Submission to the Inquiry into the Australian Education Amendment Bill 2017

Gary N. Marks
Australian Catholic University

This submission reflects the views of the author and should not be read as reflecting the views of the Australian Catholic University or any other organization associated with the author or the Australian Catholic University.

Acronyms

ACARA	Australian Curriculum, Assessment and Reporting Authority
ACER	Australian Council for Educational Research
AEU	Australian Education Union
ATAR	Australian Tertiary Entrance Rank (formerly ENTER)
CIS	Centre for Independent Studies
ENTER	Equivalent National Tertiary Entrance Rank
ESCS	Economic, Social and Cultural Status. The OECD's SES measure.
HILDA	Household Income and Labour Dynamics in Australia study
IEA	International Association for the Evaluation of Educational Achievement
LSAC	Longitudinal Study of Australian Children
LSAY	Longitudinal Surveys of Australian Youth
NAPLAN	National Assessment Program—Literacy and Numeracy
NDIS	National Disability Insurance Scheme
NLSY	National Longitudinal Survey of Youth (US)
PISA	Programme for International Student Assessment
SEIFA	SocioEconomic Indices for Areas
SES	Socioeconomic Status
SRS	Schooling Resource Standard
TIMSS	Trends in Mathematics and Science Studies

Contents

Executive Summary

I submit to the Senate Standing Committee that the Gonski funding model on which Gonski 2.0 is based is severely flawed.

1. The funding model very much overemphasizes the importance and impact of socioeconomic status (SES) on educational outcomes. The Gonski model assumes that students' socioeconomic status is the most important influence affecting student outcomes. SES dominates the Gonski funding model. It appears twice in the funding model, first as part of the loadings for disadvantaged schools and second, to reduce the per capita recurrent funds to non-government schools based on the ability of parents to pay. Half of all students are deemed low SES, but only small minorities are Indigenous, immigrant students or have disabilities.

2. The model is based on a serous misunderstanding of what influences educational performance with its emphasis on SES. SES is only *moderately* correlated with educational outcomes, note the emphasis on the word moderately, because SES is correlated with parental ability, parental ability is correlated with their children's ability and student ability is easily the strongest influence on educational performance, most strongly for performance in cognitive tests such as PISA and NAPLAN which are very like ability tests. The evidence that ability, and not students' SES, is the major influence on educational outcomes is overwhelming.

3. Given that student performance is driven mainly by ability, it is not logical to expect that the Gonski model will reduce SES inequalities in education. It is almost certain that the SES-achievement/performance relationship will not change with increased Gonski funding. Therefore, the likely future scenario is further calls for even greater government funding for low-SES schools. The most likely legacy Gonski will have on the present generation of school students is higher taxation and reduced social welfare during their adult lives because of ineffective government expenditure.

4. The Gonski model assumes that low and high performing students are limited to low and high SES schools. Low and high performing students are found throughout the education system and almost all schools have their share of high and low performing students. Low achievers attending schools not deemed as disadvantaged may not be helped in the Gonski model.

5. The Gonski model intends to use poor measures of SES. The proposal to use error prone area-measures as a proxy for SES (which is in turn is only weakly related to student performance) is particularly problematic

6. The evidence from the literature is that school resources have a negligible or very small impact on student outcomes especially in developed countries such as Australia. Real per student expenditure has increased substantially over the last 3 decades with no increase in student performance. It is not logical that further expenditure will improve performance.

7. The Schooling Resource Standard is arbitrary.

8. Contrary to the frequently made assertion that the Gonski model is sector blind, government funds to non-government schools are reduced by a convoluted, inaccurate and supposedly temporary procedure to estimate parents' capacity to pay. In contrast, there is no reduction in funding to government schools even if they are in Australia's wealthiest localities serving privileged communities.

9. There is no detail on how best to best spend the Gonski money. The report provides no discussion on what educational programs should be implemented; this is left for Gonski 2.0. Given that for over 3 decades, the educational effectiveness literature has not established "what works", it is unlikely the Gonski 2.0 will establish what education programs are most effective for Australian students over the next 6 months.

10. Even if in the unlikely event that the 2017 Gonski report establishes unambiguously "what works" for Australian education, many of programs that already have support from the literature—greater emphasis on phonics, more frequent testing and feedback, direct instruction, academic press, to name a few—would be unpalatable to large sections of education industry generating further political battles. Furthermore, can the federal government really expect that bureaucrats located in Canberra or in state and territory capital cities can control what teachers do in front of classes. It would be far better to allow schools autonomy to develop programs tailored to their students. Of course, there should be some evaluation of whether these interventions are successful but that should be left to the jurisdictions so avoid costly double-handling.

11. Expenditure of government money on this scale needs to be properly evaluated using state-of-the-art evaluation methods. Given the dearth of proper Australian evaluation studies and the absence of accurate, complete longitudinal across all jurisdictions, the

likelihood that programs and interventions using Gonski funding will be properly evaluated is low.

These criticisms of the Gonski funding model are supported by this report which details the empirical evidence relevant to Australian education: student achievement, SES, cognitive ability, schools and teachers.

1 Student Achievement

National assessments of educational achievement aim to provide evidence about the levels of student achievement in identified curriculum areas (for example, in the areas of reading and mathematics) for particular grade levels (Postlethwaite & Kellaghan, 2008). The National Assessment Program in Literacy and Numeracy (NAPLAN) is the Australian national assessment.

International assessments of student achievement allow comparisons between different educational systems aiming to identify what policies promote both higher achievement and equity. The major international assessments are Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS) directed by the IEA (International Association for the Evaluation of Educational Achievement) and the OECD's Program for International Student Assessment (PISA). PIRLS focuses on reading in Year 4, TIMSS on mathematics and science in grades 4 and 8, and PISA on student achievement in reading, mathematics and science of 15-year-olds. PIRLS has been conducted on a five-year cycle since 2001, TIMSS on a four-year cycle since 1995 and PISA a three-year cycle since 2000. In TIMSS, PIRLS and PISA, student achievement is measured on scales set at a mean of 500 and a standard deviation of 100.

1.1 Achievement has important consequences

Not only does student achievement serve to monitor education systems, student achievement is strongly predictive of subsequent high-stakes educational and post-school outcomes. Prior achievement is the strongest influence on a range of high-stakes educational outcomes that are consequential for young people's labour market prospects.

- Ainley and Sheret (1992, p. 146) report a correlation of 0.71 between a composite measure of achievement in Years 9 and 10 and Tertiary Entrance Rank. Hemmings and Kay (2010) reported a correlation of 0.77 between Year 7 numeracy scores and performance in mathematics in the Year 10 NSW school certificate examinations. Year 9 achievement correlates at around 0.7 with ATAR its effects on ATAR score are much stronger than that for SES including (Marks, 2015b). Earlier studies on tertiary entrance also show substantially stronger effects of achievement measured in Year 9 or at 15 years of age than SES (Marks, 2009a, 2010b, 2010c).

- Similarly, prior achievement is the strongest influence on university entrance/participation (Marks, 2010a; Marks, Underwood, Rothman, & Brown, 2011; OECD, 2010b)
- Low achievement is the strongest influence on early school leaving, much stronger than SES effects (Marks, 2007, 2014c).
- Achievement is easily the strongest influence on subject choice in senior secondary school, much stronger than the effects of SES (Marks, 2013).

The importance of student achievement fits within the wider literature of the effect of test scores collected in adolescence on subsequent educational and labour market outcomes (Jencks et al., 1979; Korenman & Winship, 2000; Marks & McMillan, 2003; McIntosh & Munk, 2007; McNiece, Bidgood, & Soan, 2004).

The greatest failing of international achievement studies is that they are cross-section in design so cannot include prior performance. However, Zimmer and Toma (2000) analysed longitudinal IEA data from eight countries that conducted pre-school year math tests in addition to the end of-year exams. The effects of prior achievement were very strong equivalent to a standardized effect of 0.81 (author's calculation) with weak effects for SES measures and school-SES.

The average cognitive skills of the population, or the "knowledge capital" of a nation, as measured by international achievement studies is associated with higher GPD growth, increased productivity, wealth creation and countries with higher levels of cognitive skills make a greater contribution to human knowledge, e.g. noble prize winners, inventions (Hanushek & Woessmann, 2016; Rindermann & Thompson, 2011).

In Gonski and much of Australia educational research the strong influence of achievement of high-stakes educational outcomes is simply ignored because it undermines the predominant SES narrative.

1.2 Student achievement in Australia is declining

Over time, average scores in student achievement in Australia have either remained steady or declined.

An analysis of students' literacy and numeracy 14-year-olds in LSAY surveys between 1975 and 1998 found no change (Rothman, 2002). The only positive change in average achievement scores over time in Australia was an increase in Year 4 average mathematics scores in TIMMS between 1996 and 2007. However, there was no significant change in Year 4 science although the average scores declined from 527 in 2007 to 516 in 2011 (Thomson, Hillman, Wernert, Schmid, & Buckley, 2012, pp. vii-viii).

The following information is from departmental responses to a question on notice from a member of the Senate standing committee on education (Department of Education and Training, 2017).

> Australian students' absolute performance in PISA has significantly declined in every domain:
>
> - performance in scientific literacy has declined by 17 score points, or seven months of schooling, since 2006 (527 to 510)
> - performance in mathematical literacy has declined by 30 score points, or one year of schooling, since 2003 (524 to 494)
> - performance in reading literacy has declined by 25 score points, or ten months of schooling, since 2000 (528 to 503).
>
> Australian students' relative performance has also significantly declined between 2012 and 2015:
>
> - in scientific literacy Australia was significantly outperformed by nine countries in 2015, compared to seven countries in 2012
> - in mathematical literacy Australia was significantly outperformed by 19 countries in 2015, compared to 16 countries in 2012
> - in reading literacy Australia was significantly outperformed by 11 countries in 2015, compared to nine countries in 2012.
>
> This decline has been accompanied by real increases in government expenditure on education (see section 4.5).
>
> Australia's TIMSS results have remained largely the same in absolute terms. However, other countries have improved and Australia's relative performance declined between 2011 and 2015:
>
> - in Year 4 mathematics Australia was significantly outperformed by 21 countries in 2015, compared to 17 countries in 2011

- in Year 8 mathematics Australia was significantly outperformed by 12 countries in 2015, compared to six countries in 2011
- in Year 8 science Australia was significantly outperformed by 14 countries in 2015, compared to nine countries in 2011

The only assessment to see an improvement in relative performance was Year 4 science, in which Australia was significantly outperformed by 17 countries in 2015, compared to 18 countries in 2011.

The CIS present graphically changes in changes in student achievement in PISA with changes in expenditure (Reproduced in Figure 1).

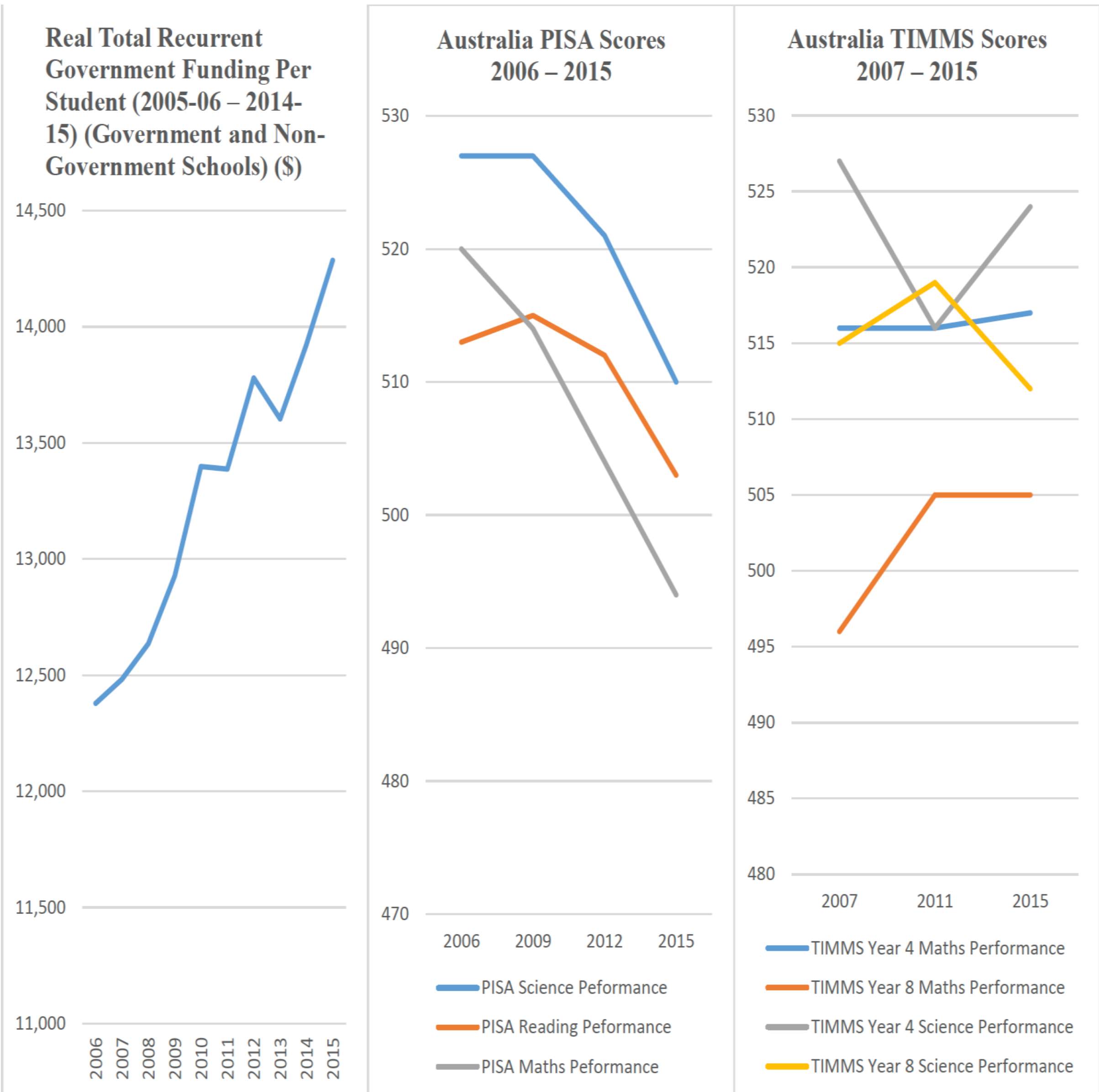

Figure 1: Changes in Student Funding and Student Achievement. From Joseph (2017, p. 15)

1.3 Student achievement is very stable

Within individual students, student achievement is very stable. The over-time correlations of achievement range from 0.5 to nearly 0.9, depending on the age or grade level of the students, the achievement domain, the number of years between tests and the reliability of the test. In the US, the correlations of test scores of the same students measured at age 8 to 10 and age 18 were between 0.7 and 0.8 (Jencks et al., 1972, pp. 59-60). Reynolds and Walberg (1992, p. 318) reported a correlation of 0.73 between mathematics achievement in grades 7 and 8. Marks (2016a) documents same-domain over-time correlations from more recent studies. Armor (2003, p. 33) presents correlations for combined reading and math achievement for New York City students from Grades 2 to 3. For adjacent grades the correlations are range from 0.8 at lower grade levels to nearly 0.9 at higher grade. The correlation in Grade 3 and Grade 8 scores was 0.73. Armor notes that the correlations at higher grades are so high, that very little true change occurs (after correcting for reliability).

For Australia, the correlations between NAPLAN achievement and same-domain prior achievement are also high: ranging from about 0.6 for writing to above 0.8 for numeracy, reading and spelling. For numeracy and spelling, the correlations are close to 0.9. These strong relationships are very much stronger (0.5 to 0.8) than the correlations between SES and student performance (0.2 to 0.4). These correlations are documented for both Victoria NAPLAN data the national data LSAC study (Marks, 2016a; Marks, 2014a).

The high-level of stability of student achievement and the strong effects of prior achievement are because student achievement has a large genetic component (see section 1.5)

The high-level of stability of student achievement and the strong effects of prior achievement has been well-documented in Australia and elsewhere for several decades. The stability of student achievement is well-known, at least, since at least the early 1970s (see Suter, 2000). There is no excuse for researchers and especially policymakers to ignore the evidence and its important policy implications (see section 2.10).

1.4 Prior achievement is not a function of SES and schools

The standard, but naïve, response to the finding that SES effects are trivial when considering prior achievement (see section 2.10) is that prior achievement is simply determined by SES justifying the emphasis on SES. This argument is not supported empirically. Since SES has bivariate correlations of only around 0.3 (up to 0.4) with achievement (see section 0), it cannot account for the much larger effects of prior achievement.

If prior achievement was simply a function of SES, the effects of prior achievement would disappear when controlling for SES. This is not the case. Marks (2017b) reported standardized estimates for prior achievement are over 0.7 compared to around 0.1 for SES. Even in the presence of early childhood ability, the effects of prior achievement are large.

Nor can the strong effects of prior achievement be accounted for by differences between schools. Analysing combined literacy and numeracy NAPLAN scores for New South Wales, Lu and Rickard (2014, p. 32) report very large standardized effects of "at least 0.8 standard deviations" for prior achievement, net of SES, demographic variables and random effects for schools. For Victoria, the within-domain standardized effects of prior achievement ranged from 0.5 (for writing) to over 0.8 (for spelling) with only small reductions in the magnitudes of its effects when controlling for SES and schools (Marks, 2014a).

1.5 Student achievement has a large genetic component

Adoption, sibling and twin studies allow estimates of the heritability (h^2) of an attribute or outcome: that is the variation attributable to genetic factors as a proportion of the total variance. The classic measure of heritability is twice the difference in the within-pair correlations between identical twins and non-identical twins or siblings (Jensen, 1998, p. 200; Nielsen, 2006, p. 201).

A meta-analysis of 61 twin studies from 11 cohorts of primary school children showed the average heritability estimates of around 0.7 for reading, 0.5 for reading comprehension, 0.6 for mathematics, 0.6 for language, 0.4 for spelling and 0.7 for general educational achievement. The contributions of the shared environment were substantially smaller with estimates mostly around 0.10 (de Zeeuw, de Geus, & Boomsma, 2015).[1] The heritability of student achievement in primary school is greater than that for cognitive ability (Kovas et al., 2013). There are sizeable genetic correlations between achievement domains with cognitive ability indicating common sets of genes (Hart, Petrill, Thompson, & Plomin, 2009; Petrill, 2016; Wainwright, Wright, Luciano, Geffen, & Martin, 2005). However, student achievement is not simply cognitive ability, there are other genetic traits involved (Krapohl et al., 2014). It is not clear what are relative contributions to student achievement of general cognitive ability, innate specific skills (for example in numeracy or spelling) and schooling and if these contributions change over the school career.

Australian twin studies also show strong genetic components to student achievement. The heritabilities for student performance in NAPLAN performance are between 0.6 and 0.7 for reading (except for Year 5 for girls), 0.7 to 0.8 for spelling, 0.50 to 0.66 for grammar, 0.4 to 0.5 for writing and 0.4 to 0.8 for numeracy. In contrast, the proportion of the variation due to the common (family) environment was much smaller, generally lower than 0.2 but often much less (Grasby, Coventry, Byrne, Olson, & Medland, 2016). They estimate that 75 to 80% of the covariation in NAPLAN performance across years and domains is due to common genetic factors. A second study concluded that much of stability in NAPLAN performance from Year 3 to Year 9 is due to genes (Grasby & Coventry, 2016). An earlier Australian study found that more than half of the correlation between general intelligence and student performance is due to common genes (Wainwright et al., 2004). These high estimates further undermine the assumption of the Gonski report that student performance is mostly about SES.

Genetic heritability increases once the children undergo formal instruction especially at school (Asbury & Plomin, 2014, pp. 22-30; Samuelsson et al., 2007). New sets of genes come into play as children grow older and there are specific (e.g. for reading or numeracy) as well general genetic factors (Byrne et al., 2009; Calvin, Fernandes, Smith, Visscher, & Deary, 2010; Kovas et al., 2013). Heritability tends to increase with age for reading (Soden et al., 2015; Wadsworth, Corley, Hewitt, & DeFries, 2001; Wadsworth, Corley, Plomin, Hewitt, & DeFries, 2006) but not for science (Haworth, Dale, & Plomin, 2009). The higher heritabilities for student achievement as children age accounts for the higher same-domain correlations for students at higher grades.

Studies showing strong genetic component to student achievement stretch back to the 1990s for Australia and earlier for other countries. There is no excuse for researchers and especially policymakers to ignore the evidence and its policy implications. There are several discussions on the policy implications on the importance of genetics to education (see Asbury & Plomin, 2014; Kovas, Malykh, & Gaysina, 2016).

1.6 Educational Attainment also has a large genetic component

Educational attainment is usually measured by the highest level of education completed; measured in years of formal education.[2] It also has a sizable genetic component, although not as large as that for student achievement.

After reviewing a number of twin and adoption studies on educational attainment, Sacerdote (2011, p. 12) concludes "Genetic effects play a large role, while there is only a small role for family environment". Using Italian parent and child data, Lucchini, Della Bella and Pisati (2013) estimate a heritability for educational attainment of 0.5 and question the usefulness of traditional sociological accounts used to explain inequalities in education.

For Australia, Baker et al. (1996) estimated a heritability for educational attainment of 0.58 and proportion of the variance attributed to the shared environment at 0.24. In later studies, heritability estimates range from 0.50 to 0.65 and estimates for the common environment range from 0.15 to 0.26, again with no gender differences (Le, Miller, Slutske, & Martin, 2011, p. 132; Miller, Mulvey, & Martin, 2001). The most recent study estimates a heritability of 0.55 for education and sizable genetic correlations between education, occupation and income (Marks, Forthcoming).

1.7 Prior achievement is necessary for estimating school and teacher effects

In other to assess the extent that schools, teachers, interventions and programs influence student outcomes (discussed in sections 4 and 5), it is necessary to control for students' prior achievement.

Gray, Goldstein and Thomas (2001) conclude that prior achievement is the most appropriate measure to assess school effectiveness. Willms (1992, p. 58) argued that prior achievement was essential for the analysis of school and program and effects and that simply controlling for social background factors was inadequate. Similarly, Lenkeit (2013, p. 53) concludes that "prior achievement scores in the prediction of achievement status considerably contributes to explaining differences between students and schools in comparison to a model with family background characteristics only". Raudenbush and Willms (1995, p. 313) argue that "highly reliable premeasures of achievement or aptitude" are required for defensible estimates for school effects. Scheerens and Bosker's (1997a, p. 54) initial value-added model, unpredicted student achievement, includes aptitude as well as social background and demographic variables. Their learning gain model adjusts only for prior achievement and the "unpredicted learning gain model" model includes prior achievement in addition to aptitude, socioeconomic and sociodemographic variables.

More complex models use not just students' same-domain prior achievement but their test scores in different domains across the educational career. The multivariate response model (MRM) is a multivariate, longitudinal, linear mixed model which analyses the complete set of observed test scores belonging to each student across multiple years and domains with categorical variables for teacher, school, district and year. The univariate response model (URM) is a more limited model that analyses student scores in a particular subject and grade but includes all previous test scores (Wright, White, Sanders, & Rivers, 2010; Wright, 2010). The rationale for including multiple prior test scores is that students' performance correlates across test domains and multiple test scores reduce the error component.

Including prior achievement changes the conclusions about teacher and school effects. Ballou, Sanders and Wright (2004) found that student SES and demographics had a negligible impact on teacher effects in models with multiple measures of prior achievement. Analyzing Russian data from students that participated in both PISA and the Grade 8 TIMSS study so includes a student-level measure of prior achievement, Carnoy et al. (2016) conclude "OECD claims about raising students' PISA scores by improving school/classroom resources are overstated".

Few Australian studies of school and teacher effects incorporate prior achievement, so their results will most often be spurious; suggesting school, teacher and program effects are much stronger than they are.

2 SES

In Australian education, students' socioeconomic status (SES) has a central place in both research and policy circles. There is a pervasive view in academic publications, commissioned reports and commentary that SES is the most important driver of students' educational outcomes (for a recent example see Lamb, Jackson, Walstab, & Huo, 2015). The Executive Summary of the 2011 Gonski (2011, p. xiv) report when justifying its recommendations argues that additional funding is needed to ensure that "outcomes are not determined by socioeconomic status or the type of school the child attends". The argument is that SES has strong causal relationships with almost every academic outcome: from children's reading and numeracy in primary school to their university entrance rank. Research demonstrates that it clearly does not.

2.1 Gonski provides no theoretical rationale why SES is important

Although the Gonski report emphasizes SES, it does not provide a plausible evidence-based theoretical explanation of the SES-education relationship. The general assumption from advocates of SES funding models (such as Gonski) is that the relationships between SES and student performance are because of financial resources, parenting, cultural capital, attitudes to education, socialization etc. (there is a plethora of possible explanations, none of which enjoys a consensus among researchers or much empirical support, see Marks, 2014b).

The contention that student performance in specific subject areas is not about the provision of teaching and learning but is really about SES does not make sense theoretically. Take for example, student performance in a foreign language which logically is a function of students' innate language abilities and how and what they have been taught at school. Unless parents are fluent speakers of that language, parental effects must be small or negligible. Similarly, student performance in solving quadratic equations, doing calculus, balancing chemical reaction equations, learning ancient history, comprehending and make inferences from complex text, and writing persuasively, is overwhelming a function of teachers and schools, not SES. The same argument applies to students' performance in cognitive tests such as PISA, TIMMS, PIRLS and NAPLAN. At most parents can supplement what students learn at school but the bulk of student learning occurs at school by teachers.

2.2 SES is only moderately associated with student performance

The Gonski report places much emphasis on the importance of SES for student outcomes to justify the proposed SES funding model. However, the empirical evidence unambiguously shows that the bivariate relationship between SES and education moderate. Note that this is the *bivariate* relationship; the effects of SES are much weaker when considering other stronger influences on SES (i.e. prior achievement, cognitive ability) in a multivariate context.

The magnitude of the relationship between students' SES and student achievement is, at best, moderate. White's (1982) meta-analysis of over 200 mainly US studies found a mean correlation of only 0.22 between SES, measured various ways, and student achievement. A later study by Sirin (2005) with better SES measures reported an average correlation of 0.27, equivalent to about 9% of explained variance.

Australian studies show that the relationship between SES and achievement is only moderate. The correlations between measures of SES and student performance are round 0.3 ranging between 0.2 and 0.4 depending on the measure used. I have documented many Australian studies in a recent journal article (Marks, 2017b). Correlations of 0.3 translate to only 9% of the variation in student achievement explained by SES. This contrasts with the very much stronger effects of prior achievement with correlations between 0.5 and 0.9 (see section 1.3)

The only moderate correlations between SES and student achievement has been well-known both internationally (White, 1980, 1982; White, Reynolds, Thomas, & Gitzlaff, 1993) and for Australia (Ainley, Graetz, Long, & Batten, 1995) for several decades. There is no excuse for researchers and especially policymakers to ignore the evidence and pretend that SES is the dominant influence.

The OECD's PISA study uses a composite measure of socioeconomic background, Economic Social and Cultural Status (ESCS). It includes cultural items and possessions as well as parents' education and occupation reports that 12% of the variation in Australian students' PISA scores can be attributed to the OECD's broad measure of SES (OECD, 2016b). Despite what some commentators try to argue 12% is not an enormous number. Some commentators have pretended that the ability of SES to account for variation in student performance is somehow cumulative: 12% for one year then 24% for two years and presumably over 100% by Year 12. This is statistically very wrong and such an argument reflects poorly on the competence on the academics who mistakenly overemphasise the significance of SES.

It is important to emphasize that when considering the impact of prior achievement, the variance explained by SES is closer to 1-2% (from standardized effects of 0.1-0.2, see section 2.10); too weak a relationship for SES to figure prominently in funding models.

2.3 SES effects are not stronger in Australia than in other countries

Commentators often claim that SES inequalities in Australian education are large compared to other countries. This claim is not at all supported by evidence. Table 1 presents the variance explained (R Square values) in the PISA tests by the OECD's measure of Economic, Social and Cultural status. The estimate for Australia is just below the OECD average. Note that the true effects of ESCS are much weaker than that suggested by the Table 1 since the estimates do not consider the strong effects of prior achievement (or ability) which correlate strongly with achievement (see sections 1.3). A correlation of 0.7 between achievement and prior achievement is equivalent to an R square of 49% which dwarfs the R square of 12% for the ESCS-PISA relationship in Australia.

Contrary to what is often claimed, the intergenerational transmission of socioeconomic inequality is not stronger in Australia than that in other developed countries. In fact, the opposite is true; Australia exhibits one of the lowest intergenerational correlations for socioeconomic outcomes in industrialized countries (Blanden, 2013; de Broucker & Underwood, 1998, p. 35; OECD, 2010a, pp. 181-198; Ranasinghe, 2015, p. 1904). According to Ranasinghe (2015, p. 1904), Australia is ranked 20 among 20 countries in the size of the parent-child correlation for education.

Table 1: Percentage of Variance Accounted in PISA Performance 2000-2015 (R Square Values)

	2000 Reading	2003 Mathematics	2006 Science	2009 Reading	2012 Mathematics	2015 Science
Australia	17	13.7	11.3	12.7	12.3	11.7
Austria	14 [a]	16.0	15.4	16.6	15.8	15.9
Belgium	21	24.1	19.4	19.3	19.6	19.3
Canada	11	10.5	8.2	8.6	9.4	8.8
Chile					23.1	16.9
Czech Republic	20	19.5	15.6	12.4	16.2	18.8
Denmark	15	17.6	14.1	14.5	16.5	10.4
Estonia					7.6	7.8
Finland	9	10.8	8.3	7.8	9.4	10.0
France	22	19.6	21.2	16.7	22.5	20.3
Germany	22	22.8	19.0	17.9	16.9	15.8
Greece	15	15.9	15.0	12.5	15.5	12.5
Hungary	26	27.0	21.4	26.0	23.1	21.4
Iceland	5	6.5	6.7	6.2	7.7	4.9
Ireland	13	16.2	12.7	12.6	14.6	12.7
Israel					17.2	11.2
Italy	11 [a]	13.6	10.0	11.8	10.1	9.6
Japan	6[a]	11.6	7.4	8.6	9.8	10.1
Korea	9	14.2	8.1	11.0	10.1	10.1
Latvia					14.7	8.7
Luxembourg	24	17.1	21.7	18.0	18.3	20.8
Mexico					10.4	10.9
Netherlands	15	18.6	16.7	12.8	11.5	12.5
New Zealand	16	16.8	16.4	16.6	18.4	13.6
Norway	13	14.1	8.3	8.6	7.4	8.2
Poland	14	16.7	14.5	14.8	16.6	13.4
Portugal	20	17.5	16.6	16.5	19.6	14.9
Slovak Republic	-	22.3	19.2	14.6	24.6	16.0
Slovenia					15.6	13.5
Spain	16	14.0	13.9	13.6	15.8	13.4
Sweden	11	15.3	10.6	13.4	10.6	12.2
Switzerland	19	16.8	15.7	14.1	12.8	15.6
Turkey	-	22.3	16.5	19.0	14.5	9.0
United Kingdom	19	19.7	13.9	13.7	12.5	10.5
United States	22	19.0	17.9	16.8	14.8	11.4
OECD Average[b]	20	20.3	14.9	14.0	14.8	12.9

Note from OECD (2001, p. 308; 2004, p. 399; 2007a, p. 184; 2010c, p. 55; 2013, p. 15; 2016b, p. 402)

a: Questionable estimate due to sample or missing data issues

b: Number of OECD countries increased between 2000 and 2016.

2.4 SES inequalities in Australian education are declining

In the heated debates about Gonski and education, it is often claimed that socioeconomic inequalities in education are stable, if not increasing. Such claims are made to support the argument that the Gonski type SES funding models should be implemented, in full, as soon as possible. The research evidence indicates that SES inequalities in Australian education are declining.

Table 1 shows that the explanatory power of their composite measure of SES has declined between 2000 and 2015.

Furthermore, the research evidence finds that in Australia socioeconomic inequalities in educational attainment, university entrance, occupational status and earnings have declined over time (Marks, 2009b; Marks & McMillan, 2007; Marks & Mooi-Reci, 2016; Ranasinghe, 2015, p. 1904). According to Ranasinghe (2015, p. 1905) the parent-child correlation for education in Australia has declined from 0.38 for the cohort born 1942-1946 to 0.24 for the cohort born 1987-1991.

2.5 Family income has little or no relationship with educational outcomes

One implication of the SES model of educational outcomes is that low SES families lack the economic resources to enhance their children's education. On the other hand, wealthier families can afford to send their children to high-fee independent schools or buy a residence in the catchment area of a high-performing government school. However, family income has very weak or negligible relationships with educational outcomes.

The bivariate relationship between family income and NAPLAN performance is weak; weaker than that for other aspects of SES such as, parental occupation and parental education. Furthermore, net of cognitive ability or prior achievement the effects of family income are close to zero (Marks, 2017a). When taking into account stable differences between students (such as intelligence, motivation, personality etc.), changes in family income have no impact on student performance in any of the five NAPLAN achievement domains (Marks, 2016b). Dockery, Seymour and Koshy (2016) analysing HILDA data study found no significant effects for real household disposable income (accurately measured) on the probability of a 17-year old going to university, net of other influences.

The conclusion that in Australia the effect of family income on student achievement are small and its direct effects are very small or negligible is consistent with overseas studies. For the United States, Mayer (1997, pp. 90-91) estimated conventional standardized effects of 0.13 for family income on Peabody Picture Vocabulary Test scores, 0.06 for mathematics scores in the Peabody Individual Achievement Test (PIAT) and 0.14 for PIAT reading. Her 'true effects' of family income were usually smaller and not statistically significant.[3] Analysing mathematics achievement, Orr (2003, pp. 291, 293) reported no effect for family income (averaged over 5 years) on mathematics achievement, net of father's occupational status, mother's education, other variables and importantly mother's ability (see section 2.11). Carlson and Corcoran (2001, p. 789) analysis of reading scores in children aged 7 to 10 found that family income had an impact but it was relatively weak. A doubling of family income increased child's reading score by about 3.2%. For Britain, Violato et al. (2011) concluded "a weak or absent direct effect of family economic resources on child development". Similarly, Aughinbaugh and Gittleman (2003, p. 429) analysis of children's test scores in the United States and Britain found that the effects of family income on test scores were quite small, the maximum effect was 0.08 of a standard deviation, net of other predictors including mother's ability. Analysing data from South Africa, Cherian and Malehase (1998, p. 431) concluded there was "no relationship between financial conditions at home and scholastic achievement of children from single-parent and two-parent families". Analysing student achievement in PISA in Denmark, Humlum (2011, p. 994) noted that the effects of family income were small and statistically insignificant. A substantial change in permanent income of 100,000 Danish Krone (equivalent to about $US15,000) was associated with a difference of only 2.6 PISA score points.[4]

Therefore, there is no evidence to support that the financial resources of students' families are important to their educational outcomes.

2.6 Area-based measures of SES weakly relate to student achievement

Gonski model of school funding recommends that area based SES measures should continue to be used to allocate funds to non-government school until a better measure is developed (Gonski et al., 2011, p. xi). Area-based SES measures only weakly relate to student achievement, substantially weaker than the only moderate SES-achievement correlations among students (see section 0).

Area-based measures of SES show weaker associations with student achievement than individual-level measures. Ainley and Long (1995a, p. 74) reported correlations of SEIFA indexes with student achievement ranging from 0.16 to 0.23, about one-third less than the correlations between composite individual-level measures of SES and student achievement. In non-urban areas the correlations between census district (CD) based SEIFA measures and student achievement are very low, at around 0.07 (Ainley & Long, 1995b, p. 80). The correlations between area-based areas and student achievement are also weak when using geographical areas larger than CDs (Postcode areas and Statistical Local Areas) ranging from 0.13 to 0.22 (Ainley & Long, 1995b, pp. 83-84).

Marks (2017b) demonstrates that the correlations of SEIFA measures with NAPLAN performance are generally low ranging between 0.08 and 0.28. The "economic resources" measure exhibits particularly weak correlations. This is consistent with the very weak effects of family income on education (see section 2.5). It is only the more general SEIFA measure of education and occupation that shows moderate correlations with student achievement. These correlations cannot be considered as casual in the sense that a students' location affects their performance in NAPLAN. The more likely explanation is that students with characteristics that enhance performance in cognitive tests tend to live in areas with higher proportions of professional and more highly educated persons, but this tendency is not very strong. The tendency for such students to live in areas with greater economic resources is particularly weak.

Given the weak associations between area-based measures of SES and student achievement funding models that rely on area-based SES measures, as does Gonski, poorly target low achieving students and will do little, if anything, to raise standards.

2.7 Area based measures of SES weakly relate to Student SES

It is inappropriate to use area based measures as proxy measures for student SES. Marks (2017b) also shows that area-based measures are poor proxies for individual level measures of SES. SIEFA measures correlate between 0.2 and 0.5 with the individual-level measures of SES. The SEIFA measure "Economic Resources" displays the weakest correlation with the SES measures, although this is the preferred measure for economic analyses. In a comparison of area and individual SES measures for higher education using HILDA data, Dockery, Seymour and Koshy (2016) conclude "The results demonstrate the tendency for area measures to misclassify individuals' higher education opportunity and the associated potential for perverse policy outcomes".

2.8 The SES measures advocated by Gonski are poor

Even if it assumed that SES has a determining influence on educational performance, which it does not (see sections 0 and 2.10), the measurement of SES is very poor.

For the loadings, enrolment data is used to collect information on parents' education and occupation, but not income. Enrolment data is typically collected in the first year of primary or secondary school so for many students will be out of date. Furthermore, there is a considerable amount of missing data, the accuracy of the data is not well-established and the measurement of parental occupation is by four or five heterogeneous occupational groupings. For the downward adjustments to the *per capita* funds to non-government schools the SES measure is even worse as it is based on geocoding students' addresses to small areas of around 400 people (Statistical Area 1) from the last census.

A detailed detailed analysis prepared by the Catholic Education Commission Victoria (2017, p. 20) identify four serious problems with using census data to estimate capacity to contribute. These are:

- Area data misclassifies individuals and families
- The lack of observations greatly limits the reliability and usefulness of income data from the census
- The infrequent collection means that census data quickly becomes out-of-date
- The collection of income data in household units can be misleading

Given that billions of dollars of public money are about to be spent, surely the measures on which funds are allocated need to be very much better than area-based measures of SES.

2.9 Gonski is not sector blind

Contrary to the repeated claim that the Gonski funding model is sector blind given the fact that government and non-government schools are treated differently when calculating the amount of the SRS received by each school this is clearly not true. Whereas the level of funding received by Catholic and independent schools is adjusted according to 'parents' capacity to pay' no such impost is placed on wealthy parents sending their children to well-resourced and privileged government schools.

Approximately 36% of students across Australia attend non-government schools (ABS, 2016). It is a clear example of discrimination to financially penalise parents simply because of the choice they make as to where their children go to school. It is not fair that families earning over $200,000 annually who send to their children to a high-SES local school or a selective school are not considered to have a capacity to pay, but a family with an annual income of 80,000 who send their child to a non-government school (with a lower SES score) are expected to contribute.

The Catholic Education Commission Victoria (2017, p. 3) report argues that the Gonski SES model is not truly needs based and that it financially discriminates against low-fee paying schools serving disadvantaged communities. The report concludes:

- the limitations make SES scores biased in favour of high-income and affluent families. SES scores underestimate the financial means of these families. Conversely, they overestimate the financial means of lower and middle income families. The practical impact on non-government schools is that independent schools (especially high-fee independent schools) appear to benefit from SES scores while Catholic schools (and probably low-fee independent schools also) appear to be disadvantaged.

2.10 SES Effects are small, when considering prior achievement

The influence of SES on student performance is very small when considering prior achievement. In the presence of prior achievement, the effects of students' SES are quite small. For the US, Benner, Boyle and Sadler (2016, p. 1059) reported standardized effects on students' GPA of 0.44 for achievement score compared to 0.09 for family SES. For Germany, Baumert et al. (2010, pp. 159-160) report no significant effects for the International Socio-Economic Index (a measure of occupational status) on mathematics score and only one significant (but trivial) effect for parental education, net of prior achievement in mathematics (from PISA) and cognitive ability.

Using NAPLAN data, two Australian studies found standardized effects, which can be understood as correlations, for measures of socioeconomic background ranged from 0.05 to 0.15 when controlling for prior achievement (Lu & Rickard, 2014, pp. 31-32; Marks, 2014a, p. 241). A recent publication on the education production function in NSW secondary school students conclude: "socio-economic background of students attending a school has no significant effect on their academic performance, whereas higher prior academic achievements have a positive and statistically significant impact on student achievement" (Haug & Blackburn, 2017).

The importance of prior achievement and the weak effects of SES has important policy implications. It means that during the school career SES has little effect on student achievement. This is contrary to a key assumption of the Gonski model. To largely fund schools according to SES, when it has little or no impact on student performance during their school career, will therefore have little or no discernible impact on student performance.

2.11 SES effects largely reflect parental ability

Parents' ability is correlated with the most commonly used SES indicators. Scarr and Weinberg (1978, p. 678) reported correlations of 0.56, 0.37 and 0.38, for father's IQ with father's education, father's occupational status and family income. The correlations for mother's IQ with mother's education (0.46) and family income (0.19) were lower. Torres (2013, p. 166) reported a correlation of 0.53 between mother's AFQT score and a composite measure of family SES. According to Strenze's (2007, p. 411) meta-analysis, ability measured between ages 3 and 23 correlates at 0.56 for educational attainment, 0.45 for occupational status and 0.23 for income during adulthood. Focusing on college education in the US, failing to account for the mother's ability seriously overestimates the relationship between parents' economic resources and children's postsecondary attainments (Doren & Grodsky, 2016).

There are no Australia studies I known of that have examined the effects of parental ability on student achievement.

The point here is that the SES-education relationships are less about economic resources, parenting, socialization etc., but much more about the transmission of cognitive ability (and non-cognitive attributes) from parents to their children. This explanation is consistent with large heritabilities for student achievement (see section 1.5), educational attainment (see section 1.6) and cognitive ability (see section 3.3) and the strong relationships between cognitive ability and educational outcomes (see section 3.2).

3 Cognitive Ability

Cognitive ability or intelligence is defined as "a very general mental capability that, among other things, involves ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (Gottfredson, 1997a, p. 13). Similarly, Neisser et al. (1996, p. 77) define intelligence as the "ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought".

3.1 Achievement tests are very like ability tests

General tests of literacy, numeracy and problem solving, and tests of cognitive ability are conceptually similar. Rindermann (2008, p. 128) maintains there is no important theoretical difference between student achievement and ability tests since they both assess "thinking and knowledge". In PISA, literacy is defined generally as "concerned with the capacity of students to apply knowledge and skills in key subject areas and to analyse, reason and communicate effectively as they pose, solve and interpret problems in a variety of situations" (OECD, 2007b, p. 16). 'Adult Literacy' closely relates to general cognitive ability since "both assess skills that appear to represent verbal comprehension and reasoning, or the ability to understand, analyse, interpret, and evaluate written information and apply fundamental principles and concepts"(Baldwin cited by Gottfredson, 1997b, p. 109). These definitions closely resemble operative definitions of intelligence. Baumert (2009, pp. 3-5) points out that like intelligence tests, reading and mathematical assessments involve reasoning and making logical inferences. Jencks et al. (1979; 1972) had no hesitation in focusing on test scores in chapters entitled 'Inequality in Cognitive Skills' and 'The Effects of Academic Ability'. Scores in the U. S. adult literacy test are referred to as 'cognitive skill' by Kerckhoff, Raudenbush and Glennie (2001). According to the OECD, the 2012 adult literacy study collects an unprecedented amount of information on 'basic cognitive skills' (Thorn, 2009, p. 19). Armor (2003, p. 19) notes the similarities between achievement tests and IQ tests. Both include subset scores for distinct types of mental skills: vocabulary, reading comprehension, mathematical concepts, numerical skills etc. Armor (2003, p. 21) suggests that the substantial overlap between IQ and achievement scores indicates they are measuring something in common, general reasoning skills. The main difference is that achievement tests specific subject knowledge (Baumert et al., 2009).

The logic of modern test theory used in PISA, NAPLAN, PIRLS and TIMSS is that the probability of a student correctly answering a test item is a function of student ability and the difficulty of the item. It is surprising that the concept of "ability" is almost never mentioned in Australian reports or academic articles based on PISA, TIMSS or NAPLAN data.

3.2 Ability has a much stronger relationship with education than SES

Cognitive ability and education are closely linked (Mayer, 2000). Cognitive ability has stronger relationships with student achievement than students' SES. Walberg (1984, p. 23) computed an average correlation of 0.71 between various IQ measures and academic achievement. Duckworth, Quinn and Tsukayama (2012, p. 443) reported correlations between 0.7 and 0.8 for IQ measured in grade 4, and grade 5 and 9 achievement tests. For New Zealand, the correlation between IQ at measured at ages 8 and 9 with academic performance at age 13 was 0.83 (Fergusson, Horwood, & Boden, 2008, p. 285). Herrnstein and Murray (1994, p. 584) show sizable correlations of over 0.7 and up to 0.9 between the ability measure in the AFQT Armed Force Qualification Test (AFQT, used in the NLSY) and standard ability tests with a median correlation of 0.8. Because of the high correlation, AFQT scores adjusted for age are used as a measure of ability (see Zagorsky, 2007, p. 491). Kaufman, Reynolds, Liu, Kaufman and McGrew (2012) calculated a mean correlation of 0.8 between latent factors of cognitive ability and student achievement. The correlations increased with student's age which may reflect the increasing heritability of intelligence. The US Scholastic Assessment Test (SAT) and the American College Readiness Assessment (ACT) are highly correlated with ability (Coyle, 2015, p. 18; Frey & Detterman, 2004; Koenig, Frey, & Detterman, 2008, p. 156).

In a study of 178,599 pupils attending English state schools the correlation between general factors derived from a cognitive ability test and attainment scores on national Key Stage 2 tests in English, mathematics and science of 11 year olds was 0.83 (Calvin et al., 2010). In a study of over 80,000 16 year-old students, Deary, Strand, Smith and Fernandes (2007) calculated a correlation of 0.81 between a latent intelligence trait measured at 11 years of age with a latent trait of subject performance in the GCSE. For the Netherlands, the correlations between IQ and CITO achievement at increase with age: 0.41, 0.50, 0.60, and 0.63, at ages 5, 7, 10, and 12 respectively (Bartels, Rietveld, Van Baal, & Boomsma, 2002). For Slovenia, the correlation between intelligence and grades in the last four years of the 9 years of primary school was 0.48 (Flerea, Krajnca, Klanjšeka, Musila, & Kirbiša, 2010, p. 54). Roth et al.'s (2015) cross-national meta-analysis of over 100,000 students calculated a correlation of 0.54 between intelligence and student performance. Again, the correlations increased with level of schooling, 0.45, 0.54 and 0.58 for elementary, middle and high school students, respectively (Roth et al., 2015, p. 123).

Detterman (2016, p. 1) in a review article about the relatively small influence of schools and teachers on students' educational outcomes writes:

> I further argue that the majority of the variance in educational outcomes is associated with students, probably as much as 90% in developed economies. A substantial portion of this 90%, somewhere between 50% and 80% is due to differences in general cognitive ability or intelligence. Most importantly, as long as educational research fails to focus on students' characteristics we will never understand education or be able to improve it.

He (2016, p. 6) concludes that "Human intelligence or general cognitive ability accounts for at least half and probably more of academic achievement attributable to student characteristics."

The importance of student ability for student performance has been known since at least the 1960s. It is difficult to believe that researchers and policymakers (especially members of the Gonski committee) have not heard of Jensen's (1969) controversial paper "How Much Can We Boost IQ and Scholastic Achievement?". Instead, Gonski et al. have decided to ignore the voluminous literature on ability and student performance and carry on regardless.

The importance of ability has implications for educational practice. Brunner (2008, pp. 160-161) (2008, pp. 160-161) advocates investing in educational programs that foster reasoning general abilities. Adey, Csapo, Demetriou, Hautamäki and Shayer (2007, pp. 92-94) list proposals for raising general ability including learning that stimulates cognitive functioning and encouraging students to make connections between concepts in different domains. Asbury and Plomin (2014) advocate personalized learning schemes and a stronger focus on early childhood development.

3.3 Student Ability also has a large genetic component

Like student achievement and educational attainment (see sections 1.5 and 1.6), cognitive ability has a large genetic component.

A variety of studies generate heritabilities for cognitive ability of between 50 to 80% with a much smaller proportion of the variance, typically less than 20%, attributed to the shared environment (Deary, Johnson, & Houlihan, 2009; Nielsen, 2006; Plomin, Fulker, Corley, & DeFries, 1997; Plug & Vijverberg, 2003; Rowe, Vesterdal, & Rodgers, 1999; van Leeuwen, van den Berg, & Boomsma, 2008). Even critics of *Bell Curve* estimate inheritabilities for IQ of around 50% and a much smaller effect for the (non-womb) environment at around 17% (Daniels, Devlin, & Roeder, 1997, pp. 54-58). Nielsen (2006) estimates the variance in IQ attributed to genetic, environmental and unique factors at 53%, 14%, and 33%. Plomin et al., (1997, p. 445) conclude that environmental transmission of cognitive ability from parent to child is negligible. Similarly, Nielsen (2006) concludes that the impact of shared environmental factors is relatively small although there a sizable unshared environmental component (that is, unique to the individual).

In early adulthood, the heritability of cognitive ability is even higher around 0.8 (Bouchard, 2013; Deary et al., 2009, p. 217).

3.4 Cognitive ability is not a function of SES or social background

Gottfredson (2016) notes that the argument that family environments determine children's IQs have been disconfirmed because of sizable IQ differences between siblings, the high heritability of IQ and non-cognitive behavioral traits, and identification of genes that contribute to IQ. Other evidence comes from raised-apart identical twins whose IQ's correlate at around 0.7 (Segal, 2012, p. 107). These are identical twins brought up by different families who completed IQ tests as adults. The correlation is only slightly lower than that for twins brought up together. There are also prominent books on the inability of socialization and parenting to substantially influence children's cognitive ability (Harris, 2009; Pinker, 2011; Rowe, 1994).

3.5 Cognitive ability is not a function of schools

Deterrman (2016) points to an interesting study on schools in Poland. I quote:

> During World War II, the city of Warsaw, Poland was completely destroyed. After the war, Warsaw came under the control of a communist government which decided it would assign residents to the reconstructed city randomly to avoid social segregation. The city government felt that this would eliminate differences in cognitive development due to social segregation. Firkowska et al. (1978) studied the effects of this social experiment. Though the distribution of people in the city was not completely random it was very close to that. They obtained Raven's Matrices tests for a large portion of the students born in 1963 for much of the city. In addition, they also collected parents' education and occupation which was used to form a 13-point index of social class. The expectation, of course, was that the correlation between IQ and the social class index of the child's home would be 0.0. Instead, r2 = 0.97, almost perfect. More interesting, the differences between schools was reduced from 10% to 2.1%. In other words, student variance accounted for nearly 98% of the outcome. Since Raven's scores are generally predictive of academic achievement, it can be assumed that a similar finding would apply to academic achievement tests. But what it certainly shows is that a large part of even school effects can be accounted for by the non-random distribution of students across school districts. It is rather counter-intuitive that a more equitable and equal geographic distribution of people across school districts would make differences between students even more apparent.

Firkowska et al. (1978) conclude "that an egalitarian social policy executed over a generation failed to override the association of social and family factors with cognitive development that is characteristic of more traditional industrial societies". They note (1978) "for further elucidation of the determinants of cognitive abilities, we need to turn our attention to intrinsic factors." These intrinsic factors have a large genetic component

4 Schools

A causal role for schools in determining educational outcomes is a fundamental aspect of the logic behind the Gonski funding model. The argument is that high SES families send their children to 'good' schools (including private schools) characterised by good facilities, excellent teachers and enthusiastic students. In contrast, low SES families are limited to 'bad' schools, characterized by poor resources, low morale staff and difficult students.

The assumptions for the role of schools are that:

1. Between-school differences in educational outcomes in Australia are large
2. School resources are in important influence on student outcomes
3. Between-school differences in educational outcomes are mainly a function of SES

These assumptions are not supported empirically.

4.1 Between-school differences are small

One of the most startling findings of the Coleman report (1966, p. 297) *Equality of Educational Opportunity* was that in the U. S. the variation in student performance is much greater *within schools* (about 80%) than between schools. The report concludes that "that schools bring little influence to bear on a child's achievement that is independent of his background and general social context" (1966, p. 325). This undermines the argument that schools and differences in school resources are responsible for socioeconomic and other social background differences in student outcomes.

In other Anglo-Saxon countries, also with largely comprehensive school systems—Australia, Canada, New Zealand and the United Kingdom—the between school variation in PISA test score (as a percentage of the total variation) is between 16 and 22% (Marks, 2006, p. 36).

Where there are larger differences between schools in average student performance and other student outcomes, much of difference can be attributed to differences in the intake characteristics of students rather than what schools do. When considering individual students' socioeconomic background (SES) and other aspects of their social backgrounds, schools contribute less than 10 per of the variation in student performance in the U. S. (Coleman, Hoffer, & Kilgore, 1982, p. xxvi). Hauser (1969) concluded that school differences in educational outcomes are small and may be largely an artefact of compositional differences.

In Australia, the intra-class correlations for schools—that is the proportion of between school variation as a proportion of the total variation—in the Australian context are not high: between 0.1 and 0.2 for Victorian primary school students, 0.25 for tertiary entrance and just over 0.3 for PISA (Hill & Rowe, 1996; Marks, 2010c, 2014d; OECD, 2013, p. 196). That means that at least 70% of the variation of student performance is within, not between, schools.

So, there are not strong school differences in Australia even without considering prior achievement. Again, the evidence from literature is clear. But despite this evidence the Gonski model assumes that between-school differences are large.

4.2 School Effects are even smaller when considering stable student attributes or prior achievement

Scheerens and Bosker's (1997a) meta-analysis settled on a figure of around 8% for the between-school variance in student outcomes, net of student characteristics. Jencks et al. (1972, p. 93) concluded that "differences between high schools contribute almost nothing to the overall level of cognitive inequality" that is, inequality in test scores. They point out that if all schools were equally effective in raising student achievement during the final years of secondary school, inequality among twelfth graders would fall by less than one per cent (1972, p. 90).

Cullen, Jacob and Levitt (2006) compared the outcomes of a random lottery of students in Chicago. Lottery winners attended schools that were "better in a number of dimensions, including peer achievement and attainment levels". Lottery losers went to their local schools. However, they found "little evidence that winning a lottery provides any systematic benefit across a wide variety of traditional academic measures".

For Australia, between-school differences decline substantially when controlling for prior achievement (Hill & Rowe, 1996, p. 21; Marks, 2010c). Analysing NAPLAN data for both primary and secondary students attending Victorian government school, Marks (2014d) found that, net of prior achievement, schools effects were very small, unstable over time and inconsistent across NAPLAN achievement domains.

The much smaller between-school variance compared to the much larger within-school variance has been well-established for many decades. Despite this, researchers and policymakers insist that schools are the major source of variation. This inability to understand the empirical reality that low achievers are not confined to particular schools, they are found in almost every school. Therefore, the Gonski model would ignore low achievers not attending those schools defined as "disadvantaged" schools.

Recently, the AEU union run a TV advertisement about a concerned mother whose daughter was performing poorly at school. Apparently, her poor performance was remedied by a program funded by the Gonski model. But if the daughter attended one of the schools not deemed as "disadvantaged" then most likely less or no funding would be available and presumably the daughter would remain a low achiever.

4.3 School-SES effects are a furphy

The correlations of school-SES and school-achievement at the aggregate level are spurious as it represents what is known as an ecological fallacy which dates to 1950. Australian studies that emphasize the effects of school-SES, especially regarding school sector differences, do not take account of students' prior achievement (for example Perry & McConney, 2010a; Perry & McConney, 2010b; Thomson, De Bortoli, & Buckley, 2013, pp. 34-35, 144, 183; Thomson, De Bortoli, Nicholas, Hillman, & Buckley, 2010, pp. 63, 188, 232).

Theoretically, the effects of school-SES rely on some kind incredible (in the true sense of the word) generalised school effect where the SES of students spreads through the student body, enhancing or reducing student performance.

There are serious statistical problems with school-SES effects. The ecological fallacy is well known: statistical relationships at the aggregate-level cannot be extended to the individual-level (Robinson, 1950). Measures of both SES and achievement aggregated at the school-level exaggerate the SES- achievement relationship by a factor of 4 (White et al., 1993). According to Hauser (1970; 1974, p. 659), contextual effects of SES relate to the ecological fallacy in that residual differences between schools are interpreted as social processes. These differences should disappear when the analysis includes appropriate individual student-level predictors (correlated with school residuals). Nash (2003, p. 446) suggests the contextual effects of school-SES are due to unmeasured factors correlated with school-SES that affect school performance. Measures of SES aggregated at the school-level substantially exaggerate the SES-achievement relationship (White et al., 1993). These authors point out that the temptation of researchers to commit the ecological fallacy is very serious and note that "massive reorganization of students based on SES would do very little to equalize achievement levels among schools, because SES of the individual student plays too small a role in achievement" (White et al., 1993, pp. 337, 342).

Recent research has characterized school-contextual effects as 'phantom' effects because the poorer the measure of SES, the stronger the effects of school-SES (Pokropek, 2014; Televantou et al., 2015). Marks (2015a) added random error to a composite measure of SES which increased the magnitude of school-SES effects on student achievement. Most studies that emphasize school-SES as an important influence do not control for students' prior achievement, school-level prior achievement or existing differences between students. When taking into account prior achievement or unobserved differences between students (as in fixed-effects models) school-SES effects are small, often negligible (Lu & Rickard, 2014, pp. 31-32; Marks, 2015a).

Despite all the issues surrounding school-SES, it school-SES effects and advocates of SES based funding models trumpet school-SES effects as if they were unproblematic. Such advocates never consider prior achievement or that fact that school-SES effects are stronger with poorer measures of SES. This is another case of political ideology trumping proper and thoughtful social science.

4.4 Schools resources are only weakly associated with student achievement

After reviewing some 400 student achievement studies Hanushek (1997) concludes "400 studies of student achievement demonstrate that there is not a strong or consistent relationship between student performance and school resources,". Later studies come to the same conclusion (Hanushek, 2006; Hanushek & Woessmann, 2017). The most recent study concludes "Overall, the international evidence provides little confidence that quantitative measures of expenditure and class size are a major driver of student achievement, across and within countries. "The study by Hanushek and Greenwald, Hedges and Laine (1996) which is sometimes used to dismiss Hanushek's work is very carefully worded: "effect sizes large enough to suggest that moderate increases in spending may be associated with significant increases in achievement". Note the use of the words *may* and *significant*.

Although there are few Australian studies on the effects of school resources, the evidence is much the same as from US studies. The effects of school resources on student achievement in Australia are, at best, modest (Cobb-Clark & Jha, 2016; Justman & Ryan, 2013).

In response to the argument that additional funding will lead to improved results the OECD's research based on the PISA tests concludes: "Greater national wealth or higher expenditure does not guarantee better student performance(OECD, 2012, p. 1). The report goes on to conclude "The countries that are the strongest performers in PISA are not the wealthiest, nor do they allocate more money to education". A more recent report repeats the fact that increased investment is not directly related to improved performance when it concludes: "PISA has consistently found that the amount of resources spent on education – including financial, human and material resources – is only weakly related to student performance"(OECD, 2014, p. 1).

The OECD's research also concludes that once a minimal level of investment is reached it is wrong to assume that spending more will lead to improved performance. Once a satisfactory level of investment is reached the OECD argues: "additional or better-quality resources appear to have little additional impact on the incidence of low performance" (OECD, 2016c, p. 175). Significant, in relation to Australia, is the fact that the OECD's country note about Australia argues "there is no (other) OECD country where large proportions of low performing students attended schools with better educational resources (OECD, 2016a)5.

4.5 Real expenditure on education is increasing

According to the CIS, "Australian governments, both federal and state, increased real per student recurrent funding by a total of 15.4% between 2005–06 and 2014–15. This represents an average annual increase of 1.7% across 9 years (Joseph, 2017, p. 3).

School funding has increased substantially in real terms (Department of Education and Training, 2017):

School funding in Australia has been growing in real terms for several decades:

- Commonwealth and state/territory spending between 1987–88 to 2011–12 has increased by 100 per cent in real terms while student numbers grew by only 18 per cent
- between 2003–4 and 2013–14, total Commonwealth spending on schools increased by 49.2 per cent in real terms.

So, increased funding has done nothing to lift standards in Australian students' performance. Such evidence is well-known among politicians, bureaucrats and policymakers. The question must be asked why are so many stakeholders advocating increased funding knowing that it is unlikely to lift performance?

4.6 School Effects are small

Although the Gonski model has enormous faith in the ability of schools to substantially improve student outcomes and Gonski 2.0 assumes that it will uncover transformative school and teacher curriculum, pedagogy and institutional practices the evidence suggest that policymakers need to be much more cautious.

Because of the high stability of student achievement over the school career, school and teacher effects must be quite modest. In a meta-analysis of school and teacher effects, Scheerens and Bosker's (1997a, p. 79) meta-analysis offer a conservative estimate effect size of around 0.20 for schools compared to an effect size of 0.30 without adjustments. The estimate of 0.2 is unlikely to be based on many studies that include prior achievement. These effect sizes are the expected change in student achievement measured by standardized scores for a one standard deviation increase in school or teacher effects. Effect sizes below 0.2 are classified as small according to Cohen' (1988) well-known classification. Similarly, Scheerens (2016, pp. 202-203) found the average effect size for schooling variables was 0.08 ranging from 0.02 to 0.15. A preceding meta-analysis concluded that most school effectiveness indicators (e.g. orderly climate, achievement orientation, homework) should be interpreted as negligible to small (Scheerens, Witziers, & Steen, 2013).

These small school and teacher effects are in marked contrast with Hattie's (2009, p. 18) meta-analysis that found sizable effects for teachers (d=0.49), curricular (0.45) and teaching (0.42) and weaker effects for the school (d=0.31) and home (d=0.23). Although these effects appear large, Hattie (2009, p. 18) points out that these effects should be considered relative to the average effect size of 0.40. Furthermore, Hattie's research is based on bivariate relationships; the effect sizes would be much smaller if the only studies selected were those that controlled for prior achievement or student ability.

The reason for inconclusive school and teacher effects is that student performance is very stable over the school career.

4.7 The School Resource Standard is arbitrary

The original Gonski (2011) model of school funding is based on the Schooling Resource Standard (SRS) which is argued as the base funding a school needs for its students to perform at an acceptable level. The base (general recurrent funding) is supplemented by loadings reflecting the proportions of student deemed disadvantaged. For non-government schools only, the base amount of funding is reduced because of parents' capacity to pay. This reduction is to be initially based (until a suitable replacement is found) on the SES score of the school.

The SRS is essentially arbitrary as there is no convincing rationale justifying what the SRS should be. As noted in Volume 1 of the Appendix to the Commonwealth National Commission of Audit: "The base SRS is arbitrarily high and not based on a detailed analysis of the cost of delivering education. Instead, it was derived from a model that included only a small proportion of high performing schools' current expenditure. It is not clear that this represents the efficient price of delivering education" (Commonwealth of Australia, 2014, p. 266). Doubts about the validity and reliability of the SRS are also expressed by Justman and Ryan (2013) when they describe it as "essentially arbitrary", having a "veneer of technical sophistication" and lacking a "sound methodological basis' (2013, p. 6).

5 Teachers

Teachers may be the key to improve student outcomes. Although the teachers are important to what students learn and there is a common mantra about improving teacher quality, there is little Australian research on teacher effects. The Gonski report has little to say about teacher effects since the emphasis is on SES and school resources.

5.1 Teachers effects appear sizable

For the US, Hanushek (2011) estimated that teachers one standard deviation above the mean effectiveness for teachers annually generates marginal gains of over $400,000 in present value of student future earnings with a class size of 20 and proportionately higher with larger class sizes. Alternatively, replacing the bottom 5–8 percent of teachers with average teachers could move the U.S. near the top of international math and science rankings with a present value of $100 trillion. Hanushek (2009) uses estimates of teacher effects to claims that dismissing 6–10 percent of the worst teachers could have a dramatic impact on student achievement even if these were replaced (permanently) with just average teachers.

Teacher effects studies typically, estimate the distribution of teacher effects, net of prior achievement, or stable differences between students, and report the difference in student achievement for a one standard deviation difference in teacher effects.

A US study of value-added teacher effects concluded that a one standard deviation increase in the distribution of teacher effectiveness (for example from the median to the 84th percentile) translates to an increase of about 0.15 to 0.25 standard deviations of student achievement. These effects can be interpreted as the change in student achievement in standard deviation units for one standard deviation increase in teacher performance, equivalent to an effect size. Similarly, Hanushek and Rivikin (2010) report average value-added teacher effects from several US studies of 0.11 for reading (range from 0.08 to 0.26) and 0.15 for math (0.11 to 0.36).

In one of the few Australian studies on teacher effectiveness, Hill and Rowe (1996) found that that variance at the class/teacher level constitutes 40 to 55 percent of measured variance, while school-level variance constitutes just 4 to 8 per cent of the total variance. Later studies on Victorian Year 12 students estimated that 60 per cent of the variance in student achievement can be attributed to class room and teacher effects and only 5 per cent at the school level (cited by Leigh, 2010).

5.2 Value-added teacher effects are modest

Although recent policy debates on improving educational outcomes have focused on teachers, teacher effects estimates from value-added models that control for students' prior achievement are quite small undermining arguments for new policy initiatives involving teachers.

Like school effects, value-added teacher effects are not large. Summarizing US studies, Leigh (2005) notes that moving the average teacher to the 90th percentile increases student achievement by on average of 0.12 of a standard deviation. His study on Queensland teacher effects estimated that a standard deviation difference in teacher effectiveness translates to a 0.10 standard deviation in test score (Leigh, 2010). He concludes that a student with a highly effective teacher (as measured by a value-added metric) could achieve in three-quarters of a year what a student with a less effective teacher could in a full year. Ryan (2017) suggests that the consensus estimate for teacher effects is 0.10. An Australian twin study concludes the ceiling for students assigned to different classrooms (understand as teacher effects) is 0.08 (Byrne et al., 2010). As was the case for school effects, teacher effects of around 0.10 mean that the bulk of students are unaffected by the teacher assigned to them but a minority of teachers have a substantial influence on student achievement.

5.3 Issues with teacher Effects

5.3.1 Teacher effects are not stable

One concern for policymakers who wish to tie teacher performance to student test scores is the stability of teacher effect estimates over time. Like school-effects, teacher effects are not particularly stable. Models may generate unbiased estimates of teacher effectiveness but still be unstable from one year to the next. Using a large dataset of elementary and middle school math tests in Florida, McCaffrey et al. (2009) estimate several value-added models and find year-to-year correlations generally ranging from 0.2 to 0.5 in elementary schools and 0.3 to 0.6 in middle schools. Comparing teacher rankings, they find that about a third of teachers ranked in the top quintile are again in the top quintile the next year. Goldhaber and Hansen (2013) perform similar analyses using state-wide elementary school data from North Carolina over a ten-year period. They do conclude that there are stable teacher effects but they are small.

5.3.2 Fade out of teacher effects

Teacher effects tend to fade out (Raudenbush, 2014)

5.3.3 Weak correspondence between teacher effects and principal evaluations

There are only weak relationships between teacher effects and evaluations of teachers by principals (Jacob & Lefgren, 2007).

5.4 Characteristics and practices of effective teachers are difficult to isolate

Although, a consensus estimate of teacher effects is around 0.10, measurable teacher characteristics do not have consummate effects on student achievement.

Scheerens (2016, pp. 202-203) meta-analysis found an average effect size for teaching variables of only 0.10 ranging from -0.08 to 0.21.

5.4.1 Teacher qualifications have little impact

Hanushek and Rivkin (2004) summarize the research on the predictive power of master's degree completion and find little consistent evidence that graduate degree attainment can identify effective teachers. Clotfelter et al.(2007) Analysing a system wide from North Carolina found that a graduate degree exerts no statistically significant effect on student achievement and in some cases the coefficient is negative.

5.4.2 Weak effects of teacher test scores

In a review of studies of teacher effects, Darling-Hammond (2000, p. 3) notes that the relationships between academic ability and teacher effectiveness are most often small and are often not statistically significant. Where effects of teacher test scores are found to be (moderately) important, they are test scores based on knowledge and skills acquired through teacher educations courses, certification tests (e.g.Clotfelter et al., 2007) not academic performance for college entry.

Contrary the argument that raising the ATAR score required for entry to teacher education courses will strengthen teaching practice and raise standards it is also important to take note of what happens in Finland. As noted by the Finnish education expert Parsi Sahlberg (2015): "A quarter of the accepted students came from the top 20% in academic ability and another quarter came from the bottom half. This means that half of the first-year students came from the 51- to 80-point range of measured academic ability. You could call them academically average. The idea that Finland recruits the academically "best and brightest" to become teachers is a myth. In fact, the student cohort represents a diverse range of academic success, and deliberately so".

5.4.3 Professional development is probably ineffective

Although great faith in put into improving student outcomes through professional development of teachers, Harris and Saris (2011) found negative effects for professional development on student achievement.

5.4.4 But teacher experience is probably important

As is true for almost all occupations, people perform better with experience. Greater experience is associated with stronger teacher effects. Most of the gains occur within the first few years (Clotfelter et al., 2007; Hanushek & Rivikin, 2004). The OECD (2012) argues that beginning teachers require more support which would reduce attrition as well as improve their performance.

6 Concluding remarks

6.1 The dominant SES narrative

The dominant SES narrative can be summarized as follows:

- SES is the dominant influence on student outcomes.

- Prior achievement and cognitive ability are not particularly relevant because they are simply functions of SES.
- School-SES demonstrates the importance of SES to educational outcomes.
- School differences are large and reflect school differences in SES scores of students.
- Differences in average performance between school-sectors are not because of differences in teaching or the delivery of the curriculum but simply reflect SES or the miasmic process of school-SES.
- There is no need to improve teaching or any other aspect of schooling because the problem is simply one of school resources, simply solved by directing more resources to government schools.

The ACER, the Grattan institute and the Mitchell institute also adopt much of the SES narrative (Goss, Chisholm, Sonnemann, & Nelson, 2016; Lamb et al., 2015; Thomson, Bortoli, & Underwood, 2017).[6] These reports appear unaware of much of the literature cited in this report.

The dominant SES narrative is accepted *holus bolus* by large sections of the media (the ABC, the Fairfax press, online outlets such as, the Conversation) and many in the commentariat. From the evidence presented here, it is unambiguously clear the dominant SES narrative is very wrong.

Much of research and policy debate on Australian education does not consider evidence that runs counter to the dominant SES narrative as detailed extensively in this report. The SES narrative quickly falls apart when subjected to, even cursory, empirical scrutiny.

Support for the argument that educational outcomes in Australia is mostly about SES is not because of the empirical evidence or the veracity of the theoretical explanations but because it fits neatly into an existing set of beliefs: that Australian society is characterized by large socioeconomic inequalities and education is the primary mechanism for the strong and persistent reproduction of socioeconomic inequalities across generations and lack of social mobility. This is not true.

Even though policymakers and senior bureaucrats regularly mouth the mantra of evidence-based policymaking, the reality is that evidence has very little bearing on educational policymaking in this country. Educational policy is developed by political parties based on political ideology and political strategies; the so-called “evidence” is cherry-picked to support the chosen policy. Any evidence that cannot be construed to support the chosen policy is ignored or dismissed. The Gonski model and the debate surrounding the model is a good example of this.[7]

6.2 How much Gonski money will be spent on administration

When the NDIS was announced the Prime Minister at the time said that 70% will be spent on administration. My concern is that much of the Gonski money will be directed to administration. Parliamentary questions should be asked on:

- What percentage of the funding will be spent on administration.
- Are there clear distinctions between the roles of federal, state and other bureaucracies, so costly duplication is avoided?
- Are there mechanisms proposed to reduce administrative costs and duplication?

6.3 Educational evaluation in Australia are far from "state-of-the-art"

The state of educational evaluation is very poor in Australia. Evaluation of educational programs in Australia often entails evaluating implementation rather than effects. Evaluation reports often include anecdotes from the administrators, participants and stakeholders who almost invariably provide positive responses. What is required is proper evaluation studies that control for prior achievement or stable differences between students and/or include control groups. If evaluations cannot be done properly, taxpayer funds should not be wasted on inappropriate and misleading evaluation studies. In the US, evaluations of educational programs are generally rigorous and statistically sound (Wright et al., 2010).

Does the Gonski funding model stipulate state-of-the-art evaluations on programs that the Gonski money has been used for?

6.4 Proposed programs and interventions are likely to be unpalatable

There is an extensive literature of school, teacher strategies and program effects on student achievement (Bosker & Scheerens, 1989; Hattie, 2009; Korpershoek, Harms, Boer, Kuijk, & Doolaard, 2016; Scheerens et al., 2013; Scheerens & Bosker, 1997b; Thapa, Cohen, Guffey, & Higgins-D'Alessandro, 2013). There is some, but limited, consensus regarding direct instruction or explicit teaching, school autonomy, ability grouping, phonics, and formative assessment and feedback (Buckingham, Wheldall, & Beaman-Wheldall, 2013; Cheng, Hitt, Kisida, & Mills, 2017; Fuchs & Wößmann, 2007; Hempenstall, 2004; Lavy, 2015; NSW

Department of Education, 2014; Steenbergen-Hu, Makel, & Olszewski-Kubilius, 2016; Yeh, 2009).

The problem here is even if the 2017 Gonski report establishes unambiguously a list of strategies, interventions and programs that raise student performance overall and improve the performance of students lagging, many would be unpalatable to large sections of the education industry generating further political battles. Similarly, some modern practices such as, allowing students to be "knowledge navigators" and "whole language" are unlikely to receive strong empirical support (Hattie, 2009; Lavy, 2015).

Furthermore, can the federal government really expect that bureaucrats, sitting in air-conditioned offices in Canberra or in state and territory capital cities, determine what teachers do in front of classes? Even if the ways teachers teach can be monitored effectively and remotely, is this desirable or cost effective? Such an approach naively assumes that effective practices can be identified unambiguously and that they apply to all students and schools. It would be far better to allow schools and subject departments within schools, autonomy to develop programs tailored to their students.

6.5 Focus should be on lifting performance of all students

If governments really want to improve educational outcomes and ensure that as many as possible students achieve acceptable standards there are better and simpler funding models than that advocated in the 2011 Gonski report and embodied in Gonski 2.0. If the aim is to improve overall levels then the curriculum needs to move up at notch or two so that students are introduced to more difficult material earlier. This largely explains why some Asian countries perform better than Australia in international cognitive tests such as PISA and TIMMS, despite larger class sizes, lower staff salaries and less per capita resources. As argued in a recent review of the Australian Curriculum it is also vital to ensure that state and territory intended curriculum documents and classroom pedagogy reflect best practice and evidence-based research (Australian Government, 2014).

6.6 Policy should directly target low achievers

To tackle poor performance, the funding model should directly focus on student performance providing schools (and if this fails their parents) with funds so that students falling behind are provided proven educational programs that allow them to catch up to their peers. NAPLAN provides accurate identification of low achievers since well-over 90% of student cohorts participate in NAPLAN. This is contrasts with the Gonski model that identifies low achievers by the poorly measured SES score of their school. The Gonski report dismisses directly identifying low achievers:

> However, there are risks in funding schools on a direct measure of student educational achievement as it can introduce perverse incentives if schools improve their reported overall educational performance. Instead, the panel believes that it is preferable to base funding on more objective characteristics of students and schools that are known to have a significant association with low educational achievement (Gonski et al., 2011, p. 155).

As documented in this report the "significant association with low educational achievement" is weak considering the strong effects of prior achievement. And the perverse incentives argument is weak. Would school principals deliberately aim to reduce to their students' NAPLAN results? How would they do this? Would the various principal associations regard such behavior as ethical? Would the school board, parents and teachers at the school endorse such behavior by the principal? The possibility of perverse incentives can easily be rectified by rewarding schools that successful lift the performance of poor achievers evidenced by students' NAPLAN performance two years later.

It is important to note that the Victorian Department of Education and Training (2015, p. 17) proposed a scheme to directly target low achievers "regardless of their socioeconomic circumstances":

> The second measure would target additional funding for students entering secondary school who struggled with basic skills in primary school.
>
> These approaches would redefine equity in our schools, recognising that as much as we need to target additional resources to students from disadvantaged backgrounds, educational disadvantage is not only defined by students' background – students who are falling behind need support regardless of their socioeconomic circumstances.

6.7 Appropriate data is not available

Without data, proper evaluation studies are not possible. The proposal for unique transferrable student identification numbers was first raised 20 years ago, if not earlier. NAPLAN which is administered to nearly all students is nearly 10 years old. The most recent recommendation for a national student identification number was made by the Grattan institute (Goss et al., 2016, p. 3). However, national student identification numbers do not exist. ACARA is not able to provide researchers with suitable long-term student-level longitudinal data across jurisdictions. When I inquired about data availability, I was asked to contact each jurisdiction independently and there are many. Surely it is not beyond the ability and resources of Australia's large educational bureaucracies to ensure that the data is in an acceptable state. Again, I make the point that without such data proper evaluation of educational programs and monies spent is not possible.[8]

The provision of Gonski money must require jurisdictions to ensure that the data is transparent and can be linked at the individual student-level and to schools over the entire school-career (including Year 12) and across jurisdictions and be freely available for research.[9]

References

ABS. (2016). *Schools Australia 2015*. Cat No. 4221.0: Australian Bureau of Statistics.

Adey, P., Csapo, B., Demetriou, A., Hautamäki, J., & Shayer, M. (2007). Can we be intelligent about intelligence? Why education needs the concept of plastic general ability. *Educational Research Review, 2*(2), 75–97. doi: 10.1016/j.edurev.2007.05.001

Ainley, J., Graetz, B., Long, M., & Batten, M. (1995). *Socioeconomic status and school education*. Canberra: Department of Employment Education and Training.

Ainley, J., & Long, M. (1995a). Measuring student socioeconomic status. In J. Ainley, B. Graetz, M. Long, & M. Batten (Eds.), *Socioeconomic status and school education* (pp. 52-76). Canberra: Australian Government Publishing Service.

Ainley, J., & Long, M. (1995b). Other aspects of area-based indexes. In J. Ainley, B. Graetz, M. Long, & M. Batten (Eds.), *Socioeconomic status and school education* (pp. 77-90). Canberra: Australian Government Publishing Service.

Ainley, J., & Sheret, M. (1992). *Progress through high school: A study of senior secondary schooling in New South Wales* (No. 23). Hawthorn: ACER.

Armor, D. J. (2003). *Maximizing intelligence*. New Brunswick and London: Transaction Publishers.

Asbury, K., & Plomin, R. (2014). *G is for genes: The impact of genetics on education and achievement*. Chichester, West Sussex: Wiley-Blackwell.

Aughinbaugh, A., & Gittleman, M. (2003). Does money matter? A comparison of the effect of income on child development in the United States and Great Britain. *Journal of Human Resources, 38*(2), 416-440.

Australian Government (2014). *Review of the Australian curriculum. Final report*. Canberra: Australian Government Department of Education.

Baker, L. A., Treloar, S. A., Reynolds, C. A., Heath, A. C., & Martin, N. G. (1996). Genetics of educational attainment in Australian twins: Sex differences and secular changes. *Behavior Genetics, 26*(2), 89-102.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29*(1), 37-65. doi: 10.3102/10769986029001037

Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M., & Boomsma, D. I. (2002). Heritability of educational achievement in 12-year-olds and the overlap with cognitive ability. *Twin Research, 5*(6), 544-553.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133-180. doi: 10.3102/0002831209345157

Baumert, J., Lüdtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review, 4*(3), 165-176. doi: 10.1016/j.edurev.2009.04.002

Benner, A. D., Boyle, A. E., & Sadler, S. (2016). Parental involvement and adolescents' educational success: The roles of prior achievement and socioeconomic status. *Journal of Youth and Adolescence, 45*(6), 1053-1064. doi: 10.1007/s10964-016-0431-4

Blanden, J. (2013). Cross-country rankings in intergenerational mobility: A comparison of approaches from economics and sociology. *Journal of Economic Surveys, 27*(1), 38-73. doi: 10.1111/j.1467-6419.2011.00690.x

Bosker, R. J., & Scheerens, J. C. J. (1989). Issues in the interpretation of the results of school effectiveness research. *International Journal of Educational Research, 13*(4), 741-751.

Bouchard, T. J., Jr. (2013). The Wilson effect: The increase in heritability of IQ with age. *Twin Research and Human Genetics, 16*(5), 923-930.

Brunner, M. (2008). No *g* in education? *Learning and Individual Differences, 18*(2), 152-165. doi: 10.1016/j.lindif.2007.08.005

Buckingham, J., Wheldall, K., & Beaman-Wheldall, R. (2013). Why poor children are more likely to become poor readers: The school years. *Australian Journal of Education, 57*(3), 190-213. doi: 10.1177/0004944113495500

Byrne, B., Coventry, W. L., Olson, R. K., Samuelsson, S., Corley, R., Willcutt, E. G., Wadsworth, S., et al. (2009). Genetic and environmental influences on aspects of literacy and language in early childhood: Continuity and change from preschool to grade 2. *Journal of Neurolinguistics, 22*(3), 219-236.

Byrne, B., Coventry, W. L., Olson, R. K., Wadsworth, S. J., Samuelsson, S., Petrill, S. A., Willcutt, E. G., et al. (2010). “Teacher effects” in early literacy development: Evidence from a study of twins. *Journal of Educational Psychology, 102*(1), 32-42. doi: 10.1037/a0017288

Calvin, C. M., Fernandes, C., Smith, P., Visscher, P. M., & Deary, I. J. (2010). Sex, intelligence and educational achievement in a national cohort of over 175,000 11-year-old schoolchildren in England. *Intelligence, 38*(4), 424-432.

Carlson, M. J., & Corcoran, M. E. (2001). Family structure and children's behavioral and cognitive outcomes. *Journal of Marriage and Family, 63*(3), 779-792. doi: 10.1111/j.1741-3737.2001.00779.x

Carnoy, M., Khavenson, T., Loyalka, P., Schmidt, W. H., & Zakharov, A. (2016). Revisiting the relationship between international assessment outcomes and educational production: Evidence from a longitudinal PISA-timss sample. *American Educational Research Journal, 53*(4), 1054-1085. doi: 10.3102/0002831216653180

Catholic Education Commission Victoria (2017). *Capacity to contribute and school SES scores*. Melbourne: Catholic Education Commission Victoria.

Cheng, A., Hitt, C., Kisida, B., & Mills, J. N. (2017). “No excuses” charter schools: A meta-analysis of the experimental evidence on student achievement. *Journal of School Choice,* 1-30. doi: 10.1080/15582159.2017.1286210

Cherian, V. I., & Malehase, M. C. (1998). Relationship between family income and achievement in english of children from single- and two-parent families. *Psychological Reports, 83*(2), 431-434. doi: 10.2466/pr0.83.6.431-434

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682. doi: http://dx.doi.org/10.1016/j.econedurev.2007.10.002

Cobb-Clark, D. A., & Jha, N. (2016). Educational achievement and the allocation of school resources. *Australian Economic Review, 49*(3), 251-271. doi: 10.1111/1467-8462.12159

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.

Coleman, J. S., Campbell, E. S., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, D.C.: US Government Printing Office.

Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High school achievement. Public, catholic, and private schools compared.* New York: Basic Books.

Commonwealth of Australia. (2014). Appendix to the report of the national commission of government.Towards Responsible Government. Canberra: Commonwealkth of Australia. Retrieved from http://www.ncoa.gov.au/report/docs/appendix_volume%201.pdf.

Coyle, T. R. (2015). Relations among general intelligence (g), aptitude tests, and gpa: Linear effects dominate. *Intelligence*, *53*, 16-22.

Cullen, J. B., Jacob, B. A., & Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica*, *74*(5), 1191-1230.

Dalliard, M. (2014). Is IQ heritability moderated by race? An analysis of the CNLSY sample. *Human Varieties*. Retrieved from: http://humanvarieties.org/2014/03/18/is-iq-heritability-moderated-by-race-an-analysis-of-the-cnlsy-sample/

Daniels, M., Devlin, B., & Roeder, K. (1997). Of genes and IQ. In B. Devlin, S. E. Fienberg, D. P. Resnick, & K. Roeder (Eds.), *Intelligence, genes, and success: Scientists respond to the Bell Curve* (pp. 45-70). New York: Springer-Verlag.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, *8*(1).

de Broucker, P., & Underwood, C. (1998). International education mobility: An international comparison with a focus on post-secondary education. *Education Quarterly Review*, *5*(2), 30-51.

de Zeeuw, E. L., de Geus, E. J. C., & Boomsma, D. I. (2015). Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends in Neuroscience and Education*, *4*(2015), 69–76.

Deary, I. J., Johnson, W., & Houlihan, L. M. (2009). Genetic foundations of human intelligence. *Human Genetics*, *126*(1), 215-232. doi: 10.1007/s00439-009-0655-4

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13-21.

Department of Education and Training. (2017). *Timms and PISA results: Answer to question on notice sq17-000543*. Canberra: Australian Government.

Detterman, D. K. (2016). Education and intelligence: Pity the poor teacher because student characteristics are more significant than teachers or schools. *The Spanish journal of psychology*, *19*.

Dockery, A. M., Seymour, R., & Koshy, P. (2016). Promoting low socio-economic participation in higher education: A comparison of area-based and individual measures. *Studies in Higher Education*, *41*(9), 1692-1714. doi: 10.1080/03075079.2015.1020777

Doren, C., & Grodsky, E. (2016). What skills can buy: Transmission of advantage through cognitive and noncognitive skills. *Sociology of Education*, *89*(4), 321-342. doi: 10.1177/0038040716667994

Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What no child left behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, *104*(2), 439-451. doi: 10.1037/a0026280

Fergusson, D. M., Horwood, L. J., & Boden, J. M. (2008). The transmission of social inequality: Examination of the linkages between family socioeconomic status in childhood and educational achievement in young adulthood. *Research in Social Stratification and Mobility*, *26*(3), 277-295.

Firkowska, A., Ostrowska, A., Sokolowska, M., Stein, Z., Susser, M., & Wald, I. (1978). Cognitive development and social policy. *Science*, *200*(4348), 1357-1362.

Flerea, S., Krajnca, M. T., Klanjšeka, R., Musila, B., & Kirbiša, A. (2010). Cultural capital and intellectual ability as predictors of scholastic achievement: A study of Slovenian secondary school students. *British Journal of Sociology of Education, 31*(1), 47-58.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or *g*? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*(6), 373-378. doi: 10.1111/j.0956-7976.2004.00687.x

Fuchs, T., & Wößmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. *Empirical Economics, 32*(2-3), 433-464

Goldhaber, D., & Hansen, M. (2013). Is it just a bad class? Assessing the long-term stability of estimated teacher performance. *Economica, 80*(319), 589-612.

Gonski, D., Boston, K., Greiner, K., Lawrence, C., Scales, B., & Tannock, P. (2011). *Review of funding for schooling: Final report*. Canberra: Australian Government.

Goss, P., Chisholm, C., Sonnemann, J., & Nelson, L. (2016). *Widening gaps: What NAPLAN tells us about student progress*. Melbourne: Grattan Institute.

Gottfredson, L. S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence, 24*(1), 13-23.

Gottfredson, L. S. (1997b). Why g matters: The complexity of everyday life. *Intelligence, 24*(1), 79-132.

Gottfredson, L. S. (2016). Hans eysenck's theory of intelligence, and what it reveals about him. *Personality and Individual Differences, 103*, 116-127. doi: http://dx.doi.org/10.1016/j.paid.2016.04.036

Grasby, K. L., & Coventry, W. L. (2016). Longitudinal stability and growth in literacy and numeracy in Australian school students. *Behavior Genetics*, 1-16. doi: 10.1007/s10519-016-9796-0

Grasby, K. L., Coventry, W. L., Byrne, B., Olson, R. K., & Medland, S. E. (2016). Genetic and environmental influences on literacy and numeracy performance in Australian school children in grades 3, 5, 7, and 9. *Behavior Genetics*, 1-22. doi: 10.1007/s10519-016-9797-z

Gray, J., Goldstein, H., & Thomas, S. (2001). Predicting the future: The role of past performance in determining trends in institutional effectiveness at a level. *British Educational Research Journal, 27*(4), x-405.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361-396. doi: 10.3102/00346543066003361

Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis, 19*(2), 141-164.

Hanushek, E. A. (2006). School resources *Handbook of the economics of education* (Vol. 2, pp. 865-908).

Hanushek, E. A. (2009). Teacher deselection. In D. Goldhaber & J. Hannaway (Eds.), *In creating a new teaching profession* (pp. 165–180). Washington, DC: Urban Institute Press.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466-479. doi: http://dx.doi.org/10.1016/j.econedurev.2010.12.006

Hanushek, E. A., & Rivikin, S. G. (2004). How to improve the supply of high-quality teachers. *Brookings Papers on Education Policy*, 7-44.

Hanushek, E. A., & Rivikin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review, 100*(2), 267–271.

Hanushek, E. A., & Woessmann, L. (2016). *The knowledge capital of nations: Education and the economics of growth*. Cambridge, MA: MIT Press.

Hanushek, E. A., & Woessmann, L. (2017). School resources and student achievement: A review of cross-country economic research. In M. Rosén, K. Yang Hansen, & U. Wolff (Eds.), *Cognitive abilities and educational outcomes: A festschrift in honour of jan-eric gustafsson* (pp. 149-171). Cham: Springer International Publishing.

Harris, D. N., & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics, 95*(7–8), 798-812. doi: http://dx.doi.org/10.1016/j.jpubeco.2010.11.009

Harris, J. R. (2009). *The nurture assumption: Why children turn out the way they do*.

Hart, S. A., Petrill, S. A., Thompson, L. A., & Plomin, R. (2009). The abcs of math: A genetic analysis of mathematics and its links with reading ability and general cognitive ability. *Journal of Educational Psychology, 101*(2), 388-402. doi: 10.1037/a0015115

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement* London and New York: Routledge.

Haug, A. A., & Blackburn, V. C. (2017). Government secondary school finances in New South Wales: Accounting for students' prior achievements in a two-stage DEA at the school level. *Journal of Productivity Analysis*, 1-15. doi: 10.1007/s11123-017-0502-x

Hauser, R. M. (1969). Schools and the stratification process. *American Journal of Sociology, 74*(6), 587-611. doi: 10.2307/2776111

Hauser, R. M. (1970). Context and consex: A cautionary tale. *American Journal of Sociology, 75*, 645-664.

Hauser, R. M. (1974). Contextual analysis revisited. *Sociological Methods and Research, 2*(3), 365-373.

Haworth, C. M. A., Dale, P. S., & Plomin, R. (2009). The etiology of science performance: Decreasing heritability and increasing importance of the shared environment from 9 to 12 years of age. *Child Development, 80*(3), 662-673. doi: 10.1111/j.1467-8624.2009.01289.x

Hemmings, B., & Kay, R. (2010). Prior achievement, effort, and mathematics attitude as predictors of current achievement. *Australian Educational Researcher, 37*(2), 41-58. doi: 10.1007/BF03216921

Hempenstall, K. (2004). The importance of effective instruction. In N. E. Marchand-Martella, T. A. Slocum, & R. C. Martella (Eds.), *Introduction to direct instruction*. Needham Heights, MA: Allyn and Bacon.

Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and class structure in American life*. New York: The Free Press.

Hill, P., & Rowe, K. J. (1996). Multilevel modelling in school effectiveness research. *School Effectiveness and School Improvement, 7*(1), 1-34.

Humlum, M. K. (2011). Timing of family income, borrowing constraints, and child achievement. *Journal of Population Economics, 24*(3), 979-1004. doi: 10.1007/s00148-010-0309-9

Jacob, B., & Lefgren, L. (2007). Principals as agents: Subjective performance assessment in education. *Journal of Labor Economics, 26*(1), 101-136.

Jencks, C., Bartlett, S., Corcan, M., Crouse, J., Eaglesfield, D., Jackson, G., McClelland, K., et al. (1979). *Who gets ahead ahead? The determinants of economic success in America*. New York: Basic Books.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., et al. (1972). *Inequality. A reassessment of the effect of family and schooling in America*. New York: Basic Books.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39*(1), 1-123.
Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, Conn.: Praeger.
Jensen, B., Sandoval-Hernández, A., Knoll, S., & Gonzalez, E. J. (2012). *The experience of new teachers results from talis 2008*: OECD.
Joseph, B. (2017). *The fantasy of Gonski funding: The ongoing battle over school spending* (CIS Research Reports No. 26). Sydney: Centre for Independent Studies.
Justman, M., & Ryan, C. (2013). *What's wrong with the Gonski report: Funding reform and student achievement?* (Melbourne Institute Policy Briefs No. 2/13). Melbourne: Melbourne Institute for Applied Economic and Social Research.
Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive *g* and academic achievement *g* one and the same *g*? An exploration on the Woodcock–Johnson and Kaufman tests. *Intelligence, 40*(2), 123-138.
Kerckhoff, A. C., Raudenbush, S. W., & Glennie, E. (2001). Education, cognitive skill, and labour force outcomes. *Sociology of Education, 74*(January), 1-24.
Koenig, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT and general cognitive ability. *Intelligence, 36*(2), 153-160. doi: 10.1016/j.intell.2007.03.005
Korenman, S., & Winship, C. (2000). A reanalysis of the Bell Curve: Intelligence, family, background and schools. In K. Arrow, S. Bowles, & S. Durlauf (Eds.), *Meritocracy and economic inequality* (pp. 137-178). Princeton: Princeton University Press.
Korpershoek, H., Harms, T., Boer, H. d., Kuijk, M. v., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research, 86*(3), 643-680. doi: doi:10.3102/0034654315626799
Kovas, Y., Malykh, S., & Gaysina, D. (Eds.). (2016). *Behavioural genetics for education*: Palgrave Macmillan UK.
Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and numeracy are more heritable than intelligence in primary school. *Psychological Science, 24*(10), 2048-2056.
Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J. B., Asbury, K., et al. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences of the United States of America, 111*(42), 15273-15278.
Lamb, S., Jackson, J., Walstab, A., & Huo, S. (2015). *Educational opportunity in Australia 2015: Who succeeds and who misses out*. Melbourne: Centre for International Research on Education Systems, Victoria University.
Lavy, V. (2015). What makes an effective teacher? Quasi-experimental evidence. *CESifo Economic Studies*. doi: 10.1093/cesifo/ifv001
Le, A. T., Miller, P. W., Slutske, W. S., & Martin, N. G. (2011). Opportunity and educational outcomes in Australia. *Economic Record, 87*(SUPPL. 1), 125-135.
Leigh, A. (2010). Estimating teacher effectiveness from two-year changes in students' test scores. *Economics of Education Review, 29*(3), 480-488.
Leigh, A., & Mead, S. (2005). *Lifting teacher performance* (Policy Report April 2005): Progressive Policy Institute.
Lenkeit, J. (2013). Effectiveness measures for cross-sectional studies: A comparison of value-added models and contextualised attainment models. *School Effectiveness and School Improvement, 24*(1), 39-63.
Lu, L., & Rickard, K. (2014). *Value added models for NSW government schools*. Sydney: Centre for Education Statistics and Evaluation, Education and Communities.

Lucchini, M., Della Bella, S., & Pisati, M. (2013). The weight of the genetic and environmental dimensions in the inter-generational transmission of educational success. *European Sociological Review, 29*(2), 289-301. doi: 10.1093/esr/jcr067

Marks, G. (2016a). Explaining the substantial inter-domain and over-time correlations in student achievement: The importance of stable student attributes. *Educational Research and Evaluation*. doi: 10.1080/13803611.2016.1191359

Marks, G. N. (2006). Are between- and within-school differences in student performance largely due to socio-economic background? Evidence from 30 countries. *Educational Research, 48*(1), 21-40. doi: 10.1080/00131880500498396

Marks, G. N. (2007). Do schools matter for early school leaving? Individual and school influences in Australia. *School Effectiveness and School Improvement, 18*(4), 429-450.

Marks, G. N. (2009a). Accounting for school-sector differences in university entrance performance. *Australian Journal of Education, 53*(1), 19-38.

Marks, G. N. (2009b). Modernization theory and changes over time in the reproduction of socioeconomic inequalities in Australia. *Social Forces, 88*(2), 917-944.

Marks, G. N. (2010a). School sector and socioeconomic inequalities in university entrance in Australia: The role of the stratified curriculum. *Educational Research and Evaluation, 16*(1), 23-37. doi: doi: 10.1080/13803611003711310

Marks, G. N. (2010b). What aspects of schooling are important? School effects on tertiary entrance performance. *School Effectiveness and School Improvement, 21*(3), 267-287. doi: 10.1080/09243451003694364

Marks, G. N. (2010c). What aspects of schooling are important? School effects on tertiary entrance performance in Australia. *School Effectiveness and School Improvement, 21*(3), 267-287.

Marks, G. N. (2013). Evaluating Effectively Maintained Inequality: School and post-school transitions, socioeconomic background, academic ability and curricular placement. *Social Science Research, 42*(6), 1635-1649. doi: http://dx.doi.org/10.1016/j.ssresearch.2013.07.004

Marks, G. N. (2014a). Demographic and socioeconomic inequalities in student achievement over the school career. *Australian Journal of Education, 58*(3), 223–247. doi: 10.1177/0004944114537052

Marks, G. N. (2014b). *Education, social background and cognitive ability: The decline of the social*. Abingdon, Oxon; New York, NY: Routledge.

Marks, G. N. (2014c). Reaching year 12 in Victoria, Australia: Student and school influences. *Educational Research and Evaluation, 20*(5), 333-347. doi: 10.1080/13803611.2014.964262

Marks, G. N. (2014d). The size, stability, and consistency of school effects: Evidence from Victoria. *School Effectiveness and School Improvement, 26*(3), 397-414. doi: 10.1080/09243453.2014.964264

Marks, G. N. (2015a). Are school-SES effects statistical artefacts? Evidence from longitudinal population data. *Oxford Review of Education, 41*(1), 122-144. doi: 10.1080/03054985.2015.1006613

Marks, G. N. (2015b). Do catholic and independent schools "add-value" to students' tertiary entrance performance? Evidence from longitudinal population data. *Australian Journal of Education, 59*(2), 133-157. doi: DOI: 10.1177/0004944115586658

Marks, G. N. (2016b). The relative effects of socio-economic, demographic, non-cognitive and cognitive influences on student achievement in Australia. *Learning and Individual Differences, 49*, 1-10. doi: http://dx.doi.org/10.1016/j.lindif.2016.05.012

Marks, G. N. (2017a). Family income has only weak effects on cognitive scores in Australia: A comment on khanam and nghiem. *Demography, 54*(2), 801-807. doi: 10.1007/s13524-017-0557-3

Marks, G. N. (2017b). Is SES really that important for educational outcomes in Australia? A review and some recent evidence. *The Australian Educational Researcher, 44*(2), 191-211. doi: 10.1007/s13384-016-0219-2

Marks, G. N. (Forthcoming). The contribution of genes and the environment to educational and socioeconomic attainments in Australia. *Twin Research and Human Genetics*. doi: 10.1017/thg.2017.32

Marks, G. N., & McMillan, J. (2003). Declining inequality? The changing impact of socioeconomic background and ability on education in Australia. *British Journal of Sociology, 54*(4), 453-471.

Marks, G. N., & McMillan, J. (2007). Changes in socioeconomic inequalities in university participation in Australia. In Y. Shavit, R. Arum, & A. Gamoran (Eds.), *Stratification in higher education: A comparative study* (pp. 351-373). Stanford, California: Stanford University Press.

Marks, G. N., & Mooi-Reci, I. (2016). The declining influence of family background on educational attainment in Australia: The role of measured and unmeasured influences. *Social Science Research, 55*(January), 171-185.

Marks, G. N., Underwood, C., Rothman, S., & Brown, J. (2011). *Career moves: Expectations and destinations of New South Wales senior secondary students*. NSW Board of Vocational Education and Training, Sydney.

Mayer, R. E. (2000). Intelligence and education In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 519−533). Cambridge, UK: Cambridge University Press.

Mayer, S. E. (1997). *What money can't buy: Family income and children's life chances*. Cambridge, MA: Harvard University Press.

McCaffrey, D. F., Sass, T. R., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy, 4*(4), 572-606. doi: 10.1162/edfp.2009.4.4.572

McIntosh, J., & Munk, M. D. (2007). Scholastic ability vs family background in educational success: Evidence from Danish sample survey data. *Journal of Population Economics, 20*(1), 101-120.

McNiece, R., Bidgood, P., & Soan, P. (2004). An investigation into using national longitudinal studies to examine trends in educational attainment and development. *Educational Research, 46*(2), 119-136. doi: 10.1080/0013188042000222412

Miller, P. W., Mulvey, C., & Martin, N. G. (2001). Genetic and environmental contributions to educational attainment in Australia. *Economics of Education Review, 20*(2), 211–224.

Nash, R. (2003). Is the school composition effect real? A discussion with evidence from the uk PISA data. *School Effectiveness and School Improvement, 14*(4), 441-457.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77-101. doi: 10.1037/0003-066X.51.2.77

Nielsen, F. (2006). Achievement and ascription in educational attainment: Genetic and environmental influences on adolescent schooling. *Social Forces, 85*(1), 193-216.

NSW Department of Education (2014). *What works best: Evidence-based practices to help improve NSW student performance*. Sydney: Centre for Education Statistics and Evaluation, NSW Department of Education and Communities.

OECD. (2001). *Knowledge and skills for life. First results from the OECD programme for international student assessment*. Paris: Organisation for Economic Co-operation and Development.
OECD (2004). *Learning for tomorrow's world. First results from PISA 2003*. Paris: Organisation for Economic Co-operation and Development.
OECD. (2007a). *Science competencies for tomorrow's world.* (Vol. 2). Paris: Organisation for Economic Co-operation and Development.
OECD. (2007b). *Science competencies for tomorrow's world. Volume 1.* (Vol. 1). Paris: Organisation for Economic Co-operation and Development.
OECD (2010a). *Economic policy reforms: Going for growth*. Paris: Organisation for Economic Co-operation and Development.
OECD (2010b). *How does academic ability affect educational and labour market pathways in Canada* (No. 30). Paris: Organisation for Economic Co-operation and Development.
OECD. (2010c). *PISA 2009 results: Overcoming social background. Equity in learning opportunities and outcomes* (Vol. II). Paris: Organisation for Economic Co-operation and Development.
OECD (2012). *PISA in focus 13*. Paris: Organisation for Economic Co-operation and Development.
OECD. (2013). *PISA 2012 results: Excellence through equity giving every student the chance to succeed* (Vol. II). Paris: Organisation for Economic Co-operation and Development.
OECD (2014). *PISA in focus 44*. Paris: Organisation for Economic Co-operation and Development.
OECD. (2016a). Country note AustraliaPISA Low-Performing Students Why They Fall Behind And How To help Them Succeed. Paris: Organisation for Economic Co-operation and deevelopment. Retrieved from https://www.oecd.org/australia/PISA-2012-low-performers-Australia-ENG.pdf.
OECD. (2016b). *PISA 2015 results: Excellence and equity in education* (Vol. I). Paris: Organisation for Economic Co-operation and Development.
OECD (2016c). *PISA low-performing students why they fall behind and how to help them succeed*. Paris: Organisation for Economic Co-operation and deevelopment.
Orr, A. J. (2003). Black-white differences in achievement: The importance of wealth. *Sociology of Education, 76*(4), 281-304.
Perry, L., & McConney, A. (2010a). School socio-economic composition and student outcomes in Australia: Implications for educational policy. *Australian Journal of Education, 54*(1), 72-85.
Perry, L. B., & McConney, A. (2010b). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teachers College Record, 112*(4), 1137-1162.
Petrill, S. A. (2016). Behavioural genetic studies of reading and mathematics skills. In S. B. Malykh, Y. Kovas, & D. Gaysina (Eds.), *Behavioural genetics for education* (pp. 60-76). UK: Palgrave Macmillan.
Pinker, S. (2011). *The blank slate: The modern denial of human nature* New York: Penguin.
Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1997). Nature, nurture, and cognitive development from 1 to 16 years: A parent-offspring adoption study. *Psychological Science, 8*(6), 442-447.
Plug, E., & Vijverberg, W. (2003). Schooling, family background, and adoption: Is it nature or is it nurture? *Journal of Political Economy, 111*(3), 611–641.

Pokropek, A. (2014). Phantom effects in multilevel compositional analysis: Problems and solutions. *Sociological Methods & Research, 44*(4), 677-705. doi: 10.1177/0049124114553801
Postlethwaite, T. N., & Kellaghan, T. (2008). *National assessments of educational achievement*. Paris: UNESCO.
Ranasinghe, R. (2015). The transmission of education across generations: Evidence from Australia. *B E Journal of Economic Analysis & Policy*.
Raudenbush, S. W. (2014). What do we know about the long-term impacts of teacher value added? *Carnegie Foundation for the Advancement of Teaching Knowledge Brief*. Retrieved from: http://www.carnegieknowledgenetwork.org/briefs/long-term-impacts/
Raudenbush, S. W., & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics, 20*(4), 307-335. doi: 10.3102/10769986020004307
Reynolds, A. J., & Walberg, H. J. (1992). A process model of mathematics achievement and attitude. *Journal for Research in Mathematics Education, 23*(4), 306-328.
Rindermann, H. (2008). Relevance of education and intelligence at the national level for the economic welfare of people. *Intelligence, 36*, 127–142.
Rindermann, H., & Thompson, J. (2011). Cognitive capitalism: The effect of cognitive ability on wealth, as mediated through scientific achievement and economic freedom. *Psychological Science, 22*(6), 754-763. doi: 10.1177/0956797611407207
Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*(3), 351-357.
Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*(November–December), 118-137. doi: 10.1016/j.intell.2015.09.002
Rothman, S. (2002). *Achievement in literacy and numeracy by Australian 14 year-olds, 1975-1998* (Longitudinal Surveys of Australian Youth Research Reports No. 29).
Rowe, D. C. (1994). *The limits of family influence: Genes, experience, and behavior*. New York: Guilford Press.
Rowe, D. C., Vesterdal, W. J., & Rodgers, J. L. (1999). Herrnstein's syllogism: Genetic and shared environmental influences on IQ, education, and income. *Intelligence, 26*(4), 405-423.
Ryan, C. (2017). *Secondary school teacher effects on student achievement in Australian schools* (Melbourne Institute Working Paper Series No. 11/17). Melbourne: Melbourne Institute of Applied Economic and Social Research, University of Melbourne.
Sacerdote, B. (2011). Nature and nurture effects on children's outcomes: What have we learned from studies of twins and adoptees? In J. Benhabib, M. Jackson, & A. Bisin (Eds.), *Handbook of social economics* (Vol. 1, pp. 1-30): North-Holland.
Sahlberg, P. (2015, March 31, 2015). What makes Finnish teachers so special? A: It's not brains. *The Guardian*. Retrieved from https://www.theguardian.com/education/2015/mar/31/finnish-teachers-special-train-teach
Samuelsson, S., Olson, R., Wadsworth, S., Corley, R., DeFries, J. C., Willcutt, E., Hulslander, J., et al. (2007). Genetic and environmental influences on prereading skills and early reading and spelling development in the United States, Australia, and Scandinavia. *Reading and Writing, 20*(1-2), 51-75. doi: 10.1007/s11145-006-9018-x
Scarr, S., & Weinberg, R. A. (1978). The influence of "family background" on intellectual attainment. *American Sociological Review, 43*(5), 674-692.

Scheerens, J., & Bosker, R. (1997a). *The foundations of educational effectiveness*. Oxford: Pergamon.

Scheerens, J., Witziers, B., & Steen, R. (2013). A meta-analysis of school effectiveness studies. *Revista de Educacion*(361), 619-645. doi: 10.4438/1988-592X-RE-2013-361-235

Scheerens, J. C. J. (2016). *Educational effectiveness and ineffectiveness: A critical review of the knowledge base*. Dordretch: Springer.

Scheerens, J. C. J., & Bosker, R. J. (1997b). *The foundations of educational effectiveness*. Oxford: Pergamon.

Segal, N. L. (2012). *Born together—reared apart: The landmark minnesota twin study*. Cambrige, MA: Harvard University Press.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytical review of research. *Review of Educational Research, 75*(3), 417-453. doi: 10.3102/00346543075003417

Soden, B., Christopher, M. E., Hulslander, J., Olson, R. K., Cutting, L., Keenan, J. M., Thompson, L. A., et al. (2015). Longitudinal stability in reading comprehension is largely heritable from grades 1 to 6. *PLoS ONE, 10*(1). doi: 10.1371/journal.pone.0113807

Steenbergen-Hu, S., Makel, M. C., & Olszewski-Kubilius, P. (2016). What one hundred years of research says about the effects of ability grouping and acceleration on k–12 students' academic achievement. *Review of Educational Research, 86*(4), 849-899. doi: doi:10.3102/0034654316675417

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*(5), 401-426.

Suter, L. E. (2000). Is student achievement immutable? Evidence from international studies on schooling and student achievement. *Review of Educational Research, 70*(4), 529-545.

Televantou, I., Marsh, H. W., Kyriakides, L., Nagengast, B., Fletcher, J., & Malmberg, L.-E. (2015). Phantom effects in school composition research: Consequences of failure to control biases due to measurement error in traditional multilevel models. *School Effectiveness and School Improvement, 26*(1), 75-101. doi: 10.1080/09243453.2013.871302

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research, 83*(3), 357-385. doi: doi:10.3102/0034654313483907

Thomson, S., Bortoli, L. D., & Underwood, C. (2017). *PISA 2015: Reporting australia's results*. Melbourne: Australian Council for Educational Research.

Thomson, S., De Bortoli, L., & Buckley, S. (2013). *PISA 2012: How Australia measures up*. Melbourne: Australian Council for Educational Research.

Thomson, S., De Bortoli, L., Nicholas, M., Hillman, K., & Buckley, S. (2010). *Challenges for Australian education: Results from PISA 2009*. Melbourne: Australian Council for Educational Research.

Thomson, S., Hillman, K., Wernert, N., Schmid, M., & Buckley, S. (2012). *Monitoring Australian year 4 student achievement internationally: Timss and pirls 2011*. Melbourne: Australian Council for Educational Research.

Thorn, W. (2009). *International adult literacy and basic skills surveys in the OECD region*. Paris: Organisation for Economic Co-operation and Development.

Torres, D. D. (2013). Understanding how family socioeconomic status mediates the maternal intelligence-child cognitive outcomes relationship: A moderated mediation analysis.

Biodemography and Social Biology, 59(2), 157–177. doi: 10.1080/19485565.2013.833804

van Leeuwen, M., van den Berg, S. M., & Boomsma, D. I. (2008). A twin-family study of general IQ. *Learning and Individual Differences, 18*(1), 76-88.

Victorian Department of Education and Training (2015). *The education state: Schools consultation paper*. Melbourne: State of Victoria (Department of Education and Training)

Violato, M., Petrou, S., Gray, R., & Redshaw, M. (2011). Family income and child cognitive and behavioural development in the United Kingdom: Does money matter? *Health Economics, 20*(10), 1201-1225. doi: 10.1002/hec.1665

Wadsworth, S. J., Corley, R. P., Hewitt, J. K., & DeFries, J. C. (2001). Stability of genetic and environmental influences on reading performance at 7, 12, and 16 years of age in the colorado adoption project. *Behavior Genetics, 31*(4), 353-359. doi: 10.1023/a:1012218301437

Wadsworth, S. J., Corley, R. P., Plomin, R., Hewitt, J. K., & DeFries, J. C. (2006). Genetic and environmental influences on continuity and change in reading achievement in the colorado adoption project. In A. Huston & M. N. Ripke (Eds.), *Developmental contexts in middle childhood: Bridges to adolescence and adulthood* (pp. 87-106). Cambridge: Cambridge University Press.

Wainwright, M. A., Wright, M. J., Geffen, G. M., Geffen, L. B., Luciano, M., & Martin, N. G. (2004). Genetic and environmental sources of covariance between reading tests used in neuropsychological assessment and IQ subtests. *Behavior Genetics, 34*(6), 673.

Wainwright, M. A., Wright, M. J., Luciano, M., Geffen, G. M., & Martin, N. G. (2005). Multivariate genetic analysis of academic skills of the Queensland core skills test and IQ highlight the importance of genetic *g*. *Twin Research and Human Genetics, 8*(6), 602-608. doi: 10.1375/183242705774860259

Walberg, H. J. (1984). Improving the productivity of America's schools. *Educational Leadership, 41*(8), 19–27.

White, K. R. (1980). Socio-economic status and achievement. *Evaluation in Education, 4*, 79-81.

White, K. R. (1982). The relationship between socio-economic status and academic achievement. *Psychological Bulletin, 91*(3), 461-481. doi: 10.1037/0033-2909.91.3.461

White, S. B., Reynolds, P. D., Thomas, M. M., & Gitzlaff, N. J. (1993). Socioeconomic status and achievement revisited. *Urban Education, 28*(3), 328-343. doi: 10.1177/0042085993028003007

Willms, J. D. (1992). *Monitoring school performance: A guide for educators*. Lewes: Falmer Press.

Wright, P., White, J. T., Sanders, W. L., & Rivers, J. C. (2010). *SAS EVAAS statistical models* (SAS White Paper).

Wright, S. P. (2010). *An investigation of two nonparametric regression models for value-added assessment in education*: SAS Institute Inc.

Yeh, S. S. (2009). The cost-effectiveness of raising teacher quality. *Educational Research Review, 4*(3), 220-232. doi: 10.1016/j.edurev.2008.06.002

Zagorsky, J. L. (2007). Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress. *Intelligence, 35*(5), 489-501. doi: 10.1016/j.intell.2007.02.003

Zimmer, R. W., & Toma, E. F. (2000). Peer effects in private and public schools across countries. *Journal of Policy Analysis and Management, 19*(1), 75–92.

Notes

1 Extended kinship designs also show high heritabilities for student achievement (Dalliard, 2014).

2 For example, 10 for students who leave school after year 10, 12 for school completers and 15 for Bachelor degree holders.

3 Mayer's (1997) 'true' effects of family income are the effects of income net of unobserved parental characteristics calculated from longitudinal data.

4 In that sample the standard deviation of reading test scores was 96 (Humlum, 2011, p. 989).

5 Presumably, in the country note about Australia, the authors left out the word "other".

6 The Grattan institute report incorporates prior achievement but the focus is SES and other forms of disadvantage. Their new time-based measure, 'years of progress' is particularly problematic for both theoretical and statistical reasons, which I will not detail here. The bottom line is that the "years of progress" measure is likely to produce misleading estimates.

7 In the 1980s there were rational debates about policy and sensible, often bipartisan, policies were implemented. However, those days are long gone and current policy debates are reduced to which party will spend the most money under the assumption that a social problem is solved once a government increases spending and recruits bureaucrats.

8 It would not be difficult for schools and systems to collect the student identification number of students who attended another school in the year previously. These would enable appropriate linkages to be made.

9 Of course, the proviso is that individual students or individual schools cannot be identified.