SiMERR National Research Centre's STATEMENT

This is SiMERR's response to the Evaluation Report (Report) of the Evidence for Learning Randomised Controlled Trial of QuickSmart Numeracy (the Trial) evaluated by the Teachers and Teaching Research Centre, University of Newcastle (the Evaluator).

Executive Summary

Data collected by the Evaluator in the Trial and analysed by SiMERR using approved Australian Council for Educational Research (ACER) procedures showed students with 75%+ QuickSmart lesson participation achieved on average 2.5-to-3 times expected one-year's growth. These results are consistent with that reported by SiMERR, education jurisdictions, principals, schools and parents over the past two decades.

students with 75%+ QuickSmart lesson participation achieved on average 2.5-to-3 times expected one-year's growth

Using Trial data, ACER procedures also show national percentile progress in the 14 months of the Trial for students who achieved:

(i) 90%+ lesson participation in QuickSmart, progressed 38 percentiles from the 11th percentile to the 49th percentile; and

(ii) 75%+ lesson participation progressed 33 percentiles from the 16th percentile to the 49th percentile.

The critical concern of the SiMERR/QuickSmart Team of the Trial is that important data are not clearly revealed in the Report. The analysis in the Report, while technically correct, does not include within-group analysis, which reveals crucial information not made evident in the Report.

***at least* 25%-to-38% of the school year was not available to schools for QuickSmart setup and lesson participation**

Also, not revealed fully or discussed carefully in the Report, are important details on how *implementation* aspects of the Trial design, including obtaining parent permissions for student participation, pre-and post-testing using PATMaths and randomising students into groups, meant that *at least* 25%-to-38% of the school year was not available to schools for QuickSmart setup and lesson participation.

This loss of school weeks to RCT implementation precluded the possibility for schools to offer 30 weeks of QuickSmart instruction. For schools, this meant a late start, early finish, no setup time and no flexibility for them to cover student or Instructor absences, or any competing in-school activities or excursions. The impact of these normal/typical school functions, with no room for schools to manoeuvre, resulted in further reductions in QuickSmart lesson rates that were beyond the control of schools to address.

Because of concerns with these limitations to the Trial, SiMERR conducted additional analyses of the Trial data (i) using the instrument-based national Australian norms of the PAT-Maths test, and (ii) incorporating lesson participation using sub-groups defined by their levels of lesson participation. We argue that a more appropriate analysis procedure is through the Australian Council for Educational Research (ACER), developers of the PAT-Maths test series. ACER's approach is robust, statistically valid, and widely used and understood throughout Australia and Internationally. The ACER approach is employed currently across Australia involving many 100,000s of students, and operates at national, state, school, sub-school group, and individual-student levels.

Specific Responses

1. **Implementation Challenges of the Trial Design**

There were significant challenges in the implementation of the Trial design resulting in low levels of Intervention Group student exposure to QuickSmart, including:

- School recruitment and establishing randomised control trial aspects of the Trial;
- Challenges associated with randomised control trial aspects meant randomisation data, which were a pre-requisite to schools setting up student QuickSmart timetables and automaticity testing before QuickSmart lessons could start, were not available until the last weeks of Term 1 and early Term 2; and
- Intervention Group students stopped QuickSmart 3-to-4 weeks before the end of the school year with data collection taking place in the last 2 weeks of November.

Typically, QuickSmart implementation, as per the Trial Protocol, requires schools to allocate a 2-week setup and close-down period and 30 weeks of instruction. Outside of the Trial, a further 'buffer' of 5 weeks is available for schools to address any in- or out-of-school disruptions, such as student and instructor absences, school-wide events carnivals/concerts, school excursions–short and long, school examination periods and Instructor attendance at QuickSmart workshops.

Typically, QuickSmart implementation…requires… a 2-week setup and close-down period and 30 weeks of instruction… a further 'buffer' of 5 weeks is available for schools to address any in- or out-of-school disruptions

In the Trial, at least 9-to-15 weeks of the 39 weeks available to schools in 2017 were devoted to randomised control trial aspects. A further 2-week setup period for QuickSmart was required by schools. A school's setup phase could not commence until the complete set of randomisation data had been received by the school. A direct consequence of the late start to QuickSmart was the decision, communicated to schools, that the timing of the second QuickSmart Workshop would be delayed by 6 weeks.

… majority of students attended less than 22 weeks and received less than 66 lessons.

Together, these actions resulted in schools' inability to offer the intended program as designed. As a result, very few students reached an acceptable 90%+ number of lessons, i.e., 28+ of the expected 30 weeks or 81+ of the expected 90 lessons. The great majority of students attended less than 22 weeks and received less than 66 lessons. It was common practice in the Diocese for non-Trial schools in the same year as the Trial, as in previous years, to achieve high-levels of QuickSmart lesson participation.

2. **ACER Analytical Procedure and Results**

When the analysis of the PAT-Maths Pre-test results for the Intervention Group and Control Group (data collected in March 2017) and the Post-test undertaken in May 2018 (six months after the end of QuickSmart lessons offered to the Intervention students), the results for the primary students showed:

- Students (*n*=7, 10%) who participated in 90% or more of the 30-week QuickSmart program:
  - (i) progressed from the 11th percentile to the 49th percentile based on ACER National normed achievement data;
  - (ii) achieved growth of 15.94 scale units, approximately three times expected yearly growth in the 14-month period since pre-test data was collected; and
  - (iii) exceeded the growth of the Control Group by 5.16 scaled units, approximately one-year's expected yearly growth.

- Students (n=23, 34%) who participated in 75% or more of the 30-week QuickSmart program:
  - (i) progressed from the 16th percentile to the 49th percentile based on ACER National normed achievement data;
  - (ii) achieved growth of 12.50 scale units, approximately 2.5 times expected yearly growth in the 14-month period since the pre-test data was collected; and
  - (iii) exceeded the growth of the Control Group students by 1.59 scale units, approximately a third of one-year's expected yearly growth.
- Students *(n*=40, 60%) who participated in 50% to 74% of the 30-week QuickSmart program:
  - (i) progressed from the 16th percentile to the 44th percentile based on ACER National normed achievement data;
  - (ii) achieved growth of 11.35 scale units, of approximately 2 times expected yearly growth in the 14-month period since the pre-test data was collected;
  - (iii) slightly exceeded the growth of the Control Group students by 0.56 scale units; and
  - (iv) required more QuickSmart lessons to be able to apply the newly acquired automaticity knowledge/skills to higher-order and extended questions as offered in PAT-Maths tests.

it is not acceptable that ... Trial protocol analyses are limited to an examination of single groups in terms of their mean performance.

The analyses presented in the Report were specified in the agreed and published Trial Protocol. However, this agreement assumed implementation of higher levels of exposure of the Intervention Group to QuickSmart lessons. As this condition was not met, it is not acceptable that the only analysis of Trial data can be what was specified. The Trial protocol analyses are limited to an examination of differences of single groups in terms of their mean performance.

The alternative ACER analytical procedure enabled not only a direct mean-score comparison of the Intervention Group students' performance with that of the Control Group students but data from significant sub-groups can be explored and interrogated as well. Using the National normed data-base of ACER applied to PAT-Maths tests, this approach offered:

(i) Australian school-year population percentile-band development for students and groups of students;

(ii) the "reasonable expectation for improvement from one year to the next is about five scale units at primary school" (ACER PATMaths Progressive Achievement Tests in Mathematics, Teacher Manual, Third Edition, 2012, p. 9); and

(iii) "if the pre-test was difficult for the students being tested, then the same test or the next test is the suitable option as the post-test." (ACER PATMaths Progressive Achievement Tests in Mathematics, Teacher Manual, Third Edition, 2012, p. 9).

"reasonable expectation for improvement from one year to the next is about five scale units at primary school" ACER PATMaths...

The use of ACER normed tables was not part of the agreed Trial Protocol, but the low-level exposure to QuickSmart lessons, for which the schools could not compensate, made essential our use of the ACER Australian population normed data to better understand and interpret the data.

Analysis of the Trial data with this method identified how the Intervention and Control groups performed down to an individual level; and, identified outstanding outcomes of students in the Trial linked to lesson exposure which were not evident in the analysis undertaken by the Evaluation Team.

In the case of students in Year 4, for example, the concern was to measure student progress through the Year 4 curriculum using the appropriate PAT-Maths test for that purpose. A PAT-Maths test has to be fair for a wide range of student abilities. ACER tests include items that enable both low performing and high performing students to demonstrate what they know and can do. Consequently, scaled scores were used on the instrument at pre-test, post-test and delayed post-test to measure progress through the Year 4 curriculum.

Importantly, SiMERR's additional analysis of the data collected in the Trial followed the ACER guidelines on how we can interpret PAT-Maths mean score for individual-to-national levels. (PATMaths Progressive Achievement Tests in Mathematics, Teacher Manual, Third Edition, Lindsey, Stephanou, Urbach, & Sadler, 2012; PATMaths Progressive Achievement Tests in Mathematics, Teacher Manual, Fourth Edition, Stephanou & Lindsay, 2013)

3. Practices in QuickSmart Implementation

Attendance data are available from non-Trial schools in the same Diocese, who were offering QuickSmart lessons, in the same calendar year as the Trial – finishing in December 2017. Table 1 below summarises this information and allows comparisons of the lesson attendance of students in the Trial and non-Trial schools.

Table 1: QuickSmart Lesson Attendance in 2017 from Trial schools and non-Trial schools in Sydney Catholic Schools

Percentage Lesson participation	Number of Trial students	Percentage of Trial students	Number of non-Trial students in same Diocese	Percentage of non-Trial students in same Diocese
75–100% 67-90 lessons	26	18.57	205	74.28
50–74% 45-66 lessons	78	55.71	43	15.58
25–49% 23-44 lessons	29	20.72	24	8.69
0–24% 0-22 Lessons	7	5.00	4	1.45
Totals	**140**	**100%**	**276**	**100%**

Previous years' data from the Diocese records equivalent lesson attendance rates to that achieved by non-Trial schools shown for 2017 in Table 1.

In the case of the protocol of 90%+ attendance, only students in the Primary schools in the Trial managed this level of instruction. In this case, there were 7 students (5% of the Trial students). In the case of the non-Trial schools there were 100 students (36.23% of non-Trial students) who achieved 90%+ levels of participation.

On the QuickSmart web site from 2013 to 2016, SiMERR reported lesson-attendance data received from schools across the country. Overtime this created some confusion in the field, and the practice of reporting these data was stopped in 2016. There were several reasons for the confusion and this centred around how QuickSmart was used in different school contexts. Consequently, working in terms of averages drawn from these data, as described in the Report, is highly problematic. To get an accurate picture of the meaning of the data requires knowledge of the practices in each school in more detail than is available. Some examples include:

(a) Schools select students who they believe need QuickSmart. In some cases not all students chosen might be in the bottom 30% of the achievement spectrum. It is likely that these students will require less than 30 weeks to complete the program. Students are able to graduate from QuickSmart once they have demonstrated the skills/knowledge required. In the cases of students not being in the bottom 30% of the national achievement spectrum, the number of lessons required is very much dependent on the student, and the need for 30-weeks lesson attendance does not stand. This situation is quite different from the case of students in the bottom 30% of achievement in Australia undertaking QuickSmart.
(b) When students complete the program earlier than 30 weeks, schools typically enrol new students so as to capitalise on, and maximise, the support that QuickSmart provides to students in the school.
(c) Some schools use the opportunity of additional lessons to give students the opportunity to start the program earlier and finish earlier in the following year or simply offer the remainder of the year to students as a top-up.
(d) Some schools start new cohorts of students after NAPLAN and go across years. They still submit data at the end of the school year so that they can report independent analysed achievement to the school Executive, knowing these data will improve in the following year.

4. Analytical Approach Used Despite Low-lesson Attendance

A complicating feature of the low-lesson attendance rate was the use in the Report of a linear mixed-model and Instrumental Variable analysis approach. Whilst acknowledging that these approaches were in accordance with the agreed Trial Protocol, both are based on a comparison of group mean performance and do not take explicit account of the non-linear trajectory of cognitive growth, which suggests considerably improved performance towards the end of a program (see Fischer, K.W. (2008). Dynamic cycles of cognitive and brain development: Measuring growth in mind, brain and education. In Battro A.M., Fischer K.W., & Lena P. (Eds), *The educated brain* (pp. 127-150). Cambridge University Press). As a result, findings in this Trial should not be based solely on these approaches, especially when lesson participation scores are low in the Trial compared to real-world experiences.

attendance … below 74% lesson exposure … results in mean scores of these students having a disproportionate influence on … the group mean score

In practical terms, the impact of most students' QuickSmart attendance being below 74% lesson exposure (below 66 lessons) results in mean scores of these students having a disproportionate influence on the calculation of the group mean score and hence overall growth. There is little difference in the mean value calculated for students who attended 50%-74% of lessons to that of the total Intervention Group. Because few students were able to experience high levels of exposure, it is misleading to base comparisons solely on overall mean scores. In short, the low levels of exposure for the Intervention Group and the limited analysis undertaken are such that definitive and generalisable conclusions using these data should not be made.

…low levels of exposure … and the limited analysis undertaken are such that definitive and generalisable conclusions using these data should not be made

5. Automaticity and its Relation to QuickSmart and Neuroscience

The identification of 'automaticity' scores for improvements in accuracy and response times for Intervention and Control students was not measured as part of the Trial, as the primary outcome was to measure the program's effect on maths achievement. However, automaticity data were collected by staff in schools to identify base-line data and where best to start each student on QuickSmart. The data were collected using the Cognitive Aptitude Assessment System, Australian version (OZCAAS). This is an adaption of the original CAAS program, which operationalises automaticity by capturing student's response speed and accuracy, that was developed by researchers in the Department of Psychology at the University of Massachusetts under Professor James (Mike) Royer.

The results found that the Intervention and Control Groups of students recorded equivalent scores using OZCASS on all measures of automaticity in the Pre-test. In the case of the OZCASS Post-test of automaticity, undertaken once the QuickSmart lessons ceased, on most measures Intervention Group students were significantly ($p<0.001$) better than the Control Group students. As expected, Intervention Group students demonstrated greatly improved accuracy scores and marked reductions in response times.

...Intervention Group students demonstrated greatly improved accuracy scores and marked reductions in response times.

Overall, the ACER and OZCAAS analysis confirm the last four-of-the-five stages (points (ii) to (v) below) of neural development operating when students undertake the QuickSmart program, namely: the brain:

(i) begins to learn new automatic processes associated with fundamental skills by developing an alternative to previous effortful practice;
(ii) builds automatic processes that can be applied accurately across a range of fundamental skills to speed up responses;
(iii) uses automaticity of fundamental skills consistently on basic-fact recall;
(iv) begins applying automatic skills/knowledge as a tool in problem solving; and
(v) seamlessly uses automaticity of basic skills in higher-order and extended questions, such as offered in PAT-Maths tests.

6. Recognition of QuickSmart Instructor Agency

The term teacher 'agency' refers to the ability of teachers to understand the context of student learning to such an extent that they can act adequately with intention throughout a lesson, particularly, when the situation appears novel. Encouraging Instructor agency is at the heart of the six-days of QuickSmart professional learning in the first year and the additional days offered in subsequent years. The aims of this support are to develop skills and understandings about QuickSmart, the brain and its functioning, instruction approaches, and alternative ways to support students as issues arise.

Encouraging Instructor agency is at the heart of ... QuickSmart professional learning

Instructors need to be in a position, when faced with new challenges, to have a strong theoretical and practical basis upon on which to guide actions. Also, it is not uncommon for QuickSmart Instructors to conduct workshops on QuickSmart with teachers in their school. Hence, Instructors achieving, strong knowledge of, and insights about, QuickSmart are very important.

The Report identifies between-school differences in the practices of some Instructors. The information is presented as if there exist numbers of set routines developed for Instructors. Also, these examples suggest, negatively, that Instructors are not following approved routines. This is not the case. In the light of Instructor agency development, specific differences in approaches are expected and seen as commendable, and part of Instructor ownership of QuickSmart. The only caveats are that the actions and decisions taken are consistent with the philosophic and neural bases of QuickSmart, and any school-specific requirements. The practices identified in the Report are consistent with these caveats.

QuickSmart qualitative data indicate that poor examples of Instructor actions in QuickSmart lessons are more prominent early in the first year and decline progressively over the year. The more experienced Instructors become, the more nuanced are expectations on how Instructors proceed.

There is no doubt that poor interpretations of appropriate actions occur on occasions. There is no doubt that the same finding could be made of trained teachers in normal classrooms across the country. Supporting student learning is a difficult task whether working with a small group or whole class.

Not made clear in the Report, is that QuickSmart has put in place many features to minimise the occurrence of various forms of poor practice. First, the program of professional learning builds, from the ground up, a strong and robust theoretical basis linked closely to actual practice of QuickSmart – no information about the program and how aspects come together is held back from participants.

Not made clear in the Report, is that QuickSmart has put in place many features to minimise the occurrence of various forms of poor practice

Second, each school has a QuickSmart Coordinator who is a trained teacher (and often a school executive member). This person undertakes the same training, and hears the same messages, as the Instructor. The QuickSmart Coordinator is the first port-of-call for Instructors. There is also a help-line with trained and experienced QuickSmart experts available to discuss issues and provide advice. This support is also available through email. Finally, although less direct, there is extensive documentation in the *QuickSmart User Guide*, *Frequently Asked Questions*, and other support materials that provides detailed advice on a large range of specific and general issues.

7. National Finding of Impact of QuickSmart

QuickSmart, a University of New England numeracy and literacy program for schools, achieved top ratings for engagement and impact in the first national assessment of Australian research impact.

The Australian Research Council (ARC), the peak national body in research and research funding for Australia, in the 2018 assessment of research undertaken by all universities in Australia covering the period 2011-2016, included for the first time a national analysis of the *Impact* of research conducted by all Australian universities. The ARC described 'impact' as the contribution that research makes to the economy, society, environment or culture, beyond the contribution to academic research.

The results of the national **Engagement and Impact** 2018 assessment, were based on a rating scale of low, medium and high across three key facets of impact: 'engagement', 'impact' and 'approach to impact'.

- **Engagement** – the interaction between researchers and research end-users outside of academia, for the mutually beneficial transfer of knowledge, technologies, methods or resources.
- **Impact** – the contribution that research makes to the economy, society, environment or culture, beyond the contribution to academic research.
- **Approach to Impact** – the mechanisms used to encourage the translation of research into impacts beyond academia.

The process involved the development of case studies that showcased the benefits that research programs have delivered to the broader community and the economy. Each submission from an Australian university to the ARC was limited to a single case study for each broad field of research. The Education case study submitted by the University of New England (UNE) detailed the ***QuickSmart* program**.

UNE was only one of four universities in Australia that received a ranking of 'high' across each of the three Impact indicators for Education.

The ARC award to UNE through the QuickSmart program stands in contrast to the outcome of the Trial and reported findings. This award by the ARC should assist the reader to place into perspective the identified challenges associated with the implementation of the randomised control features of the Trial and the limitations of the broad statistical analysis undertaken.

Conclusions

Trial data, when considered through the ACER analysis framework and a focus on lesson participation, confirmed very strong results recorded by students who achieved high levels of QuickSmart lesson participation. This is consistent with findings of previous analyses of QuickSmart and reported by education jurisdictions, principals, teachers and parents.

The ACER analysis … confirmed very strong results … (and) is consistent with findings of previous analyses

The ACER analysis also provided further supportive insights into the nature of neural development of low-achieving students as they first acquire automaticity of fundamental skills and then as they applied this knowledge/skill when addressing more difficult questions or problem situations (see for example Tronsky, L.N. & Royer, J.M. (2002). Relationships among basic computational automaticity, working memory and complex mathematical problem solving. In Royer J.M. (Ed), *Mathematical Cognition*, (pp. 117-146). Information Age Publisher).

Several important issues were identified that were relevant to this Trial, and of concern to the application of randomised control trials within schools that require a much wider discussion. These include:

(i) ways to ensure that implementation and results of intervention programs are not impacted by school-time loss through the process of implementing randomised control trial conditions and expectations;
(ii) the consequences of placing students, randomised to the intervention and control groups, into the same class and with the same teacher;
(iii) the potential impact of intervention and control students being known to one another, class students and teachers;
(iv) associated with point (iii) above are, the possibilities of unexpected achievements of students in the Control Group related to the possibility of the Hawthorne effect or other confounding factors;
(v) identifying the effect-size measure and the meaning which should be applied in an Australian context;
(vi) ways to analyse and interpret the non-linear nature of learning where the gains accrue steadily but are most likely evident towards the latter part of an intervention program; and
(vii) how best to ensure the interpretations of qualitative data collected in randomised control trials are consistent with philosophical and theoretical positions that underpin a program.