Institute for Social Science Research
**10 February 2023**

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
CREATE CHANGE

# Building Capacity for Quality Teaching in Australian Schools: Queensland Replication Study Final Report

**Lead:** Assoc. Professor Jenny Povey

**Project Manager:** Lizzie Kennedy

**ISSR Project Team:** Sophie Austerberry, Melindy Bellotti, Steffi Cook, Dr Iryna Kolesnikova, Dr Azhar Hussain Potia, Dr Macarena San Martin Porter, Anne Schirmann, Wojtek Tomaszewski.

**Authors of this report:** Assoc. Professor Jenny Povey, Dr Macarena San Martin Porter, Lizzie Kennedy, Dr Azhar Hussain Potia, Melindy Bellotti, Steffi Cook and Sophie Austerberry.

# Contents

# Tables



ISBN: 978-0-7259-0946-8

# 1. Executive summary

This project Building Capacity for Quality Teaching in Australian Schools – Queensland Replication Study examined the efficacy of Quality Teaching Rounds (QTR) in a new jurisdiction.

Early findings based on the baseline assessments from both the Reading and Maths components of the ACER Progressive Assessment Testing (PAT-R and PAT-M) and school implementation fidelity data for Cohort 2 were presented in Progress Report 3. This report details the project developments since the Progress Report delivered in September 2022.

Data presented in this final report includes findings from the baseline and post-intervention assessments for the Reading and Maths components of the ACER Progressive Assessment Testing (PAT-R and PAT-M) and school implementation fidelity data for Cohort 2. The online teacher and student survey data for Cohort 2 is also included in this report.

A summary of the key findings is presented below:

**Implementation fidelity**

- The implementation fidelity of the QTR intervention was high for Cohort 2, with a slightly lower mean score recorded for the observed (M = 8.31) compared with the self-reported (M = 8.70) fidelity. Seven out of the 13 schools achieved 100% for the observed fidelity check and 9 out of the 13 schools achieved 100% for the self-reported fidelity check (achieving 9/9).

**Teacher survey outcomes**

- There were no significant differences in the teacher survey outcomes between QTR and Control groups.

**Student survey outcomes**

- There were no significant differences in the student survey outcomes between QTR and Control groups.

**PAT data**

- There were no significant differences in the Maths PAT assessments over time between students from the QTR and Control groups.
- There was a significant improvement in Reading PAT assessment over time for students from the QTR group when compared to students from the Control group.

# 2. Introduction

Researchers at the Teachers and Teaching Research Centre (TTRC) at The University of Newcastle (UoN) have been investigating the impact of Professional Development (PD) on student outcomes. Teaching practice is an important in-school factor impacting on student outcomes. As well as recruiting and training 'better' teachers, improving initial teacher education, and evaluating the quality of teaching; providing professional development to build the capacity of practising teachers is a key strategy employed globally to improve teaching practice.

The TTRC research team has developed, and rigorously tested, a professional development intervention—Quality Teaching Rounds (QTR)—that aims to support teachers to improve their teaching practice. At the core of the approach is a Quality Teaching Model that focuses on 3 key concepts:

- the need for intellectual quality, rigour or challenge in every learning experience
- the need to create classroom environments that support not only students but also their learning
- the need to increase the significance of student learning so they can see its connection to the world beyond the classroom.

The QTR process involves teachers collaborating in professional learning communities, to observe and analyse each other's teaching (one lesson per teacher) over a period of weeks and reach consensus on the quality of pedagogy. This practice has proven successful in New South Wales (NSW) where it has been evaluated to have positively impacted on both student achievement and quality of teaching; as well as having broader impacts on teacher morale, school culture and collaboration.

The TTRC research team also investigated the impact of this PD in jurisdictions outside of NSW. In 2020 a Pilot study of QTR was conducted in Queensland. Following on from the pilot, The University of Queensland Institute of Social Science Research (ISSR) was engaged as independent evaluators in a replication study (randomised controlled trials) in Queensland state primary schools. Fifty-eight Queensland state primary schools participated in the Queensland QTR study in 2021 (Cohort 1). With the exception of the student survey data, findings from Cohort 1 are presented in Progress Reports 1 and 2, which showed no significant differences between the QTR and Control group (we note the disruptions to schooling caused by the COVID-19 pandemic at this time). Cohort 1 student survey results are presented in a separate report. In 2022, 42 Queensland state primary schools participated in the study (Cohort 2).

# 3. Methodology

The Queensland replication study (comprising Cohort 1 and Cohort 2) follows the research design specified by the TTRC team. The TTRC team provided support to replicate the research design and protocols followed by the study in NSW. An overview of the Cohort 2 research design is presented in Table 1.

At the beginning of 2022 the timing of baseline PATs and student surveys was impacted by COVID-19 related absences, the administration of NAPLAN, and in some cases, school closures due to flooding or severe weather events. Subsequently, during Terms 3 and 4, staff shortages in Queensland schools further impacted data collection. As a result, all research activities were completed but not always undertaken within the suggested timeframe.

Although all participating schools completed post-intervention PAT data collection, some schools could not complete in the suggested timeframe of Term 4 weeks 3–6. Data collection was thus extended into the beginning of week 8, until 15th November 2022. This extended timeframe accommodated those teachers who could not complete due to end of year activities including school camps, assessments, and personal leave. The post-intervention student survey was also extended from weeks 3–6, until Tuesday 22nd November (Term 4 week 8). Similarly, the post-intervention teacher survey was also extended until Tuesday 29th November 2022.

*Table 1: Research design (Cohort 2)*

| Activity | Timeframe | Data reported in this progress report |
|---|---|---|
| Baseline assessment | Term 1 and Term 2, 2022 | |
| • PAT-R (ACER) | | ✓ |
| • PAT-M (ACER) | | ✓ |
| • Teacher survey | | ✓ |
| • Student survey | | ✓ |
| Randomisation and group allocation | Term 1, 2022 | |
| • QTR | | ✓ |
| • Control | | ✓ |
| QTR workshop | Term 1, 2022 | ✓ |
| Teacher survey (Interim T2) | Term 2, 2022 | ✓ |
| Fidelity checks | Term 2 and Term 3, 2022 | ✓ |
| Teacher survey (Interim T3) | Term 3, 2022 | ✓ |

| | | |
|---|---|---|
| Post-intervention assessment | Term 4, 2022 | |
| • PAT-R (ACER) | | ✓ |
| • PAT-M (ACER) | | ✓ |
| • Teacher survey | | ✓ |
| • Student survey | | ✓ |

## 3.1 Study sample (Cohort 2)

The recruitment of schools for this study was undertaken by the TTRC team.

### 3.1.1 Schools

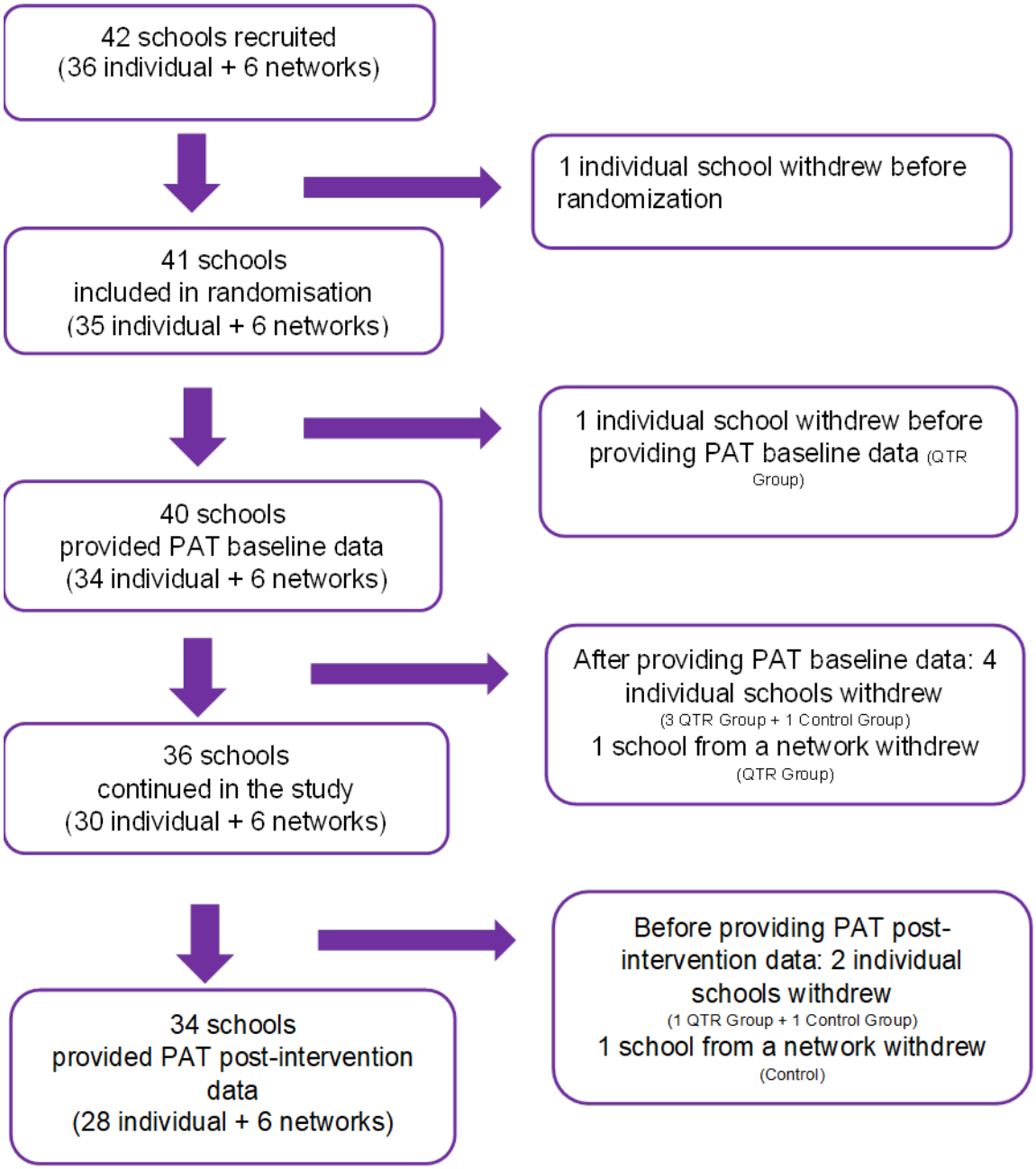
At the commencement of 2022, 42 schools agreed to participate in the study, this included 6 clusters of schools known as 'networks'. For randomisation and analysis purposes, each network, regardless of the number of schools contained within, is counted as 'one school' (as described further in Section 5: Implementation Fidelity). Over the course of 2022, 8 schools withdrew and an additional 2 schools that were part of a network also withdrew.

#### PAT data collection

Instructions on how to administer the PATs, including the links, were sent to each school during Term 1 for baseline and Term 4 for post-intervention PAT assessment. Two QTR schools used their own PAT licence for baseline data collection and 5 QTR schools used their own PAT licence for post-intervention data collection. The TTRC allocated the test levels for each student. Schools that used their own links, allocated their own tests.

None of the 8 individual (non-network) schools that withdrew agreed to collect the post-intervention PAT data. However, one network school that withdrew prior to collecting post-intervention data agreed and proceeded to collect their post-intervention PAT data. At the time of post-intervention data collection, there were 34 schools in the study (this included the 6 networks). See Figure 1 for an overview.

*Figure 1: Overview of school withdrawals throughout the research period*



### 3.1.2 PAT sample

The sample comprised 106 teachers who collected PAT baseline assessment data from their Year 5 or Year 6 class and 87 teachers who collected post-intervention PAT data from their Year 5 or Year 6 class. An overview of the total number of schools, teachers and students at both PAT baseline and PAT post-intervention is presented in Table 2.

*Table 2: Sample of schools, teachers and students (n) who completed PAT assessments*

| Baseline | Schools (n) | Teachers (n) | Students (n) |
|---|---|---|---|
| QTR group | 19 (including 2 networks) | 52 | 590 |
| Control group | 21 (including 4 networks) | 54 | 632 |
| **TOTAL** | **40 (including 6 networks)** | **106** | **1,222*** |
| **Post-intervention** | **Schools (n)** | **Teachers (n)** | **Students (n)** |
| QTR group | 15 (including 2 networks) | 40 | 434 |
| Control group | 19 (including 4 networks) | 46 | 511 |
| **TOTAL** | **34 (including 6 networks)** | **86** | **945*** |

*Notes: ^1 network school did not complete PAT baseline collection as had no consenting students from Year 5 or 6.*

**Students with disabilities were excluded from this total. Student numbers differ between Maths and Reading PAT scores. Please see Appendix A: Flow diagram for students who completed PAT assessments for more details.*

### 3.1.3 School location

The location of schools that either participated in Quality Teaching Rounds professional development immediately after the QTR workshops—referred to as the QTR group—or continued their usual teacher professional development—the Control group—is presented in Table 3. School locations have been defined using the Australian Bureau of Statistics Remoteness Structure.

**QTR group**

Schools that participated in the Quality Teaching Rounds were mainly located in a major city or outer regional/remote. The total includes one network of schools.

**Control group**

Schools that continued their usual teacher professional development were also mainly located in a major city, or outer regional/remote.

*Table 3: Sample by region (n) of schools that completed PAT assessment*

| Baseline Region | QTR group (n)* | Control group (n)* |
|---|---|---|
| Major city | 9 | 9 |
| Inner regional | 3 (1 network=2 schools) | 3 |
| Outer regional and Remote | 7 (1 network=2 schools) | 9 (4 networks n=11 schools) |
| **TOTAL** | **19** | **21** |
| **Post-intervention Region** | **QTR group (n)*** | **Control group (n)*** |
| Major city | 7 | 8 |
| Inner regional | 2 (1 network=2 schools) | 3 |
| Outer regional and Remote | 6 (1 network=2 schools) | 7 (3 networks n=9 schools) |
| **TOTAL** | **15^** | **19** |

*Notes: ^15 control schools completed reading and 14 Maths PAT assessments*

# 4. Data integrity

## 4.1 Data collection

Throughout the project, the ISSR team closely followed the RCT protocols used for the study in NSW and consulted with the UoN team about issues and concerns related to the online survey administration, and PAT data collection. The ISSR team also made a number of recommendations to improve the data collection and these recommended changes were approved by the UoN team prior to implementing. The recommendations made to improve the online survey administration and subsequent responses from the UoN were outlined in the September progress report (Progress Report 3).

## 4.2 Randomisation and group allocation

As a first step, excluding Cohort 1 schools that were part of the RCT in 2021, all other State primary schools across Queensland were invited to participate in the QTR RCT. Schools were informed that, if they were randomly allocated to the Control group, they would receive the QTR intervention later, towards the end of the year. Fifty-two schools consented to be part of the study (one withdrew prior to randomisation). Sixteen schools were classified as outer regional and remote schools, with composite classes and a small number of Year 5 and 6 students. To accommodate these challenges, these 16 schools were clustered into 6 ‘school network’ groups. Counting these ‘school network’ groups and other participating schools, the total number of schools involved in this RCT was 42 (including one school that withdrew).

For this study, the randomisation process was replicated following the procedure used in the QTR trial conducted in NSW (Miller et al., 2019). The randomisation process took place at the school level rather than the classroom level, in the interest of maintaining an unbiased study and anonymity between teachers and students involved in the Control and QTR groups. The schools that consented to be part of the RCT were

stratified based on 2 criteria: (1) urban or outer regional or inner regional school location (3 groups), and (2) socio-economic status using the continuous variable Index of Community Socio-Educational Advantage (ICSEA). These criteria are recognised by the research community, used nationally to classify schools, and publicly available on the Australian Government's My School website (https://www.myschool.edu.au/).

### 4.2.1 Randomisation steps

Randomisation included 4 steps, as follows and summarised in Table 4:

**Step 1**: Stratify schools based on location (i.e., outer regional or inner regional or urban)

**Step 2**: Rank ICSEA in ascending order by location

**Step 3**: Separate schools into blocks with 2 schools in each block

**Step 4**: Use a computerised random number generator to randomly allocate one school per block to the Control group and the other school per block to the QTR (treatment) group.

For school location, 3 strata were created:

1) 'Outer regional' for schools classified as Remote, Very Remote and Outer Regional
2) 'Inner regional' for schools classified as Inner Regional
3) 'Urban' for schools classified as Major Cities.

For Cohort 2, 3 location groups were created because there were more schools located in outer regional and remote locations (n=17, including one school that withdrew).

Next, schools within each location strata were ranked (in ascending order) by ICSEA. Schools were then grouped into blocks with 2 schools per block. Using a computerised random number generator, schools within each block were allocated to one of 2 conditions: QTR or Control. Where blocks had just one school, a dummy school was included to complete the block and maintain the same probability of schools being allocated to one of the 2 conditions. Schools were notified of their group allocation as soon as possible after baseline data collection. This aligns with the preferred method of randomisation in cluster randomised controlled trials (Murray et al., 2004; Miller et al., 2019).

*Table 4: Randomisation and group allocation process - fictitious example*

| School | Step 1 | Step 2 | Step 3 | Step 4 |
|---|---|---|---|---|
| | **Location strata** | **ICSEA** | **Block** | **Groups** |
| School 1 | Inner regional | 989 | 1 | QTR |
| School 2 | Inner regional | 992 | | Control |
| School 3 | Inner regional | 1010 | 2 | Control |
| School 4 | Inner regional | 1020 | | QTR |
| School 5 | Outer regional | 960 | 3 | Control |
| School 6 | Outer regional | 972 | | QTR |
| School 7 | Outer regional | 991 | 4 | Control |
| School 8 | Outer regional | 996 | | QTR |
| School 9 | Outer regional | 1002 | 5 | QTR |
| School 10 | Outer regional | 1005 | | Control |
| School 11 | Outer regional | 1010 | 6 | Control |
| School 12 | Outer regional | 1015 | | QTR |
| School 13 | Urban | 990 | 7 | Control |
| School 14 | Urban | 993 | | QTR |
| School 15 | Urban | 1004 | 8 | QTR |
| School 16 | Urban | 1008 | | Control |
| School 17 | Urban | 1010 | 9 | QTR |
| School 18 | Urban | 1012 | | Control |

Following the randomisation process, Table 5 summarises the group characteristics. There were no significant differences in ICSEA between the QTR and Control groups. Further statistical tests, including the variance ratio test, two-sample Kolmogorov-Smirnov test for equality of distribution functions, and the Mann-Whitney test, confirmed the lack of significant differences between the 2 groups.

*Table 5: Group characteristics post-randomisation*

| Characteristics | QTR group | Control group | Overall | Difference |
|---|---|---|---|---|
| **Schools, N** | **20** | **21** | **41** | |
| ICSEA, mean (SD) | 991.80 (75.89) | 995.05 (81.94) | 993.46 (78.08) | 3.25 t-test (p-value = 0.896) |
| NAPLAN reading, mean (SD) | 503.65 (34.15) | 506.10 (36.53) | 504.90 (34.97) | 2.45 t-test (p-value = 0.826) |
| NAPLAN numeracy, mean (SD) | 483.60 (33.33) | 486.33 (30.54) | 485 (31.56) | 2.73 t-test (p-value = 0.786) |
| NAPLAN numeracy, mean (SD) | 483.60 (33.33) | 486.33 (30.54) | 485 (31.56) | 2.73 t-test (p-value = 0.786) |
| Consenting students, mean (SD) As of 25/03/2022 | 33.60 (18.92) | 36.91 (19.87) | 35.29 (19.25) | 3.30 t-test (p-value = 0.589) |
| School location Outer regional [Inner regional] (Urban) | 9 [3] (8) | 9 [3] (9) | | |

The demographic characteristics of students who consented to be included in the study are reported in Table 6. The demographic proportions are fairly equal between the QTR and Control groups. There are no significant differences in proportions between the QTR and Control groups for students characterised as female, LBOTE and/or with a disability. However, there is a significantly higher share of students who identify as Indigenous in the Control group. The demographic characteristics of students who completed student surveys and PAT assessments are presented in their respective sections (6.2 and 6.3).

*Table 6: Student characteristics at baseline*

| Characteristics | QTR group | Control group | Overall | Proportion test (p-values) |
|---|---|---|---|---|
| **Students, N** | **685** | **774** | **1496** | |
| Female, N (%) | 384 (56.06%) | 410 (53.04%) | 794 (54.46%) | 0.248[2] |
| Indigenous, N (%) | 35 (5.11%) | 61 (7.88%) | 96 (6.58%) | 0.033[2] |
| LBOTE, N (%) | 88 (12.85%) | 90 (11.63%) | 178 (12.20%) | 0.478[2] |
| ICP, N (%) | 0 (0%) | 0 (0.0%) | 0 (0%) | 0.00 |
| Disability, N (%) | 33 (4.83%) | 33 (4.26%) | 66 (4.52%) | 0.611[2] |

*Notes: 1 Characteristic of students who consent to participate in the study*

*2 Two-sample test of proportions*

## 4.3 Data collection activities completed

The data collection schedule was guided by the TTRC research team. An overview of the specific data collection activities is presented in Table 7.

*Table 7: Schedule of data collection activities*

| Time frame | Activity | Responsibility | Key dates and details |
|---|---|---|---|
| Term 1, 2022 | Commencement of Cohort 2 | ISSR | 24/01 - ISSR team members attended Queensland RCT Cohort 2 discussion with UoN, delivered by TTRC |
| Term 1, 2022 | Baseline assessment<br>• PAT-R (ACER)<br>• PAT-M (ACER) | ISSR | 14/02 – 04/03 - PAT licenses set up |
| | | | 14/03 - PAT instructions sent to schools |
| | | | 14/03 – 09/05 - PAT administration (batches 1–15) sent out to schools |
| | | | 15/03 – 17/05 - PAT-R & PAT-M baseline assessments conducted |
| Term 1, 2022 | Randomisation and group allocation<br>• QTR group<br>• Control group | ISSR | 28/03 – 01/04 |
| Term 1, 2022 | QTR workshop | TTRC | 26-27 / 28–29/04 - All new ISSR staff undertaking fidelity checks attended |
| Term 1, 2022 | Student survey (Baseline) | ISSR | 14–15/02 - Surveys uploaded into Qualtrics and tested |
| | | | 14–15/03 - Student survey instructions sent with PAT email and distributed in batches |
| Term 1, 2022 | Teacher survey (Baseline) | ISSR | 2– 5/03 - Surveys uploaded into Qualtrics and tested |

| Time frame | Activity | Responsibility | Key dates and details |
|---|---|---|---|
| | | | Teacher baseline survey distributed sent (reminders sent to all non-responders) |
| Term 1, 2022 | Fidelity checks | TTRC | 30/03 - Qualtrics scheduling survey link sent to QTR schools |
| Term 2, 2022 | Fidelity checks | ISSR | 20/04 – 28/07 - Fidelity check scheduling |
| | | | 09/05 – 28/07 - Confirmation emails sent to schools, and follow up calls to confirm fidelity check details |
| | | | 2/5 – 28/07 - Fidelity checks undertaken*<br>RA fidelity check surveys completed<br>Reminders sent to schools to complete fidelity check PLC survey |
| Term 2, 2022 | Teacher survey (Interim T2) | ISSR | 30/05 - Teacher interim survey distributed (reminders sent to all non-responders) |
| Term 3, 2022 | Teacher survey (interim T3) | ISSR | 29/08 - Teacher T3 survey distributed<br>Reminders sent to all non-responders |
| Term 4, 2022 | Post-intervention PATs | ISSR | 13/10 - PAT instructions sent to schools<br>17/10 – 15/11 - PAT-R & PAT-M post-intervention assessments conducted<br>20/10 – 15/11 - PAT progress checked |
| Term 4, 2022 | Student post-intervention survey | ISSR | 13/10 - Student survey instructions sent with PAT email<br>1/11 – 22/11 - Student survey follow up |
| Term 4, 2022 | Teacher survey (T4) | ISSR | 31/10 - Teacher T4 survey distributed<br>29/11 - Teacher survey closed |
| Term 4, 2022 | Teacher access to PAT results | ISSR | 24/11 - PAT results access information sent to teachers |

*Notes: *One school conducted their reading discussions separately and this part was not observed by ISSR until Thursday 25th August 2022*

## 4.4 Quality Teaching Rounds workshop

The 2-day QTR workshop was attended by all staff from ISSR who conducted the fidelity checks. ISSR staff reported that attendance was beneficial to their understanding of QTR, and was important prior learning for the fidelity check activity. In addition, staff observed that the teachers attending the workshop were engaged with the content and could see the benefit of undertaking QTR in their school.

# 5. Implementation fidelity

Implementation fidelity refers to the extent to which the teachers in each PLC adhered to the protocols for implementing QTR as described in the workshop. When implementing QTR in their schools, a self-appointed volunteer from each PLC (self-appointed) completed a survey to provide details about the activities they engaged in, as part of each Round. In addition to this self-reported data, a research assistant observed all PLC activities at the QTR group schools in either Round 1 or Round 2. Seven schools were scheduled to be visited face-to-face, while the remaining schools (n=12) were to be observed remotely, via Microsoft Teams. The research assistants recorded the activities undertaken against the fidelity criteria but did not provide any assistance with the QTR process, acting solely as an observer or 'fly on the wall'.

These implementation fidelity checks were captured in an online survey format programmed in Qualtrics. The data from the implementation fidelity checks were assessed using the fidelity checklist (Table 8) for all schools in the QTR group.

A stratified randomisation procedure was used to select the schools (QTR group) that would be observed face-to-face (n=7) following the steps below:

- Allocated the schools into 2 strata: a) accessible by car (Major City) and b) accessible by plane (Inner and Outer Regional)
- Calculated distance in km from nearest airport
- Excluded schools > 40km from nearest airport (one school)
- Excluded distance education schools
- Numbered each remaining school
- Used a formula in Excel to generate a random number: =RANDBETWEEN(1,10)
- As per the contract, selected 2 schools from strata a) and 5 schools from strata b)
- Two of the schools selected for face-to-face observation withdrew, the randomisation process was repeated to select the replacement schools. One of the replacement schools withdrew towards the end of fieldwork and could not be replaced. Consequently, 6 schools were observed face-to-face.
- Two of the schools selected for observation via Microsoft Teams withdrew, leaving 10 schools that were observed via Teams.

*Table 8: Implementation fidelity criteria*

| Implementation fidelity criteria |
|---|
| 1. Was a professional reading session conducted? |
| 2. Was a full lesson observed? |
| 3. Were all PLC members in attendance throughout the lesson? |
| 4. Did all PLC members individually code prior to discussion for this Round? |
| 5. Did all PLC members provide their codes and justification (using lesson evidence) for each QT element? |
| 6. Did PLC members take turns leading the discussion of elements during this Round? |
| 7. Was the QT Classroom Practice Guide a consistent point of reference throughout the discussion? |
| 8. Were PLC members (including the observed teacher) present throughout the discussion? |
| 9. How long was the post lesson discussion? (> 60 minutes required for fidelity) |

From the 16 schools observed, only 13 provided at least one self-reported fidelity check survey (2 provided one incomplete survey, and one school did not complete any of the 4 fidelity check surveys). For fidelity check analysis of schools that were observed but did not provide their own, self-reported data, please see Appendix B. The fidelity check analysis includes only those 13 schools that provided self-reported fidelity check data. Twelve schools were observed in Round 1 and one school was observed in Round 2.

Table 9 presents the mean scores for observed and self-reported fidelity checks, and the proportion of schools (observed and self-reported) that coded 100% fidelity (9/9). The mean score was slightly lower for the observed (M = 8.31, SD = 0.95) than for the self-reported (M = 8.70, SD = 0.80). Seven out of the 13 schools achieved 100% for the observed fidelity check and 9 out of the 13 schools achieved 100% for the self-reported fidelity check (achieving 9/9) (Table 9). It should be noted that the distribution differs between the 2 sets of fidelity data—the self-reported fidelity data had a lower minimum (6 for the observed vs. 5 for the self-reported) and the proportion of those meeting 8 out of 9 criteria was higher in the observed sessions (4 for the observed vs. 2 for the self-reported), thus while the means are quite similar, the proportion of sessions meeting 100% fidelity is not. If any PLC member did not provide their codes or justifications for any of the 18 elements, this was considered not to be met when applying the pre-defined fidelity cut points. In both the self-reported and observed data, the criteria "*Were all PLC members in attendance throughout the lesson?*" was most frequently unmet.

*Table 9: Fidelity of implementation*

| Outcome | QTR group |
|---|---|
| **Fidelity Score** | |
| Observed, mean (SD) | 8.3 (.95) |
| Self-reported, mean (SD) | 8.7 (.80) |
| Fidelity 9/9 (100% fidelity) | |
| Observed, % | 54% |
| Self-reported, % | 69% |

# 6. Data analysis

## 6.1 Teacher survey data

The teacher survey aimed to measure change over time in Teacher Morale, Teacher Appraisal and Recognition, Efficacy for Student Engagement, Teacher's Self-Efficacy and School Connectedness. Teachers completed 4 surveys over the 2022 school year. Overall, 155 teachers completed at least one survey and 108 completed both baseline and post-intervention surveys. Only teachers who completed both baseline and post-intervention surveys (balanced panel) were included in the following analysis. A sensitivity analysis was run which included all teachers with valid responses at baseline, even if they didn't answer both surveys (unbalanced panel) with no significant differences found (Appendix C: Sensitivity analyses teacher survey). Table 10 presents the balanced panel (baseline and post-intervention) over the 4 time points, by condition.

*Table 10: Number of teachers **included in the analyses** who completed teacher survey by time and condition*

| Group | Teacher baseline survey (n) | Teacher T2 survey (n) | Teacher T3 survey (n) | Teacher post-intervention survey (n) |
|---|---|---|---|---|
| **QTR** | 49 | 43 | 40 | 49 |
| **Control** | 59 | 51 | 48 | 59 |
| **Total** | 108 | 94 | 88 | 108 |

### Teacher Morale

Teacher Morale positively affects student academic achievement (Abazaoglu & Aztekin, 2016). Teacher Morale was examined using 5 items (Hart et al., 2000), as follows: (1) There is good team spirit in this school, (2) The morale in this school is high, (3) Teachers go about their work with enthusiasm, (4) Teachers take pride in this school, (5) There is a lot of energy in this school. Teachers were asked to respond to each item on a 5-point scale, ranging from *strongly disagree* to *strongly agree*. All items loaded onto one factor and a Teacher Morale index was then constructed, calculating the mean of the 5 items (alpha= 0.94[1]).

[1] Cronbach's alpha test

Teachers who had missing information on any items of the Teacher Morale scale did not receive an index score.

As presented in Table 11, teachers from the Control group reported a higher mean for Teacher Morale at baseline compared to teachers from the QTR group. Morale of teachers declined for both groups over the school year, and this decline was significant for the Control group, but not the QTR group. There were no significant differences in Teacher Morale between the QTR and Control groups in the difference in difference analyses.

### Teacher Appraisal and Recognition

Teacher appraisal is an important element in improving teaching quality (Elliott, 2015) and teacher recognition has been identified as a motivating factor in job performance (Mertler, 2016). Teacher Appraisal and Recognition was examined using 6 items (Hart et al., 2000), as follows: (1) I am encouraged in my work by praise, thanks or other recognition, (2) I have the opportunity to discuss and receive feedback on my work performance, (3) I am regularly given feedback on how I am performing my role, (4) There is a structure and process that provides feedback on my work performance, (5) I am happy with the quality of feedback I receive on my work performance, (6) Staff receive recognition for good work. Teachers were asked to respond to each item on a 5-point scale, ranging from *strongly disagree* to *strongly agree*. All items loaded onto one factor and a Teacher Appraisal and Recognition index was then constructed, calculating the mean of the 6 items (alpha= 0.93[1]). Teachers who had missing information on any items for Teacher Appraisal and Recognition did not receive an index score.

As shown in Table 11, a higher mean of Teacher Appraisal and Recognition was observed among teachers from the Control group at baseline, when compared to teachers from the QTR group. Teacher Appraisal and Recognition declined for both groups over the school year. The Control group showed a greater decline of Teacher Appraisal and Recognition during the school year than the QTR group, and this reached significance for the Control group but not the QTR group. There were no significant differences in teacher appraisal between the QTR and Control groups in the difference in difference analyses.

### Efficacy for Student Engagement

Teacher Efficacy for Student Engagement is important because students who are more engaged have better school attendance and achievement levels (Shoulders & Krei, 2015). Efficacy for Student Engagement was examined using 4 items (Tschannen-Moran & Woolfolk Hoy, 2001) as follows: (1) How much can you do to motivate students who show low interest in school work?, (2) How much can you do to get students to believe they can do well in school work?, (3) How much can you do to help your students value learning?, (4) How much can you do to assist families in helping their children do well in school? Teachers were asked to respond to each item on a 9-point scale, ranging from *nothing* to *a great deal*. All items loaded onto one factor and an Efficacy for Student Engagement index was then constructed, calculating the mean of the 4 items (alpha= 0.90). Teachers who had missing information on any items for Efficacy for Student Engagement did not receive an index score.

As shown in Table 11, teachers from the Control group reported a higher mean for Efficacy for Student Engagement index at baseline compared to teachers from the QTR group. The mean of this index increased for both groups between baseline and post-intervention. There was a greater increase observed in the QTR group when compared to the Control group, however this increase was not significant for either group. There were no significant differences in Efficacy for Student Engagement between the QTR and Control groups in the difference in difference analyses.

### Teachers' Self-Efficacy

Teacher self-efficacy has been widely studied in education and has been recognised as an influence on student achievement and behaviour (Yoo, 2016). Teachers' Self-Efficacy was examined using 4 items (Mankin et al., 2017). Teachers were asked to answer questions about their experiences as a teacher, and to choose the one response that best described how they felt in the past month: (1) I am a successful teacher, (2) I am good at helping students learn new things, (3) I have accomplished a lot as a teacher (4) I feel like my teaching is effective and helpful. Teachers were asked to respond to each item on a 4-point Likert-type scale ranging from *almost never* to *almost always*. All items loaded onto one factor and a teachers' self-efficacy index was then constructed, calculating the mean of the 4 items (alpha= 0.91[1]). Teachers who had missing information on any items for the Teachers' Self-Efficacy did not receive an index score.

As shown in Table 11, teachers from the Control group reported a higher mean for the Teachers' Self-Efficacy at baseline compared to teachers from the QTR group. The mean of this index increased between baseline and post-intervention for both groups, however this increase was not significant for either group. There were no significant differences in Teachers' Self-Efficacy between the QTR and Control groups in the difference in difference analyses.

### School Connectedness

A sense of connectedness and strong social relationships in school environments has been demonstrated to positively affect the health of children and adolescents and contributes to academic achievement and engagement (Rowe et al., 2007). School Connectedness was examined using 4 items (Mankin et al., 2017), as follows: (1) I feel like I belong at this school, (2) I can really be myself at this school, (3) I feel like people at this school care about me, (4) I am treated with respect at this school. Teachers were asked to respond to each item on a 4-point Likert-type scale ranging from *almost never* to *almost always*. All items loaded onto one factor and a school connectedness index was then constructed, calculating the mean of the 4 items (alpha= 0.93[1]). Teachers who had missing information on any items for School Connectedness did not receive an index score.

Teachers from the Control group reported a higher mean for School Connectedness at baseline when compared to teachers from the QTR group, as presented in Table 11. The mean of this index decreased for both groups over the school year but was not significant for either group. There were no significant differences in School Connectedness between the QTR and Control groups in the difference in difference analyses.

*Table 11: Teacher outcome by condition and time*

| Group | Time | | | | | | | | T-test | | Difference in difference Clustered by school | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Baseline | | T2 | | T3 | | Post-intervention | | Dif. between post-intervention and baseline | Paired t-test | Coef. (CI 95%) | P-value |
| | n | Mean (SD) | n | Mean (SD) | n | Mean (SD) | n | Mean (SD) | | | | |
| **Teacher Morale** | | | | | | | | | | | | |
| QTR | 49 | 4.02 (0.94) | 43 | 3.94 (0.94) | 40 | 3.94 (0.98) | 49 | 3.84 (1.03) | -0.184 | 0.160 | 0.02 (-0.35, 0.39) | 0.899 |
| Control | 59 | 4.14 (0.72) | 51 | 3.81 (0.98) | 48 | 3.99 (0.88) | 59 | 3.94 (0.99) | -0.203 | **0.016** | | |
| **Teacher Appraisal and Recognition** | | | | | | | | | | | | |
| QTR | 49 | 3.77 (1.04) | 43 | 3.73 (0.98) | 40 | 3.68 (1.03) | 49 | 3.67 (1.09) | -0.099 | 0.466 | 0.17 (-0.19, 0.52) | 0.346 |
| Control | 59 | 3.80 (0.89) | 51 | 3.50 (1.08) | 49 | 3.70 (0.89) | 59 | 3.55 (1.07) | -0.254 | **0.006** | | |
| **Efficacy for Student Engagement** | | | | | | | | | | | | |
| QTR | 49 | 6.82 (1.20) | 43 | 6.77 (1.23) | 40 | 7.07 (1.19) | 49 | 6.98 (1.10) | 0.105 | 0.126 | 0.13 (-0.21, 0.47) | 0.451 |
| Control | 59 | 7.05 (1.22) | 51 | 6.93 (1.36) | 49 | 7.00 (1.17) | 59 | 7.11 (1.35) | 0.064 | 0.649 | | |
| **Teachers' Self-Efficacy** | | | | | | | | | | | | |
| QTR | 49 | 3.13 (0.50) | 43 | 3.09 (0.59) | 40 | 3.24 (0.55) | 49 | 3.19 (0.53) | 0.061 | 0.336 | 0.01 (-0.14, 0.15) | 0.944 |
| Control | 59 | 3.20 (0.60) | 51 | 3.20 (0.62) | 49 | 3.23 (0.56) | 59 | 3.27 (0.61) | 0.068 | 0.322 | | |
| **School Connectedness** | | | | | | | | | | | | |
| QTR | 49 | 3.41 (0.69) | 43 | 3.42 (0.66) | 40 | 3.36 (0.73) | 49 | 3.24 (0.76) | -0.173 | 0.072 | -0.03 (-0.30, 0.24) | 0.818 |
| Control | 59 | 3.45 (0.62) | 51 | 3.36 (0.65) | 49 | 3.32 (0.63) | 59 | 3.31 (0.68) | -0.140 | 0.098 | | |

## 6.2 Student survey analysis (Cohort 2)

To measure change over time, a student survey was administered at 2 timepoints, baseline and post-intervention, that were between 5 and 8 months apart. The survey included questions related to self-efficacy and enjoyment of schoolwork overall, as well as specifically related to the learning areas of English and Maths. Eight outcomes were measured: Quality of School Life–Achievement, Reading (frequency), Reading Self-Efficacy, Reading Enjoyment, Reading Comprehension, Maths Self-Efficacy, Maths Interest, Maths Anxiety. To determine whether there were changes in the outcomes over time, difference in difference models were fitted to compare student outcomes between students from the QTR group and students from the Control group. Both models clustered standard errors by school, the first model was not adjusted for any covariates and the second model was adjusted for student characteristics (gender, Indigenous status and LBOTE).

Only students who completed both surveys were included in the analyses. Table 12 presents student characteristics for Cohort 2. No significant differences in gender, Indigenous status and LBOTE were found between students who completed both surveys and those who completed only one. Sensitivity analyses were run in the unbalanced panel and no significant differences were found between the unbalanced and balanced panels (Appendix D: Sensitivity analyses student survey).

*Table 12: Summary characteristics at student, class and school level (balanced panel)*

| Characteristics | QTR group | Control group | Overall | Dif. Baseline between groups T-test (p-values) |
|---|---|---|---|---|
| **Students, N** | **261** | **324** | **585** | |
| Female, N (%) | 150 (57.47) | 173 (53.40) | 323 (55.21) | 0.3244[1] |
| Indigenous, N (%) | 11 (4.21) | 21 (6.48) | 32 (5.47) | 0.2307[1] |
| LBOTE, N (%) | 38 (14.56) | 32 (9.88) | 70 (11.97) | 0.0828[1] |
| **Class, N** | **33** | **40** | **73** | |
| Students per class, means (SD) | 7.90 (5.62) | 8.1 (4.42) | 8.01 (4.97) | 0.6602[2] |
| Classrooms per school, means (SD) | 2.2(1.01) | 2.11 (0.99) | 2.15 (0.99) | 0.7783[2] |
| **Schools, N** | **15** | **19** | **34** | |
| Students per school, means (SD) | 17.4 (10.57) | 17.05 (12.69) | 17.21 (11.63) | 0.6768[2] |

*Notes: [1] Two-sample test of proportions*

*[2] Two-sample Mann Whitney test*

** The class size and school size are calculated with the students that provided valid responses in student survey and not the total of students per class.*

### 6.2.1 Outcomes

#### Quality of School Life–Achievement

Quality of School Life–Achievement was examined using 5 items, as follows: (1) I am success as a student, (2) I know how to cope with the work, (3) I am good at school work, (4) I know I can keep up with the work, (5) I achieve a satisfactory standard. Students were asked to respond to each item on a 4-point scale, ranging from *definitely disagree* to *definitely agree* (Ainley & Bourke, 1992). All items loaded onto one factor, a mean of these items was used to calculate the Quality of School Life–Achievement Index score (alpha=0.86[2]).

Indices that use similar items are commonly referred to in the literature as Academic self-efficacy. Academic self-efficacy defines student's self-judgements about their ability to attain their educational goals and is positively related to academic performance (Honicke & Broadbent, 2016).

Students from the QTR group reported a higher mean for this outcome at baseline compared to students from the Control group. Looking at change over time, no significant changes were found in the mean Quality of School Life–Achievement index when comparing post-intervention with baseline mean scores for either group. For the difference in difference analyses, there were no significant differences in Quality of School Life–Achievement index between the QTR and Control groups (Table 13).

#### Reading (frequency)

Research has shown a positive relationship between reading frequency, reading enjoyment and attainment (Clark, 2011; Clark & Douglas, 2011). Reading was examined using one question: How many books do you read each year (please give your best guess). Students were asked to respond on a 5-point scale, ranging from *none* to *more than 50*.

Students from the QTR group reported a lower mean for the amount of books they read each year at baseline compared to students from the Control group. Examining the change over time, both groups (QTR and Control) reported a significant decrease in the amount of books read each year. For the difference in difference analyses, there were no significant differences in Reading between the QTR and Control groups (Table 13).

#### Reading Self-Efficacy[3]

Reading self-efficacy is positively related to reading achievement (Retelsdorf et al., 2014). Reading Self-Efficacy was examined by asking the students: How good are you at reading? Students were asked to respond to on a 10-point scale, ranging from *not good at all*, to *really good*, and *not at all*, to *very much*, respectively.

Students from the QTR group reported similar means for Reading Self-Efficacy at baseline compared to students from the Control group. There were no significant changes over the year in relation to Reading Self-

[2] Cronbach's alpha test

Efficacy for either group. For the difference in difference analyses, there were no significant differences in Reading Self-Efficacy between the QTR and Control groups (Table 13).

### Reading Enjoyment[3]

Researchers have found a positive association between reading enjoyment and performance (Malanchini et al., 2017). Reading Enjoyment was examined by asking students: How much do you enjoy reading? Students were asked to respond to each item on a 10-point scale, ranging from *not at all*, to *very much*.

Students from the QTR group reported a lower mean for Reading Enjoyment at baseline compared to students from the Control group. Examining the change over time, both QTR and Control groups had a significantly lower mean at post-intervention. For the difference in difference analyses, there were no significant differences in Reading Enjoyment between the QTR and Control groups (Table 13).

### Reading Comprehension

Students with better reading comprehension are more likely to choose to read, which can further improve reading and academic performance (Žolgar-Jerkovic et al., 2018). Reading Comprehension was examined using 4 items, as follows: (1) How easily do you understand what you are reading? (2) How good are you at figuring out key points in what you read? (3) How good are you at understanding the hidden meaning in texts? (4) How good are you at writing persuasively? Students were asked to respond to each item on a 10-point rating scale, ranging from *not easily* to *very easily* for item one, and *not good at all* to *really good* for items 2 to 4. All items loaded onto one factor and a Reading Comprehension index was then constructed using the mean of the 4 items (alpha=0.86[2]). Students who had missing values on any items were not included in the analysis.

Students from the QTR group reported a lower mean for Reading Comprehension at baseline compared to students from the Control group. Looking at change over time, the mean score for the QTR group significantly increased at post-intervention. For the difference in difference analyses, there were no significant differences in Reading Comprehension index between the QTR and Control groups (Table 13).

### Maths Self-Efficacy

Maths Self-Efficacy has been positively linked to Maths achievement levels (Parker et al., 2014; Schöber et al., 2018). Maths Self-Efficacy was examined using 8 items, as follows: How confident do you feel about having to do the following Maths tasks: (1) adding 2 numbers in the hundreds, (2) subtracting 2 numbers in the hundreds, (3) multiplying any number by 2, (4) multiplying any number by 7, (5) understanding graphs presented in newspapers, (6) changing measuring units from centimetres to metres, (7) identifying shapes by the by the number of sides they have (for example a triangle or hexagon), (8) calculating the decimal value of a simple fraction like ¾. Students were asked to respond to each item on a 4-point scale, ranging from *not at all confident* to *very confident*. All items loaded onto one factor and a student Maths Self-Efficacy

[3] Reading **self-efficacy** and enjoyment were not able to load in one factor (alpha <0.7)

index was then constructed using the mean of the 8 items (alpha=0.85[2]). Students who had missing values on any items were not included in the analysis.

Students from the QTR group reported a similar mean index for Maths Self-Efficacy at baseline compared to students from the Control group. Examining the change over time, both the QTR and Control groups had a significantly higher mean index at post-intervention. For the difference in difference analyses, there were no significant differences in the Maths Self-Efficacy index between the QTR and Control groups (Table 13).

### Maths Interest

Students who enjoy Maths are more likely to achieve well in Maths (Pinxten et al., 2014). Maths Interest was examined using 3 items, as follows: (1) I look forward to my Mathematics lessons, (2) I do Mathematics because I enjoy it, (3) I am interested in the things I learn in Mathematics. Students were asked to respond to each item on a 4-point scale, ranging from *strongly disagree* to *strongly agree*. All items loaded onto one factor and a student Maths Interest index was then constructed using the mean of the 3 items (alpha=0.90[2]). Students who had missing information on any items were not included in the analysis.

Students from the QTR group reported a similar mean index for Maths Interest at baseline as students from the Control group. Examining the change over time, both the QTR and Control groups had a significantly lower mean index at post-intervention. For the difference in difference analyses, there were no significant differences in Maths Interest index between the QTR and Control groups (Table 13).

### Maths Anxiety

Students with Maths anxiety are more likely to achieve lower scores on Maths tests than those who do not have Maths anxiety (Barroso et al., 2020). Maths Anxiety was examined using 5 items, as follows: (1) I often worry that it will be difficult for me in Mathematics classes, (2) I get very stressed when I have to do Mathematics homework, (3) I get very nervous doing Mathematics problems, (4) I feel helpless when doing a Mathematics problem, (5) I worry that I will get poor marks in Mathematics. Students were asked to respond to each item on a 4-point scale, ranging from *strongly disagree* to *strongly agree*. All items loaded into one factor and a student's Maths Anxiety index was then constructed using the mean of the 4 items (alpha=0.90[2]). Students who had missing values on any items were not included in the analysis.

Students from the QTR group reported a lower mean index for Maths Anxiety at baseline compared to students from the Control group. Examining the change over time, there were no significant changes over the year for Maths Anxiety. For the difference in difference analyses, there were no significant differences for the Maths Anxiety mean index between the QTR and Control groups (Table 13).

*Table 13: Student survey outcomes by condition and time*

| 2022 | Time | | | | T-test | | Difference in difference Clustered by school | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Group | Baseline | | Post-intervention | | Dif. between post-intervention and baseline | Paired t-test | Coef. (CI 95%) | P-value | Coef. ^(CI 95%) adjusted | P-value |
| | n | Mean (SD) | n | Mean (SD) | | | | | | |
| **Quality of School Life–Achievement** | | | | | | | | | | |
| QTR | 261 | 3.23 (0.57) | 261 | 3.26 (0.58) | 0.028 | 0.533 | 0.010 (-0.093, 0.114) | 0.839 | 0.007 (-0.095, 0.109) | 0.891 |
| Control | 324 | 3.20 (0.64) | 324 | 3.22 (0.64) | 0.649 | 0.204 | | | | |
| **Reading (frequency)** | | | | | | | | | | |
| QTR | 260 | 3.59 (1.13) | 260 | 3.40 (1.14) | -0.185 | **0.002** | -0.010 (-0.158, 0.138) | 0.890 | -0.025 (-0.177, 0.127) | 0.741 |
| Control | 321 | 3.70 (1.10) | 321 | 3.53 (1.13) | -0.174 | **0.001** | | | | |
| **Reading Self-Efficacy** | | | | | | | | | | |
| QTR | 260 | 7.70 (1.97) | 260 | 7.79 (1.85) | 0.088 | 0.432 | 0.129 (-0.146, 0.404) | 0.348 | 0.120 (-0.128, 0.398) | 0.387 |
| Control | 321 | 7.71 (2.13) | 321 | 7.67 (2.19) | -0.040 | 0.704 | | | | |
| **Reading Enjoyment** | | | | | | | | | | |
| QTR | 260 | 6.95 (2.92) | 260 | 6.61 (2.92) | -0.338 | **0.034** | 0.347 (-0.131, 0.825) | 0.150 | 0.361 (-0.137, 0.859) | 0.149 |
| Control | 321 | 7.41 (2.79) | 321 | 6.73 (3.07) | -0.685 | **<0.001** | | | | |
| **Reading Comprehension** | | | | | | | | | | |
| QTR | 260 | 8.43 (1.96) | 260 | 8.68 (1.88) | 0.253 | **0.021** | 0.271 (-0.063, 0.605) | 0.109 | 0.268 (-0.078, 0.614) | 0.125 |
| Control | 320 | 8.50 (2.04) | 320 | 8.48 (2.03) | -0.018 | 0.863 | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Maths Self-Efficacy** | | | | | | | | | | |
| QTR | 259 | 3.21 (0.55) | 259 | 3.32 (0.50) | 0.119 | **<0.001** | 0.051 (-0.034, 0.137) | 0.230 | 0.051 (-0.037, 0.138) | 0.248 |
| Control | 319 | 3.22 (0.52) | 319 | 3.28 (0.54) | 0.067 | **0.004** | | | | |
| **Maths Interest** | | | | | | | | | | |
| QTR | 259 | 2.94 (0.80) | 259 | 2.81 (0.85) | -0.127 | **0.011** | 0.023 (-0.155, 0.201) | 0.794 | 0.028 (-0.157, 0.213) | 0.758 |
| Control | 319 | 2.93 (0.77) | 319 | 2.78 (0.82) | -0.150 | **<0.001** | | | | |
| **Maths Anxiety** | | | | | | | | | | |
| QTR | 259 | 1.96 (0.81) | 259 | 2.03 (0.81) | 0.062 | 0.171 | 0.079 (-0.071, 0.228) | 0.292 | 0.093 (-0.045, 0.232) | 0.180 |
| Control | 319 | 2.01 (0.79) | 319 | 1.99 (0.79) | -0.017 | 0.646 | | | | |

*Notes: ^Adjusted for gender, Indigenous status and LBOTE *p<0.05*

## 6.3 Progressive Achievement Test (PAT) data

### 6.3.1 Exclusion criteria

PAT scores from students who reported a disability were excluded from this report as they may not respond to the intervention due to these factors. Students who scored ceiling at baseline (i.e., scored in the 99th percentile) were removed from the analysis, due to the limitation to measure their improvements over time. Students who only presented post-intervention PAT scores were removed from the analyses. Details on the numbers of students removed from the analysis are presented in Appendix A: Flow diagram for students who completed PAT assessments.

### 6.3.2 Summary characteristics at student, class and school level

Students who completed baseline and post-intervention (balanced panel) were included in all analyses (excluding schools that withdrew from the intervention). At a student level there were no significant differences in gender, Indigenous status and LBOTE between students from schools that withdrew from the study and from those schools that continued.

Table 14 summarises student, class and school level summary statistics of students from the balanced panel by condition. At a student level there were no significant differences in Indigenous status and LBOTE between the Control and QTR groups. There was a significant difference in gender with a greater proportion of females in the QTR group compared to the Control group. There were no significant differences in class and school characteristics between the Control and QTR groups. Adjusting or controlling for key demographic characteristics in data analysis was required to account for gender differences.

*Table 14: Summary characteristics at student, class and school level (Maths and Reading balanced panel)*

| Maths | Maths | | | Dif. Baseline between groups | Reading | | | Dif. Post-intervention between groups |
|---|---|---|---|---|---|---|---|---|
| **Characteristics** | **QTR group** | **Control group** | **Overall** | **T-test (p-values)** | **QTR group** | **Control group** | **Overall** | **T-test (p-values)** |
| **Students, N** | **361** | **440** | **801** | | **355** | **423** | **778** | |
| Female, N (%) | 213 (59.00) | 228 (51.82) | 441 (55.06) | **0.042**[1] | 209 (58.87) | 219 (51.77) | 428 (55.01) | **0.047**[1] |
| Indigenous, N (%) | 20 (5.54) | 36 (8.18) | 56 (6.99) | 0.145[1] | 16 (4.51) | 33 (7.80) | 49 (6.30) | 0.060[1] |
| LBOTE, N (%) | 45 (12.47) | 50 (11.36) | 95 (11.86) | 0.631[1] | 51 (14.37) | 51 (12.06) | 102 (13.11) | 0.342[1] |
| **Class, N**[3] | **37** | **46** | **83** | | **38** | **45** | **83** | |
| Students per class, means (SD) | 9.76 (5.10) | 9.57 (5.13) | 9.65 (5.09) | 0.905[2] | 9.34 (5.18) | 9.4 (4.95) | 9.37 (5.03) | 0.766[2] |
| Classrooms per school, means (SD) | 2.64 (1.15) | 2.42 (0.90) | 2.52 (1.00) | 0.588[2] | 2.53 (1.06) | 2.37 (0.96) | 2.44 (0.99) | 0.548[2] |
| **Schools, N**[4] | **14** | **19** | **33**[5] | | **15** | **19** | **34** | |
| Students per school, means (SD) | 25.79 (14.75) | 23.16 (14.24) | 24.27 (14.29) | 0.662[2] | 23.67 (15.43) | 22.26 (14.04) | 22.88 (14.46) | 0.822[2] |

*Notes: [1] Two-sample test of proportions*

*[2] Two-sample Mann Whitney test*

*[3]Students per class represent the number of students in the class that consented to participate in this research and sat the PAT test.*

*[4]Students per school represent the number of students in the school that consented to participate in this research and sat the PAT test.*

*[5]Networks count as one school.*

### 6.3.3 PAT scale score – Maths and Reading

Student achievement in Maths and Reading was measured at baseline and post-intervention assessment. Test scaled scores were used for the analyses, as per ACER protocol/recommendations. These scaled scores adjusted for schools who allocated the same test level at baseline and post-intervention. Table 15 presents the mean scores for Maths and Reading. When examining the average differences in PAT scores at baseline (Table 17), there is no significant difference between the QTR and Control groups for either Maths or Reading PAT scores. Examining the change over time for all students, it was found that mean Maths PAT scores at post-intervention improved significantly when compared to baseline Maths PAT scores. However, mean Reading PAT scores improved significantly for all year 5 students and year 6 students from the QTR group and decreased significantly for year 6 students from the Control group.

In order to determine whether there were significant differences between the 2 groups (QTR and Control) over time, difference in difference and ANCOVA models were fitted to compare PAT scores between students from the QTR intervention group and students from the Control group. For the difference in difference analyses, 4 different models were run (Table 16). The first model clustered standard errors by school, the second model clustered by school and controlled for student characteristics (gender, Indigenous status and LBOTE), the third model clustered by school and controlled for student characteristics and school characteristics (location and ICSEA scores), the fourth and last model consisted of a multilevel analysis by school and class controlled by student and school characteristics. There were no significant differences in Maths PAT scores between the QTR and Control groups for the difference in difference analyses (Table 16). However, students from the QTR group present a significant greater improvement over time for all models when compared to those from the Control group. ANCOVA analyses showed no significant differences between Maths PAT scores at post-intervention across groups (QTR and Control) when taking baseline scores into consideration. However, significant differences were found in Reading PAT scores for ANCOVA analysis (Table 17). Schools were then categorised by ICSEA scores and difference in difference analysis models were run clustered by school (Table 18). No significant differences were found for Maths or Reading PAT scores between the QTR and Control groups for the difference in difference analyses by ICSEA groups. Similar results were found in sensitivity analyses ran for all students who completed baseline (using an unbalanced panel), including baseline PAT scores of schools that withdrew from the study (Table 19).

*Table 15: Maths and Reading PAT scores*

| Outcome | Year level | Group | N | Baseline mean (SD) | Post-intervention mean (SD) | Difference at baseline (QTR-Control) | Difference between post-intervention and baseline Difference and paired T-test p(values) |
|---|---|---|---|---|---|---|---|
| **Maths** | | | | | | | |
| Scale score | All | QTR | 361 | 124.12 (11.43) | 128.64(11.55) | 1.07 | 4.52 p-value < 0.0001 |
| | All | Control | 440 | 123.05 (11.09) | 127.31 (10.86) | p-value = 0.182 | 4.26 p-value < 0.0001 |
| Scale score | 5 | QTR | 154 | 121.70 (9.44) | 124.92 (10.24) | 0.44 | 3.22 p-value < 0.0001 |
| | 5 | Control | 202 | 121.26 (10.07) | 124.72 (10.96) | p-value = 0.672 | 3.46 p-value < 0.0001 |
| Scale score | 6 | QTR | 207 | 125.91 (12.43) | 131.42 (11.72) | 1.34 | 5.51 p-value < 0.0001 |
| | 6 | Control | 238 | 124.57 (11.69) | 129.51 (10.30) | p-value = 0.241 | 4.94 p-value < 0.0001 |
| **Reading** | | | | | | | |
| Scale score | All | QTR | 355 | 123.75 (12.22) | 128.06 (12.22) | 0.61 | 4.31 p-value < 0.0001 |
| | All | Control | 423 | 123.14 (11.69) | 124.37 (12.00) | p-value = 0.488 | 1.23 p-value = 0.019 |
| Scale score | 5 | QTR | 143 | 119.91 (10.11) | 126.01 (11.28) | -0.42 | 6.1 p-value < 0.001 |
| | 5 | Control | 194 | 120.33 (11.27) | 124.95 (11.42) | p-value 0.727 | 4.62 p-value < 0.001 |
| Scale score | 6 | QTR | 211 | 126.37 (12.85) | 129.46 (12.66) | 0.85 | 3.09 p-value < 0.001 |
| | 6 | Control | 229 | 125.52 (11.53) | 123.87 (12.48) | p-value = 0.468 | -1.65 p-value = 0.022 |

*Table 16: Maths and Reading PAT scores analyses – balanced panel*

| Outcome | n | Difference in difference Clustered by school Coef. (CI 95%) | Difference in difference Clustered by school Controlling for student characteristics* Coef. (CI 95%) | Difference in difference Clustered by school Controlling for student and school characteristics** | Difference in difference Multilevel analysis school and class Controlling for student and school characteristics** |
|---|---|---|---|---|---|
| **Maths** | | | | | |
| All | 801 | 0.27 (-1.32, 1.86) p = 0.733 | 0.28 (-1.33, 1.88) p = 0.729 | 0.28 (-1.33, 1.88) p =0.729 | 0.27 (-1.59, 2.13) p =0.777 |
| Year 5 | 356 | -0.24 (-2.15, 1.66) p= 0.795 | -0.24 (-2.15, 1.66) p = 0.796 | -0.24 (-2.15, 1.66) p = 0.796 | -0.24 (-2.89, 2.40) p = 0.857 |
| Year 6 | 445 | 0.57 (-1.71, 2.85) p = 0.616 | 0.57 (-1.72, 2.86) p = 0.616 | 0.57 (-1.72, 2.86) p = 0.617 | 0.57 (-1.94, 3.08) p = 0.657 |
| **Reading** | | | | | |
| All | 778 | 3.08 (0.17, 6.00) p =**0.039** | 3.08 (0.16, 6.01) p =**0.039** | 3.08 (0.16, 6.01) p = **0.039** | 3.08 (1.02, 5.14) p =**0.003** |
| Year 5 | 338 | 1.48 (-1.57, 4.53) p = 0.331 | 1.48 (-1.58, 4.53) p = 0.332 | 1.48 (-1.59, 4.54) p = 0.332 | 1.48 (-1.58, 4.53) p = 0.343 |
| Year 6 | 440 | 4.74 (1.88, 7.61) p = **0.002** | 4.74 (1.87, 7.61) p = **0.002** | 4.74 (1.87, 7.62) p= **0.002** | 4.74 (2.00, 7.48) p= **0.001** |

*Notes: *gender, LBOTE, Indigenous status*

*** gender, LBOTE, Indigenous status, location ICSEA*

*Table 17: Maths PAT scores ANCOVA analysis*

| ANCOVA | Effect size | P-value | [95% conf. | interval] |
|---|---|---|---|---|
| **Maths n=801** | | | | |
| Condition/group | 0.001 | 0.306 | 0.000 | 0.62 |
| **Reading n=778** | | | | |
| Condition/group | 0.031 | **<0.001** | 0.012 | 0.059 |

*Table 18: Maths and Reading PAT scores analyses by ICSEA groups – balanced panel*

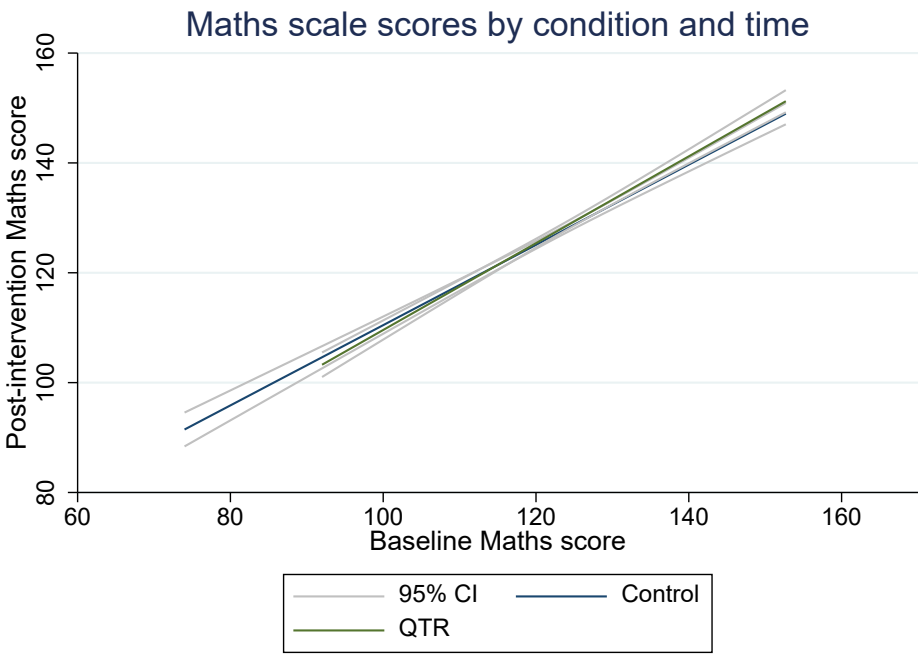
| Outcome | n | Difference in difference Clustered by school ICSEA <950 n=123 Coef. (CI 95%) | n | Difference in difference Clustered by school ICSEA 950-1050 n=352 Coef. (CI 95%) | n | Difference in difference Clustered by school ICSEA >1050 n=326 |
|---|---|---|---|---|---|---|
| **Maths** | 123 | -0.63 (-5.19, 3.93) p-value= 0.757 | 352 | -0.31 (-2.41, 1.78) p-value=0.758 | 326 | 1.25 (-1.79, 4.29) p-value=0.362 |
| **Reading** | 118 | 2.88 (-2.95, 8.70) p-value=0.281 | 340 | 4.18 (-0.37, 8.73) p-value=0.070 | 320 | 2.49 (-2.56, 7.53) p-value=0.282 |

*Table 19: Maths and Reading PAT scores – sensitivity analyses unbalanced*

| Outcome | n | Difference in difference Clustered by school Coef. (CI 95%) | Difference in difference Clustered by school Controlling for student characteristics* Coef. (CI 95%) | Difference in difference Clustered by school Controlling for student and school characteristics** |
|---|---|---|---|---|
| **Maths** | 1188 | 1.13 (-2.00, 4.26) p-value=0.472 | 1.19 (-1.88, 4.27) p-value=0.439 | 0.76 (-0.99, 2.52) p-value=0.384 |
| **Reading** | 1160 | 3.97 (0.24, 7.70) p-value=**0.038** | 3.84 (0.27, 7.41) p-value=**0.035** | 3.14 (0.42, 5.86) p-value=**0.024** |

*Notes: *gender, LBOTE, Indigenous status*

*** gender, LBOTE, Indigenous status, location ICSEA*



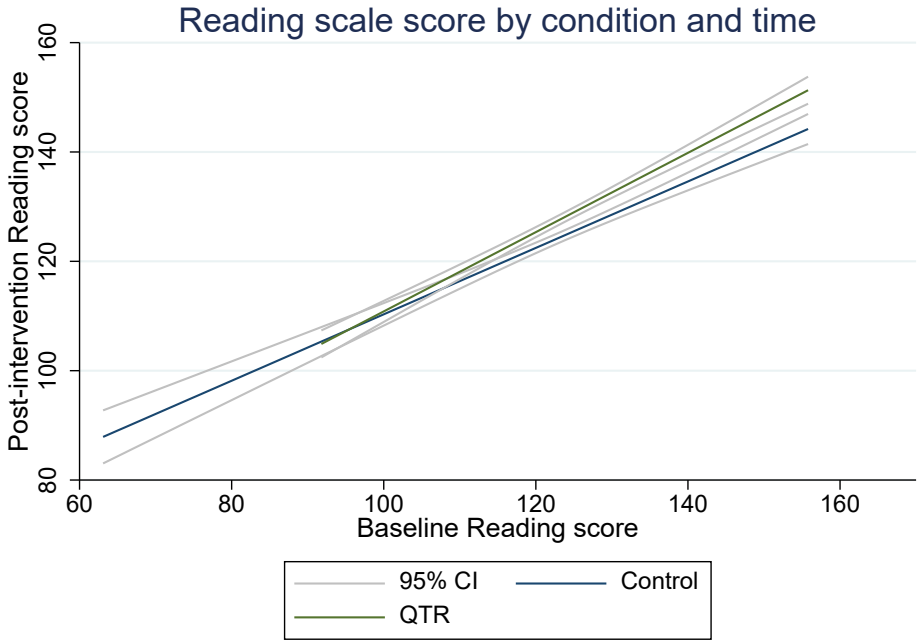




*Figure 2: Maths and Reading fitted lines with confidence intervals: post-intervention and baseline for Control and QTR group*

# 7. References

Abazaoglu, I., & Aztekin, S. (2016). The Role of Teacher Morale and Motivation on Students' Science and Math Achievement: Findings from Singapore, Japan, Finland and Turkey. *Universal Journal of Educational Research*, *4*(11), 2606–2617. https://doi.org/10.13189/ujer.2016.041114

Ainley, J., & Bourke, S. (1992). Student views of primary schooling. Research Papers in Education, 7(2), 107–128. https://doi.org/10.1080/0267152920070202

Australian Bureau of Statistics, Remoteness Structure. https://www.abs.gov.au/websitedbs/d3310114.nsf/home/remoteness+structure Accessed 10 June, 2021

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2020). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, *147*(2), 134. https://doi.org/10.1037/BUL0000307

Clark, C. (2011). Setting the Baseline: The National Literacy Trust's first annual survey into reading - 2010. London: National Literacy Trust.

Clark, C. & Douglas, J. (2011). Young People's Reading and Writing An in-depth study focusing on enjoyment, behaviour, attitudes and attainment. National Literacy Trust

Elliott, K. (2015). Australian Journal of Teacher Education Teacher Performance Appraisal: More about Performance or Development? Teacher Performance Appraisal: More about Performance or Development? *Australian Journal of Teacher Education*, *40*(9). https://doi.org/10.3316/INFORMIT.490666162557912

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, *17*, 63–84. https://doi.org/10.1016/J.EDUREV.2015.11.002

Malanchini, M., Wang, Z., Voronin, I., Schenker, V. J., Plomin, R., Petrill, S. A., & Kovas, Y. (2017). Reading self-perceived ability, enjoyment and achievement: A genetically informative study of their reciprocal links over time. *Developmental Psychology*, *53*(4), 698. https://doi.org/10.1037/DEV0000209

Mankin, A., von der Embse, N., Renshaw, T. L., & Ryan, S. (2017). Assessing teacher wellness: Confirmatory factor analysis and measurement invariance of the Teacher Subjective Wellbeing Questionnaire. Journal of Psychoeducational Assessment, Advance online publication. https://doi.org/10.1177/0734282917707142

Mertler, C. A. (2016). Should I Stay or Should I Go? Understanding Teacher Motivation, Job Satisfaction, and Perceptions of Retention among Arizona Teachers. *International Research in Higher Education*, *1*(2). https://doi.org/10.5430/irhe.v1n2p34

Miller, A., Gore, J., Wallington, C., Harris, J., Prieto-Rodriguez, E., & Smith, M. (2019). Improving student outcomes through professional development: protocol for a cluster randomised controlled trial of quality teaching rounds. *International Journal of Educational Research*, 98, 146–158.

Murray, D. M., Varnell, S. P., & Blitstein, J. L. (2004). Design and analysis of group-randomized trials: a review of recent methodological developments. *American journal of public health*, 94(3), 423–432.

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, *34*(1), 29–48. https://doi.org/10.1080/01443410.2013.797339

Pinxten, M., Marsh, H. W., de Fraine, B., van den Noortgate, W., & van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology*, *84*(1), 152–174. https://doi.org/10.1111/BJEP.12028

Retelsdorf, J., Köller, O., & Möller, J. (2014). Reading achievement and reading self-concept – Testing the reciprocal effects model. *Learning and Instruction*, *29*, 21–30. https://doi.org/10.1016/J.LEARNINSTRUC.2013.07.004

Rowe, F., Stewart, D., & Patterson, C. (2007). Promoting school connectedness through whole school approaches. *Health Education*, *107*(6), 524–542. https://doi.org/10.1108/09654280710827920/FULL/PDF

Schöber, C., Schütte, K., Köller, O., McElvany, N., & Gebauer, M. M. (2018). Reciprocal effects between self-efficacy and achievement in mathematics and reading. *Learning and Individual Differences*, *63*, 1–11. https://doi.org/10.1016/J.LINDIF.2018.01.008

Shoulders, T. L., & Krei, M. S. (2015). Rural High School Teachers' Self-Efficacy in Student Engagement, Instructional Strategies, and Classroom. In *Source: American Secondary Education*, *44*(1).

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher Efficacy: Capturing an Elusive Construct. Teaching and Teacher Education, 17, 783-805. http://dx.doi.org/10.1016/S0742-051X(01)00036-1

Yoo, J. H. (2016). The effect of professional development on teacher efficacy and teachers' self-analysis of their efficacy change. *Journal of Teacher Education for Sustainability*, *18*(1), 84–94. https://doi.org/10.1515/jtes-2016-0007

Žolgar-Jerkovic, I., Jenko, N., & Lipec-Stopar, M. (2018). Affective Factors and Reading Achievement in Different Groups of Readers. International Journal of Special Education, 33(1), 201–211.

# Appendix A: Flow diagram for students who completed PAT assessments

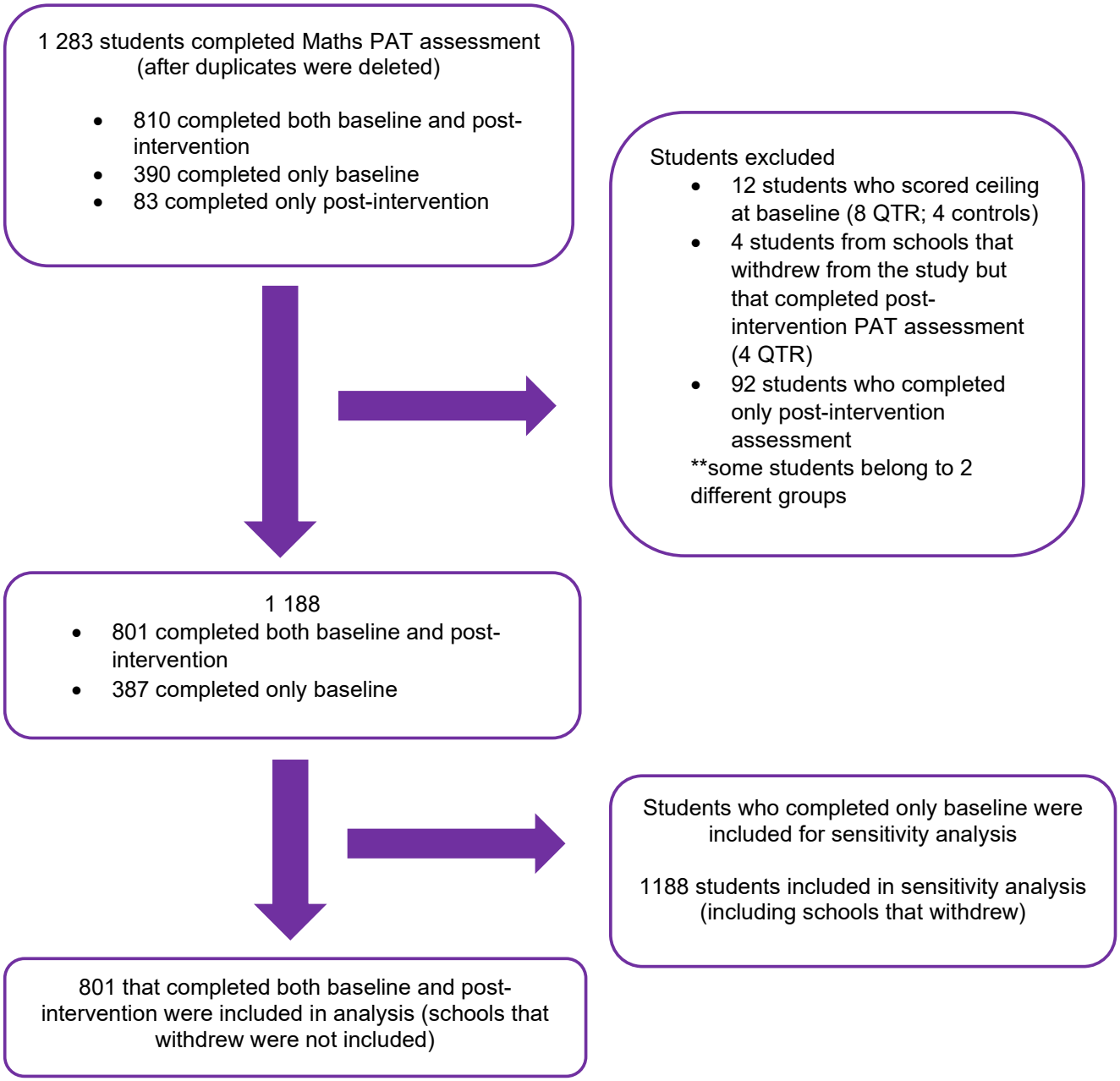



*Figure 3: Flow diagram of students who completed Maths PAT assessments*

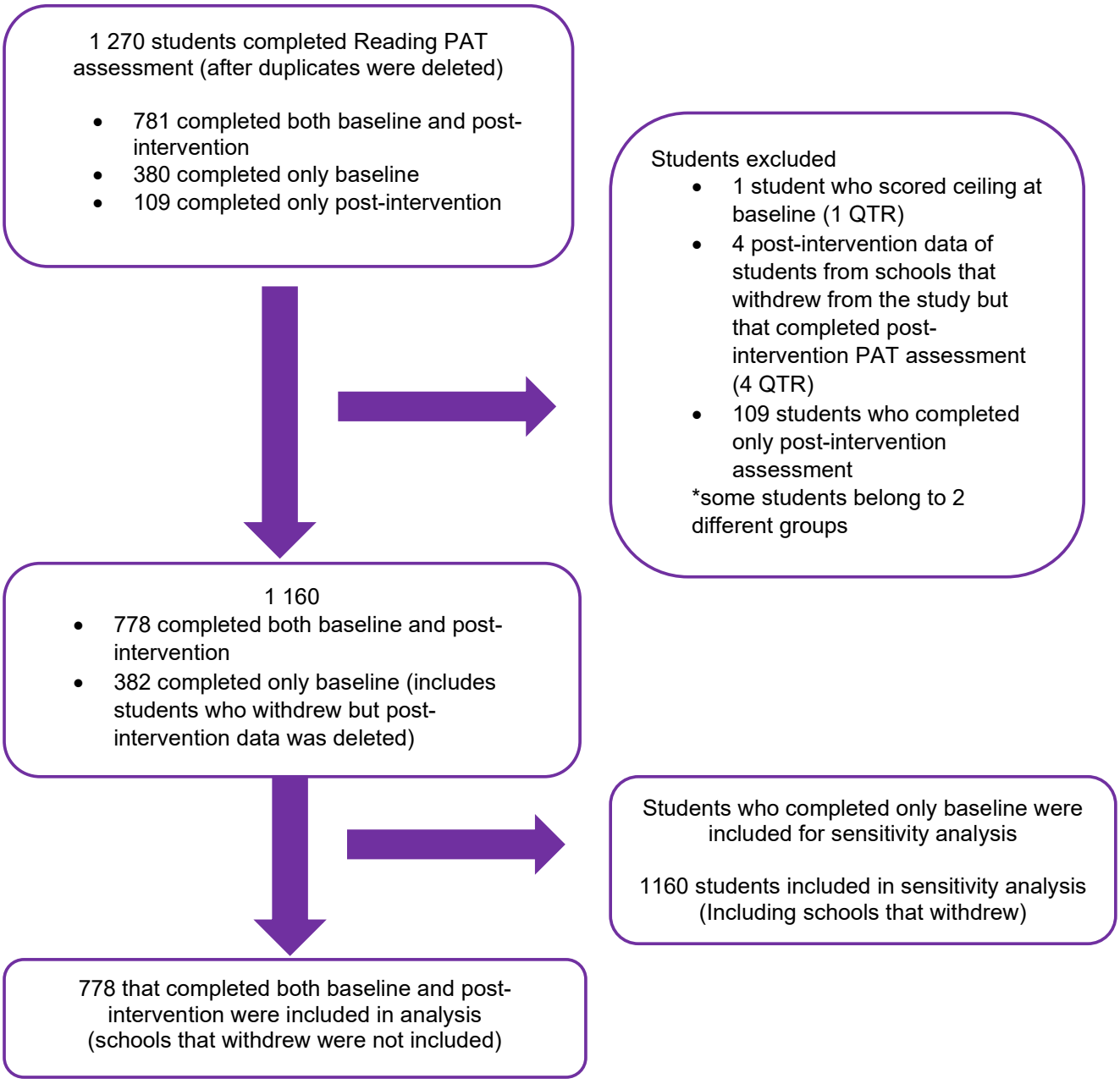



*Figure 4: Flow diagram of students who completed Reading PAT assessments*

# Appendix B: Supplementary fidelity check analyses

The fidelity implementation data in Table 20 pertains to the (n=3) who were observed but did not submit their own, self-reported data via the PLC checklist survey. The mean score was 8.66 indicating high levels of implementation fidelity. Two out of 3 schools achieved 100% for the observed fidelity check (achieving 9/9).

*Table 20: Fidelity of implementation of schools with observed data only*

| Outcome | QTR group |
|---|---|
| **Fidelity Score** | |
| Observed, mean (SD) | 8.66 (0.58) |
| Fidelity 9/9 (100% fidelity) | |
| Observed, % | 67% |

# Appendix C: Sensitivity analyses teacher survey

*Table 21: Sensitivity analyses teacher survey*

| Difference in difference Clustered by school | | | | |
|---|---|---|---|---|
| **Group** | **Baseline** | **Post-intervention** | **Coef. (CI 95%)** | **P-value** |
| | n | **n** | | |
| **Teacher Morale** | | | | |
| QTR | 77 | 49 | 0.02 (-0.39, 0.44) | 0.919 |
| Control | 74 | 62 | | |
| **Teacher Appraisal and Recognition** | | | | |
| QTR | 77 | 49 | 0.048 (-0.36, 0.45) | 0.816 |
| Control | 74 | 62 | | |
| **Efficacy for Student Engagement** | | | | |
| QTR | 77 | 49 | 0.048 (-0.30, 0.40) | 0.784 |
| Control | 74 | 62 | | |
| **Teachers' Self-Efficacy** | | | | |
| QTR | 77 | 49 | 0.02 (-0.14, 0.17) | 0.837 |
| Control | 74 | 62 | | |
| **School Connectedness** | | | | |
| QTR | 77 | 49 | -0.04 (-0.36, 0.28) | 0.800 |
| Control | 74 | 62 | | |

# Appendix D: Sensitivity analyses student survey

*Table 22: Sensitivity analyses student survey*

| *Difference in difference Clustered by school* | | | | |
|---|---|---|---|---|
| **Group** | **Baseline** | **Post-intervention** | **Coef. (CI 95%)** | **P-value** |
| | **n** | **n** | | |
| **Quality of School Life-Achievement** | | | | |
| QTR | 443 | 261 | 0.05 (-0.77, 0.17) | 0.457 |
| Control | 479 | 324 | | |
| **Reading (frequency)** | | | | |
| QTR | 439 | 260 | 0.09 (-0.31, 0.13) | 0.417 |
| Control | 476 | 321 | | |
| **Reading Self-Efficacy** | | | | |
| QTR | 439 | 260 | 0.03 (-0.30, 0.37) | 0.852 |
| Control | 476 | 321 | | |
| **Reading Enjoyment** | | | | |
| QTR | 439 | 260 | 0.25 (-0.23, 0.73) | 0.295 |
| Control | 476 | 321 | | |
| **Reading Comprehension** | | | | |
| QTR | 439 | 260 | 0.27 (-0.17, 0.69) | 0.229 |
| Control | 476 | 320 | | |
| **Maths Self-Efficacy** | | | | |
| QTR | 437 | 259 | 0.09 (-0.03, 0.22) | 0.129 |
| Control | 471 | 319 | | |
| **Maths Interest** | | | | |
| QTR | 437 | 259 | 0.06 (-0.11, 0.23) | 0.503 |
| Control | 471 | 319 | | |
| **Maths Anxiety** | | | | |
| QTR | 437 | 259 | 0.05 (-0.10, 0.21) | 0.469 |
| Control | 471 | 319 | | |

THE UNIVERSITY OF NEWCASTLE AUSTRALIA

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

CREATE CHANGE

## Contact details

Institute for Social Science Research
T +61 7 3346 7471
E issr@uq.edu.au
W uq.edu.au

CRICOS Provider Number 00025B

ISBN: 978-0-7259-0946-8