

Briefings

Thought leadership for the independent schooling sector

Volume 22 Issue 1 • January/February 2018

COMMENTARY ON SCHOOLING ON THE WRONG TRACK

From the Executive Director

The commencement of the 2018 school year has seen the predictable public commentary about schools funding, enrolments and fees.

The 2014 promise by then Education Minister Julia Gillard that the introduction of the Gonski schools funding model would end the public versus private funding debate has not come to pass. In fact, despite five years of Gonski (with the fifth year being the modified Gonski 2.0), the debate over the funding of schools appears to be louder and more divisive than ever.

Unfortunately, a new “front” has opened up in the debate with the Catholic Education Commission of Victoria (CECV) openly attacking the funding arrangements for independent schools. A 50-year collaboration between the Catholic and independent sectors to advocate for the funding entitlements of parents who choose non-government schools is seriously under threat.

The latest CECV attack¹ on high-fee independent schools ignores the long-accepted right of every parent, no matter the school they choose for their children or their socio-economic background, to receive a basic entitlement in terms of government funding support.

These parents are taxpayers contributing to government expenditure that supports Australia’s education system. However, as is their right, they have chosen to invest from their after-tax income in a non-government school education for their children. It seems CECV wants to penalise a group of these

parents further (those who choose a school where private income exceeds the legislated Schooling Resource Standard) when it might be better to acknowledge the investment these parents make to school education and the savings they present to taxpayers.

The implications for all non-government schools of the CECV position that students in some schools deserve no government funding should not be underestimated. This includes Catholic schools noting that recent ISQ research² reveals 28% of families in the three highest income deciles (those earning more than \$2,346 per week) send their children to Catholic schools (compared to 22% for independent schools and 50% for government schools).

Memories of Mark Latham’s 2004 “hit-list” as part of the ALP’s federal election policies come to mind with the obvious question of how and where is the line drawn in respect of non-government schools which might or might not receive government funding.

The current National School Resourcing Board (NSRB)³ review of the socio-economic status (SES) score methodology used to determine non-government school funding from the Australian Government will be important in the resolution of this current funding debate. The SES measure, which determines a school community’s capacity to contribute to the costs of schooling, has been utilised

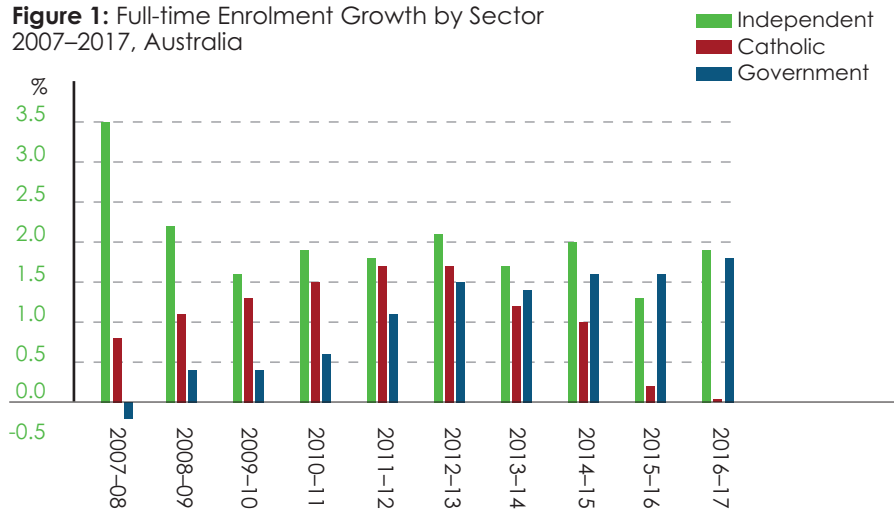
1 See *The need to rethink need – How the Gonski Review got it wrong on funding non-government schools* Catholic Education Commission of Victoria February 2018 at www.cecv.catholic.edu.au/Publications

2 *Income Levels of Families with Students in Queensland Schools* (November 2017) available at https://rms.isq.qld.edu.au/files/Income_Levels_of_Families_2017.pdf

3 For further details go to <https://www.education.gov.au/review-socio-economic-status-ses-score-methodology>

COMMENTARY ON SCHOOLING ON THE WRONG TRACK

Figure 1: Full-time Enrolment Growth by Sector 2007–2017, Australia



(with broad acceptance) since 2001 for independent schools and since 2004 for Catholic schools.

At a time when the Australian Government is implementing a fairer and more equitable funding model, let's hope the NSRB affirms the critical principle that schools funding for non-government schools should be based on capacity to contribute, rather than willingness to contribute (which might be measured by school fee levels) and dismisses any notion that the wealth of a school (which may have been built-up over many years with funding from parents) should be used to determine an individual school's funding.

There is much at stake, with the most recent Productivity Commission report on Government Services⁴ revealing total government recurrent expenditure on school education was \$55.7 billion in 2015–16. This comprised \$42.4 billion on government schools and \$13.3 billion on non-government schools.

In addition to this recurrent expenditure, government investment in government school infrastructure was \$6.1 billion in 2015–16.

Non-government schools received 23.9% of government recurrent funding (down 0.2% on the previous year) compared to 76.1% for government schools. There appears to be a clear trend emerging of a higher proportion of government funding going to government schools, perhaps driven by the implementation of Gonski from 2014.

The Productivity Commission report notes that government recurrent funding accounted for only 57.2% of total non-government school recurrent funding in 2016, with the remaining 42.8% sourced from private sources. Further, the average government recurrent expenditure per FTE student in government schools was \$17,275 in 2015–16 compared to average government recurrent expenditure per FTE student in the

non-government sector of \$10,147. On the Queensland data, if every one of the more than 260,000 non-state school students in Queensland took up a fully taxpayer funded place at a state school, the Australian and Queensland Governments would face an additional \$1.7 billion annual education bill⁵.

Data released by the Australian Bureau of Statistics⁶ about enrolments by school sector also resulted in media commentary focusing on a minor decrease in the proportion of students in non-government schools (34.6% in 2016 to 34.4% in 2017). This commentary ignored the fact that Australia continues to have one of the world's highest proportion of participation in non-government schooling based on the long-accepted principle of school choice. Can parents be that wrong?

Independent schools in Queensland continue to be strongly supported by parents with enrolments growing by 1.3% in 2017 to reach a record high of 118,942⁷. The independent sector maintained its share of total enrolments at 14.7% and of the additional 10,949 students who enrolled in Queensland schools in 2017, 14% (1,526) were in the independent sector.

Given the financial pressures facing many families, slow wages growth and economic uncertainty in regional areas, the independent sector's 2017 growth confirms parental confidence in independent schools.

During the period 2007 to 2017, the independent sector across Australia has consistently experienced the highest growth rates (see Figure 1). In 1997, the independent sector accounted for 10% of all school enrolments in Australia – today it is 14.5%, with an increase of 240,000 students from 1997.

4 Report on Government Services 2018 available at <http://www.pc.gov.au/research/ongoing/report-on-government-services>

5 See ISQ Media Release *New Report Sets the Record Straight on Public Funding for Schools* at <https://www.isq.qld.edu.au/media-resources/new-report-sets-the-record-straight-on-public-funding-for-schools>

6 Schools Australia 2017 available at <http://www.abs.gov.au/ausstats/abs@nsf/0/9448F2F814FA0311CA2579C700118E2D?Opendocument>

7 See ISQ Media Release *Parental Confidence in Qld Independent Schools Continues* (2 February 2018) available at <https://www.isq.qld.edu.au/members/parental-confidence-in-qld-independent-schools-continues>

On the Queensland data, if every one of the more than 260,000 non-state school students in Queensland took up a fully taxpayer funded place at a state school, the Australian and Queensland Governments would face an additional \$1.7 billion annual education bill.

School fees have also featured heavily in media coverage since the start of 2018, with the Australian Scholarships Group annual survey of education costs⁸ provoking misleading headlines about the costs faced by parents educating their children in independent schools.

Claims by the group that parents face costs of nearly \$500,000 to educate a child in an independent school were strongly rebuked by the Independent Schools Council of Australia (ISCA) and Independent Schools Queensland (ISQ)⁹. In Queensland, the average annual tuition fee for an independent school is approximately \$8,000 a year. Data released by ISCA shows a median Australian metropolitan independent school fee of \$6,441 per annum. Nationally, 70% of metropolitan independent schools are charging below \$10,000 a year.

The media continues to fuel the public perception that independent schools only cater for wealthy families. However, given the increasing enrolments and parental confidence in independent schooling, we know that parents are more discerning than these misconstrued public perceptions.

The final word goes to Centre for Independent Studies education policy analyst Blaise Joseph whose excellent piece in *The Spectator*¹⁰ titled "Private schools aren't only for the rich" provides a compelling alternative view to the mainstream media's perception of independent schools. Mr Joseph sums it up with the following statement: "The truth is non-government schools generally don't charge high fees, and also save taxpayers' money, while often having more disadvantaged students than some government schools".

He concludes by saying, "all parents should be supported, no matter what school they choose".



DAVID ROBERTSON
Executive Director

⁸ See <http://education.asg.com.au/cost-index/>

⁹ See <http://isca.edu.au/wp-content/uploads/2018/01/Media-Release-16-January-2017-Independent-schools-challenge-misleading-fee-portrayal.pdf>

¹⁰ At www.spectator.com.au/2018/private-schools-arent-only-for-the-rich/

COMPARATIVE JUDGEMENT: AN OPTION FOR REDUCING MARKING LOADS



MARK NEWHAM
Director (Education
Performance & Improvement)

Comparative judgement, at its simplest, consists of teachers being presented with two pieces of student work and then simply and quickly choosing which of the two is better.

Intensification of workload, particularly driven by assessment demands is an issue for many teachers and schools. This intensification can include the need to report on a five-point scale, new requirements as a result of changes to senior assessment in Queensland, pressure to ensure the consistency of marking results through moderation and the hours devoted to setting assessment tasks and marking them. Given this, perhaps now is the time to look at what comparative judgement might be able to offer.

Comparative judgement could allow for reduced workload for teachers with increased reliability of the professional judgements they make regarding primarily summative student work.

Understanding comparative judgement

Comparative judgement, at its simplest, consists of teachers being presented with two pieces of student work and then simply and quickly choosing which of the two is better. The theory behind comparative judgement is not new. Thurstone's law of comparative judgement, established in 1927, states that all judgements are comparative (Thurstone, 1927). The idea being that judgements are easier to make when comparing one thing against another. So, we can state which of two dogs is larger, but it might be harder to determine a single dog's exact size.

With comparative judgement then, a teacher would skim read, for less than 30 seconds, two pieces of work and identify which is better. Then, another pair of responses is judged and so on until all the pieces of work are ranked. If using software, after a few rounds of comparing pairs an algorithm can sort all the pieces into a rank order. This use of an algorithm, called adaptive comparative judgement allows each piece of work to also be given a score on a quantitative interval scale. With adaptive comparative judgement, the score of each piece of work is re-estimated after each judgement. In subsequent rounds, each script is compared to another script with a similar score to make a finer-grained judgement and a more efficient rank order. This is all completed without the use of a scoring rubric.

Currently, teachers spend many hours attempting to set tasks that are relevant and linked to the real world. These tasks then require the development of detailed criteria rubrics. The necessity of having to qualify and sometimes quantify those aspects of the task that can be identified in a rubric by extension reduces the real-world aspect of the task and can limit the originality able to be produced by the students. This is illustrated in Figure 1.

Rubrics can also be problematic when vague words are used to discriminate between levels of achievement. Use of words such as *fluent*, *consistent* and *extensive* are subjective and can tend to give the veneer of objectivity and the ability to provide fine grained judgements when it isn't necessarily the case. The use of a detailed rubric offers the promise of a higher degree

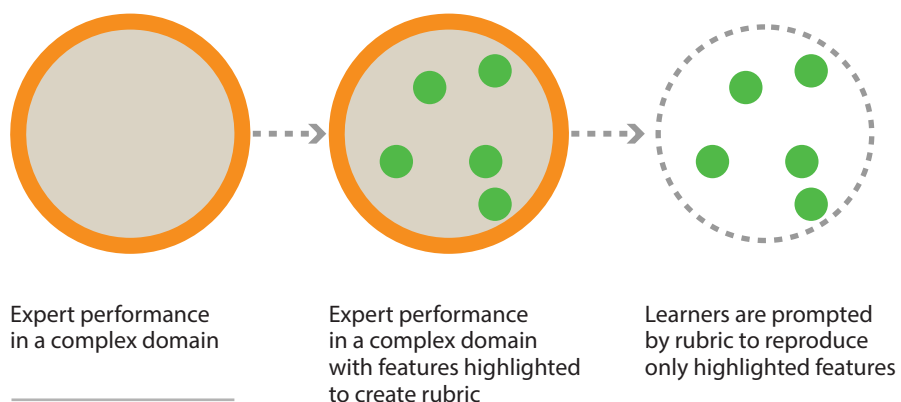
of reliability of any mark assigned by the teacher but the rubric might not always allow us to measure what is important. Therefore, the task's validity may be reduced.

A multiple-choice question could provide reliability but again, the validity might suffer. Essays offer more validity of assessing what you might be interested in but are not necessarily reliable. It seems therefore that the setting and marking of assessment tasks is a constant trade-off between reliability and validity. This work of teachers is further frustrated when the many hours spent by teachers providing written feedback to students is often ignored (William, 2014).

No More Marking, a company that offers support for comparative judgement, completed a study of 1,600 teachers who judged more than 8,500 portfolios and found that the participants showed a high consistency in their judgements – a reliability in excess of 0.84. This high reliability is explained through the involvement of multiple judges. However, because each judgment only takes 30 seconds, it is also very efficient. This efficiency is further enhanced by including several pre-marked essays amongst the scripts. This allows the rank order to show the number of scripts above and below the set benchmarks and further test the reliability of the pre-marked scripts.

A number of Australian studies have also applied comparative judgement to the teacher assessment of students' work. Hedsinger and Humphry (2010) completed a study in Western Australia where teachers judged 30 narrative texts from students aged six to 12 years old. They found high reliability of the rank order (0.982) and strong concurrent validity. This was based on a high correlation ($r = 0.921$) between the paired comparisons results and independent estimates obtained from an experienced examiner using a well-established mark scheme (Whitehouse and Pollitt, 2012).

Figure 1: How rubrics fail



Adapted from Greg Ashman (2015).

Comparative judgement has also been used to assess mathematical understanding (Jones et al., 2015). Their study *The problem of assessing problem solving: can comparative judgement help?* found that the examination writers, when freed from the constraint of producing a marking scheme, designed questions that were less structured and more problem-based than is typical in many current school maths examination papers. They also found the comparative judgement approach to assessing the student work proved successful in terms of inter-rater reliability and validity.

Beyond aspects of reliability, validity and efficiency, the use of comparative judgement for peer assessment is offering promise. Allowing students to judge, and justify their judgements of which script is better, could be a more powerful learning experience for students than many other assessment and feedback activities. Other uses could include students completing a pre-assessment at the start of the year, then using some of those scripts as benchmarks in a later assessment and measuring student progress over time.

Comparative judgement can also be used by individual teachers to identify highly ranked examples of aspects such as fluency, relevance, analysis and introductions that can be used as exemplars with students. Using

comparative judgement within an entire subject department would enable teachers to gain a perspective across different classes, identify inconsistencies of judgement and encourage a professional conversation around desired attributes and teaching approaches. This might take on greater urgency under the changes to Senior assessment to be implemented that require schools to complete internal quality assurance processes, within or across syllabuses cohorts, to ensure validity, reliability and fairness of results.

Possible areas of contention

As promising as comparative judgement might appear as an efficient and effective alternative to marking, there are several areas of contention. A concern is that of the perceived limited role of feedback. The outcome of comparative judgement is that a particular script is ranked at a particular relative position. There is no automatic explanation of why and what needs to be worked on. With a rubric, a mark against particular criteria could, by default, be considered feedback.

Further, if feedback was desired using comparative judgement, a teacher would need to provide reasons for the result and would in effect perhaps have to do the work that comparative

COMPARATIVE JUDGEMENT: AN OPTION FOR REDUCING MARKING LOADS

Tomas Needham, Head of English at Trinity School, outlines a trial of comparative judgement:

“Like many other schools, our summative process involved several stages. Firstly, teachers used to grade their own class’ submissions. In the case of an essay based subject like English, this may mean marking 30 scripts of several A4 sides of writing, a process that could take several hours. Following the isolated marking, teachers would then meet to ‘moderate’ their grades. In my experience, a moderation meeting involved several teachers looking over a piece of work and arguing the toss about whether it was a 6a or a 7c. More often than not, the final grade was reached based on nothing more than the quality of rhetoric espoused by the most articulate, the most passionate or usually the most senior staff member in attendance. As well as the obvious lack of rigour and objectivity, the moderation stage took another hour or so, and within that short time, it was only possible to moderate a small sample of the entire year group, meaning that most scripts went completely unmoderated. So, in total, a department of six would have spent 13 hours on summative assessment. And this is for just one year group!

We ran an initial No More Marking pilot with 120 Year 7 scripts and each teacher took 30–40 minutes to complete their judgments. As there was no need for moderation, the comparative judgment programme performing this role as part of the process, the entire time spent by everyone was two hours!

judgement is aimed at reducing, ie. marking. Again, this assumes that the written feedback is found useful by the student. Also, most feedback is provided, and is perhaps most useful, in the drafting stages rather than after the piece is completed.

Another issue is the assumption that marking schemes limit a student’s ability to respond in original ways. Teachers might believe that a rubric provides a useful scaffold for students and any constraint on originality would only be for the very brightest students and therefore not worth discarding for the vast majority of students.

The advantages of comparative judgement in terms of reliability are also not universally accepted. All teachers want to feel confident that the marks assigned to their students are consistent, both given

from themselves and from other teachers. The reliability of marking for essay-based examinations is usually considered to be between 0.6-0.7. That is, there is a 30%-40% probability that a different marker would assign a different mark. Many comparative judgement studies have found a reliability of 0.8-0.97. However, some academics have raised issues with the accuracy of adaptive comparative judgement. Bramley (2015) argues that there are strong concerns that adaptive comparative judgement overstates the accuracy it claims due to potential flaws in identifying batches of pairs for comparison. Bramley suggests that “other indicators of reliability, such as correlations with measures obtained from comparisons among a different group of judges or correlations with relevant external variables, should be used instead.”

Perhaps a response to this concern is the use of calibrated or pre-marked exemplars that refer to any relevant achievement standards. In a study by Heldsinger and Humphry (2013) teachers collected 60 performances across the range of ability of four to seven-year olds. In Stage 1 of the study the comparative judgement process was used to calibrate the performances of students by developing a scale. These performances were then used as calibrated exemplars. In Stage 2 teachers assessed the student work by judging which calibrated exemplar a performance was most like. Separately, two experienced markers assessed another set of 118 writing performances using both a criterion-based rubric and the calibrated exemplars.

The study found judgements made by experienced markers with the calibrated exemplars correlated well with judgements made using the criterion-based rubric. The findings suggest that using calibrated exemplars has potential as a method of teacher assessment in contexts where extensive training and moderation is not possible or desirable (Heldsinger and Humphry, 2013).

Another area of concern is that in order for comparative judgement to provide validity it requires an assessment piece to be short and narrow in scope, given that the judgement is completed so quickly. Whitehouse and Pollit (2012) found “there is a need for further work to identify the assessments that are most suitable for adaptive comparative judgement”.

Related to validity is the extent to which experts can recognise examples of what they are seeking to measure. Obviously, comparative judgement relies on a teacher’s tacit, internalised knowledge of quality and the successful communication of that to students. It assumes that there is agreement amongst knowledgeable teachers of what quality looks like, even if it can’t be

captured in a standard marking guide. This knowledge might be lacking in inexperienced teachers and without a rubric, the tacit knowledge of the judges becomes even more important. This concern is emphasised by van Daal et al (2017) who make the point that “differences between judges in discriminating ability should be taken into account in the set-up of comparative judgement assessments and in the development of algorithms to distribute pairs of representations.”

Assuming these differences are taken into account, comparative judgement would appear to be popular with many teachers who have trialled it. In one study, Pollit (2012) found that when asked if they would like to use adaptive comparative judgement or marking in future:

- 19 preferred adaptive comparative judgement
- 4 preferred adaptive comparative judgement but with ‘marker training’
- 2 preferred adaptive comparative judgement but expressed some reservations
- 2 stated that both should be used
- 0 chose marking.

Conclusion

So, in conclusion, there are some areas of contention but when compared to rubric scoring, comparative judgement offers some potential advantages. There is less need to moderate and greater likelihood of accurate judgements. Comparative judgement also offers the potential for less time preparing assessment items, allowing teachers extra time for individual interventions. As with any potential activity, teachers will need to ask ‘under what circumstances is comparative judgement useful’ but as Tarricone and Newhouse (2016) state, “comparative judgement delivered by online technologies is a viable, valid and highly reliable alternative to traditional analytical marking”.

Bibliography

- Ashman, G. (2015). Ouroboros. <https://www.ejunkie.com/ecom/gb.php?c=cart&ejc=2&cl=306804&i=1470898>
- Bramley, T. (2015). Investigating the reliability of Adaptive Comparative Judgement. Cambridge Assessment Research Report. Cambridge, UK: Cambridge Assessment.
- Christodoulou, C. (2016). Making Good Progress? United Kingdom: Oxford University Press
- Christodoulou, C. (2016). Comparative judgment: 21st century assessment. <https://thewingtoheaven.wordpress.com/2015/11/15/comparative-judgment-21st-century-assessment/>
- Heldsinger, S., & Humphry, S. (2010). Using the Method of Pairwise Comparison to Obtain Reliable Teacher Assessments. *Australian Educational Researcher*, 37, 1–19. <http://doi.org/10.1007/BF03216919>
- Heldsinger, S., & Humphry, S.M. (2013). Using calibrated exemplars in the teacher-assessment of writing: an empirical study. *Educational Research*, 55, 219–235. <https://doi.org/10.1080/00131881.2013.825159>
- Humphry, S. M., & McGrane, J. A. (2015). Equating a large-scale writing assessment using pairwise comparisons of performances. *Australian Educational Researcher*, 42, 443–460. <https://doi.org/10.1007/s13384-014-0168-6>
- Jones, I., & Inglis, M. (2015). The problem of assessing problem solving: can comparative judgement help?
- Educational Studies in Mathematics*, 89, 337–355. <https://doi.org/10.1007/s10649-015-9607-1>
- Needham, T. (2016). <https://blog.nomoremarking.com/judging-controlled-assessment-3dd152b25118>
- Pollitt, A. (2012). The method of Adaptive Comparative Judgement. *Assessment in Education: Principles, Policy & Practice*, 19, 281–300. <http://doi.org/10.1080/0969594X.2012.665354>
- Tarricone, P. & Newhouse C. (2016). Using comparative judgement and online technologies in the assessment and measurement of creative performance and capability. *International Journal of Educational Technology in Higher Education*, 13. <https://doi.org/10.1186/s41239-016-0018-x>
- Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review*, 34, 273–286. Retrieved from <http://dx.doi.org/10.1037/h0070288>
- van Daal T., Lesterhuis M., Coertjens L., van de Kamp M. T., Donche V. & De Maeyer S. (2017). The Complexity of Assessing Student Work Using Comparative Judgment: The Moderating Role of Decision Accuracy. <https://doi.org/10.3389/educ.2017.00044>
- Whitehouse, C., & Pollitt, A. (2012). Using adaptive comparative judgement to obtain a highly reliable rank order in summative assessment. Manchester: CERP.
- William, D. (2014). <http://www.dylanwilliamcenter.com/is-the-feedback-you-are-giving-students-helping-or-hindering/>

ISQ supports schools to think and learn more about assessment and understanding educational data in a number of ways in projects, face to face workshops and online. A number of courses are available for self-paced learning through [ISQ Connect&Learn](#).

BUILDING ASSESSMENT COMMUNITIES

- promoting capacity building of assessment expertise amongst school staff
- designing quality assessment tasks
- developing effective rubrics
- establishing positive social moderation processes to ensure consistency of teacher judgments.

EDUCATIONAL DATA

- A suite of six modules designed to help learners to understand the education data available and support them in evidence-based decision making for individual students and cohorts as well as strategic planning.
- Data informed pedagogy and whole of school educational data projects.

ISQ thanks its 2018 Alliance Partners

Gold+ Alliance

**ngs
Super**

Your fund.
Your wealth.
Your future.



CommonwealthBank

.smart.
TEACHERS

part of **tes**

Bronze Alliance



**Willis
Towers
Watson**

Disclaimer:

The information contained in this publication is to the best of our knowledge and belief correct at the date of publication. However, no warranty or guarantee is or can be given by Independent Schools Queensland or any member of its staff, and no liability is or can be accepted for any loss or damage resulting from any person relying on or using the information contained in this publication.

FEBRUARY 2018
Licensed under NEALS

Independent Schools Queensland Ltd
ACN 614 893 140 ABN 88 662 995 577



FIFTY YEARS
SUPPORTING
CHOICE &
DIVERSITY



**Independent Schools
Queensland**

Head Office
1st Floor, 96 Warren Street, Spring Hill Q 4000
PO Box 957, Spring Hill Q 4004
P (07) 3228 1515 E office@isq.qld.edu.au

Professional Learning Centre
Level 5, 500 Queen Street, Brisbane Q 4000
P (07) 3228 1507 E events@isq.qld.edu.au

www.isq.qld.edu.au