32.7
– Regional Cities	123	8.3	121	8.9	29	9.8	26	7.7	2	5.7	3	6.1
Remote Proctoring	737	49.5	706	51.7	165	55.7	196	57.8	22	62.9	30	61.2
Total	1488	100.0	1365	100.0	296	100.0	339	100.0	35	100.0	49	100.0

2.2 Demographic Characteristics of Candidates by Test Windows

In 2016, the number of candidates was greatest in test window 1. The reverse was true in 2017, most likely because remote proctoring was the only medium offered. In 2017, for both literacy and numeracy, the number of candidates rose from nearly 4300 in test window 1 to an average of nearly 7000 across the other three test windows. The actual numbers for test windows 1–4 were: 4264, 7508, 6668 and 6766 for literacy and 4287, 7540, 6662 and 6730 for numeracy. Table 19 and Table 20 show the number and proportion of candidates who participated in the test by demographic characteristics for each test window.

The distribution of candidates by demographic characteristics for literacy was similar to the distribution for numeracy. For example, Table 19 shows that for literacy, across four test windows, the majority of candidates were female (in the range of 74% to 77%), and candidates were mostly in the age group 17–25 (in the range of 60% to 65%). The majority of candidates sitting the literacy component in each test window were enrolled in an undergraduate course. In 2017, undergraduate candidates were a larger proportion in test windows 1 and 4 (63% and 67%) compared to test windows 2 and 3 (54% and 60%). Undergraduates in their fourth year were by far the largest group (42%) in test window 1 in 2017. The demographic characteristics of candidates in each test window were reasonably consistent with the demographic characteristics of candidates overall, except for international students and candidates for whom English was not their first language. These groups were under-represented in test window 1.

Over test windows 1 to 4, the proportion of fourth-year undergraduates sitting the literacy component decreased from 42% to 14% while the proportion of third-year undergraduates increased from 8% to 26%. The opposite trend was observed in postgraduate candidates taking the literacy component. The proportion of second-year postgraduate candidates dropped from 14% in test window 1 to 7% in test window 4. The largest group of postgraduate candidates was those in their first year, for whom, apart from test window 1 (13%), the proportion averaged was about 20% across test windows 1 to 3.

For course category, the proportions across test windows were quite consistent apart from candidates in secondary courses, where the proportions declined steadily from test window 1 (41%) to test window 4 (32%). The opposite trend was observed for those in Other courses where the proportions increased from test window 1 (12%) to test window 4 (17%).

Table 17: Demographic characteristics of candidates by test windows – literacy

Characteristic	Category	TW1		TW2		TW3		TW4	
		N	%	N	%	N	%	N	%
Gender	Female	3245	76.1	5635	75.1	4946	74.2	5178	76.5
	Male	1016	23.8	1868	24.9	1719	25.8	1585	23.4
	Indeterminate/intersex	3	0.1	5	0.1	3	0.0	3	0.0
Age	17–25	2620	61.4	4857	64.7	4254	63.8	4089	60.4
	26–30	637	14.9	1114	14.8	1009	15.1	1063	15.7
	31–35	341	8.0	537	7.2	500	7.5	582	8.6
	36–40	287	6.7	443	5.9	401	6.0	446	6.6
	41–45	214	5.0	299	4.0	279	4.2	334	4.9
	46+	165	3.9	258	3.4	225	3.4	252	3.7
International Students	No	4153	97.4	6860	91.4	6201	93.0	6426	95.0
	Yes	111	2.6	648	8.6	467	7.0	340	5.0
English as a First Language	Yes	3835	89.9	6328	84.3	5686	85.3	5853	86.5
	No	429	10.1	1180	15.7	982	14.7	913	13.5
Indigenous	No	4078	95.6	7253	96.6	6429	96.4	6515	96.3
	Yes	86	2.0	100	1.3	98	1.5	123	1.8
	Not disclosed	100	2.3	155	2.1	141	2.1	128	1.9
Residential Area	Metropolitan areas	3533	82.9	6435	85.7	5698	85.5	5661	83.7
	Provincial areas	681	16.0	1018	13.6	926	13.9	1053	15.6
	Remote areas	31	0.7	28	0.4	29	0.4	35	0.5
	International	14	0.3	18	0.2	11	0.2	13	0.2
	Invalid or Missing	5	0.1	9	0.1	4	0.1	4	0.1
Program Type	Undergraduate	2702	63.4	4052	54.0	3996	59.9	4500	66.5
	Postgraduate	1562	36.6	3456	46.0	2672	40.1	2266	33.5
Program Type by Year Level	Undergraduate first year	95	2.2	596	7.9	515	7.7	662	9.8
	Undergraduate second yr	94	2.2	215	2.9	479	7.2	716	10.6
	Undergraduate third year	340	8.0	608	8.1	1214	18.2	1751	25.9
	Undergraduate fourth yr	1784	41.8	2105	28.0	1429	21.4	974	14.4
	Undergraduate fifth yr +	189	4.4	258	3.4	137	2.1	85	1.3
	Undergraduate graduated	200	4.7	270	3.6	222	3.3	312	4.6
	Postgraduate first year	532	12.5	1839	24.5	1435	21.5	1245	18.4
	Postgraduate second year	590	13.8	860	11.5	597	9.0	458	6.8
	Postgraduate third year	51	1.2	86	1.1	80	1.2	97	1.4
	Postgraduate fourth year	86	2.0	175	2.3	138	2.1	113	1.7
	Postgraduate fifth year +	105	2.5	191	2.5	114	1.7	76	1.1
	Postgraduate graduated	198	4.6	305	4.1	308	4.6	277	4.1
Course Category	Early childhood	251	5.9	467	6.2	471	7.1	573	8.5
	Primary	1726	40.5	2681	35.7	2658	39.9	2833	41.9
	Secondary	1738	40.8	3224	42.9	2458	36.9	2141	31.6
	Special education	27	0.6	46	0.6	39	0.6	45	0.7
	Other	522	12.2	1090	14.5	1042	15.6	1174	17.4

The observations and patterns described above for literacy candidates across the four test windows are also pertinent for numeracy, as shown in Table 20.

Table 18: Demographic characteristics of candidates by test windows – numeracy

Characteristic	Category	TW1		TW2		TW3		TW4	
		N	%	N	%	N	%	N	%
Gender	Female	3300	77.0	5684	75.4	4977	74.7	5186	77.1
	Male	984	23.0	1852	24.6	1683	25.3	1542	22.9
	Indeterminate/intersex	3	0.1	4	0.1	2	0.0	2	0.0
Age	17–25	2628	61.3	4872	64.6	4236	63.6	3995	59.4
	26–30	641	15.0	1104	14.6	1031	15.5	1080	16.0
	31–35	348	8.1	540	7.2	492	7.4	582	8.6
	36–40	295	6.9	443	5.9	394	5.9	445	6.6
	41–45	213	5.0	310	4.1	277	4.2	360	5.3
	46+	162	3.8	271	3.6	232	3.5	268	4.0
International Students	No	4184	97.6	6897	91.5	6284	94.3	6509	96.7
	Yes	103	2.4	643	8.5	378	5.7	221	3.3
English as a First Language	Yes	3875	90.4	6376	84.6	5794	87.0	5944	88.3
	No	412	9.6	1164	15.4	868	13.0	786	11.7
Indigenous	No	4105	95.8	7285	96.6	6431	96.5	6482	96.3
	Yes	87	2.0	95	1.3	95	1.4	127	1.9
	Not disclosed	95	2.2	160	2.1	136	2.0	121	1.8
Residential Area	Metropolitan areas	3576	83.4	6470	85.8	5671	85.1	5624	83.6
	Provincial areas	660	15.4	1015	13.5	943	14.2	1049	15.6
	Remote areas	32	0.7	27	0.4	32	0.5	39	0.6
	International	14	0.3	19	0.3	11	0.2	14	0.2
	Invalid or Missing	5	0.1	9	0.1	5	0.1	4	0.1
Program Type	Undergraduate	2728	63.6	4055	53.8	4025	60.4	4463	66.3
	Postgraduate	1559	36.4	3485	46.2	2637	39.6	2267	33.7
Program Type by Year Level	Undergraduate first year	92	2.1	593	7.9	490	7.4	621	9.2
	Undergraduate second yr	96	2.2	214	2.8	491	7.4	702	10.4
	Undergraduate third year	346	8.1	589	7.8	1217	18.3	1717	25.5
	Undergraduate fourth yr	1802	42.0	2112	28.0	1460	21.9	1004	14.9
	Undergraduate fifth yr +	192	4.5	268	3.6	150	2.3	103	1.5
	Undergraduate graduated	200	4.7	279	3.7	217	3.3	316	4.7
	Postgraduate first year	527	12.3	1846	24.5	1393	20.9	1208	17.9
	Postgraduate second year	591	13.8	870	11.5	600	9.0	472	7.0
	Postgraduate third year	50	1.2	88	1.2	80	1.2	102	1.5
	Postgraduate fourth year	90	2.1	173	2.3	145	2.2	116	1.7
	Postgraduate fifth year +	110	2.6	197	2.6	111	1.7	82	1.2
	Postgraduate graduated	191	4.5	311	4.1	308	4.6	287	4.3
Course Category	Early childhood	249	5.8	453	6.0	485	7.3	584	8.7
	Primary	1749	40.8	2723	36.1	2686	40.3	2877	42.7
	Secondary	1728	40.3	3217	42.7	2420	36.3	2092	31.1
	Special education	23	0.5	44	0.6	43	0.6	41	0.6
	Other	538	12.5	1103	14.6	1028	15.4	1136	16.9

Table 21 and Table 22 show the number and proportion of candidates participating in test centres and by remote proctoring in each test window for literacy and numeracy respectively. It can be seen that the proportion of candidates sitting in test centres in capital cities ranged between 56% and 63% for test windows 2 to 4. For test centres in regional cities, the proportion stayed relatively steady, between 9% and 12%. For remote proctoring, the proportion fluctuated from 100% in test window 1 to between 25% and 35% for test windows 2 to 4. A more detailed breakdown by test centre can be found in Appendix 2.

Table 19: Number and proportion of candidates in test centres and by remote proctoring by test window – literacy

Location of Testing	TW1 ⁸		TW2		TW3		TW4	
	N	%	N	%	N	%	N	%
Test Centres	NA	NA	4877	65.0	4958	74.4	4961	73.3
– Capital Cities	NA	NA	4204	56.0	4168	62.5	4215	62.3
– Regional Cities	NA	NA	673	9.0	790	11.8	746	11.0
Remote Proctoring	4264	100	2631	35.0	1710	25.6	1805	26.7
Total	4264	100.0	7508	100.0	6668	100.0	6766	100.0

Table 20: Number and proportion of candidates in test centres and by remote proctoring by test window – numeracy

Location of Testing	TW1 ⁹		TW2		TW3		TW4	
	N	%	N	%	N	%	N	%
Test Centres	NA	NA	4903	65.0	4967	74.6	4955	73.6
– Capital Cities	NA	NA	4229	56.1	4174	62.7	4209	62.5
– Regional Cities	NA	NA	674	8.9	793	11.9	746	11.1
Remote Proctoring	4287	100.0	2637	35.0	1695	25.4	1775	26.4
Total	4287	100.0	7540	100.0	6662	100.0	6730	100.0

2.3 Accessibility and Accommodations

In 2017, 179 candidates (0.7%) required special testing conditions.

Table 23 indicates the number of accommodations made for the eight most common conditions. A complete list of conditions follows the table.

Table 21: Largest accommodation groups

Condition	Literacy		Numeracy	
	2016	2017	2016	2017
Anxiety disorder (including panic attacks)	3	28	5	31
Dyslexia	9	26	9	25
Diabetes	11	8	10	9
Epilepsy/Seizures	1	5	2	6
Attention Hyper Activity Disorder (ADHD)	1	5	1	5
Hearing impairment	2	5	1	4
Visual impairment	6	4	6	5
Dyscalculia	NA	NA	2	7

Types of conditions:

- Acquired Brain Injury
- Adrenal Insufficiency, Panhypopituitarism and Benign Intracranial Hypertension
- Agoraphobia
- Anxiety, Depression, Panic Attacks
- Asperger Syndrome (High functioning – ASD)
- Asthma
- Auditory Processing Disorder and Visual-Perceptual Dysfunction (Scotopic Sensitivity/Irlen Syndrome)
- Auditory–Verbal Memory Disorder / Language Learning Disability
- Autism Spectrum Disorder
- Bilateral Sensorineural Hearing Loss

⁸ For test window 1, the test was available only by remote proctoring.

⁹ For test window 1, the test was available only by remote proctoring.

- Brain Tumour
- Chronic Autoimmune Urticaria and Angio-Edema
- Chronic Soft Tissue Inflammation with mild Dupuytren's Contracture of the Hands
- Chronic Back Pain
- Congenital Nystagmus (Eye Disorder)
- Cranial Diabetes Insipidus
- Cystic Fibrosis
- Degenerative Spondylosis
- Diabetes
- Dyslexia, Dyscalculia, ADD
- Epilepsy
- Fibromyalgia
- Hashimoto's Autoimmune Disease
- Herniated Discs and Fractured Vertebrae
- High Blood Pressure and Hypoglycaemia
- Hip Dysphasia
- Hypothalamic Pituitary Dysfunction secondary to Craniopharyngioma
- Hypothyroidism
- Idiopathic Intracranial Hypertension
- Low Working Memory
- Migraine
- Obsessive Compulsive Disorder and Pronounced Exam Phobia
- Pigmentary Retinopathy with High Myopia and Astigmatism
- Profoundly Deaf / Auslan User
- PTSD
- Recovery from Stroke
- Rheumatoid Arthritis
- Severe Language Disorder
- Supraventricular Tachycardia
- Systemic Lupus Erythematosus
- Temporary physical conditions – e.g. broken leg, broken wrist, back injury
- Tinnitus
- Visual Impairment / Legally Blind

Types of accommodations granted:

- Emergency Action Plan (for Epilepsy – seizures)
- Extra time (20 minutes per test component)
- Management of hearing impairment for test sessions conducted by remote proctoring (communication via chat box only)
- Permission to bring blood-insulin monitor, food and drink relating to medical condition
- Permission to bring support aids (cushion and mobility aids)
- Permission to take medication (e.g. Ventolin inhaler and diabetes/glucose monitoring kit)
- Permission to stand and stretch
- Permission to magnify text and to wear Irlen Spectral Filters / coloured glasses
- Permission to use listening device
- Permission to wear headphones and listen to music during the test session
- Provision of additional working out paper
- Provision of an Auslan interpreter
- Provision of a small group test environment (no more than 5 candidates per test room)

- Provision of text to speech software
- Rest breaks (5 mins per test component)
- Seated near bathroom
- Seated at the front of the test room (for hearing loss) and other special seating requests for the front and back of the test room, and near the aisle
- Seated in a quiet room
- Special support for candidates with limited mobility (i.e. limit time standing in the registration queue)
- Test supervisor to provide written assistance during the instructions component of the test sessions.

During 2017, a consultant reviewed the updated online test versions (one for each of literacy and numeracy) that are accessible to candidates using screen readers. Further modifications were made based upon the consultant's findings, mainly to aid navigation. Accessible versions of paper and online practice items were also created for each component of the test.

During test window 3, the accessible paper version of the accessible test form was utilised by one candidate with visual impairment. The candidate sat both the literacy and numeracy components (one component per day) at the ACER office in Camberwell. The candidate was given extended test time and extended and multiple rest breaks, was permitted to bring other items such as personal calculator, highlighters, ruler and clock, necessitating individual supervision for each of the tests.

The online accessible versions of the test were used twice for literacy and once for numeracy in 2017.

3. CANDIDATE PERFORMANCE

This section describes the performance of candidates who participated in the test in 2017. The analysis breaks the cohort of candidates into two groups: first-attempt candidates (the majority) and those who did not achieve the standard at their first attempt and resat the test. It presents the distributions of candidate performance overall, by subscale and by candidates' collected demographic information: gender, age group, program type, program type by year level, course category, and location of testing.

3.1 Scale Score Distributions

Figure 1 and Figure 2 present the distributions of first-attempt candidate performance on the literacy component and numeracy component respectively. The vertical lines in each figure represent the standards for that component of the test: the dashed vertical line is the provisional standard used in 2016 and test windows 1 and 2 of 2017; the solid vertical line is the revised standard used in test windows 3 and 4 in 2017.

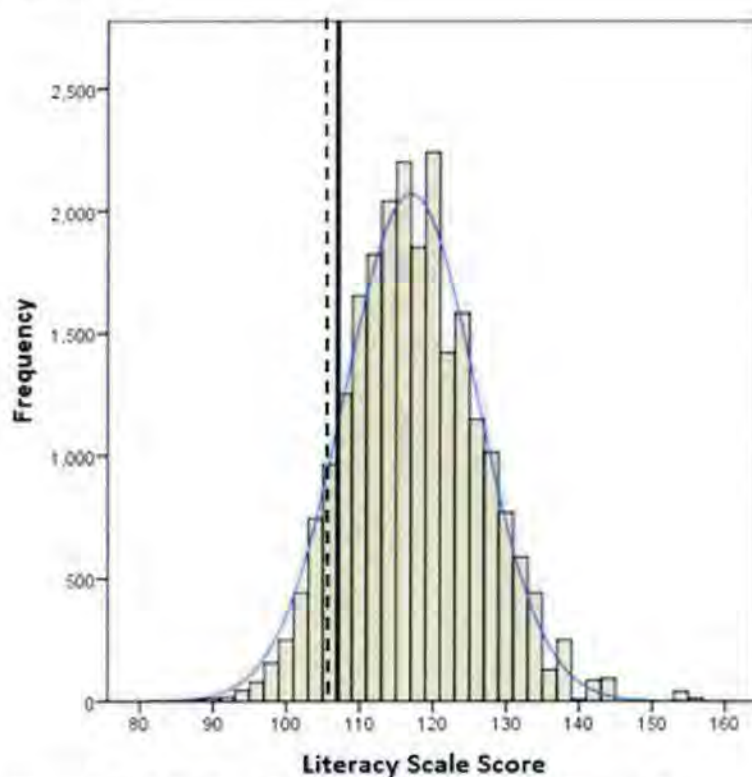


Figure 1: Distribution of candidate scale scores for literacy¹⁰

¹⁰ In 2017 test windows 1–2, the literacy standard scale score was 106. From 2017 test windows 3–4 onwards, the literacy scale score is 107.

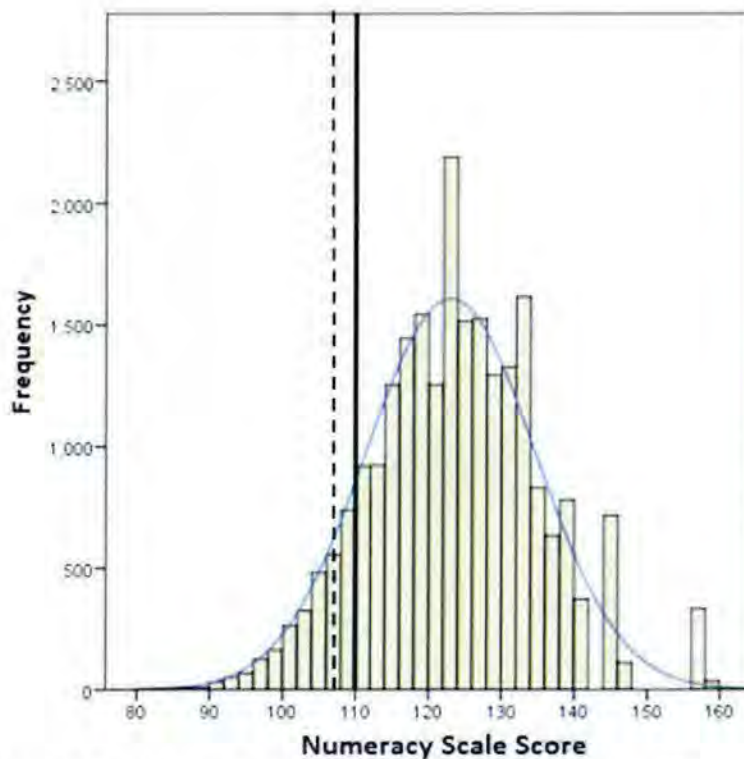


Figure 2: Distribution of candidate scale scores for numeracy¹¹

Figure 1 and Figure 2 show that scores in both tests are approximately normally distributed and that the tests spread candidates acceptably across the score scales. For both literacy and numeracy, the majority of candidates achieved scale scores above the standard at their first attempt. It can be seen that a proportion of candidates achieved scale scores below the standard at their first attempt.

¹¹ In 2017 test windows 1–2, the numeracy standard scale score was 107. From 2017 test windows 3–4 onwards, the numeracy standard scale score is 110.

3.2 Candidate Scale Scores by Subscales and Subgroups

Table 24 shows the performance of first-attempt candidates in 2017. It shows the number of candidates, the mean scale scores and standard deviation of the scale scores, overall and by subscale. The pass rates for the literacy and numeracy components for this cohort of candidates are also shown in this table. The overall mean scale score for literacy was 117.0 (very similar to 117.5 in 2016) with a pass rate of 89.2% (down from 93.3% in 2016). The overall mean scale score for numeracy was 123.0 (similar to 122.4 in 2016) with a pass rate of 90.0 (down from 92.4% in 2016). This table also shows the performance of candidates on each subscale. The performance of candidates by subscale was similar to the performance of candidates on the whole test, except that, as for 2016, the average performance of the numeracy subscale 'calculator not available' was lower (3.5 scale score points) than the average performance of the numeracy subscale 'calculator available'. Scale score frequency distributions for the candidates who participated in the tests are shown in Appendix 3.

Table 22: Candidate performance overall and by subscale for first-attempt candidates

Component	Whole test and subscale	Valid N	Mean	S.D.	Pass Rate
Literacy	Overall	23387	117.0	9.0	89.2
	Reading		117.1	9.5	
	Technical skills of writing		116.9	10.5	
Numeracy	Overall	23465 ¹²	123.0	11.6	90.0
	Number & algebra		122.3	12.8	
	Measurement & geometry		122.8	12.1	
	Statistics & probability		123.0	11.9	
	Calculator available		123.2	11.2	
	Calculator not available		119.7	14.0	

Table 25 shows the number of candidates, mean scale score, standard deviation of the scale scores, and pass rate by demographic characteristics. Performance of any subgroup with a sample size of less than 10 was not reported.

Table 23: Performance by demographic characteristics for first-attempt candidates

Characteristic	Category	Literacy			Numeracy		
		N	Mean	Pass Rate	N	Mean	Pass Rate
Gender	Female	17484	116.5	88.1	17537	121.4	88.1
	Male	5892	118.4	92.4	5918	127.7	95.7
	Indeterminate/intersex	11	115.1	81.8	10	125.0	90.0
Age	17–25	14541	115.7	88.1	14617	122.4	90.0
	26–30	3569	118.4	91.2	3562	124.1	90.3
	31–35	1868	119.0	91.4	1850	124.6	91.8
	36–40	1484	119.6	91.2	1499	124.2	90.8
	41–45	1073	120.1	91.8	1079	124.0	89.5
	46+	852	119.4	88.5	858	122.1	83.8
International Students	No	22093	117.3	90.1	22160	122.9	89.7
	Yes	1294	111.6	73.8	1305	124.8	94.6
English as a First Language	Yes	20475	117.7	91.5	20532	123.1	90.4
	No	2912	111.8	73.3	2933	122.2	87.5
Indigenous	No	22562	117.0	89.4	22636	123.1	90.1
	Yes	356	113.8	80.9	356	118.0	83.7
	Not disclosed	469	117.4	88.5	473	124.0	90.5
Residential Area	Metropolitan areas	19792	116.9	89.0	19857	123.1	90.0
	Provincial areas	3400	117.0	90.4	3413	122.2	90.0
	Remote areas	120	118.6	90.8	119	122.9	83.2
	International	55	123.8	100.0	56	127.6	92.9
	Invalid or Missing	20	119.4	90.0	20	121.9	90.0

¹² One candidate who had their numeracy result cancelled due to misconduct was excluded from scale score aggregates.

Program Type	Undergraduate	14030	115.2	86.6	14086	121.1	87.9
	Postgraduate	9357	119.6	93.2	9379	125.8	93.2
Program Type by Year Level	Undergraduate first year	1709	112.5	75.2	1739	120.3	83.8
	Undergraduate second year	1437	115.2	83.9	1459	121.7	86.1
	Undergraduate third year	3748	115.2	87.3	3746	121.6	88.9
	Undergraduate fourth year	5753	116.0	90.3	5758	121.2	89.2
	Undergraduate fifth year or above	606	116.9	91.6	613	121.4	88.9
	Undergraduate graduated	777	113.6	81.2	771	118.8	84.3
	Postgraduate first year	4867	119.8	93.8	4868	127.1	94.8
	Postgraduate second year	2387	120.6	94.6	2396	125.8	93.7
	Postgraduate third year	292	120.4	94.5	295	123.9	88.5
	Postgraduate fourth year	469	117.8	90.0	469	122.2	87.0
	Postgraduate fifth year or above	449	119.1	93.8	451	124.7	91.8
	Postgraduate graduated	893	116.7	87.6	900	122.0	88.3
Course Category	Teacher education: early childhood	1584	113.0	77.2	1575	117.1	77.3
	Teacher education: primary	9189	116.5	89.5	9243	121.8	89.2
	Teacher education: secondary	9047	119.1	93.3	9053	126.2	94.5
	Teacher education: special education	144	116.4	92.4	144	121.8	88.2
	Teacher education: other	3423	114.4	83.2	3450	120.7	86.0

The t-test and Cohen's *d* for effect size were used to determine if group mean scale scores were significantly different for first-attempt candidates. Only differences where $p \leq 0.05$ and $d > 0.2$ are reported here as significant.

In 2017, there were significant differences in mean scale scores for the demographic characteristics as follows:

- Language background and Indigeneity for both literacy and numeracy
- Residential area (international or domestic) for both literacy and numeracy
- Program type and course category for both literacy and numeracy
- Candidate age for literacy but not for numeracy
- Gender for numeracy but not for literacy.

Table 25 shows that female candidates performed similarly to the male candidates in literacy, but male candidates significantly outperformed female candidates in numeracy. As for 2016, the numeracy mean scale score of male candidates (127.7) was significantly higher (effect size 0.6) than the mean scale score of female candidates (121.4). For literacy, while the mean scale score of males was also higher than that of females, the difference was not significant. In literacy, the pass rate of female candidates (88.1%) was marginally lower than that of male candidates (92.4%). In numeracy, the pass rate of male candidates (95.7%) was higher than the pass rate of female candidates (88.1%).

For literacy and numeracy, achievement on the test tended to increase with the age of the candidates, but the pattern was much more evident for literacy. For literacy, the youngest group of candidates aged 17–25 (mean scale score 115.7) achieved significantly lower (effect size 0.4) than candidates aged over 25 (119.0). For numeracy, while the mean scale score of candidates aged over 25 was also greater than that of those aged 17–25 (124.0 to 122.4), the difference was not significant.

While the mean scale score of international candidates (111.6) for literacy was significantly lower (effect size 0.6) than the mean scale score of other candidates (117.3), the same was not true for numeracy. For numeracy, while the mean scale score of international candidates (124.8) was higher than the mean scale score of other candidates (122.9), the difference was not significant.

The mean scale score for literacy of candidates for whom English was a first language (117.7) was significantly higher (effect size 0.7) than the mean scale score for literacy of other candidates (111.8). However, the same was not true for numeracy where the difference (123.1 and 122.2) was not significant.

The mean scale score of candidates who identified as Aboriginal or Torres Strait Islander was significantly lower (effect size 0.4) than for other candidates, for both literacy and numeracy. For literacy, the mean scale scores were 113.8 and 117.0 respectively; and for numeracy, 118.0 and 123.1 respectively. However, it is worth noting that the pass rates of first-attempt candidates who identified as Aboriginal or Torres Strait Islanders were still relatively high at 81% for literacy and 84% for numeracy. For literacy, the pass rate of candidates identifying as Aboriginal or Torres Strait Islander was higher than that of international candidates (74%) and candidates for whom English was not a first language (73%).

Residential postcode data were used to place candidates into three main categories: metropolitan, provincial and remote. Where postcodes could not be matched to an indicator they were categorised as international, missing or invalid. For both literacy and numeracy, there was little difference in achievement by Australian residential areas (metropolitan, provincial and remote). However, for both literacy and numeracy, the mean scale scores of candidates with an international residential postcode were significantly higher than those of candidates who gave an Australian residential postcode (effect size 0.7 for literacy and 0.4 for numeracy). While the size of the cohort with international residential postcodes is small (55 for literacy and 56 for numeracy), their achievement is notable. For example, in literacy, the pass rate of first-attempt candidates with an international postcode was 100% compared to 89% for those with a metropolitan residential postcode. The mean scale scores of each group were significantly different, 123.8 and 116.9 respectively.

As for 2016, for both literacy and numeracy, the mean scale scores of postgraduate candidates were significantly higher (effect size 0.5 and 0.4 respectively) than for undergraduate candidates. The difference in mean scale scores was 4 to 5 scale score points for both components: 119.6 and 115.2 for literacy and 125.8 and 121.1 for numeracy respectively.

For literacy, the mean scale scores of undergraduate candidates increased with the year of the course, ranging from 112.5 for first-year undergraduates to 116.9 for fifth-year undergraduates. The pattern was less evident for numeracy. For postgraduate candidates, there were no obvious patterns in achievement by year of course for both literacy and numeracy.

As for 2016, for both literacy and numeracy, the mean scale scores of candidates in the secondary education course category were significantly higher (effect size 0.3 and 0.4 respectively) than those of candidates in the primary and special education course categories. In turn, the mean scale scores of candidates in the primary and special education course categories were significantly higher (effect size 0.4) than those of candidates in the early childhood course category. Candidates in the 'other' course category achieved relatively better in numeracy (equivalent to candidates in the primary and special education categories) than in literacy (equivalent to candidates in the early childhood category). Candidates in the other course category achieved significantly higher (effect size 0.3) on numeracy than candidates in the early childhood category.

Table 26 shows the performance by test centres and remote proctoring. As for 2016, the candidates at the Hobart test centre had the highest mean score (120.8 compared to 123.2 in 2016) in literacy. And for numeracy, the candidates at the Perth CBD test centre again had the highest mean score (126.3 compared to 127.8 in 2016).

For both literacy and numeracy, the mean scale scores of candidates who used the test centres at Perth CBD and at Sunshine Coast / Maroochydore were relatively high (4 to 5 scale score points better) compared to the mean scaled scores of candidates at the Melbourne CBD, Wagga Wagga, Ballarat and Darwin test centres.

Out of the 23 test centres, the mean scale scores of candidates using the Adelaide CBD and Bathurst test centres were ranked relatively high on the numeracy component of the test compared to their ranks on the literacy component. For example, Adelaide CBD was ranked second for numeracy and ninth for literacy.

In contrast, the mean scaled scores of candidates using the Canberra CBD test centre were ranked relatively high on the literacy component of the test compared to their rank on the numeracy component (seventh for literacy and sixteenth for numeracy).

It can be seen from the last three rows of Table 26 that the performance of candidates using remote proctoring was very similar to the performance of candidates who took tests in capital city test centres and regional city test centres, with only 0.3 scale score points separating the mean scale scores of the three groups for both test components. The small differences were not significant.

Table 24: Performance by test centres and remote proctoring

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Test Centre	Adelaide CBD	349	117.7	9.5	88.8	349	125.8	11.5	92.8
	Albury	111	115.9	7.8	87.4	110	121.5	11.0	90.0
	Armidale (NSW)	108	117.6	8.9	95.4	106	122.6	11.5	89.6
	Ballarat	231	115.3	8.3	90.5	234	121.0	9.5	88.9
	Bathurst	52	115.6	9.1	80.8	57	123.8	13.6	86.0
	Brisbane CBD	1662	117.7	8.4	92.4	1656	123.5	11.1	92.0
	Bundaberg	55	116.7	6.8	94.5	56	123.4	10.9	96.4
	Cairns	57	119.2	9.2	93.0	56	124.3	9.5	92.9
	Canberra CBD	211	117.9	9.2	88.6	208	122.3	11.1	90.4
	Darwin	32	115.2	7.6	90.6	32	117.6	8.4	87.5
	Gold Coast	99	116.2	8.3	87.9	99	121.8	11.5	87.9
	Hobart	50	120.8	9.5	98.0	49	123.6	12.9	89.8
	Melbourne CBD	5380	115.5	9.3	84.6	5450	121.8	11.8	86.8
	Mildura	36	115.6	7.5	86.1	35	122.5	11.9	88.6
	Newcastle	452	118.2	8.2	93.8	456	124.7	11.2	93.0
	Parramatta	341	116.3	9.5	87.7	339	121.8	12.0	87.0
	Perth CBD	976	120.1	8.9	94.5	980	126.3	11.5	93.8
	Special testing conditions	3	-	-	-	2	-	-	-
	Sunshine Coast / Maroochy	134	119.2	8.3	97.8	133	124.6	10.9	92.5
	Sydney CBD	3183	117.4	9.0	89.4	3215	123.9	11.8	90.8
	Townsville	75	117.6	7.0	93.3	76	124.5	10.6	94.7
	Wagga Wagga	94	115.4	7.3	92.6	94	121.2	9.9	91.5
	Warrnambool	39	118.4	8.4	94.9	36	122.9	10.7	86.1
	Wollongong	171	117.9	7.8	91.8	176	124.9	10.4	96.0
	Remote Proctoring	9486	117.1	8.9	90.1	9461	122.8	11.6	90.5
	Capital Cities	11846	116.8	9.2	88.1	11941	123.1	11.7	89.4
	Regional Cities	2055	117.1	8.4	91.4	2063	123.1	11.1	90.9

Figure 3 to Figure 6 display scale score distributions of demographic characteristics for first-attempt candidates in 2017. The left panel of each figure shows literacy scale score distributions and the right panel of each figure shows numeracy scale score distributions. The horizontal lines in each figure represent the standard scale score for each component of the test.

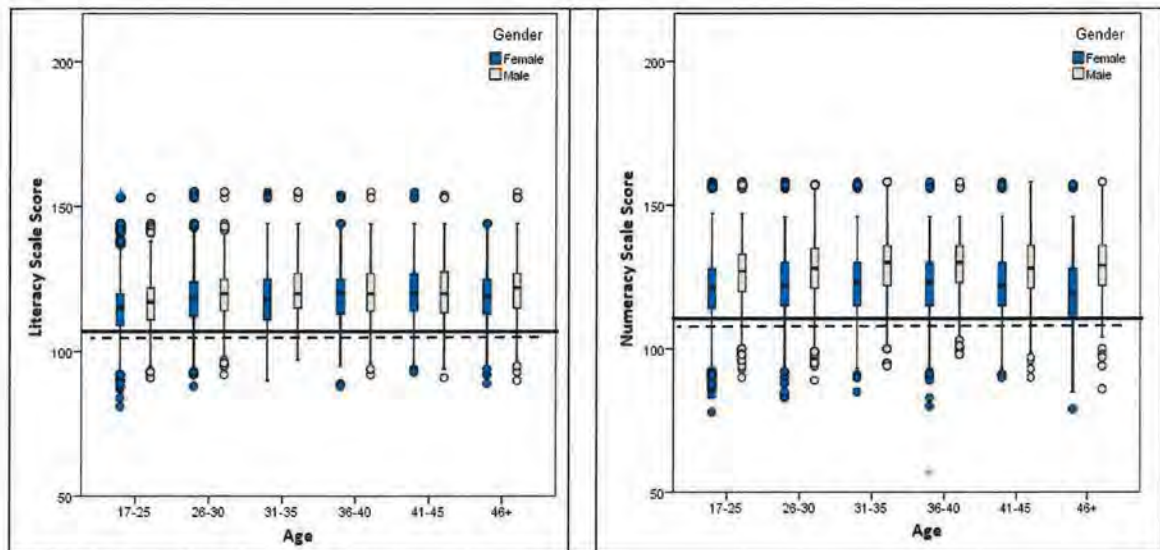


Figure 3: Score distribution by gender and age

Figure 3 shows that, for all age groups, the difference in achievement between male candidates and female candidates is more pronounced for numeracy than for literacy. However, in each age category and for male or female, there are candidates who achieve well above the standard.

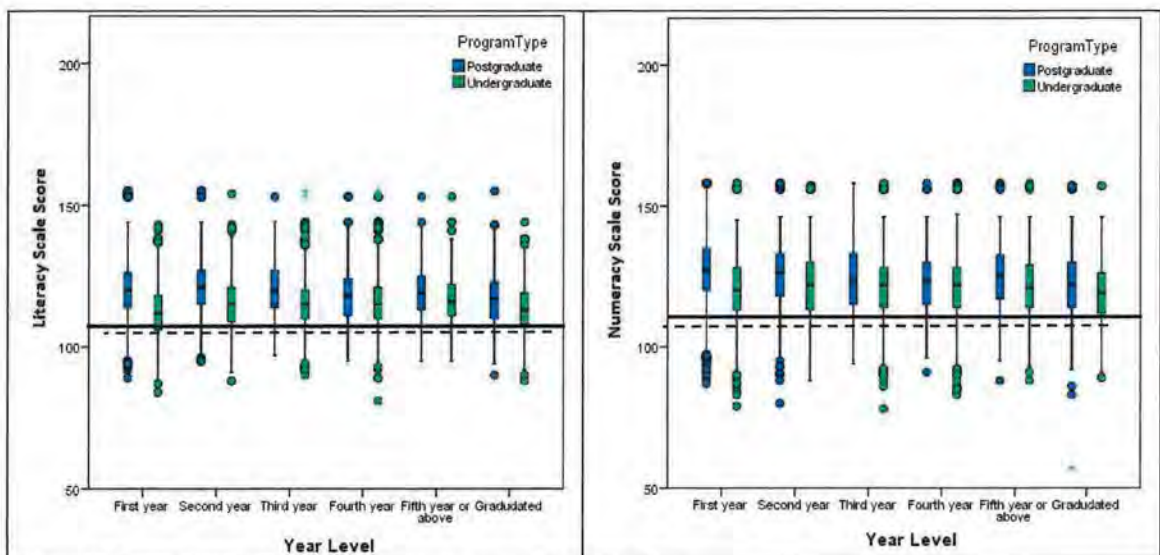


Figure 4: Score distribution by program type and year level

Figure 4 shows that, for all year levels, the difference in achievement between postgraduate candidates and undergraduate candidates is similar for literacy and numeracy. As for gender, at each year level and in each program type, there are candidates who achieve well above the standard.

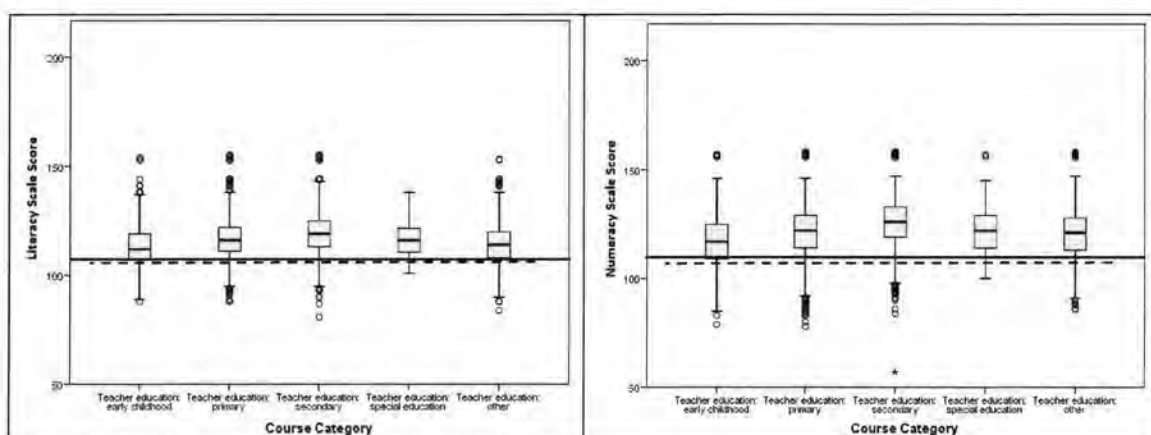


Figure 5: Score distribution by course category

Figure 5 shows that, although candidates in secondary education courses achieve highest in both literacy and numeracy, there are candidates who achieve well above the standard in each of the other courses. For example, for both literacy and numeracy, the achievement of the top 25% of candidates in the early childhood category is broadly equivalent to the top 50% of candidates in the secondary category.

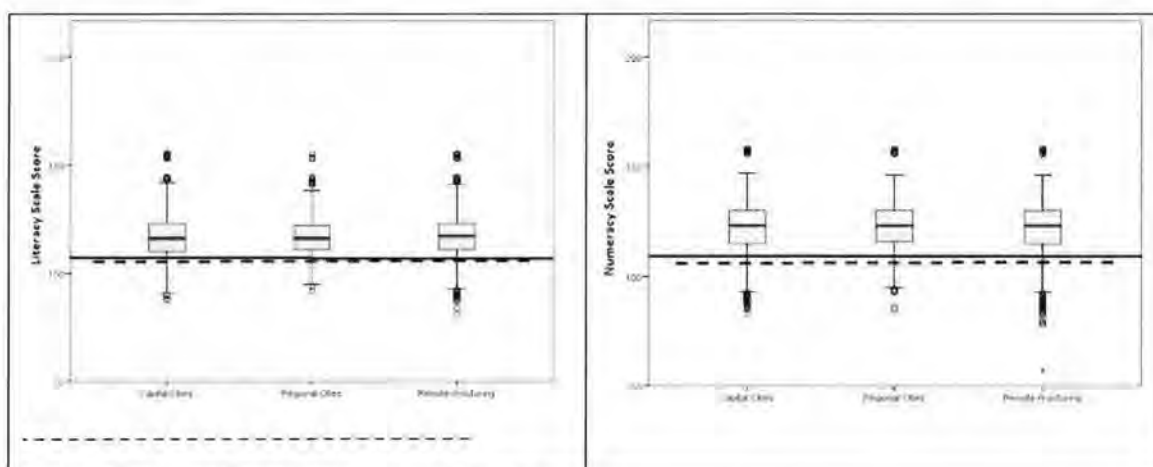


Figure 6: Score distribution by location of testing

Figure 6 shows that for test centres in capital cities and regional cities, and for remote proctoring, for both literacy and numeracy, there are candidates who achieve well above the standard. Figure 6 also shows that for all three categories, there are candidates who did not achieve the standard at their first attempt.

Table 27 shows the number and proportion of candidates who did not achieve the test standard in 2017 after one attempt. The proportion of first-attempt candidates who did not achieve the standard on the literacy component was 10.8% (up from 6.7% in 2016) and 10.0% (up from 7.6% in 2016) for numeracy. The percentage of first-attempt candidates who did not achieve either standard in 2017 was 4.5% (up from 2.7% in 2016).

Table 25: Number and proportion of first-attempt candidates who did not achieve the standard

	Number	Percentage
Candidates who did not achieve the literacy standard	2523	10.8
Candidates who did not achieve the numeracy standard	2350	10.0
Candidates who did not achieve the literacy and the numeracy standard	1034	4.5

3.3 Performance of Resit Candidates

Table 28 shows the performance of candidates who resat the test, at their second attempt and third attempt, overall and by subscale. Table 28 includes 408 candidates who first attempted literacy in 2016 and also 492 candidates who first attempted numeracy in 2016. As would be expected, the performance of resit candidates was significantly lower than the performance of the majority of candidates who achieved the standard at their first attempt. For example, for the 1488 candidates who were making their second attempt for literacy, their overall mean scale score was 106.5 with a pass rate of 53.6% (compared to 117.0 and 89.2% for the first-attempt candidates). For the 1365 candidates who were making their second attempt for numeracy, their overall mean scale score for numeracy was 108.1 with a pass rate of 50.0% (compared to 123.0 and 90.0% for the first-attempt candidates).

In 2016, the mean pass rates of resit candidates in numeracy were much lower than the mean pass rates in literacy. However, this cannot be said for 2017, where the mean pass rates of resit candidates are similar for both components. The pass rate of candidates in their third attempt was 39.9% for literacy (compared to 76.0% in 2016) and 41.3% for numeracy (compared to 37.5% in 2016). The change in the pass rates of third-attempt candidates is most likely due to the relatively small numbers in 2016 (25 for literacy and 40 for numeracy). In 2017, approximately 50% of candidates achieved the literacy and numeracy standards on their second attempt, and approximately 40% on their third or fourth attempt.

For the literacy subscales, the mean scale score of resit candidates was similar for reading and technical skills of writing. However, for the numeracy subscales, the mean scores of resit candidates on the 'number & algebra' and 'calculator not available' subscales were lower than the mean scores of resit candidates on the other three numeracy subscales, suggesting these are the numeracy skills where resit candidates need most support.

Table 26: Resit candidate performance overall and by subscale

Component	Whole test and subscale	Second Attempt (Resit 1)			Third Attempt (Resit 2)			Fourth Attempt (Resit 3)		
		N	Mean	Pass Rate	N	Mean	Pass Rate	N	Mean	Pass Rate
Literacy	Overall	1488	106.5	53.6	296	105.3	39.9	35	106.1	40.0
	Reading		106.6			105.5			105.5	
	Technical skills of writing		106.1			105.1			107.2	
Numeracy	Overall	1365	108.1	50.0	339	107.2	41.3	49	105.9	34.7
	Number & algebra		106.4			105.7			104.0	
	Measurement & geometry		109.1			108.2			107.0	
	Statistics & probability		110.0			108.8			108.2	
	Calculator available		109.4			108.2			106.7	
	Calculator not available		103.4			103.5			103.5	

Despite greater familiarity with the format of the test and time for additional study, the resit cohorts showed little or no change in the mean scale scores by resit for literacy and a decline by resit for numeracy. However, at the individual candidate level, a different picture emerged.

Most second-attempt candidates (73% for literacy and 69% for numeracy) had their second attempt within 4 months. Almost all second-attempt candidates (90% for literacy and 87% for numeracy) had their second attempt within 6 months. A very small number of candidates waited until 12 or 18 months after the first attempt before resitting. Regression analysis was performed to understand the effect of resitting the test after different time periods. The analysis indicated that there was a very weak *negative* relationship between change in test scores and the length of time between first and second attempts for literacy, and no relationship for numeracy. It can be seen from Table 29 that for both literacy and

numeracy, but particularly for numeracy, there is very little difference in the mean score change between first and second attempts regardless of the time taken between the attempts.

Table 27: Change in scale score between first and second attempts by time

Component	Mean score change (scale score points)				
	Less than 2 mths	From 2 to <4 mths	From 4 to <6 mths	More than 6 mths	All
Literacy	6.2	5.2	4.4	4.6	5.0
Numeracy	6.9	6.3	6.3	6.5	6.4

Overall, each second-attempt cohort improved their scales scores (5.0 points for literacy and 6.4 points for numeracy). However, after taking performance (Band level *after* second attempt) into account it can be seen from Table 30 that the change in scale scores was not uniform. The mean score change of the least able cohort (those with second-attempt scores below Band 1) was negative (by almost 5 scale score points) for literacy and zero (no change) for numeracy. In general, the higher the performance of the second-attempt candidates, the more they were able to improve their scores between their first and second attempts.

Table 28: Change in scale score between first and second attempts by performance (Band)

Component	Mean score change (scale score points)				
	Below Band 1	Band 1	Band 2	Band 3 and above	All
Literacy	-4.8	1.0	7.1	21.7	5.0
Numeracy	0.0	3.0	9.1	28.3	6.4

The findings above suggest that it is more likely that the change in score between first and second attempts is more likely to be explained by performance than it is by the time between testing.

Additional analysis investigated the impact of resit candidates on pass rates. Table 31 shows that the pass rates in 2017 for no-resit candidates and for all candidates (including resit candidates) were 93.8% and 91.6% for literacy, and 93.9% and 91.6% for numeracy, respectively. Therefore, performance of resit candidates reduced the pass rate by 2.2% in literacy and 2.3% in numeracy. The impact of resit candidates on the pass rate was greater in 2017 than in 2016. The performance of resit candidates in 2016 reduced the pass rate by 0.2% (98.5% to 98.3%) in literacy and 0.4% in numeracy (98.3% to 97.9%).

Table 29: Candidate performance by number of test sittings in 2016 and 2017

Component	Year	Number of Test Sittings	Number of Unique Candidates	Standard Achieved	Standard Not Achieved	Pass Rate
Literacy	2016	1 (no resits)	12391	12210	181	98.5
		2	263	231	32	87.8
		3	20	19	1	95.0
		All	12674	12460	214	98.3
	2017	1 (no resits)	22249	20864	1385	93.8
		2 ¹³	1245	797	448	64.0
		3	266	118	148	44.4
		4	35	14	21	40.0
		All	23795	21793	2002	91.6
Numeracy	2016	1 (no resits)	12286	12082	204	98.3

¹³ The n-attempt candidates included candidates who did their n-1 attempt in 2016 and resat in 2017.

		2	277	225	52	81.2
		3	26	15	11	57.7
		4	2	2	0	100.0
		All	12591	12324	267	97.9
	2017	1 (no resits)	22492	21114	1378	93.9
		2	1115	683	432	61.3
		3	302	140	162	46.4
		4	49	17	32	34.7
		All	23958 ¹⁴	21954	2004	91.6

¹⁴ Total excludes a resitting candidate who registered in 2016 and whose numeracy result was cancelled due to misconduct.

4. CANDIDATE PERFORMANCE BY TEST WINDOWS

This section presents candidate performance by test window. The performance of resit candidates is also described by test window.

4.1 Distributions of Candidate Scale Scores by Subscales and Subgroups

Table 32 shows the performance of all candidates (first-attempt and resits) who sat the test for each test window in 2017. The overall mean scale scores for literacy were similar across the four test windows, in the range of 115.6 to 116.8. The same was true for numeracy, where the range was 121.1 to 122.7.

However, for both literacy and numeracy, the pass rates in test windows 3 and 4 were lower than in test windows 1 and 2, primarily due to the revised standards. For literacy the overall pass rates dropped from 90.9% and 90.0% to 83.6% and 82.6%. For numeracy they dropped from 90.3% and 91.0% to 85.2% and 82.5%.

The performance of candidates by subscale was similar to the average performance of candidates on the whole test, except that the average performance on the 'calculator not available' subscale was consistently significantly lower than the overall numeracy score, indicating that the candidates found this subscale relatively more difficult.

Table 30: Candidate performance by test window, overall and by subscale

Component	Test Window	Whole test and subscale	Valid N	Mean	S.D.	Pass Rate
Literacy	TW1	Overall	4264	116.8	8.6	90.9
		Reading		116.9	9.3	
		Technical skills of writing		117.2	10.1	
	TW2	Overall	7508	116.4	8.7	90.0
		Reading		116.4	9.4	
		Technical skills of writing		116.6	10.3	
	TW3	Overall	6668	116.2	9.8	83.6
		Reading		116.3	10.0	
		Technical skills of writing		115.7	11.1	
	TW4	Overall	6766	115.6	9.6	82.6
		Reading		115.7	9.9	
		Technical skills of writing		115.2	11.0	
Numeracy	TW1	Overall	4286	121.9	11.7	90.3
		Number & algebra		121.9	13.0	
		Measurement & geometry		121.1	12.3	
		Statistics & probability		121.9	12.1	
		Calculator available		122.3	11.2	
		Calculator not available		116.7	13.4	
	TW2	Overall	7540	122.2	11.6	91.0
		Number & algebra		121.4	12.6	
		Measurement & geometry		121.4	11.8	
		Statistics & probability		122.7	12.1	
		Calculator available		122.4	11.3	
		Calculator not available		119.0	13.2	
	TW3	Overall	6662	122.5	12.2	85.2
		Number & algebra		121.6	13.4	
		Measurement & geometry		123.0	12.5	
		Statistics & probability		122.3	12.2	

		Calculator available		122.9	11.6	
		Calculator not available		119.7	15.2	
	TW4	Overall	6730	121.1	12.3	82.5
		Number & algebra		120.1	13.6	
		Measurement & geometry		121.7	12.8	
		Statistics & probability		121.1	12.2	
		Calculator available		121.5	11.7	
		Calculator not available		118.1	15.6	

Appendix 4 provides a detailed analysis of candidate performance by demographic characteristics and test windows. Candidate performance of any subgroup with a sample size less than 10 was not reported.

Appendix 5 provides a detailed analysis of performance by test centres and remote proctoring by test window. The distributions of candidate performance within a subgroup were similar across test windows.

Table 33 shows the proportion of candidates by test window who did not achieve the standard in 2017 after one or more attempts. For literacy, the percentages ranged from 9.1% in test window 1 to 17.4% in test window 4. For numeracy, the range was from 9.0% in test window 2 to 17.5% in test window 4. Again, the main reason for the increased proportion of candidates not achieving the standards in test windows 3 and 4 were the revised standards. However, it is also likely that approximately 2% of the increase is due to less able resit candidates and less confident candidates delaying their resits until later in the year, as was evident in 2016.

Table 31: Number and proportion of candidates who did not achieve the standard by test window

Test Window	Component	Number	% of Candidates
TW1	Literacy	389	9.1
	Numeracy	416	9.7
	Both ¹⁵	157	3.9
TW2	Literacy	749	10.0
	Numeracy	678	9.0
	Both	237	3.4
TW3	Literacy	1096	16.4
	Numeracy	989	14.8
	Both	369	6.1
TW4	Literacy	1179	17.4
	Numeracy	1180	17.5
	Both	500	8.4

4.2 Performance of Resit Candidates by Test Windows

Table 34 shows the performance of resit candidates overall and by subscale, and by test attempt, for each test window. In each test window, the overall mean scale scores of resit candidates who had a second attempt (resit 1) were close to the standard in both literacy (106 for test window 1 and 2 and 107 for test window 3 and 4) and numeracy (107 and 110).

For literacy, the candidates who sat the test for the third time (resit 2) had very similar overall mean scores to those candidates who sat the test for the second time except for test windows 3 and 4. In these test windows, the overall mean scale scores were 3 and 2 scale score points below the standard, indicating less able cohorts, and contributing to the lower pass rates (29% and 33%).

¹⁵ A subset of literacy and numeracy.

For numeracy, the overall mean scores of candidates who sat the numeracy component for the third time was nearly 2 scale score points higher than those candidates who sat the test for the second time in test window 1, but for test windows 3 and 4 the opposite was true. The overall mean score of these candidates in test windows 3 and 4 were not only below those for test windows 1 and 2 but also below the numeracy standard. As for 2016, it can be seen that the difference in mean scale scores for the 'calculator available' and the 'calculator not available' subscales was greater (4 to 6 scale score points) for the resit cohort than for the first-attempt candidates (3 scale score points).

Compared to 2016, where between 71% and 88% of third-attempt candidates achieved the literacy standard, in 2017 these rates declined to 57% and 62% in test windows 1 and 2 where the same standard applied. It is possible that this variation is due to the relatively small numbers of third-attempt candidates in both years.

In 2017, there were some fourth-attempt candidates; however, it is not possible to reach reliable conclusions about this cohort by test window because of the small numbers. It is worth noting that the mean scale score of the 21 candidates who attempted the numeracy component in test window 4 for the fourth time was very low (104.3 or 6 scale score points below the standard) and that only 4 of the 21 candidates achieved the standard.

It is not surprising that mean scores and pass rates should decrease as the number of resits increases. As the more able candidates progressively achieve the standard they leave the resit pool.

Table 32: Resit candidate performance by test window, overall and by subscale

Component	Test Window	Whole test and subscale	Second Attempt (Resit 1)			Third Attempt (Resit 2)			Fourth Attempt (Resit 3)		
			N	Mean	Pass Rate	N	Mean	Pass Rate	N	Mean	Pass Rate
Literacy	TW1	Overall	184	106.8	58.2	28	106.9	57.1			
		Reading		106.5			105.8				
		Technical skills of writing		107.5			110.0				
	TW2	Overall	274	106.9	62.4	55	107.1	61.8	10	106.4	50.0
		Reading		106.7			106.9			104.2	
		Technical skills of writing		107.1			108.4			111.3	
	TW3	Overall	434	105.5	45.9	63	104.1	28.6	7	-	-
		Reading		105.8			104.2			-	
		Technical skills of writing		105.0			103.8			-	
	TW4	Overall	596	106.8	53.7	150	104.9	33.3	18	106.7	38.9
		Reading		107.2			105.4			107.0	
		Technical skills of writing		106.1			103.5			105.2	
Numeracy	TW1	Overall	200	106.3	50.5	30	108.2	63.3	1	-	-
		Number & algebra		105.2			106.0			-	
		Measurement & geometry		106.0			109.9			-	
		Statistics & probability		109.0			110.1			-	
		Calculator available		107.6			109.1			-	
		Calculator not available		101.7			104.7			-	
	TW2	Overall	285	107.6	58.9	87	107.3	52.9	10	106.0	60.0
		Number & algebra		105.9			106.9			103.3	
		Measurement & geometry		108.8			107.3			107.0	
		Statistics & probability		109.7			108.5			110.0	
		Calculator available		108.6			108.0			106.6	
		Calculator not available		104.2			105.1			104.5	
	TW3	Overall	352	108.4	43.2	61	106.0	21.3	17	107.7	35.3
		Number & algebra		106.5			103.5			105.6	

		Measurement & geometry		109.9			108.3			108.4	
		Statistics & probability		109.8			108.0			110.9	
		Calculator available		109.8			107.3			108.8	
		Calculator not available		102.5			101.4			103.9	
	TW4	Overall		109.0	49.6		107.5	38.5		104.3	19.0
		Number & algebra		107.1			105.8			102.8	
		Measurement & geometry	528	110.0		161	108.3		21	105.8	
		Statistics & probability		110.6			109.1			105.2	
		Calculator available		110.2			108.6			105.0	
		Calculator not available		104.1			103.1			101.6	
		Overall		107.5			105.8			102.8	
		Number & algebra		108.3			109.1			105.2	
		Measurement & geometry		109.1			108.6			105.0	
		Statistics & probability		108.6			103.1			101.6	
		Calculator available		103.1							
		Calculator not available									

5. TRIAL ITEM ANALYSIS

5.1 In-test trialling

For each of the literacy and numeracy components, 160 literacy items and 165 numeracy items were trialled within the live tests. These items were placed in small clusters (5-item clusters for literacy, and 4-item 'calculator available' clusters and a 1-item 'calculator not available' cluster for numeracy) at the end of each test, requiring the creation of 32 tests for each component. Candidates were unaware of the location of these trial items. The trial items did not contribute to a candidate's score. The items were trialled in multiple test windows until sufficient candidates had attempted them. In this way, robust trial item estimates were obtained to enable selection of new, balanced clusters for refreshment of the tests in test window 3 and test window 4 in 2018.

5.2 Trial item analysis

Table 6 shows that the Phase 3 trial items were well-targeted by difficulty, with most items targeting Bands 2 and 3 and a small number targeting Bands 1 and above Band 3, as required by the test construct. A small number of literacy items were below Band 1, making them too easy to be of much use in refreshing the test or other uses, such as for the emergency test or as sample items.

Table 33: Distribution of Phase 3 trial items by Band

Achievable by candidates ...	Number of literacy items	Number of numeracy items
above Band 3	15	10
in Band 3: Clearly above the standard	45	54
in Band 2: At and above the standard	52	68
in Band 1: Below the standard	22	25
below Band 1	5	0
Total	139	157

5.3 Differential item functioning

During the item development and revision phase, avoiding items that might favour one subgroup of candidates over another is attempted. Despite this, it is normal for a proportion of items to show differential item functioning (DIF). DIF analysis is performed on all trial items. On many occasions, no obvious content or context bias is observable. Investigating reasons for a particular item showing DIF for a particular group involves looking for an explanatory connection between actual characteristics of the item and assumed or posited characteristics of the group.

It is not possible to withhold all items showing DIF from the live tests, so the approach is to attempt to 'balance' the tests accordingly and thereby minimise the likelihood of any test bias. This is achieved by spreading items showing DIF between test clusters. In this way, items showing DIF are spread across the tests and no candidate is required to attempt all items showing DIF.

Table 7 shows the number of Phase 3 items showing differential item functioning. There were an insufficient number (<50) of Indigenous candidates and international candidates for reliable DIF analysis for these two variables. For both literacy and numeracy, the DIF was well-balanced for most variables. For example, for the age variable, while 8 numeracy items favoured candidates aged 17–25 years, 9 favoured candidates aged 26+ years. The exception for numeracy was the gender variable where 15 of the 25 items favoured males. The exceptions for literacy were the age and residential area variables, where 11 of the 16 items favoured candidates aged 26+ years and 9 of the 14 items favoured candidates from provincial or remote areas. Items showing DIF are investigated for unfair content and where this is found to exist the items are not selected. Usually, this is not the case and the DIF is performance related; that is, the favoured subgroup is simply better at the skills being assessed for a variety of reasons. To minimise differential test functioning, DIF cancelling methodology is applied at the cluster formation stage. That is, items showing DIF are paired with items showing DIF in the opposite direction. In this way, clusters are well-balanced and the tests from which the clusters are created are fair.

Table 34: Differential item functioning

Variable	Favours	Number of literacy items	Number of numeracy items
Age	17–25 years	5	8
	26+ years	11	9
Course Category	Early childhood & primary	3	6
	Secondary	2	6
English as a First Language	Yes	6	4
	No	8	4
Gender	Female	5	10
	Male	6	15
Program Type	Postgraduate	6	5
	Undergraduate	7	5
Residential Area	Metropolitan areas	5	7
	Provincial areas or remote areas	9	6

6. DIFFERENTIAL TEST FUNCTIONING

When investigating differential test functioning for seven background variables, the difference in the expected mean raw test scores for candidates of the same ability (same scale score) was found to be small. This difference was less than or equal to 1 score point in the 10 literacy tests and the 8 numeracy tests for six of the seven variables: age (17–25 years compared to 26+ years), course category (early childhood & primary compared to secondary), gender (male compared to female), indigeneity (Indigenous compared to non-Indigenous), program type (undergraduate compared to postgraduate), and residential area (metropolitan areas compared to provincial & remote areas). These differences were not statistically significant. Only English as a first language compared to English not as a first language (and only for papers L51 and L52 for candidates with ability at the standard) showed a difference greater than 1 raw score point (Figure 21). Chi-square tests revealed that the ability differences (even the largest difference between English as a first language and English not as a first language candidates) for any raw score around the standard were not statistically significant.

The data suggest that it is possible that items based on continuous text are more likely to advantage English as a first language speakers (possibly because they have had far more experience of reading continuous text in English), and that items that assess syntax, grammar and punctuation are more likely to advantage non-English as a first language speakers (possibly because they have had more extensive formal instruction in English syntax, grammar and punctuation when learning English as a foreign language). Consequently, there is the potential to create bias in a given literacy test form if there is an imbalance between items based on continuous text and those assessing syntax, grammar and punctuation. Care must be taken to ensure that items from each of these two categories are evenly represented in any given test form to minimise the potential for bias. For a full discussion of the methodology and findings from the differential test functioning, see Appendix 7.

7. REVISION OF THE STANDARDS

For a background to the PIAAC equating methodology please see sections 6.1 to 6.3 in the 2016 technical report.

6.1 Summary of revision of the standards

The equating study conducted in 2016/17 found that for the standard to be equivalent to the 70th percentile of the Australian adult population the literacy standard needed to be raised from 106 to 107 and the numeracy standard from 107 to 110. The revision was implemented in the second half of 2017 for test windows 3 and 4.

Based upon 2016 data, modelling predicted that 91.3% of literacy candidates and 88.0% of numeracy candidates would have achieved the revised standards at their first attempt. The actual pass rates of first-attempt candidates during test windows 3 and 4 in 2017 were 88.9% for literacy and 89.8% for numeracy.

Similarly, for second-attempt candidates modelling predicted pass rates of 61.6% for literacy and 37.4% for numeracy. The actual pass rates of second-attempt candidates during test windows 3 and 4 in 2017 were 45.9% and 53.7% respectively for literacy and 43.2% and 49.6% respectively for numeracy.

6.2 Impact of revised standards on subgroups in 2017

This section presents the impact of the revised standards on the performance by demographic groups, including Aboriginal and Torres Strait Islander and English as a first language. Only the data from 2017 test windows 3 and 4, the period during which the revised standards applied, were used for this analysis.

In general, the most significant effect of the revised standards is on the less able candidates, those with abilities just above the provisional standard. The pass rates for first-attempt candidates using provisional and revised standards are shown in Table 31. Females were more affected than males by the revision of the standard for numeracy than for literacy (5.3% decline in female pass rate for numeracy compared to 2.9% decline in female pass rate for literacy). There were no discernible differences by age except for numeracy, where candidates aged 46 and above had an 8.1% lower pass rate when the revised standard was applied.

International students were more affected than other students for literacy (5.1% compared to 2.5%) but not for numeracy. Candidates for whom English was not their first language were slightly more affected than other candidates for both literacy and numeracy, but the differences in the effects were similar (1.8% for literacy and 1.5% for numeracy). For Indigenous candidates the effect was small, showing about a 1% greater effect than non-Indigenous for literacy. The reverse was true for numeracy, where the effect was about 1% less on Indigenous candidates.

The effect on candidates from different residential areas was similar except for the small number who gave an international residential postcode, where the change was 8.3% compared to 4–5% for other candidates.

The difference in the effect on the pass rates by program type was small but greater for undergraduates than postgraduates: 1.5% greater for literacy and 2% greater for numeracy.

For course category, candidates in early childhood courses were most affected in both literacy and numeracy, with pass rates reduced by 4.5% for literacy and 7.2% for numeracy. There was also a large effect on candidates in special education courses for numeracy, where the pass rate decreased by nearly 9%.

Table 35: Comparison of outcomes between the old and the revised standards for first attempt candidates (test windows 3 and 4 only)

Characteristic	Category	Literacy				Numeracy			
		N	Pass Rate (Old Standard)	Pass Rate (Revised Standard)	Difference	N	Pass Rate (Old Standard)	Pass Rate (Revised Standard)	Difference
Gender	Female	9066	88.4	85.5	-2.8	9129	90.5	85.2	-5.3
	Male	3096	92.8	90.6	-2.2	3119	96.3	94.4	-1.9
	Indeterminate/intersex	4	100.0	75.0	-25.0	4	100.0	100.0	0.0
Age	17-25	7461	88.5	85.3	-3.2	7534	92.1	87.4	-4.7
	26-30	1898	91.8	89.6	-2.2	1904	93.0	88.8	-4.1
	31-35	1011	91.4	89.8	-1.6	994	93.3	89.9	-3.3
	36-40	778	90.5	89.2	-1.3	794	91.4	88.4	-3.0
	41-45	575	90.4	89.0	-1.4	579	89.5	86.0	-3.5
	46+	443	88.5	86.5	-2.0	447	87.5	79.4	-8.1
International Students	No	11602	90.2	87.7	-2.5	11681	91.8	87.3	-4.5
	Yes	564	74.8	69.7	-5.1	571	95.4	91.9	-3.5
English as a First Language	Yes	10730	91.7	89.3	-2.4	10796	92.4	88.1	-4.3
	No	1436	72.9	68.7	-4.2	1456	89.0	83.2	-5.8
Indigenous	No	11741	89.5	86.9	-2.6	11826	92.1	87.6	-4.5
	Yes	188	83.5	79.8	-3.7	190	84.2	80.5	-3.7
	Not disclosed	237	91.6	87.8	-3.8	236	92.4	89.0	-3.4
Residential Area	Metropolitan areas	10268	89.1	86.4	-2.6	10335	91.8	87.5	-4.4
	Provincial areas	1806	91.6	88.8	-2.8	1823	92.9	88.0	-4.9
	Remote areas	62	90.3	88.7	-1.6	63	88.9	84.1	-4.8
	International	23	100.0	100.0	0.0	24	100.0	91.7	-8.3
Program Type	Invalid or Missing	7	100.0	100.0	0.0	7	100.0	100.0	0.0
	Undergraduate	7657	86.7	83.5	-3.2	7731	90.3	85.1	-5.2
Program Type by Year Level	Postgraduate	4509	94.1	92.4	-1.7	4521	94.9	91.7	-3.2
	Undergraduate first year	1020	75.2	70.8	-4.4	1055	84.0	78.3	-5.7
	Undergraduate second year	1136	86.2	83.5	-2.6	1153	91.0	85.3	-5.6
	Undergraduate third year	2822	89.4	86.1	-3.3	2826	92.6	87.5	-5.1
	Undergraduate fourth year	2075	89.7	86.8	-2.9	2089	90.7	85.6	-5.1
	Undergraduate fifth yr or above	184	89.1	87.5	-1.6	189	89.9	86.2	-3.7
	Undergraduate graduated	420	82.9	79.3	-3.6	419	86.6	82.1	-4.5
	Postgraduate first year	2505	95.3	93.5	-1.9	2501	96.0	93.6	-2.4
	Postgraduate second year	988	94.1	93.2	-0.9	988	94.7	91.7	-3.0
	Postgraduate third year	161	96.3	94.4	-1.9	166	92.8	87.3	-5.4

	Postgraduate fourth year	222	89.6	87.8	-1.8	224	89.3	83.5	-5.8
	Postgraduate fifth year or above	165	95.8	93.9	-1.8	166	95.8	91.6	-4.2
	Postgraduate graduated	468	88.5	85.9	-2.6	476	92.6	87.4	-5.3
Course Category	Early childhood	908	77.8	73.2	-4.5	922	81.1	74.0	-7.2
	Primary	5032	90.1	87.5	-2.6	5071	92.1	87.2	-4.9
	Secondary	4227	94.3	92.4	-2.0	4238	96.0	93.3	-2.8
	Special education	78	92.3	89.7	-2.6	80	92.5	83.8	-8.8
	Other	1921	82.5	79.0	-3.4	1941	88.1	82.5	-5.6

The effect of the revised standards on resit candidates was greater than for first-attempt candidates. For example, as shown in Table 38, for second-attempt candidates, pass rates declined on average 5% to 10% for literacy and 15% to 20% for numeracy across all subgroups. For third-attempt and fourth-attempt candidates, the declines in pass rates are even larger but group sizes are small and percentages become less meaningful. The analysis for third-attempt and fourth-attempt candidates can be found in Appendix 7.

Table 36: Comparison of outcomes between the old and the revised standards for second attempt candidates

Characteristic	Category	Literacy				Numeracy			
		N	Pass Rate (Old Standard)	Pass Rate (Revised Standard)	Difference	N	Pass Rate (Old Standard)	Pass Rate (Revised Standard)	Difference
Gender	Female	856	56.3	48.1	-8.2	793	61.8	46.3	-15.5
	Male	173	67.6	61.3	-6.4	87	63.2	54.0	-9.2
	Indeterminate/intersex	1	100.0	100.0	0.0	-	-	-	-
Age	17-25	721	60.1	51.5	-8.6	532	62.8	46.8	-16.0
	26-30	128	57.0	50.8	-6.3	157	66.2	52.2	-14.0
	31-35	63	54.0	49.2	-4.8	64	56.3	45.3	-10.9
	36-40	56	51.8	42.9	-8.9	39	64.1	46.2	-17.9
	41-45	32	50.0	43.8	-6.3	45	48.9	33.3	-15.6
	46+	30	50.0	46.7	-3.3	43	55.8	48.8	-7.0
International Students	No	830	61.2	53.7	-7.5	856	61.3	46.6	-14.7
	Yes	200	46.0	36.5	-9.5	24	83.3	62.5	-20.8
English as a First Language	Yes	659	65.6	58.1	-7.4	726	63.4	48.5	-14.9
	No	371	45.3	36.7	-8.6	154	55.2	40.3	-14.9
Indigenous	No	976	58.6	50.8	-7.8	842	62.0	47.4	-14.6
	Yes	26	53.8	42.3	-11.5	21	57.1	38.1	-19.0
	Not disclosed	28	50.0	42.9	-7.1	17	64.7	41.2	-23.5
Residential Area	Metropolitan areas	893	57.1	49.4	-7.7	741	61.5	46.4	-15.1

	Provincial areas	134	66.4	57.5	-9.0	131	66.4	51.9	-14.5
	Remote areas	2	0.0	0.0	0.0	5	20.0	20.0	0.0
	International	1	100.0	100.0	0.0	1	0.0	0.0	0.0
	Invalid or Missing	-	-	-	-	2	50.0	50.0	0.0
Program Type	Undergraduate	688	58.0	50.1	-7.8	577	60.0	45.8	-14.2
	Postgraduate	342	58.8	50.9	-7.9	303	65.7	49.5	-16.2
Program Type by Year Level	Undergraduate first year	136	41.2	37.5	-3.7	52	50.0	40.4	-9.6
	Undergraduate second year	57	63.2	56.1	-7.0	37	54.1	35.1	-18.9
	Undergraduate third year	131	64.1	54.2	-9.9	101	66.3	52.5	-13.9
	Undergraduate fourth year	252	63.9	55.2	-8.7	274	61.7	46.0	-15.7
	Undergraduate fifth yr or above	29	65.5	51.7	-13.8	41	43.9	36.6	-7.3
	Undergraduate graduated	83	51.8	44.6	-7.2	72	63.9	50.0	-13.9
	Postgraduate first year	153	62.1	52.3	-9.8	92	75.0	62.0	-13.0
	Postgraduate second year	55	54.5	54.5	0.0	65	64.6	44.6	-20.0
	Postgraduate third year	11	63.6	54.5	-9.1	12	83.3	66.7	-16.7
	Postgraduate fourth year	25	68.0	60.0	-8.0	29	65.5	55.2	-10.3
	Postgraduate fifth year or above	18	66.7	50.0	-16.7	23	65.2	21.7	-43.5
	Postgraduate graduated	80	50.0	42.5	-7.5	82	53.7	42.7	-11.0
Course Category	Early childhood	114	50.0	40.4	-9.6	116	55.2	40.5	-14.7
	Primary	365	61.9	53.2	-8.8	379	63.3	49.1	-14.2
	Secondary	309	61.8	55.0	-6.8	207	62.8	48.3	-14.5
	Special education	3	33.3	33.3	0.0	4	100.0	75.0	-25.0
	Other	239	52.3	45.2	-7.1	174	61.5	44.8	-16.7

8. RELEASED ITEMS

During 2017, items were released to illustrate the Described Proficiency Scale (DPS) and provide a 30-item practice test. The items were chosen to represent the test in difficulty, content and item type. The purpose of the DPS was to illustrate each of the three bands referred to on candidate reports and to help candidates below the standard (Band 1) understand the skills required to achieve the standard (Band 2 and above).

The 30-item practice test was released before test window 3 in 2017. Its purpose was to supplement the sample questions provided in 2016. Online and accessible versions of the DPS and the practice test were also created.

The released items were selected from retired test items and trial items not used to refresh the test.

9. PHASE 4 TEST DEVELOPMENT

During 2017, 105 Phase 4 literacy items and 105 Phase 4 numeracy items mapped against the Assessment Framework were written by ACER. The items are to be reviewed in February 2018 and revised based upon reviewers' feedback. A small proportion will be retired. A selection of at least 60 literacy items and at least 60 numeracy items will be in-test trialled in test windows 3 and 4 of 2018 and test window 1 of 2019. A selection of these will be used to refresh the test in 2019 and to create a secure emergency test.

10. CONCLUSION

The test was successfully administered in four test windows in all Australian states and territories in metropolitan and regional testing centres and by remote proctoring to over 24 000 candidates. Another 160 new items were successfully trialled enabling the test to be refreshed and a practice test to be released.

A successful equating exercise was conducted enabling the provisional standards to be mapped onto the PIACC scales, showing that small revisions (an increase of 1 scale score point for literacy and an increase of 3 scale score points for numeracy) were required. The revised standards were reviewed and accepted by expert groups and put into effect for test windows 3 and 4.

A described proficiency scale, describing the three bands on candidates' reports and illustrated with released items, was provided to support candidates understanding of the skills required around the standards.

Differential test functioning revealed no significant bias against any of the seven subgroups investigated; however, useful patterns were found relating to differential item functioning by subgroup to inform future test construction.

At the end of 2017, after second, third and fourth attempts were accounted for, 91.6% of candidates had achieved the literacy standard and 91.6% of candidates had achieved the numeracy standard. Over the two years of testing, 93.9% of candidates had achieved the literacy standard and 93.8% of candidates had achieved the numeracy standard.

11. REFERENCES

- Adams, R.J., Wu, M.L., and Wilson, M.R. (2015). *ConQuest 4*. [computer program] Camberwell: Australian Council for Educational Research.
- Monseur, C. and Bereznier, A. (2007). The Computation of Equating Errors in International Surveys in Education. *Journal of Applied Measurement*, 8(3), 323–335.
- Rasch, G (1980). *Probabilistic Models for Some Intelligence and Attainment Tests*. Chicago: MESA Press (original work published 1960).

Stocking, M. L. & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7(2), 201–210.

12. APPENDICES

Appendix 1: Proportion of candidates by test centre and by attempt

Table 39 shows the number and proportion of candidates participating at each test centre and by remote proctoring. In 2017, 45% of first-attempt candidates sat the test at three test centres: Melbourne CBD, Sydney CBD and Brisbane CBD, compared to 76% in 2016. The main reason for the smaller percentage of candidates at the three CBD testing centres was because remote proctoring was the only medium available in test window 1 in 2017.

Table 37: Number and proportion of candidates who participated by test centre

Test Centre	First Attempt				Second Attempt				Third Attempt				Fourth Attempt			
	Literacy		Numeracy		Literacy		Numeracy		Literacy		Numeracy		Literacy		Numeracy	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Adelaide CBD	349	1.5	349	1.5	8	0.5	5	0.4	2	0.7	0	0.0	0	0.0	0	0.0
Albury	111	0.5	110	0.5	8	0.5	6	0.4	1	0.3	1	0.3	0	0.0	0	0.0
Armidale (NSW)	108	0.5	106	0.5	6	0.4	8	0.6	4	1.4	2	0.6	0	0.0	1	2.0
Ballarat	231	1.0	234	1.0	22	1.5	21	1.5	3	1.0	3	0.9	0	0.0	1	2.0
Bathurst	52	0.2	57	0.2	6	0.4	4	0.3	0	0.0	1	0.3	0	0.0	0	0.0
Brisbane CBD	1662	7.1	1656	7.1	52	3.5	52	3.8	9	3.0	14	4.1	1	2.9	2	4.1
Bundaberg	55	0.2	56	0.2	4	0.3	4	0.3	0	0.0	1	0.3	0	0.0	0	0.0
Cairns	57	0.2	56	0.2	1	0.1	4	0.3	0	0.0	0	0.0	0	0.0	0	0.0
Canberra CBD	211	0.9	208	0.9	5	0.3	9	0.7	0	0.0	1	0.3	0	0.0	0	0.0
Darwin	32	0.1	32	0.1	2	0.1	1	0.1	0	0.0	0	0.0	0	0.0	0	0.0
Gold Coast	99	0.4	99	0.4	5	0.3	6	0.4	1	0.3	1	0.3	0	0.0	0	0.0
Hobart	50	0.2	49	0.2	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Melbourne CBD	5380	23.0	5450	23.2	354	23.8	290	21.2	58	19.6	57	16.8	6	17.1	11	22.4
Mildura	36	0.2	35	0.1	1	0.1	0	0.0	2	0.7	4	1.2	0	0.0	0	0.0
Newcastle	452	1.9	456	1.9	23	1.5	20	1.5	4	1.4	3	0.9	2	5.7	1	2.0
Parramatta	341	1.5	339	1.4	29	1.9	32	2.3	9	3.0	6	1.8	0	0.0	0	0.0
Perth CBD	976	4.2	980	4.2	23	1.5	18	1.3	3	1.0	5	1.5	0	0.0	0	0.0
Remote Proctoring	9486	40.6	9462	40.3	737	49.5	706	51.7	165	55.7	196	57.8	22	62.9	30	61.2
STC	3	0.0	2	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Sunshine Coast/Mar	134	0.6	133	0.6	6	0.4	6	0.4	0	0.0	0	0.0	0	0.0	0	0.0
Sydney CBD	3183	13.6	3215	13.7	184	12.4	163	11.9	30	10.1	40	11.8	4	11.4	3	6.1
Townsville	75	0.3	76	0.3	0	0.0	1	0.1	0	0.0	0	0.0	0	0.0	0	0.0
Wagga Wagga	94	0.4	94	0.4	4	0.3	3	0.2	2	0.7	1	0.3	0	0.0	0	0.0
Warrnambool	39	0.2	36	0.2	0	0.0	2	0.1	0	0.0	1	0.3	0	0.0	0	0.0
Wollongong	171	0.7	176	0.8	8	0.5	4	0.3	3	1.0	2	0.6	0	0.0	0	0.0
Total	23387	100.0	23466	100.0	1488	100.0	1365	100.0	296	100.0	339	100.0	35	100.0	49	100.0

Appendix 2: Proportion of candidates by test centre and by test window

Table 40 and Table 41 show the number and proportion of candidates participating in each test centre in test windows 1–4 for literacy and for numeracy respectively. It can be seen that some test centres in some locations were not available for all test windows.

Table 38: Number of candidates in test centres by test windows – literacy

Test Centre	TW1		TW2		TW3		TW4	
	N	%	N	%	N	%	N	%
Adelaide CBD	NA ¹⁶	NA	89	1.2	87	1.3	183	2.7
Albury	NA	NA	39	0.5	43	0.6	38	0.6
Armidale (NSW)	NA	NA	38	0.5	40	0.6	40	0.6
Ballarat	NA	NA	82	1.1	93	1.4	81	1.2
Bathurst	NA	NA	16	0.2	23	0.3	19	0.3
Brisbane CBD	NA	NA	657	8.8	542	8.1	525	7.8
Bundaberg	NA	NA	19	0.3	21	0.3	19	0.3
Cairns	NA	NA	NA	NA	19	0.3	39	0.6
Canberra CBD	NA	NA	60	0.8	78	1.2	78	1.2
Darwin	NA	NA	14	0.2	10	0.1	10	0.1
Gold Coast	NA	NA	NA	NA	47	0.7	58	0.9
Hobart	NA	NA	18	0.2	9	0.1	23	0.3
Melbourne CBD	NA	NA	1826	24.3	2039	30.6	1933	28.6
Mildura	NA	NA	20	0.3	NA	NA	19	0.3
Newcastle	NA	NA	156	2.1	171	2.6	154	2.3
Parramatta	NA	NA	105	1.4	143	2.1	131	1.9
Perth CBD	NA	NA	310	4.1	433	6.5	259	3.8
Remote Proctoring	4264	100.0	2631	35.0	1710	25.6	1805	26.7
STC	NA	NA	1	0.0	1	0.0	1	0.0
Sunshine Coast / Maroochy	NA	NA	49	0.7	43	0.6	48	0.7
Sydney CBD	NA	NA	1229	16.4	969	14.5	1203	17.8
Townsville	NA	NA	34	0.5	41	0.6	NA	NA
Wagga Wagga	NA	NA	38	0.5	42	0.6	20	0.3
Warrnambool	NA	NA	19	0.3	NA	NA	20	0.3
Wollongong	NA	NA	58	0.8	64	1.0	60	0.9
Total	4264	100.0	7508	100.0	6668	100.0	6766	100.0

¹⁶ NA indicates that the test centre was not available for that particular test window

Table 39: Number of candidates in test centres by test windows – numeracy

Test Centre	TW1		TW2		TW3		TW4	
	N	%	N	%	N	%	N	%
Adelaide CBD	NA ¹⁷	NA	90	1.2	80	1.2	184	2.7
Albury	NA	NA	38	0.5	42	0.6	37	0.5
Armidale (NSW)	NA	NA	37	0.5	40	0.6	40	0.6
Ballarat	NA	NA	85	1.1	92	1.4	82	1.2
Bathurst	NA	NA	19	0.3	23	0.3	20	0.3
Brisbane CBD	NA	NA	655	8.7	542	8.1	527	7.8
Bundaberg	NA	NA	19	0.3	23	0.3	19	0.3
Cairns	NA	NA	NA	NA	21	0.3	39	0.6
Canberra CBD	NA	NA	60	0.8	81	1.2	77	1.1
Darwin	NA	NA	14	0.2	10	0.2	9	0.1
Gold Coast	NA	NA	NA	NA	49	0.7	57	0.8
Hobart	NA	NA	17	0.2	9	0.1	23	0.3
Melbourne CBD	NA	NA	1826	24.2	2049	30.8	1933	28.7
Mildura	NA	NA	20	0.3	NA	NA	19	0.3
Newcastle	NA	NA	155	2.1	171	2.6	154	2.3
Parramatta	NA	NA	100	1.3	144	2.2	133	2.0
Perth CBD	NA	NA	311	4.1	433	6.5	259	3.8
Remote Proctoring	4287	100.0	2637	35.0	1695	25.4	1775	26.4
STC	NA	NA	1	0.0	1	0.0	NA	NA
Sunshine Coast / Maroochy	NA	NA	49	0.6	42	0.6	48	0.7
Sydney CBD	NA	NA	1255	16.6	969	14.5	1197	17.8
Townsville	NA	NA	36	0.5	41	0.6	NA	NA
Wagga Wagga	NA	NA	38	0.5	42	0.6	18	0.3
Warrnambool	NA	NA	19	0.3	NA	NA	20	0.3
Wollongong	NA	NA	59	0.8	63	0.9	60	0.9
Total	4287	100.0	7540	100.0	6662	100.0	6730	100.0

¹⁷ NA indicates that the test centre was not available for that particular test window

Appendix 3: Score Frequency Distribution

Table 40: Literacy score frequency distribution of first-attempt candidates

Scale Score	Frequency	Percentile
46	1	0.0
49	1	0.0
81	1	0.0
84	1	0.0
87	1	0.0
88	5	0.0
89	3	0.1
90	9	0.1
91	4	0.1
92	13	0.2
93	21	0.3
94	24	0.4
95	33	0.5
96	46	0.7
97	60	1.0
98	97	1.4
99	103	1.8
100	149	2.4
101	198	3.3
102	245	4.3
103	349	5.8
104	397	7.5
105	439	9.4
106	527	11.7
107	553	14.0
108	700	17.0
109	822	20.5
110	833	24.1
111	912	28.0
112	911	31.9
113	954	36.0
114	1088	40.6
115	1128	45.4
116	1074	50.0
117	829	53.6
118	1022	58.0
119	980	62.1
120	1261	67.5
121	487	69.6
122	934	73.6
123	1082	78.2
124	502	80.4
125	801	83.8
126	348	85.3
127	688	88.2
128	326	89.6
129	560	92.0
130	211	92.9
131	432	94.8
132	158	95.5
133	104	95.9
134	339	97.3
135	64	97.6
136	64	97.9
137	64	98.2
138	188	99.0
139	9	99.0
141	37	99.2
142	49	99.4
143	44	99.6
144	49	99.8
153	32	99.9
154	7	99.9
155	14	100.0

Standard in 2017 TW1–2
Standard in 2017 TW3–4

Table 41: Numeracy score frequency distribution of first-attempt candidates

Scale Score	Frequency	Percentile	
45	1	0.0	
57	1	0.0	
78	1	0.0	
79	1	0.0	
80	1	0.0	
83	3	0.0	
84	2	0.0	
85	3	0.1	
86	5	0.1	
87	3	0.1	
88	6	0.1	
89	4	0.1	
90	13	0.2	
91	19	0.3	
92	20	0.4	
93	30	0.5	
94	29	0.6	
95	39	0.8	
96	62	1.0	
97	64	1.3	
98	77	1.6	
99	88	2.0	
100	138	2.6	
101	124	3.1	
102	153	3.8	
103	172	4.5	
104	273	5.7	
105	209	6.6	
106	260	7.7	
107	295	8.9	Standard in 2017 TW1-2
108	374	10.5	
109	366	12.1	
110	417	13.9	Standard in 2017 TW3-4
111	502	16.0	
112	498	18.1	
113	426	19.9	
114	640	22.7	
115	615	25.3	
116	665	28.1	
117	782	31.5	
118	698	34.4	
119	847	38.0	
120	985	42.2	
121	270	43.4	
122	1058	47.9	
123	1131	52.7	
124	212	53.6	
125	1303	59.2	
126	812	62.6	
127	715	65.7	
128	974	69.8	
129	321	71.2	
130	1316	76.8	
131	12	76.9	
132	877	80.6	
133	740	83.8	
135	831	87.3	
136	621	89.9	
137	12	90.0	
139	782	93.3	
140	359	94.9	
141	12	94.9	
144	79	95.2	
145	638	98.0	
146	107	98.4	
147	4	98.4	
156	174	99.2	
157	159	99.9	
158	35	100.0	

Table 44 shows the percentage of candidates who sat the test in 2017 in each of the three bands for both literacy and numeracy. For literacy, nearly three-fifths or 60% of the candidates were located in Band 2: At or above the standard, while approximately one-third or 30% were located in Band 3: Well above the standard or above Band 3. For numeracy, approximately two-fifths or 40% were located in Band 2: At or above the standard, while just under half or 47% were located in Band 3: Well above the standard or above Band 3.

Table 42: Candidates attempting the test in 2017 by Band achievement

Component	Year of Registration	No. of Unique Candidates	Below Band 1 (%)	Band 1 (%)	Band 2 (%)	Band 3 (%)	Above Band 3 (%)
Literacy	2017	23387	0.1	8.8	58.6	29.8	2.7
	2017 plus the 2016 resitters	23795	0.1	9.3	58.6	29.4	2.6
Numeracy	2017	23466	0.6	9.3	42.8	37.3	10.0
	2017 plus the 2016 resitters	23959	0.7	10.2	42.7	36.6	9.8

Appendix 4: Performance by Demographic Characteristics and Test Windows

Table 45 to Table 48 show performance by demographic characteristics for each test window. It can be seen that differences between the mean scaled score of each group are consistent across test windows. In general, the overall findings in Section 3: Candidate Performance are also true for each test window.

Mostly, there are no major differences in the achievement of candidates between test windows. One exception was test window 3 where, for literacy, the mean scale score of female candidates was more than two scale score points lower than male candidates, and the pass rate of female candidates was 7% lower than that of male candidates. Another exception was for numeracy, where the pass rates of male candidates were 12% higher than the pass rates of female candidates in test windows 3 and 4.

Table 43: Performance by demographic characteristics in test window 1

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	3245	116.4	8.8	89.8	3299	120.2	11.4	88.4
	Male	1016	118.0	7.9	94.3	984	127.7	11.1	97.0
	Indeterminate/intersex	3	-	-	-	3	-	-	-
Age	17–25	2620	115.8	7.8	90.6	2628	121.7	11.2	90.9
	26–30	637	117.1	9.4	90.9	641	122.3	12.6	89.9
	31–35	341	118.4	9.0	92.4	348	123.2	11.3	92.5
	36–40	287	119.7	9.6	91.6	295	122.9	13.0	89.5
	41–45	214	119.5	9.5	92.5	213	122.2	13.0	88.7
	46+	165	119.7	10.1	89.1	161	120.5	13.9	82.0
International Students	No	4153	116.9	8.5	91.2	4183	121.9	11.7	90.3
	Yes	111	113.5	9.3	79.3	103	121.9	11.9	89.3
English as a First Language	Yes	3835	117.2	8.4	92.4	3874	122.1	11.6	90.8
	No	429	112.8	9.1	77.2	412	120.4	12.6	85.4
Indigenous	No	4078	116.9	8.6	91.1	4104	122.1	11.7	90.6
	Yes	86	112.5	7.3	84.9	87	115.7	12.3	83.9
	Not disclosed	100	115.5	8.7	87.0	95	120.7	13.2	82.1
Residential Area	Metropolitan areas	3533	116.9	8.5	91.2	3576	122.1	11.8	90.4
	Provincial areas	681	116.3	9.0	88.7	659	121.1	11.7	90.1
	Remote areas	31	119.7	7.3	96.8	32	124.4	11.2	90.6
	International	14	121.6	6.1	100.0	14	122.1	9.7	85.7
	Invalid or Missing	5	-	-	-	5	-	-	-
Program Type	Undergraduate	2702	115.5	8.0	89.5	2727	120.4	11.2	89.2
	Postgraduate	1562	119.0	9.1	93.2	1559	124.6	12.3	92.3
Program Type by Year Level	Undergraduate first year	95	114.3	8.2	86.3	92	121.8	11.1	90.2
	Undergraduate second year	94	114.1	9.3	80.9	96	120.7	12.7	86.5
	Undergraduate third year	340	116.0	8.0	90.3	345	121.5	10.7	91.9
	Undergraduate fourth year	1784	115.9	7.8	91.3	1802	120.7	11.1	89.8
	Undergraduate fifth yr or above	189	115.8	8.7	87.3	192	119.4	11.2	88.0
	Undergraduate graduated	200	112.3	7.7	80.5	200	116.4	10.7	81.0
	Postgraduate first year	532	120.1	8.4	95.7	527	127.3	11.5	96.0
	Postgraduate second year	590	119.9	9.4	94.6	591	124.8	11.1	93.9
	Postgraduate third year	51	118.4	9.9	90.2	50	124.3	14.9	86.0
	Postgraduate fourth year	86	116.9	10.1	88.4	90	119.5	12.3	82.2
	Postgraduate fifth yr or above	105	118.6	8.4	96.2	110	123.6	12.9	90.9
	Postgraduate graduated	198	114.7	8.6	83.8	191	119.5	14.1	84.3
Course Category	Early childhood	251	113.0	8.9	79.3	249	115.7	10.9	82.3
	Primary	1726	116.2	8.2	91.1	1748	120.6	11.2	89.3
	Secondary	1738	118.7	8.4	94.0	1728	125.2	11.7	94.3
	Special education	27	116.0	8.2	88.9	23	121.2	11.1	91.3
	Other	522	114.3	8.8	85.6	538	118.7	11.2	84.6

Table 44: Performance by demographic characteristics in test window 2 (including resits)

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	5635	115.9	8.6	89.0	5684	120.6	11.4	89.2
	Male	1868	118.1	8.8	93.1	1852	127.1	11.1	96.5
	Indeterminate/intersex	5	-	-	-	4	-	-	-
Age	17–25	4857	115.3	8.1	89.1	4872	121.5	11.2	90.8
	26–30	1114	117.7	9.3	91.4	1104	123.0	12.3	89.9
	31–35	537	118.5	9.4	91.8	540	123.9	12.4	92.4
	36–40	443	119.0	9.4	91.9	443	123.9	11.9	93.7
	41–45	299	120.3	8.8	94.3	310	124.0	11.7	92.9
	46+	258	119.1	9.5	90.3	271	122.8	13.3	89.7
International Students	No	6860	116.9	8.6	91.5	6897	121.9	11.8	90.4
	Yes	648	111.0	8.3	74.5	643	125.1	9.7	97.2
English as a First Language	Yes	6328	117.3	8.5	92.8	6376	122.3	11.6	91.2
	No	1180	111.3	8.2	75.3	1164	122.0	11.7	89.9
Indigenous	No	7253	116.5	8.7	90.3	7285	122.3	11.6	91.1
	Yes	100	111.7	7.7	77.0	95	115.5	11.6	81.1
	Not disclosed	155	116.3	9.7	85.2	160	123.5	11.7	91.9
Residential Area	Metropolitan areas	6435	116.4	8.8	89.8	6470	122.4	11.7	91.0
	Provincial areas	1018	116.4	8.1	91.7	1015	120.8	11.0	91.4
	Remote areas	28	116.2	8.8	85.7	27	117.2	12.7	70.4
	International	18	125.9	7.4	100.0	19	128.7	12.5	100
	Invalid or Missing	9	-	-	-	9	-	-	-
Program Type	Undergraduate	4052	114.7	8.0	88.2	4055	120.0	11.0	88.9
	Postgraduate	3456	118.4	9.1	92.2	3485	124.7	11.8	93.5
Program Type by Year Level	Undergraduate first year	596	113.0	8.5	81.2	593	122.6	10.8	92.7
	Undergraduate second year	215	115.4	9.0	86.5	214	121.9	11.8	89.7
	Undergraduate third year	608	114.8	7.3	90.5	589	121.1	10.1	93.5
	Undergraduate fourth year	2105	115.4	7.8	90.6	2112	119.6	10.8	88.3
	Undergraduate fifth yr or above	258	116.3	8.2	93.8	268	119.8	11.7	87.7
	Undergraduate graduated	270	111.4	7.7	75.9	279	114.7	12.0	76.0
	Postgraduate first year	1839	118.7	8.9	93.6	1846	126.4	11.3	96.0
	Postgraduate second year	860	119.6	9.1	93.5	870	124.8	11.9	93.0
	Postgraduate third year	86	120.3	9.0	91.9	88	121.4	11.6	89.8
	Postgraduate fourth year	175	116.5	8.7	90.9	173	121.5	10.7	91.9
	Postgraduate fifth yr or above	191	117.5	8.7	91.1	197	122.9	12.0	89.8
	Postgraduate graduated	305	113.8	8.9	81.0	311	118.4	12.0	84.2
Course Category	Early childhood	467	113.0	8.5	82.2	453	116.5	11.9	77.7
	Primary	2681	115.7	8.3	89.5	2723	120.3	11.0	89.9
	Secondary	3224	118.1	8.8	92.6	3217	125.2	11.4	94.4
	Special education	46	115.9	7.8	93.5	44	122.5	11.6	95.5
	Other	1090	114.5	8.5	86.9	1103	120.8	11.5	89.0

Table 45: Performance by demographic characteristics in test window 3 (including resits)

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	4946	115.6	9.8	81.8	4977	120.8	11.9	82.0
	Male	1719	117.9	9.5	88.8	1683	127.7	11.3	94.4
	Indeterminate/intersex	3	-	-	-	2	-	-	-
Age	17-25	4254	114.9	9.1	82.2	4236	122.0	11.6	85.4
	26-30	1009	118.1	10.1	86.3	1031	123.8	13.0	85.3
	31-35	500	119.0	10.3	86.6	492	124.4	13.2	85.8
	36-40	401	118.4	10.7	85.8	394	124.0	12.3	88.1
	41-45	279	118.8	11.2	85.3	277	122.9	12.9	84.8
	46+	225	117.9	11.2	84.0	232	120.5	14.0	75.0
International Students	No	6201	116.7	9.6	85.4	6284	122.4	12.2	84.7
	Yes	467	109.5	9.4	59.3	378	124.6	10.4	92.9
English as a First Language	Yes	5686	117.3	9.5	87.4	5794	122.7	12.1	85.7
	No	982	109.6	8.7	61.2	868	121.3	12.5	81.7
Indigenous	No	6429	116.2	9.8	83.8	6431	122.6	12.2	85.2
	Yes	98	112.7	8.6	72.4	95	119.0	11.5	78.9
	Not disclosed	141	116.3	10.8	78.7	136	124.0	12.2	86.0
Residential Area	Metropolitan areas	5698	116.1	9.8	83.3	5671	122.6	12.2	85.0
	Provincial areas	926	116.5	9.5	84.8	943	122.1	11.8	86.2
	Remote areas	29	119.9	10.5	86.2	32	123.2	14.5	78.1
	International	11	118.8	7.3	100.0	11	124.8	10.9	81.8
	Invalid or Missing	4	-	-	-	5	-	-	-
Program Type	Undergraduate	3996	114.3	9.0	80.4	4025	120.7	11.7	82.6
	Postgraduate	2672	119.0	10.2	88.3	2637	125.4	12.3	89.0
Program Type by Year Level	Undergraduate first year	515	111.2	9.9	65.8	490	120.5	12.0	82.7
	Undergraduate second year	479	115.3	9.7	82.3	491	122.8	12.3	85.7
	Undergraduate third year	1214	114.9	8.2	83.9	1217	121.4	10.7	87.1
	Undergraduate fourth year	1429	114.7	8.8	82.7	1460	120.0	11.9	79.9
	Undergraduate fifth yr or above	137	115.7	9.5	83.9	150	119.6	13.2	75.3
	Undergraduate graduated	222	112.4	9.2	74.3	217	117.4	11.6	73.7
	Postgraduate first year	1435	119.5	9.9	90.4	1393	127.0	11.7	93.1
	Postgraduate second year	597	121.0	10.1	91.3	600	126.1	12.7	89.3
	Postgraduate third year	80	117.3	10.5	85.0	80	121.1	13.2	78.8
	Postgraduate fourth year	138	117.2	9.9	82.6	145	121.5	13.1	77.2
	Postgraduate fifth yr or above	114	119.4	9.7	88.6	111	124.4	11.5	84.7
	Postgraduate graduated	308	114.2	9.9	76.0	308	120.1	11.9	79.9
Course Category	Early childhood	471	112.4	9.0	72.0	485	117.0	11.6	74.0
	Primary	2658	115.9	9.3	84.5	2686	121.2	11.6	83.5
	Secondary	2458	118.5	9.8	88.6	2420	126.2	12.1	90.8
	Special education	39	115.6	7.9	87.2	43	121.7	10.1	93.0
	Other	1042	113.1	9.8	74.6	1028	120.0	11.6	81.0

Table 46: Performance by demographic characteristics in test window 4 (including resits)

Characteristic	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
Gender	Female	5178	115.0	9.6	80.8	5186	119.4	11.9	79.8
	Male	1585	117.4	9.3	88.3	1542	126.7	12.1	91.4
	Indeterminate/intersex	3	-	-	-	2	-	-	-
Age	17–25	4089	114.1	8.7	80.6	3995	120.4	11.7	82.0
	26–30	1063	117.3	9.9	85.5	1080	122.2	12.7	84.2
	31–35	582	117.8	9.9	87.3	582	122.8	12.2	87.5
	36–40	446	118.0	10.7	85.2	445	122.7	13.0	84.3
	41–45	334	118.9	11.4	86.2	360	121.9	14.5	77.8
	46+	252	118.0	11.3	82.5	268	120.0	14.1	75.7
International Students	No	6426	115.9	9.5	83.7	6509	121.0	12.3	82.3
	Yes	340	109.4	9.0	60.6	221	123.1	11.8	86.9
English as a First Language	Yes	5853	116.5	9.3	86.1	5944	121.4	12.2	83.5
	No	913	109.7	9.4	59.7	786	118.9	13.1	74.3
Indigenous	No	6515	115.6	9.6	82.7	6482	121.2	12.3	82.7
	Yes	123	114.0	9.2	74.8	127	116.3	11.1	70.1
	Not disclosed	128	116.9	9.9	85.9	121	121.2	12.4	82.6
Residential Area	Metropolitan areas	5661	115.5	9.7	81.9	5624	121.1	12.3	82.4
	Provincial areas	1053	115.9	9.0	86.1	1049	121.1	12.2	83.0
	Remote areas	35	117.0	9.3	85.7	39	121.1	14.2	76.9
	International	13	126.2	8.6	100.0	14	131.1	12.5	92.9
	Invalid or Missing	4	-	-	-	4	-	-	-
Program Type	Undergraduate	4500	113.8	8.7	79.4	4463	119.3	11.7	80.1
	Postgraduate	2266	119.1	10.2	88.9	2267	124.6	12.8	87.2
Program Type by Year Level	Undergraduate first year	662	110.8	9.2	66.5	621	116.6	11.9	71.7
	Undergraduate second year	716	114.5	8.7	82.0	702	120.2	11.9	82.3
	Undergraduate third year	1751	114.8	8.0	84.9	1717	121.1	11.2	85.5
	Undergraduate fourth year	974	114.3	8.9	80.3	1004	118.6	11.8	77.4
	Undergraduate fifth yr or above	85	114.9	9.6	75.3	103	118.1	13.3	70.9
	Undergraduate graduated	312	111.6	8.8	68.3	316	115.9	10.4	73.4
	Postgraduate first year	1245	119.9	10.0	91.2	1208	126.8	12.5	91.3
	Postgraduate second year	458	119.2	10.1	89.3	472	123.4	12.6	86.2
	Postgraduate third year	97	120.6	9.4	93.8	102	124.6	12.6	89.2
	Postgraduate fourth year	113	117.2	9.9	87.6	116	120.3	11.1	80.2
	Postgraduate fifth yr or above	76	117.1	11.1	84.2	82	120.6	13.0	76.8
	Postgraduate graduated	277	115.9	10.8	77.6	287	119.9	12.8	76.7
Course Category	Early childhood	573	111.0	9.1	65.8	584	114.9	12.1	65.4
	Primary	2833	115.7	9.2	84.2	2877	120.7	11.5	83.6
	Secondary	2141	118.4	9.5	89.9	2092	125.2	12.5	89.8
	Special education	45	114.3	8.0	84.4	41	119.5	12.1	73.2
	Other	1174	112.4	9.0	73.5	1136	117.9	11.8	75.4

Figure 7 to Figure 11 show achievement distributions in literacy and numeracy by demographic characteristics for each test window.

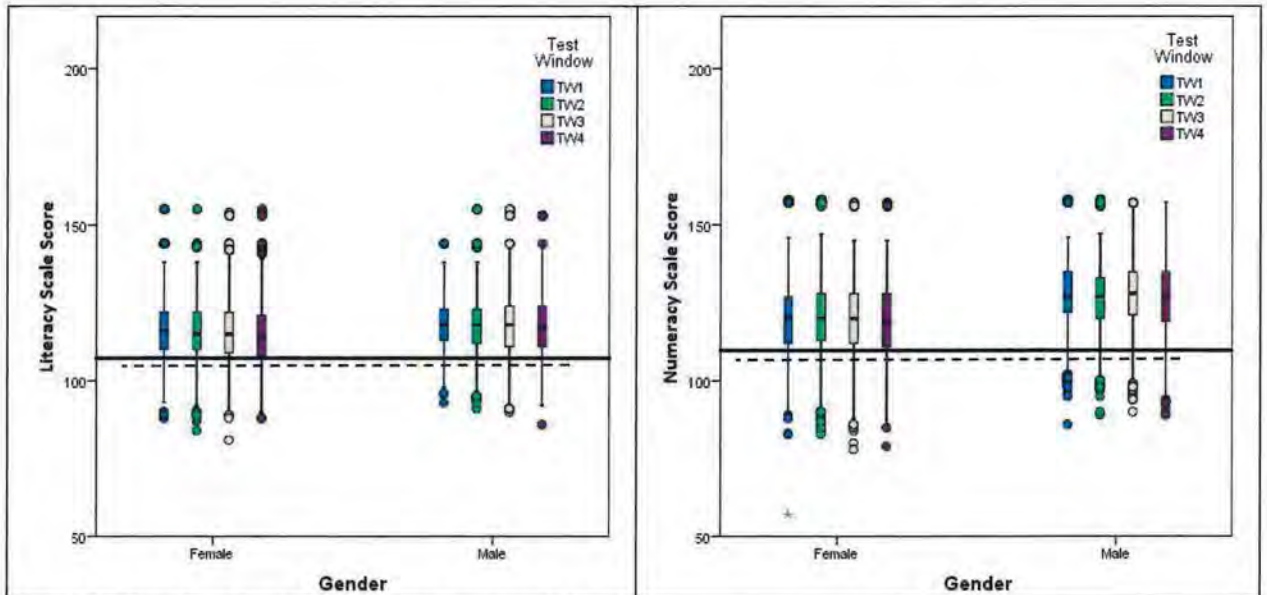


Figure 7: Score distribution by gender and test window

Figure 7 shows that, for literacy, the distributions of the scale scores of female candidates and male candidates are very similar. For numeracy, while the distributions of the scale scores of male candidates were higher up the scale than those of female candidates, it can be seen that in each test window many female candidates achieved well above the numeracy standard.

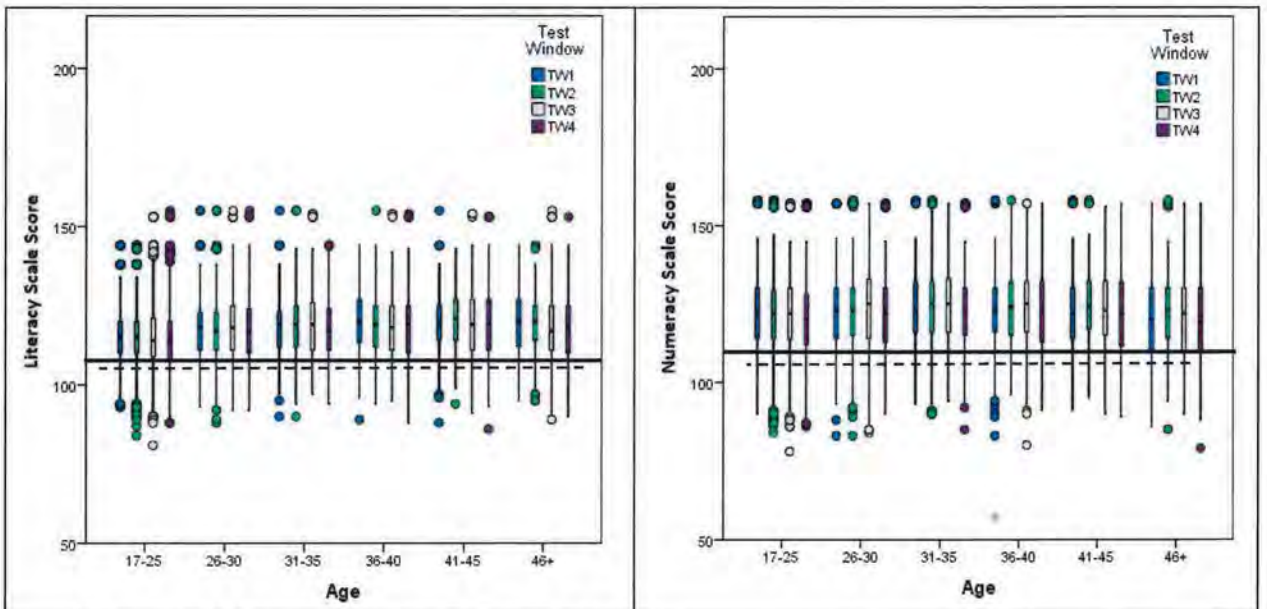


Figure 8: Score distribution by age group and test window

Figure 8 shows that, for both literacy and numeracy, the distributions of the scale scores of the age groups are very similar. The one exception is for literacy where it can be seen that the achievement of candidates in the 17–25 year age group is lower than the achievement of older candidates.

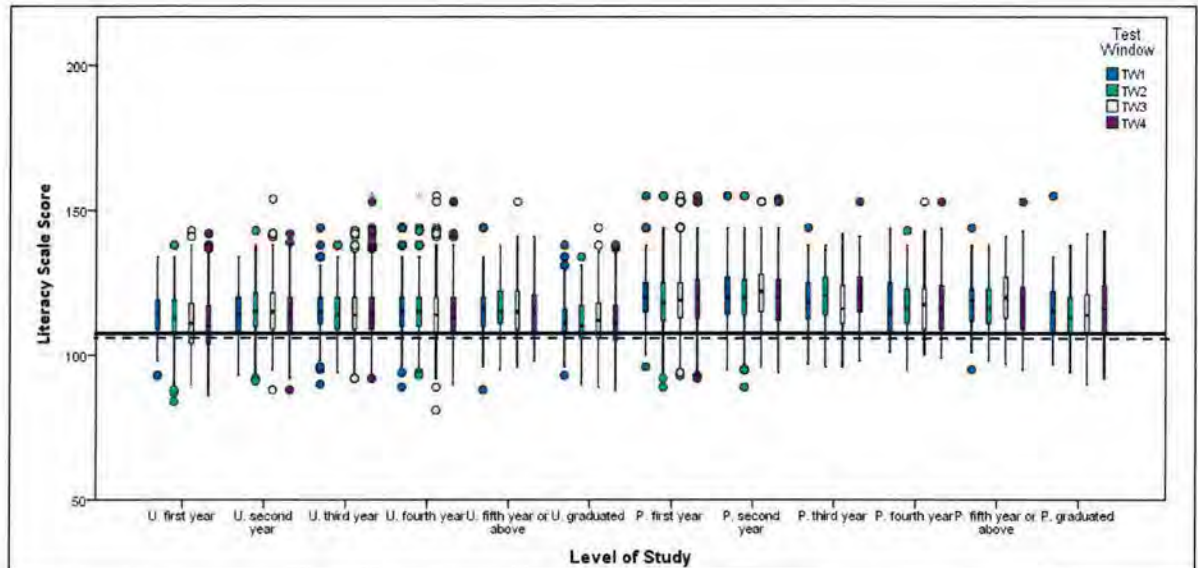


Figure 9: Score distribution by program type, year level and test window – literacy

Figure 9 shows that there are no observable patterns in the scale score distributions of the program types by year level between test windows for literacy.

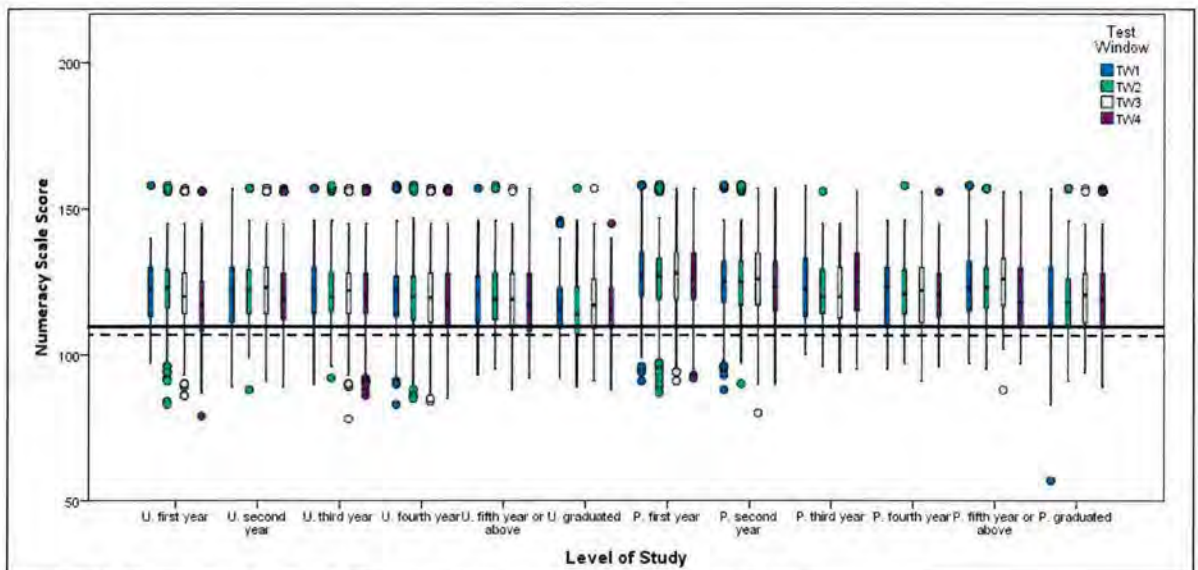


Figure 10: Score distribution by program type, year level and test window – numeracy

Similarly, Figure 10 shows that there are no observable patterns between test windows in the scale score distributions of the program types by year level for numeracy.

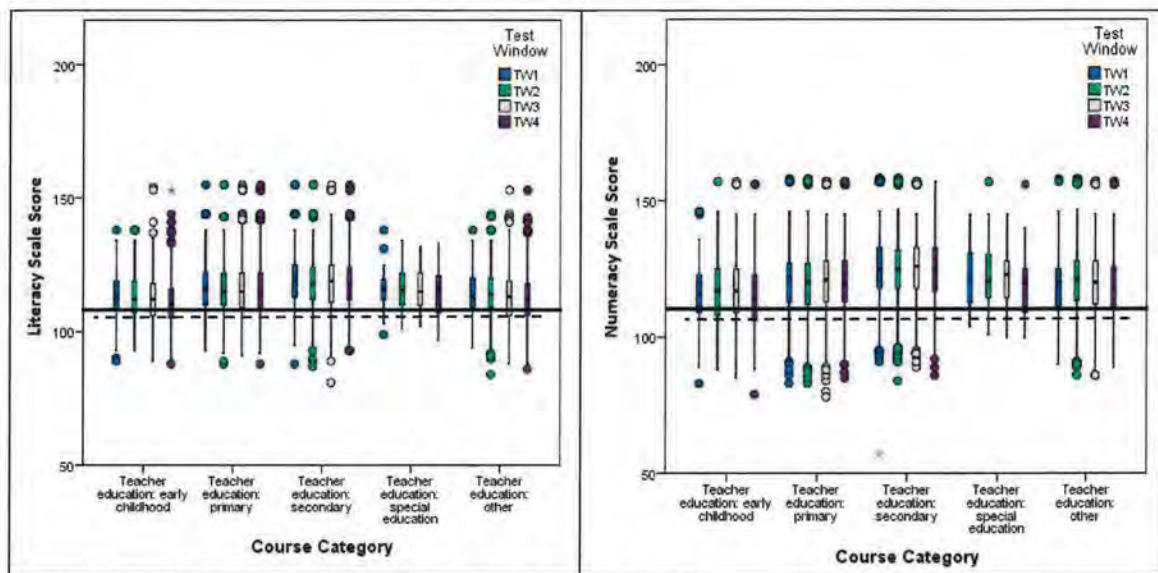


Figure 11: Score distribution by course category and test window

Figure 11 shows that there are no observable patterns between test windows in the scale score distributions of the course categories for both literacy and numeracy.

Appendix 5: Performance by Test Centres and Remote Proctoring by Test Window

Table 49 shows performance by location of test centres and remote proctoring in test windows 1 to 4.

Table 47: Performance by test centres and remote proctoring in test windows 1 to 4

Test Window	Category	Literacy				Numeracy			
		N	Mean	S.D.	Pass Rate	N	Mean	S.D.	Pass Rate
TW1	Capital Cities	NA	NA	NA	NA	NA	NA	NA	NA
	Regional Cities	NA	NA	NA	NA	NA	NA	NA	NA
	Remote Proctoring	4264	116.8	8.6	90.9	4286	121.9	11.7	90.3
TW2	Capital Cities	4204	116.6	8.7	90.5	4229	122.6	11.6	91.7
	Regional Cities	673	116.2	8.2	91.4	674	121.3	10.9	91.5
	Remote Proctoring	2631	116.1	8.8	89.0	2637	121.8	11.8	89.8
TW3	Capital Cities	4168	116.4	9.5	84.7	4174	123.0	12.0	86.7
	Regional Cities	790	116.2	8.8	86.5	793	122.4	11.5	85.8
	Remote Proctoring	1710	115.7	10.7	79.4	1695	121.5	12.7	81.0
TW4	Capital Cities	4215	115.6	9.7	81.7	4209	121.2	12.4	82.6
	Regional Cities	746	116.7	8.9	89.3	746	122.3	12.0	86.5
	Remote Proctoring	1805	115.0	9.4	81.9	1775	120.2	12.3	80.6

Comparing the mean scale scores for remote proctoring and test centres in capital cities and regional cities at the bottom of each test window table, it can be seen that while mean scale scores tended to be higher in test window 1 and lower in test windows 3 and 4, the differences were in the order of one scale score point only and not significant.

Appendix 6: Analysis of Differential Item Functioning (DIF)

Differential item functioning was performed to investigate if there is any scored item in the live tests that may be biased on gender, age group, program type, course category, Indigeneity, English as first language, and region. The data files for both 2016 and 2017 were used for this analysis. Items showing DIF were investigated for any content bias and findings are shown in the comment column of tables 50 to 56.

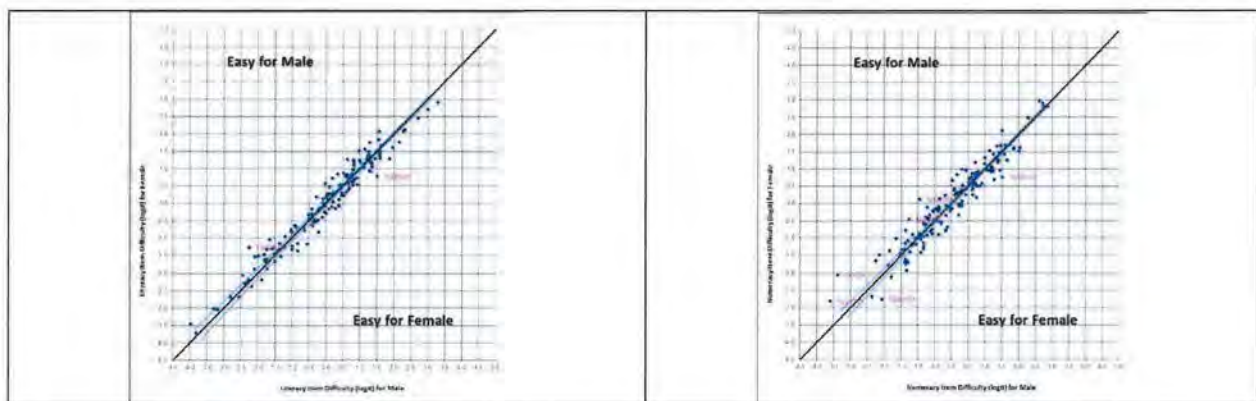


Figure 12: Gender DIF plots

Due to the large sample size, the confidence interval is very narrow and most items fall outside it. A few items are relatively distant from the boundary lines, however.

Table 50 lists potential gender DIF items with a standardised difference of 6.5 and above. The table shows that out of 160 test items, for literacy, one item significantly favoured females and one item significantly favoured males. For numeracy, two items significantly favoured females and four items significantly favoured males. A review of the items with gender bias revealed that a calculator was not available for 3 of the 4 numeracy items (*) favouring males, a possible factor. Because the items appeared in different clusters, no one test contained all items. The two literacy items do not provide any basis for tentative conclusions.

Table 48: List of potential gender DIF items

Item Label	Name	Difference in Item Difficulties (logit) (Male – Female)	Standardised Difference in Item Difficulties (logit) (Male – Female)	Chi-square	Probability	Gender favoured	Comment
L002404	Indigenous Retention Rates	-0.97	-11.6	134.4	0.00	Male	Reading (mixed text). Recognise related words in a relatively technical text.
L006201	School board	0.74	29.2	851.2	0.00	Female	TSW (Spelling). Correct misspelling of a relatively common word.
N000401(*)	Test Scores	-1.32	-16.7	279.3	0.00	Male	Identify common fraction as percentage without a calculator. No obvious content favouring males.
N023701	Aboriginal And Torres Strait Islander Students	-0.77	-6.6	43.4	0.00	Male	Non-routine graph reading. No obvious content favouring males.
N026102	Staff Meetings	0.77	20.1	402.9	0.00	Female	Solve a problem requiring knowledge of fortnight and the calendar. No obvious content favouring females.
N022101*	Geographical Distribution of Australians	-0.79	-22.4	503.8	0.00	Male	Convert common fraction to percentage without a calculator. No obvious content favouring males.
N062101	School Play	0.72	6.5	42.2	0.00	Female	Solve a multi-step problem involving calculation of break-even costs for an event. No obvious content favouring females.
N027901*	Waterfall Distance	-1.06	-20.9	436.4	0.00	Male	Convert a distance in kilometres to metres, without a calculator. No obvious content favouring males.

Figure 13 shows the DIF plots for age groups (17–25 year olds compared to 26+ year olds). The plots show that most items are close to the confidence interval limits.

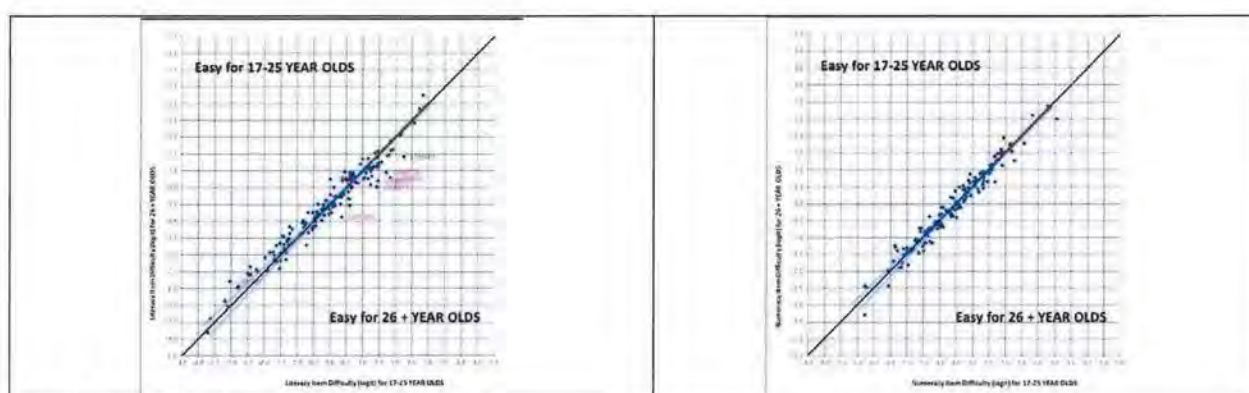


Figure 13: Age group DIF plots

Table 51 lists the items with potential DIF by age group. Of the eight literacy items, two favoured candidates aged 17–25 and six favoured older candidates. No potential age DIF was found in the pool of numeracy items. Five of the six items that favoured older candidates involved relatively sophisticated vocabulary or shades of meaning. It is possible that this involves a kind of knowledge that is extended over time, rather than early-acquired skills that can equally be applied to new situations. While there were more items favouring older candidates, because each candidate attempted only 60 out of 165 literacy items the effect was lessened at the candidate level.

Table 49: List of potential age group DIF items

Item Label	Name	Difference in Item Difficulties (logit) (17–25 / 26+)	Standardised Difference in Item Difficulties (logit) (17–25 / 26+)	Chi-square	Probability	Age favoured	Comment
L006402	Storm damage	1.06	37.4	1396.3	0.00	26+	TSW (Word usage). Recognise meaning of a relatively formal word.
L000209	English spelling	0.73	24.5	599.1	0.00	26+	Reading (continuous text). Recognise a link between reactions described in the text and a word describing them.
L006801	Discipline	0.83	18.9	355.8	0.00	26+	TSW (Word usage). Distinguish between easily mistaken, relatively sophisticated words.
L006211	School board	0.91	35.6	1269.3	0.00	26+	TSW (Word usage). Distinguish between fine shades of meaning in selecting the best synonym.
L001704	School website plan	–0.75	–11.6	134.0	0.00	17–25	Reading (Non-continuous text). Navigate nested dot-points. Possibility of benefit from more recent relevant experience?
L003502	Concussion	–0.73	–15.2	232.4	0.00	17–25	Reading (mixed text). Recognise a synonymous match between item and text.
L008803	School dance	0.77	20.9	435.0	0.00	26+	TSW (Word usage). Select a relatively sophisticated word to complete a sentence.
L008703	Playground incident	0.77	21.7	470.5	0.00	26+	TSW (SGP). Correct an error in an irregular verb.

Figure 14 shows the DIF plots for program type (undergraduate compared to postgraduate).

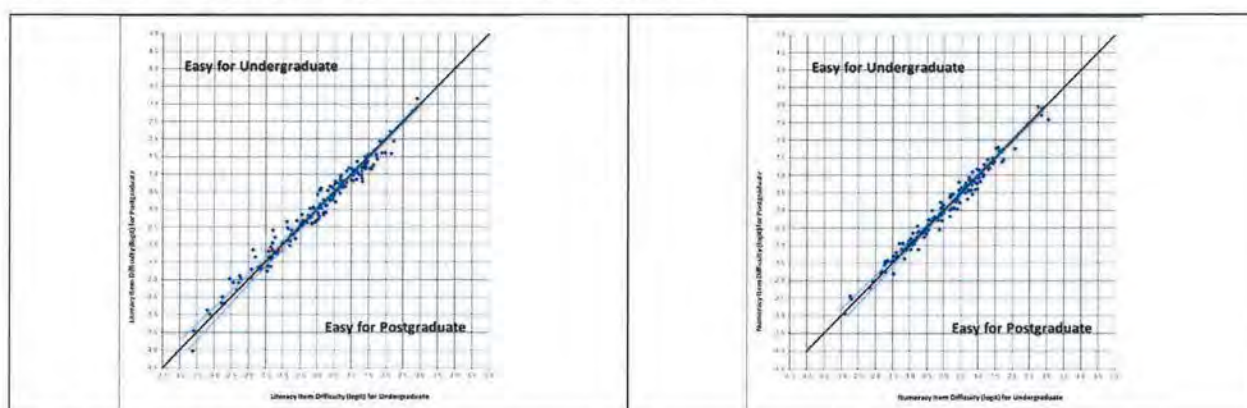


Figure 14: Program type DIF plots

Table 52 lists the item with potential DIF by program type. One literacy item favoured undergraduate candidates. No conclusions can be attempted on this evidence. It is worth noting that this item shows DIF in four of the seven categories considered here.

Table 50: List of potential program type DIF items

Item Label	Name	Difference in Item Difficulties (logit) (Undergraduate – Postgraduate)	Standardised Difference in Item Difficulties (logit) (Undergraduate – Postgraduate)	Chi-square	Probability	Program favoured	Comment
L005302	Robert	-0.71	-10.0	99.6	0.00	Undergraduate	TSW (Spelling). Correct a commonly mistaken homophone.

Figure 15 shows the DIF plots for course category (early childhood & primary compared to secondary).

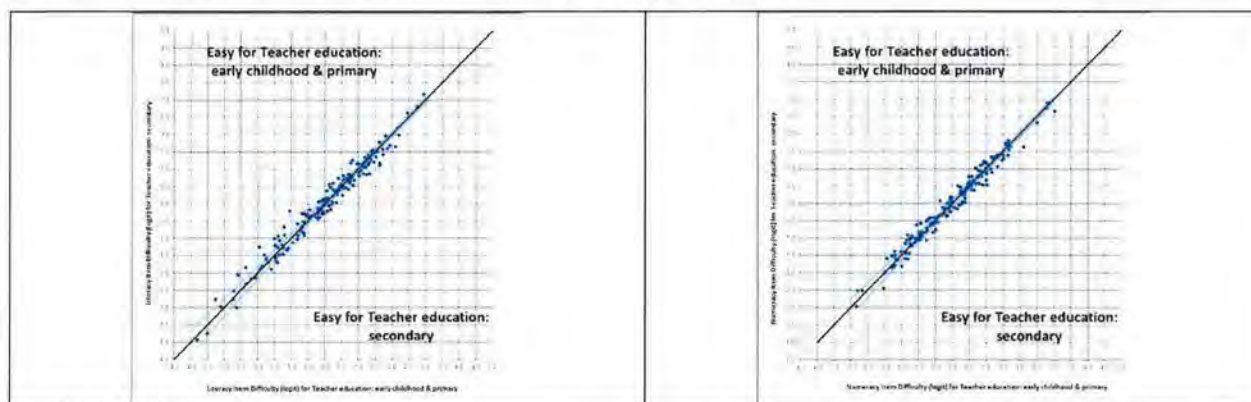


Figure 15: Course category DIF plots

Table 53 lists the item with potential course category DIF. One literacy item favoured candidates taking early childhood & primary courses. No conclusions can be attempted on this evidence.

Table 51: List of potential course category DIF items

Item Label	Name	Difference in Item Difficulties (logit) (early childhood & primary - secondary)	Standardised Difference in Item Difficulties (logit) (early childhood & primary - secondary)	Chi-square	Probability	Course favoured	Comment
L005210	Letter to parents	-0.83	-20.9	435.0	0.00	Early childhood & primary	TSW (SGP). Correct simple punctuation error, in a sentence containing a mildly colloquial usage.

Figure 16 shows the DIF plots for Indigenous (Indigenous compared to non-Indigenous).

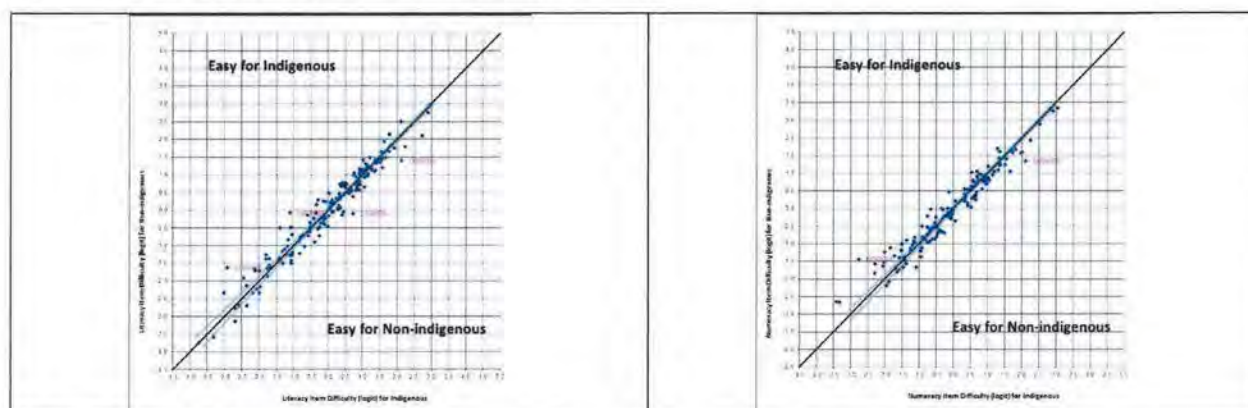


Figure 16: Indigenous DIF plots

Table 54 lists the items with potential DIF by Indigeneity. Of the two numeracy items, one favoured Indigenous candidates and one favoured non-Indigenous candidates. Of the four literacy items, two favoured Indigenous candidates and two favoured non-Indigenous candidates. One of the literacy items that favoured Indigenous candidates related to a text with an Indigenous focus and required an answer from the text that may also have coincided with likely Indigenous positions on an issue.

Table 52: List of potential Indigenous DIF items

Item Label	Name	Difference in Item Difficulties (logit) (Indigenous – Non-Indigenous)	Standardised Difference in Item Difficulties (logit) (Indigenous – Non-Indigenous)	Chi-square	Probability	Favoured	Comment
L003908	Yolngu education	-1.02	-22.5	506.6	0.00	Indigenous	Reading (narrative text with Indigenous theme). Identify attitudes referred to in text. Correct answer corresponds to likely Indigenous positions.
L008905	Saving water	0.82	17.1	291.9	0.00	Non-Indigenous	TSW (TO). Identify correct conjunction (relatively sophisticated) to complete a sentence.
L008703	Playground incident	0.73	20.7	428.0	0.00	Non-Indigenous	TSW (SGP). Correct an error in an irregular verb.
L005302	Robert	-1.29	-18.2	332.1	0.00	Indigenous	TSW (Spelling). Correct a commonly mistaken homophone.
N040001	Sprint Final	-1.32	-20.4	414.2	0.00	Indigenous	Determine the accuracy of statements about times at a sporting event, without access to a calculator. No obvious content favouring Indigenous candidates.
N026001	School Funding	0.77	20.8	434.1	0.00	Non-Indigenous	Interpret a bar graph and identify a related proportion expressed as a unitary fraction. No obvious content favouring non-Indigenous candidates.

Figure 17 shows the DIF plots for English as first language (English as a first language compared to English not as a first language).

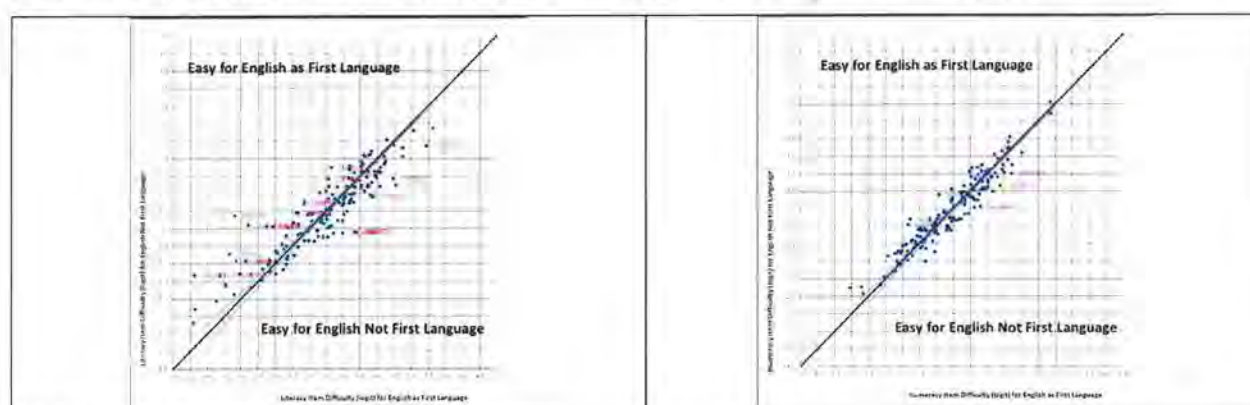


Figure 17: English as first language DIF plots

Table 55 lists the items with potential English as first language DIF. Of the 6 numeracy items, 3 favoured candidates whose first language is English, and 3 favoured candidates whose first language is not English. A review of the numeracy items with English as first language DIF revealed that a calculator was not available for all 3 of the numeracy items (*) favouring candidates whose first language is not English, a possible factor. Of the 23 literacy items, 16 favoured candidates whose first language is English, and 7 favoured candidates whose first language is not English. Only 6 of these 23 items were reading items, although two-thirds of the test items are reading items. Five of the 7 items that favoured candidates whose first language is not English tested 'Syntax, grammar and punctuation' – it is possible that these candidates learned English through a more grammar-focused approach. Five of the 16 items that favoured candidates whose first language is English related to finesse in using the language: using an appropriate style and tone, recognising an ambiguity, selecting a word to convey a precise shade of meaning. Only one of the 6 reading items involves a continuous text (a narrative); the others involved either a non-continuous or a mixed text (about half the reading items in the test relate to continuous texts.) Five of the 6 reading items favour candidates whose first language is English. It would seem that candidates whose first language is not English have greater difficulty with non-continuous or mixed texts.

Table 53: List of potential English as first language DIF items

Item Label	Name	Difference in Item Difficulties (logit) (English as FL – English Not FL)	Standardised Difference in Item Difficulties (logit) (English as FL – English Not FL)	Chi-square	Probability	Favoured	Comment
L001003	Put a shirt on	-1.89	-35.2	1235.6	0.00	English as FL	Reading (narrative text). Deduce (or know) the meaning of a relatively sophisticated word in context.
L005602	Brenda	-1.63	-26.8	717.6	0.00	English as FL	TSW (TO). Organise ideas in a note-form text.
L005606	Brenda	-0.95	-17.2	296.1	0.00	English as FL	TSW (Word usage). Identify inappropriateness of a style of writing.
L005002	Library	1.40	43.1	1855.1	0.00	English Not FL	TSW (SGP). Identify correct possessive pronoun. Explicitly taught grammar might help.
L005105	School excursion	1.32	40.6	1649.3	0.00	English Not FL	TSW (SGP). Identify correct use of apostrophes. Explicitly taught grammar might help.
L006303	Sport	1.47	80.0	6401.6	0.00	English Not FL	TSW (SGP). Identify correct use of apostrophes. Explicitly taught grammar might help.
L006502	Musical performance	-1.10	-51.8	2684.2	0.00	English as FL	TSW (Word usage). Identify most appropriate style and tone.
L006210	School board	-2.04	-14.3	203.8	0.00	English as FL	TSW (Word usage). Identify most appropriate style and tone.
L001807	Risk assessment	-1.22	-12.3	151.6	0.00	English as FL	Reading (non-continuous text). Identify a detail not included (synonymously) in a number of non-continuous pieces of information.
L001706	School website plan	1.11	56.1	3148.6	0.00	English Not FL	Reading (non-continuous text). Identify information that is/is not included (synonymously) in a non-continuous text.
L007004	Job application	-1.14	-18.3	335.1	0.00	English as FL	TSW (TO). Classify an activity according to a given scheme.
L003501	Concussion	-1.31	-8.8	77.7	0.00	English as FL	Reading (mixed text). Identify underlying questions behind a text.

L008405	Petition	-1.30	-19.5	382.0	0.00	English as FL	TSW (Word usage). Identify a more sophisticated word to convey a shade of meaning.
L003801	Care of animals	-0.83	-13.6	185.9	0.00	English as FL	Reading (non-continuous text). Apply a classification given in the text.
L008901	History excursion	-0.85	-18.8	352.0	0.00	English as FL	TSW (TO). Identify an ambiguous reference in a text.
L005504	Computer tablets	1.07	39.9	1588.6	0.00	English Not FL	TSW (Spelling). Correct the spelling of a relatively common word.
L005210	Letter to parents	-0.77	-24.7	611.0	0.00	English as FL	TSW (SGP). Correct simple punctuation error in a sentence containing a mildly colloquial usage.
L005202	Letter to parents	1.06	35.7	1276.3	0.00	English Not FL	TSW (SGP). Correct an apostrophe error.
L001701	School website plan	-0.86	-33.7	1138.8	0.00	English as FL	Reading (non-continuous text). Identify a relevant location within a dot-point text, based on understanding of a technical term.
L005302	Robert	-2.57	-33.0	1090.9	0.00	English as FL	TSW (Spelling). Correct a commonly mistaken homophone.
L005304	Robert	-0.74	-7.7	59.1	0.00	English as FL	TSW (Spelling). Correct a misspelling of a common word.
L005804	Homework	0.96	29.5	872.8	0.00	English Not FL	TSW (SGP). Correct an apostrophe error.
L008403	Petition	-0.91	-18.9	357.4	0.00	English as FL	TSW (Spelling). Correct a misspelling of a relatively sophisticated word.
N000101	Yard Duty	-0.86	-19.6	384.6	0.00	English as FL	Calculate a time duration in a school context and halve it, without access to a calculator. It is possible that the word 'duration' is unfamiliar to ENFL candidates.
N003303	Currency Conversion	-0.98	-23.1	533.3	0.00	English as FL	Round number to the nearest 1000 in a currency context. No obvious content favouring EFL candidates.
N026901*	School Day	0.70	17.7	312.7	0.00	English Not FL	Convert minutes per week to hours and minutes per day, without access to a calculator. No obvious content favouring ENFL candidates. A calculator was not available and it is possible that ENFL candidates are more competent without a calculator.

N028401*	Other Language	0.74	16.9	285.1	0.00	English Not FL	Identify the common percentage that one number is of another, without access to a calculator. It is possible that the context was more engaging to ENFL candidates. Another factor could be that a calculator was not available and that ENFL candidates are more competent without a calculator.
N053301	Pre-school Hours	-0.81	-16.8	283.6	0.00	English as FL	Determine the accuracy of statements about a delivery requirement. It is possible that the word 'duration' is unfamiliar to ENFL candidates.
N028101*	Large School	0.73	17.8	317.0	0.00	English Not FL	Calculate a fraction of a 4-digit number, without access to a calculator. A calculator was not available and it is possible that ENFL candidates are more competent without a calculator.

Figure 18 shows the DIF plots for region (metropolitan areas compared to provincial/remote areas).

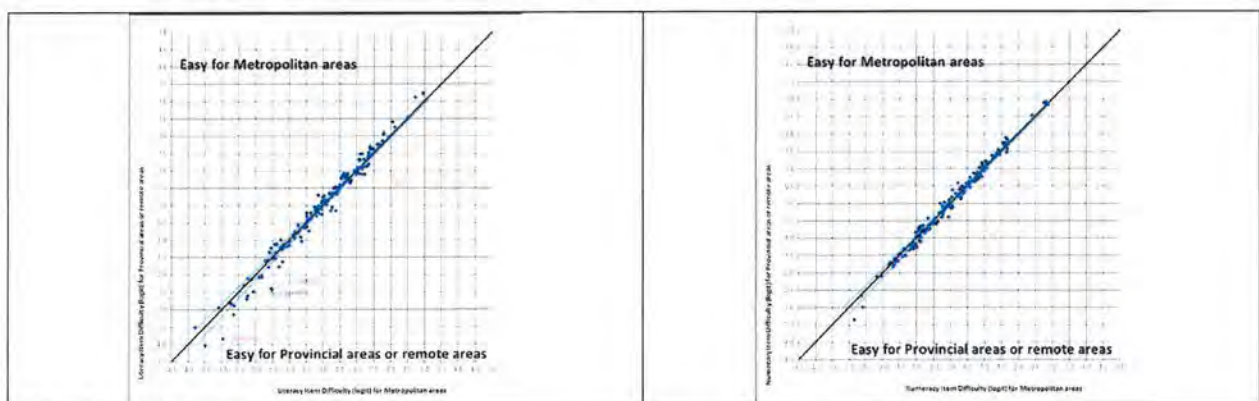


Figure 18: Location DIF plots

Table 56 lists the items with potential location DIF. All three literacy items favoured candidates whose residential areas are in provincial areas or remote area.

Table 54: List of potential region DIF items

Item Label	Name	Difference in Item Difficulties (logit) (Metropolitan – Provincial or remote)	Standardised Difference in Item Difficulties (logit) (Metropolitan – Provincial or remote)	Chi-square	Probability	Location favoured	Comment
L001003	Put a shirt on	0.87	17.1	290.8	0.00	Provincial or remote	Reading (narrative text). Deduce (or know) the meaning of a relatively sophisticated word in context.
L006210	School board	0.88	6.3	39.3	0.00	Provincial or remote	TSW (Word usage). Identify most appropriate style and tone.
L005302	Robert	0.92	13.0	168.6	0.00	Provincial or remote	TSW (Spelling). Correct a commonly mistaken homophone.

Appendix 7: Impact of differential item functioning at the test level

Differential Test Functioning (DTF) investigates if there is significant difference in expected raw scores of a test across subgroups of candidates at the same ability level. In this study, test characteristic curves (TCCs) were used to study the DTF for seven subgroups of candidates with the following demographic background variables: age (17–25 compared to 26+), course category (early childhood & primary compared to secondary), English as first language (English as first language compared to English as not first language), gender (male compared to female), indigeneity (Indigenous compared to non-Indigenous), program type (undergraduate compared to postgraduate), and residential area (metropolitan areas compared to provincial & remote areas).

A TCC describes the relationship between latent ability and expected raw scores in a test. For each variable of interest, the comparison of subgroup TCCs provides a way to evaluate the impact of differential item functioning at the test level.

To create the TCCs, a multiple-facets Rasch model was used to estimate subgroup item difficulty for each demographic variable. A total of 165 literacy items from 10 tests and a total of 156 numeracy items from 8 tests were *re-estimated* based upon the response data from each subgroup of candidates who participated in the test in 2016–17 (8 test windows). The multiple-facets Rasch model can calibrate item difficulties for each subgroup and put items from different subgroups on a new scale. For each subgroup, this enabled the abilities of candidates to be re-estimated on the new scale for each test. The new scale was then equated with the existing test scale using the common-item equating method so that estimated test raw scores across subgroups can be compared for candidate abilities on the existing test scale. Each TCC is a plot of the *re-estimated* candidate abilities (in scale score points) for each pair of subgroups versus estimated raw scores on the test. For example, Figure 19 and Figure 20 are the TCCs for the English as a first language and English not as a first language subgroups for literacy tests L47 and L52.

Rather than provide 18 sets of TCCs for each demographic variable, one summary graph for each demographic variable is provided. Each graph plots two summary statistics (difference in mean raw test score *across all abilities* and difference in raw test score *at the standard*) derived from the TCCs for each test. These can be seen in Figure 21 to Figure 27. A zero difference indicates that the estimated raw test score achieved by candidates in each subgroup with the same estimated ability is equal. A positive difference indicates that the test favours the first named subgroup.

Results for English as a First Language

Figure 19 is the TCC for the English as a first language variable for test L47. The closer the two curves are to each other, the smaller the differential test functioning for that variable. It can be seen that the tests are close together for higher ability candidates (higher scale scores) but diverge up to 2 raw score points for lower abilities. At the standard (107), English as a first language candidates are slightly favoured by nearly one raw score point.

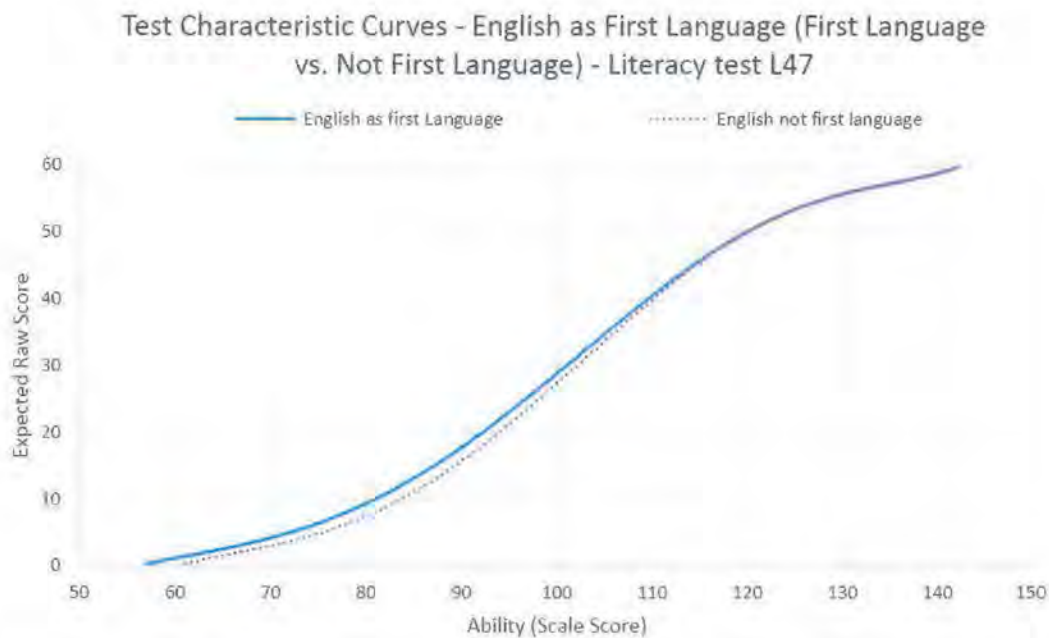


Figure 19: Comparison of Test Characteristic Curves by English as a first language – literacy test L47

Figure 20 is the TCC for the English as a first language variable for test L52. It can be seen that, unlike test L47, the curves are close for low ability candidates but diverge for candidates with ability at and above the standard (107). Unlike test L47, at the standard, it is English not as a first language candidates who are slightly favoured by 1 to 2 raw score points.

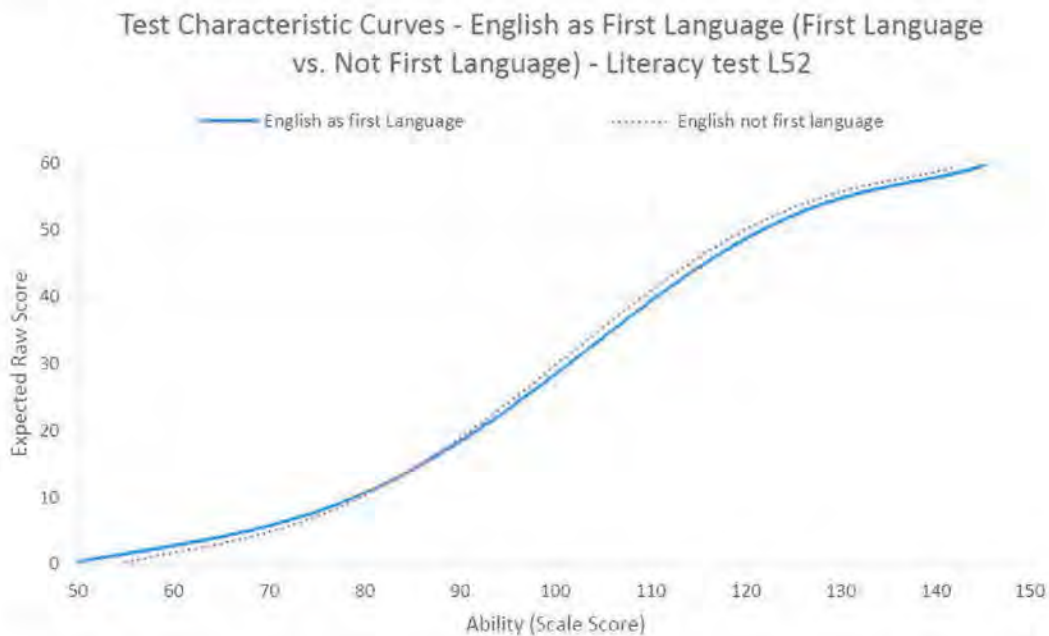


Figure 20: Comparison of Test Characteristic Curves by English as a first language – literacy test L52

It was the English as a first language variable that resulted in the greatest variation in DTF for both literacy and numeracy. While there were near zero differences in subgroup means (all abilities), it can be seen that for candidates with estimated ability at the standard some differences exceeded 1 raw score

point. This occurred for literacy tests L51 and L52. These differences across the 18 tests are quantified in Figure 21.

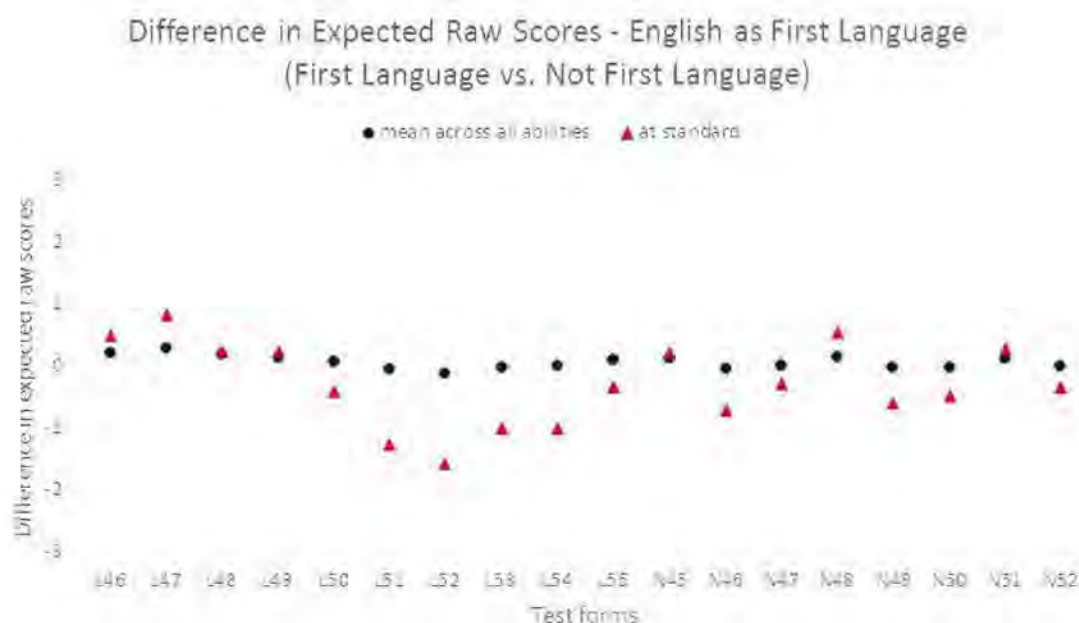


Figure 21: Difference in expected raw scores by English as a first language

Figure 21 shows the difference in expected raw scores by English as a first language for the 10 literacy tests and the 8 numeracy tests. On test L47, it was estimated that candidates from the English as a first language subgroup, with ability at the standard, would score nearly 1 raw score point higher on the test than candidates from English not as a first language background. On the other hand, the reverse was true for test L52, where it was estimated that English not as a first language candidates with ability at the standard would score approximately 1.6 raw score points higher than English as a first language candidates. Neither of these differences for candidates with ability at the standard was statistically significant.

To try and explain the reason for the different direction of the DTF for L47 and L52, a closer look at the items on each test was undertaken. On L47, there were 8 items out of 60 that significantly favoured English as a first language candidates and 8 that significantly favoured English not as a first language candidates.

On L52, there were 5 items out of 60 that significantly favoured English as a first language candidates and 10 that significantly favoured English not as a first language candidates, in agreement with the DTF finding above.

If the sensitivity of the DIF is increased to a more moderate level of significance, on L47 another 6 items moderately favour English as a first language candidates and 6 favour English not as a first language candidates, as shown in Table 57.

Table 55: Number of items showing DIF on test L47 for the English as a first language variable

L47	Reading text format		Technical Skills of Writing aspect		
	Continuous	Non-Continuous	Spelling	Syntax, Grammar and Punctuation	Word Usage
fav EFL	3	1	1	1	
fav ENFL	4		1	3	

mfav EFL	5				1
mfav ENFL	5	1			

On L52, when a more moderate level of significance is applied, another 3 items moderately favour English as a first language candidates and 7 favour English not as a first language candidates, as shown in Table 58.

Table 56: Number of items showing DIF on test L52 for the English as a first language variable

L52 DIF direction	Reading text format			Technical Skills of Writing aspect		
	Continuous	Mixed	Non-Continuous	Syntax, Grammar and Punctuation	Text Organisation	Word Usage
fav EFL		1	1		1	2
fav ENFL	3	1	2	2	1	1
mfav EFL			3			
mfav ENFL	2	1	1	1	1	1

A closer look at the nature of the items reveals that DIF is effectively cancelled in regard to reading text format on L47, which explains the smaller DTF. However, on L52, there were 5 continuous text reading items that favoured English not as a first language candidates, explaining the larger DTF favouring English not as a first language candidates.

For technical skills of writing, the majority of DIF favouring English not as a first language candidates arises from the syntax, grammar and punctuation items. Some cancelling of DIF occurred for the other aspects of technical skills of writing on both papers, but on L52, there was no cancelling effect for the syntax, grammar and punctuation aspect where 3 items were found to moderately or significantly favour English not as a first language candidates.

The difference in expected raw scores by English as a first language is less than 1 raw score point for the numeracy tests. The differences in expected raw scores in other background variables are smaller.

Gender

Figure 22 shows the summary graph for gender. It is clear that both plotted differences are small (< 1 raw score point) for each literacy test and for each numeracy test. For candidates at the standard, test L46 slightly favoured females and tests L51 and L52 slightly favoured males.

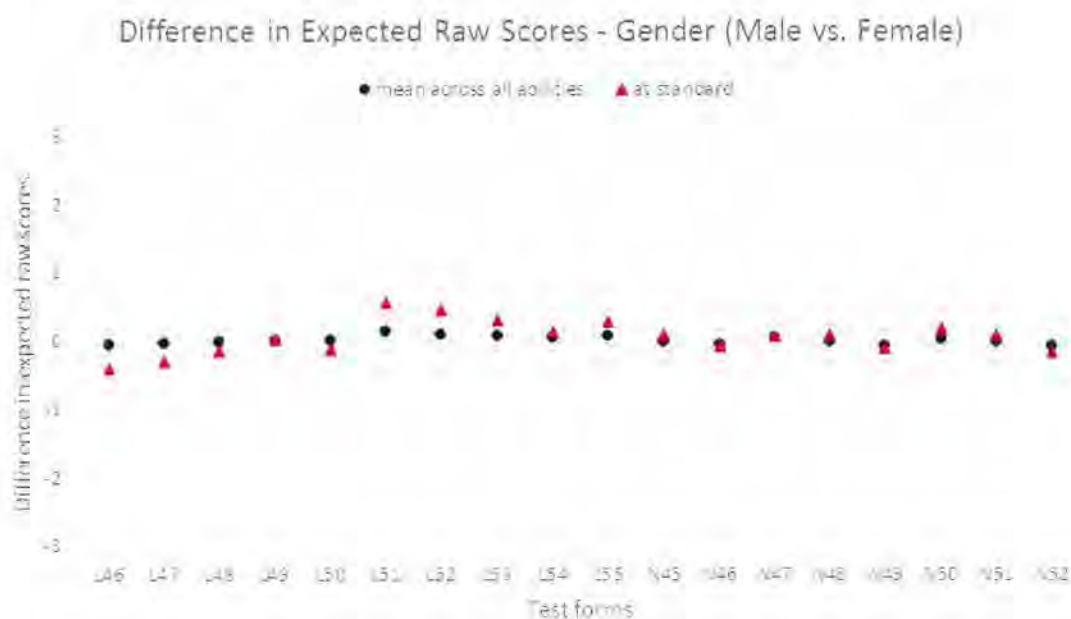


Figure 22: Difference in expected raw scores by gender

In summary, the impact of differential item functioning at the test level was found to be small. For the 8 numeracy tests, the absolute difference of the expected raw scores was found to be < 1 raw score point in all tests, across all the demographic variables studied. For the 10 literacy tests, the difference of

expected raw scores was found to be < 1 in all tests across the demographic variables studied, except for English as a first language.

Age

It can be seen from Figure 23 that all literacy tests slightly favoured older candidates (by less than one raw score point), possibly because older candidates have had more experience at literacy related tasks or a more formal grounding in the technical skills of writing.

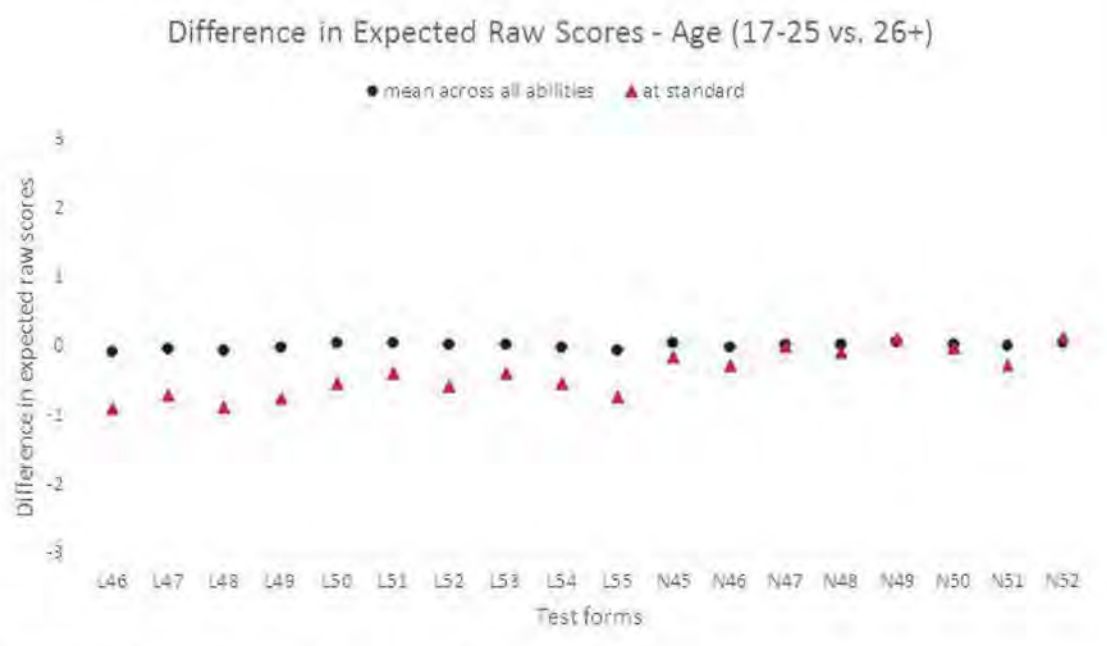


Figure 23: Difference in expected raw scores by age

Course category

It can be seen from Figure 24 that for both literacy and numeracy, there was little differential test functioning for the course category variable. Only a small difference (less than one raw score point favouring secondary candidates) was observed on literacy tests L51 to L55.

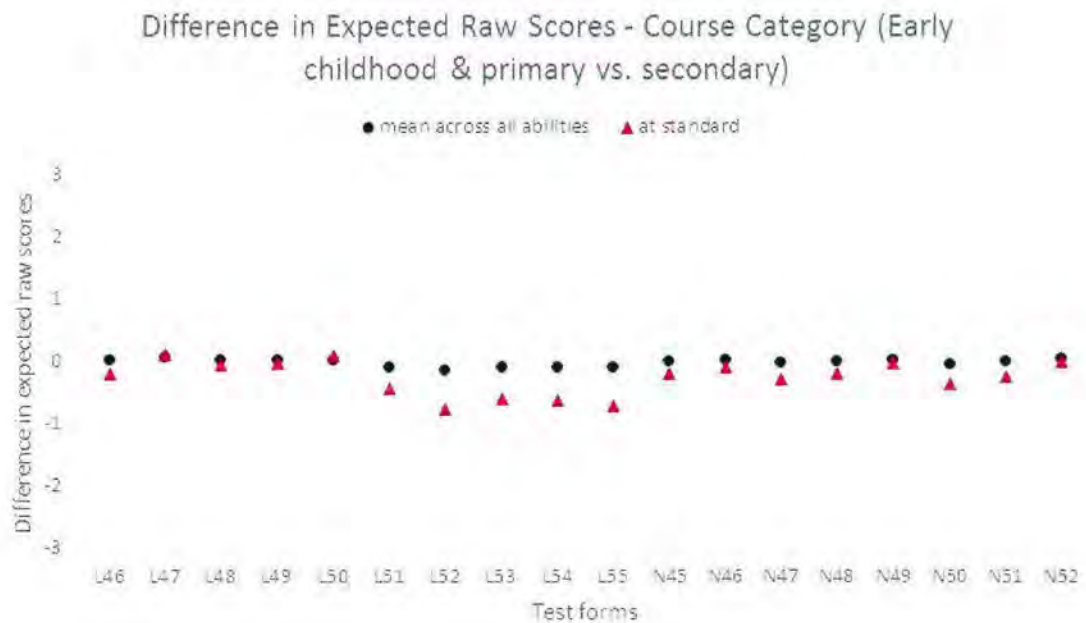


Figure 24: Difference in expected raw scores by course category

Indigeneity

It can be seen from Figure 25 that, for the Indigeneity variable, the differential test functioning was very small (half a raw score point or less) and in both directions for literacy.



Figure 25: Difference in expected raw scores by Indigeneity

Program type

It can be seen from Figure 26 that all literacy tests slightly favoured candidates in postgraduate programs (by less than one raw score point), possibly because these candidates have had more experience at literacy related tasks or a more formal grounding in the technical skills of writing.

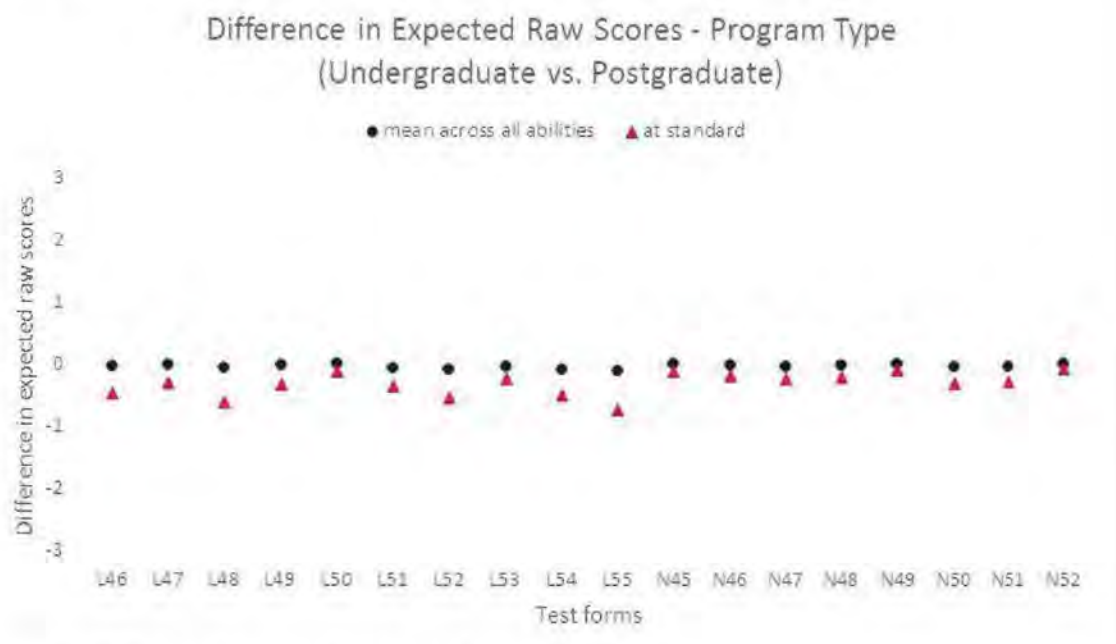


Figure 26: Difference in expected raw scores by program type

Residential area

It can be seen from Figure 27 that there was very little, if any, differential test functioning in regard to the location of candidates.



Figure 27: Difference in expected raw scores by residential area

Appendix 8: Effect of revised standards on third- and fourth-attempt candidates

Table 57: Comparison of outcomes between the old and the revised standards for third-attempt candidates

Characteristic	Category	Literacy				Numeracy			
		N	Pass Rate (Old Standard)	Pass Rate (Revised Standard)	Difference	N	Pass Rate (Old Standard)	Pass Rate (Revised Standard)	Difference
Gender	Female	181	40.3	30.4	-9.9	205	55.6	34.1	-21.5
	Male	31	58.1	38.7	-19.4	17	70.6	29.4	-41.2
	Indeterminate/intersex	1	100.0	100.0	0.0	-	-	-	-
Age	17-25	144	46.5	34.0	-12.5	139	63.3	37.4	-25.9
	26-30	42	42.9	28.6	-14.3	42	47.6	26.2	-21.4
	31-35	8	25.0	25.0	0.0	14	57.1	50.0	-7.1
	36-40	10	50.0	50.0	0.0	5	60.0	40.0	-20.0
	41-45	5	0.0	0.0	0.0	13	23.1	15.4	-7.7
	46+	4	0.0	0.0	0.0	9	44.4	11.1	-33.3
International Students	No	171	41.5	30.4	-11.1	219	56.6	33.3	-23.3
	Yes	42	50.0	38.1	-11.9	3	66.7	66.7	0.0
English as a First Language	Yes	133	45.9	34.6	-11.3	184	54.9	31.0	-23.9
	No	80	38.8	27.5	-11.3	38	65.8	47.4	-18.4
Indigenous	No	202	43.6	32.2	-11.4	208	58.7	35.1	-23.6
	Yes	7	28.6	28.6	0.0	10	40.0	20.0	-20.0
	Not disclosed	4	50.0	25.0	-25.0	4	0.0	0.0	0.0
Residential Area	Metropolitan areas	178	44.4	32.6	-11.8	187	57.2	34.2	-23.0
	Provincial areas	34	38.2	29.4	-8.8	32	53.1	31.3	-21.9
	Remote areas	-	-	-	-	3	66.7	33.3	-33.3
	Invalid or Missing	1	0.0	0.0	0.0	-	-	-	-
Program Type	Undergraduate	139	43.2	28.8	-14.4	152	57.2	32.9	-24.3
	Postgraduate	74	43.2	37.8	-5.4	70	55.7	35.7	-20.0
Program Type by Year Level	Undergraduate first year	21	42.9	28.6	-14.3	4	75.0	75.0	0.0
	Undergraduate second year	2	50.0	0.0	-50.0	3	100.0	66.7	-33.3
	Undergraduate third year	11	54.5	27.3	-27.3	7	57.1	28.6	-28.6
	Undergraduate fourth year	71	43.7	32.4	-11.3	87	55.2	32.2	-23.0
	Undergraduate fifth yr or above	9	55.6	33.3	-22.2	19	68.4	31.6	-36.8
	Undergraduate graduated	25	32.0	20.0	-12.0	32	50.0	28.1	-21.9
	Postgraduate first year	22	63.6	54.5	-9.1	7	57.1	28.6	-28.6
	Postgraduate second year	11	18.2	18.2	0.0	17	64.7	47.1	-17.6

Course Category	Postgraduate third year	3	33.3	33.3	0.0	3	66.7	33.3	-33.3
	Postgraduate fourth year	4	75.0	75.0	0.0	7	57.1	28.6	-28.6
	Postgraduate fifth year or above	7	28.6	14.3	-14.3	4	0.0	0.0	0.0
	Postgraduate graduated	27	37.0	33.3	-3.7	32	56.3	37.5	-18.8
	Early childhood	19	42.1	26.3	-15.8	26	57.7	38.5	-19.2
	Primary	85	43.5	29.4	-14.1	97	63.9	37.1	-26.8
	Secondary	58	48.3	41.4	-6.9	57	54.4	35.1	-19.3
	Special education	2	50.0	50.0	0.0	-	-	-	-
	Other	49	36.7	26.5	-10.2	42	42.9	21.4	-21.4

Table 58: Comparison of outcomes between the old and the revised standards for fourth-attempt candidates

Characteristic	Category	Literacy				Numeracy			
		N	Pass Rate (Old Standard)	Pass Rate (Revised Standard)	Difference	N	Pass Rate (Old Standard)	Pass Rate (Revised Standard)	Difference
Gender	Female	21	57.1	33.3	-23.8	36	38.9	25.0	-13.9
	Male	4	75.0	50.0	-25.0	2	50.0	50.0	0.0
Age	17-25	17	58.8	35.3	-23.5	26	34.6	19.2	-15.4
	26-30	4	75.0	50.0	-25.0	8	50.0	50.0	0.0
	31-35	-	-	-	-	2	50.0	50.0	0.0
	36-40	3	66.7	33.3	-33.3	1	100.0	0.0	-100.0
	41-45	1	0.0	0.0	0.0	-	-	-	-
	46+	-	-	-	-	1	0.0	0.0	0.0
International Students	No	24	58.3	33.3	-25.0	37	37.8	24.3	-13.5
	Yes	1	100.0	100.0	0.0	1	100.0	100.0	0.0
English as a First Language	Yes	17	64.7	41.2	-23.5	32	37.5	25.0	-12.5
	No	8	50.0	25.0	-25.0	6	50.0	33.3	-16.7
Indigenous	No	25	60.0	36.0	-24.0	37	37.8	24.3	-13.5
	Yes	-	-	-	-	1	100.0	100.0	0.0
Residential Area	Metropolitan areas	20	55.0	35.0	-20.0	32	40.6	25.0	-15.6
	Provincial areas	5	80.0	40.0	-40.0	6	33.3	33.3	0.0
Program Type	Undergraduate	12	66.7	33.3	-33.3	28	32.1	21.4	-10.7
	Postgraduate	13	53.8	38.5	-15.4	10	60.0	40.0	-20.0
Program Type by Year Level	Undergraduate third year	1	0.0	0.0	0.0	-	-	-	-
	Undergraduate fourth year	5	100.0	20.0	-80.0	14	14.3	7.1	-7.1
	Undergraduate fifth yr or above	-	-	-	-	4	50.0	50.0	0.0
	Undergraduate graduated	6	50.0	50.0	0.0	10	50.0	30.0	-20.0
	Postgraduate first year	-	-	-	-	1	100.0	100.0	0.0
	Postgraduate second year	1	100.0	100.0	0.0	2	50.0	0.0	-50.0
	Postgraduate third year	2	0.0	0.0	0.0	1	100.0	0.0	-100.0
	Postgraduate fourth year	-	-	-	-	1	0.0	0.0	0.0
Course Category	Postgraduate graduated	10	60.0	40.0	-20.0	5	60.0	60.0	0.0
	Early childhood	3	33.3	0.0	-33.3	5	40.0	40.0	0.0
	Primary	9	66.7	55.6	-11.1	16	37.5	25.0	-12.5
	Secondary	5	80.0	60.0	-20.0	10	50.0	30.0	-20.0
	Special education	1	100.0	0.0	-100.0	-	-	-	-
	Other	7	42.9	14.3	-28.6	7	28.6	14.3	-14.3