Twin Research and Human Genetics (2020), **23**, 165–173
doi:10.1017/thg.2020.49

CAMBRIDGE
UNIVERSITY PRESS

Article

The Academic Development Study of Australian Twins (ADSAT): Research Aims and Design

Sally A. Larsen[1], Callie W. Little[1], Katrina Grasby[2], Brian Byrne[1], Richard K. Olson[3] and William L. Coventry[1]

[1]School of Psychology, University of New England, Armidale, Australia, [2]Queensland Institute of Medical Research, Brisbane, Australia and [3]Psychology Department, University of Colorado, Boulder, CO, USA

Abstract

The Academic Development Study of Australian Twins was established in 2012 with the purpose of investigating the relative influence of genes and environments in literacy and numeracy capabilities across two primary and two secondary school grades in Australia. It is the first longitudinal twin project of its kind in Australia and comprises a sample of 2762 twin pairs, 40 triplet sets and 1485 nontwin siblings. Measures include standardized literacy and numeracy test data collected at Grades 3, 5, 7 and 9 as part of the National Assessment Program: Literacy and Numeracy. A range of demographic and behavioral data was also collected, some at multiple longitudinal time points. This article outlines the background and rationale for the study and provides an overview for the research design, sample and measures collected. Findings emerging from the project and future directions are discussed.

Keywords: Twin studies; behavior genetics; literacy; numeracy

(Received 28 February 2020; accepted 14 April 2020; First Published online 2 June 2020)

The Academic Development Study of Australian Twins (ADSAT) is the first nationwide, longitudinal twin study of educational achievement in Australia. Initiated in 2012, the project recruited child and adolescent twins, triplets and their nontwin siblings enrolled in school in any Australian state or territory, and attending any grade from Grade 3 to Grade 12. The overarching purpose of the project was to investigate genetic and environmental influences on reading, writing, spelling, grammar and numeracy, and the stability of genetic and environmental influence on these educational phenotypes across time in Australia. Alongside standardized testing data on literacy and numeracy achievement, a range of behavioral and environmental measures was collected biennially. Measures included demographic information at the family level, plus a series of behavioral measures previously shown to be related to educational achievement. The latter included information about twins' health, preschool attendance, attention and hyperactivity behaviors, sleep, technology use, homework behavior, and diet.

Extant research has demonstrated that academic achievement is partly heritable, with estimates for the core skills of literacy and numeracy generally reported at between 40% and 70% (see Kovas et al., 2016, for summaries of results from behavior genetic studies of educational achievement). Many studies emerging from twin projects in the USA, UK and Europe have shown a heritable component to the variance in reading and numeracy assessments from early primary school grades through to upper secondary grades. While heritability of literacy and numeracy is consistently moderate to high across studies, the estimates vary across different samples, as do the portions of variance attributed to shared and nonshared environment (Little et al., 2017). For example, while reports of the Twins Early Development Study conducted in the UK since 1996 usually emphasize high heritability of educational achievement alongside large nonshared environment estimates and negligible shared environment (e.g., Harlaar et al., 2007, 2010, 2013), other studies in the USA reveal lower heritability estimates and nonshared environment alongside higher shared environment estimates (Haughbrook et al., 2017; Logan et al., 2013). The differences in results from samples in two countries highlight that heritability estimates and environmental influences may be context- or study-dependent (Daucourt et al., 2020; Haughbrook et al., 2017; Little et al., 2017), and that behavior genetic studies of educational outcomes conducted in different environments are warranted.

ADSAT was conceptualized as a progression from the International Longitudinal Twin Study (ILTS), which followed a sample of Australian, US and Scandinavian twins ($n = 1000$ pairs) from preschool through to Grade 2, and investigated the extent to which genetic and environmental factors accounted for variation in early reading skills (Byrne et al., 2013). Between 1999 and 2010, the Australian arm of the ILTS recruited preschool-aged twin pairs ($N = 256$ pairs) from one metropolitan location in the state of New South Wales and followed them longitudinally for 4 years. ADSAT was designed to build on and extend the findings from the ILTS with a much larger sample of Australian school-aged twins in Grades 3 through 9.

One of the key findings of the ILTS supported the proposition that heritability estimates for academic abilities can vary depending on context. At the end of the kindergarten year, the heritability

Author for correspondence: Sally Larsen, Email: slarsen3@une.edu.au

Cite this article: Larsen SA, Little CW, Grasby K, Byrne B, Olson RK, and Coventry WL. (2020) The Academic Development Study of Australian Twins (ADSAT): Research Aims and Design. *Twin Research and Human Genetics* **23**: 165–173, https://doi.org/10.1017/thg.2020.49

of reading assessments varied by country, with estimates of 0.84, 0.68 and 0.33 for the Australian, US and Scandinavian samples, respectively; the parallel figures for the shared environment were 0.09, 0.25 and 0.52, and for nonshared environment, 0.08, 0.07 and 0.15 (Samuelsson et al., 2008). The authors attributed this variation to considerable differences between countries in policies and practices toward initial reading instruction. Writing about the ILTS more recently, Byrne et al. (2019) note

> The broad lesson from this set of results is that it is not appropriate to speak of the heritability of some variable. It is better to speak of the heritability of a variable under X circumstances (of sample, environmental particularities, period in history, and so on). (p. 8)

From its inception, ADSAT aimed to further contribute to the understanding of the heritability of and environmental influences on educational achievement specifically in the Australian context, and across a wider span of the school years than had been attempted formerly.

Knowledge about the specific nature of environmental influences, separate from genetic influences, also remains limited not only in the Australian context. Extant educational research in Australia suggests that environmental influences shared equally by twins in a pair (*shared environment* in the terminology of behavior genetics) do contribute to the academic performance of students. These shared environment factors are similar to those that have been found in international research and include socioeconomic status (SES) and related aspects of family life, preschool attendance, school effects and location, for example, (Buckingham et al., 2014; Marks 2015a, 2015b; Perry & McConney, 2013; Smith et al., 2019; Warren & Haisken-Denew, 2013). Nonetheless, behavior genetic research designs that investigate the extent to which shared environment contributes to variation in achievement across multiple longitudinal time points, and separate from genetic factors and nonshared environment, had not been widely applied in the Australian system before ADSAT was established.

Likewise, systematic sources of the nonshared environmental variance identified in behavior genetic research have not as yet been reliably identified (Plomin & Daniels, 2011; Tikhodeyev & Shcherbakova, 2019) There is some argument that nonshared environment variance in academic outcomes may be entirely stochastic (Plomin et al., 2016), and fairly transient (Bartels et al., 2002). A further advantage of a large, longitudinal twin project like ADSAT was the potential to further examine possible systematic contributors to nonshared environmental variance in the Australian context.

The schooling system and social policy environment in Australia is notably different to that in both the USA and the UK where much of the behavior genetic research into academic abilities has been conducted (Grasby, Coventry et al., 2019; OECD, 2009), with the exception of the ILTS and the Brisbane Adolescent Twin Study (Wright & Martin, 2011). In light of the ongoing suggestions of researchers that the results of genetic studies be applied to education policy and practice (Asbury & Plomin, 2014; Asbury & Wai, 2019; Thomas et al., 2015), it is imperative to further investigate whether the conclusions drawn from international research are comparable in different schooling systems. Establishing the ADSAT was a step toward achieving this aim.

Research Aims

The main aims of ADSAT were twofold. The first was to investigate whether the heritability of scholastic abilities in the middle years of schooling in Australia was comparable to those reported in US and UK samples. A secondary aim was to amass a wide range of measures of the shared and nonshared environment to further investigate possible environmental contributors to academic achievement proposed by existing educational research, using a genetically sensitive research design.

Sample Recruitment and Description

Beginning in 2013, the study combined both prospective and retrospective approaches in order to recruit as large and representative a sample as possible. Participants were approached via Twins Research Australia (TRA; formerly the Australian Twin Registry), an organization that holds records for over 70,000 twins residing Australia-wide (approximately 20% of the total Australian twin population; TRA, 2020). Overall, 8604 families of age-appropriate twins and triplets were contacted in annual mail approaches between 2013 and 2017. Of those approached, 2824 families enrolled in the study, a 33% response rate. Of these families, 1485 also enrolled a nontwin sibling. Figure 1 shows a complete recruitment and follow-up flow chart.

Upon enrollment, parents completed the initial Family Questionnaire, consent forms for participation in the study, and consent for researchers to request twin and sibling National Assessment Program: Literacy and Numeracy (NAPLAN) data from Australian state or territory education departments. Parents also provided information about twin and sibling school grade levels and schools attended, and twins and siblings completed assent forms. The study was approved by the Human Research Ethics Committee at the University of New England (Approval No. HE12–150 to December 31, 2017 and HE 18–163, current to June 21, 2021), by TRA, and by each state and territory Department of Education.

The study recruited twins, triplets and their closest age siblings, who had completed NAPLAN tests at any grade level in any Australian state or territory since 2008, the year the NAPLAN test program was first instituted in Australia. New participants at any school grade level were again recruited in 2014 and 2015. In 2016 and 2017, new participants were recruited only at the Grade 3 level. Participating students have birthdates ranging from 1993 (Grade 9 in 2008) to 2009 (Grade 3 in 2017). Table 1 shows numbers of participants by cohort.

Zygosity of same-sex twins was determined by parent report of DNA tests, or by parent responses to five questions about twin similarity in eye color, hair color and difficulty telling twins apart (Lykken et al., 1990). Zygosity information provided by parents to TRA was also compared with information provided to ADSAT to ensure consistency. DNA test information was recorded for 714 (56%) of monozygotic (MZ) and 413 (50%) of same-sex dizygotic (DZ) twins. Comparisons of DNA zygosity classification with questionnaire responses indicated that questionnaire responses identified 94.6% of twins as the correct zygosity (Grasby et al., 2016). Misclassifications were approximately equal for each twin type (54% DZ; Grasby et al., 2016). Numbers of MZ and DZ twin pairs can be seen in Table 2.

Data Collection Strategy

The Family Questionnaire comprised demographic information about twins and their families, and questions about the home environment. Information on twins and siblings included their birth dates, gender and the nature of the relatedness between twins and siblings (i.e., full, half, step-sibling). Information specific to

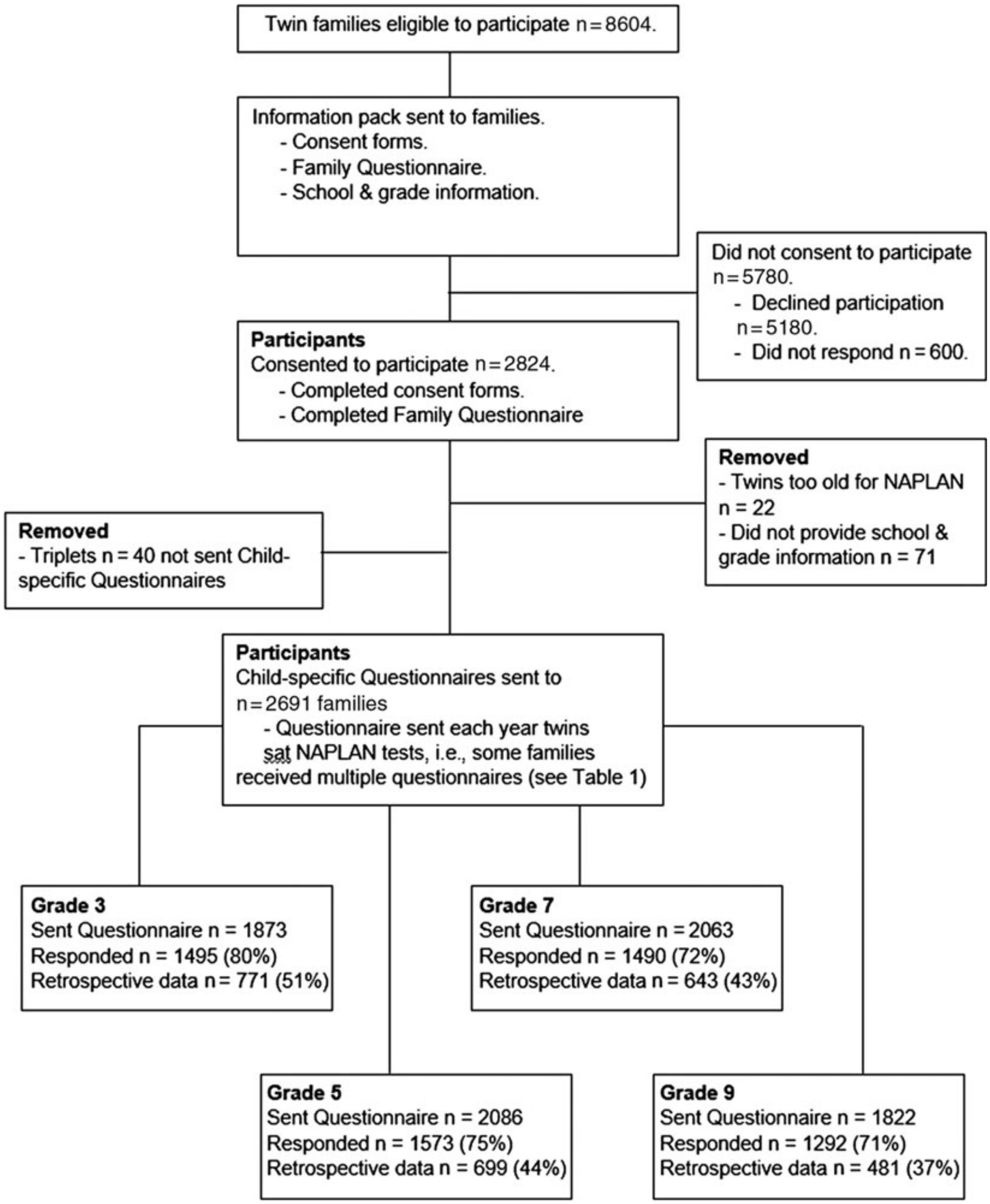

Fig. 1. Flow of subjects through the Academic Development Study of Australian Twins. Data as at end 2018.

twins included zygosity, birth weight, gestational age at birth and birth complications. Parents also answered questions on their ancestry, education levels, occupations, educational resources in the home, perceived importance of mathematics and a measure of household disorder. The full protocol for the Family Questionnaire can be found in Supplementary Table S1. Completing the Family Questionnaire was a condition of enrolling into the study, so there was a 99% response rate on this questionnaire (2764 questionnaires from 2802 families); in 95% of families, mothers completed the questionnaire.

The follow-up Child-specific Questionnaire was sent to families each year their twins sat NAPLAN tests. In early 2013, the Child-specific Questionnaire was piloted with three samples of twins who had just completed Grades 3, 7 and 9 ($N = 51$). Questionnaire length was tested to ensure the time taken to complete was no longer than 60 min. Minor alterations to formatting and wording were instituted following this pilot, but no major changes to the questionnaire were deemed necessary. Questionnaires were sent each year in June via an online survey platform or via post with follow-up messages in September and a second paper copy of the questionnaire sent to nonresponders in November. The questionnaire asked parents to respond to a series of items separately for each twin. Constructs included health information, details about school attendance and whether twins shared classes, inattention and hyperactivity behaviors, twins' enjoyment of reading and mathematics, homework behaviors, participation in extracurricular activities, sleep, screen time and dietary patterns. The protocol for the Child-specific Questionnaire can be found in Supplementary Table S2.

In total, 2221 families have returned at least one follow-up questionnaire (80%); this includes families with only one follow-up occasion, and families who were followed up multiple times.

https://doi.org/10.1017/thg.2020.49 Published online by Cambridge University Press

Table 1. ADSAT cohorts with grade-level information

	NAPLAN year				Participants		
Birth year	Grade 3	Grade 5	Grade 7	Grade 9	Twin pairs	Triplet sets	Siblings
1993–1994				2008	76	—	70
1994–1995				2009	84	—	72
1995–1996			2008	2010	137	2	71
1996–1997			2009	2011	152	1	81
1997–1998		2008	2010	2012	171	1	85
1998–1999		2009	2011	**2013**	198	5	105
1999–2000	2008	2010	2012	**2014**	200	7	131
2000–2001	2009	2011	**2013**	**2015**	174	6	104
2001–2002	2010	2012	**2014**	**2016**	207	6	124
2002–2003	2011	**2013**	**2015**	**2017**	219	3	117
2003–2004	2012	**2014**	**2016**	**2018**	204	3	131
2004–2005	**2013**	**2015**	**2017**	**2019**	202	—	91
2005–2006	**2014**	**2016**	**2018**	**2020**	199	1	80
2006–2007	**2015**	**2017**	**2019**	**2021**	158	3	60
2007–2008	**2016**	**2018**	**2020**	**2022**	154	2	20
2008–2009	**2017**	**2019**	**2021**	**2023**	151	—	17
2009–2010	**2018**	**2020**	**2022**	**2024**	5	—	24
2010–2011	**2019**	**2021**	**2023**	**2025**	—	—	21
Total					2691	40	1404

Note: Calendar years in bold indicate child-specific data collection concurrent with NAPLAN test grade, while unbolded represent retrospective reporting. A further 71 pairs did not provide school grade information.

Table 2. Monozygotic and dizygotic twin pairs by gender

	Monozygotic	Dizygotic
Male	609	410
Female	665	421
Opposite sex	—	626
Total	1274	1457

Note: 50 pairs were unable to be reliably identified as monozygotic or dizygotic. Percentage of twin pairs by zygosity is similar across all cohorts.

Table 3. Child-specific questionnaires returned by calendar year and grade

NAPLAN year	Grade 3	Grade 5	Grade 7	Grade 9
2008	161	119	106	68
2009	122	141	116	70
2010	156	161	119	106
2011	167	122	141	116
2012	165	156	161	121
2013	**169**	**167**	**122**	**141**
2014	**151**	**156**	**146**	**152**
2015	**144**	**159**	**152**	**124**
2016	**129**	**155**	**150**	**129**
2017	**128**	**122**	**142**	**129**
2018	**3**	**115**	**135**	**136**
Total:	1495	1573	1490	1292

Note: Numbers in bold indicate data collection concurrent with NAPLAN tests, while unbolded represent retrospective reporting.

Similar proportions of families responded to the questionnaire from each Australian state or territory, although families in Western Australia and Queensland had slightly lower response rates (77% and 78%, respectively) than the remaining jurisdictions (80%–85%).

Response rates have been gradually decreasing from 80% in the first round in 2013 ($N = 1940$ questionnaires returned) to a 69% response rate in 2018 ($N = 567$; Grades 5, 7 and 9 participants in this round). Similarly, response rates decline as twins' grade levels increase, with an 80% response rate on questionnaires when twins are in Grade 3 declining to a response rate of 71% when twins are in Grade 9 (see Figure 1). Response rates by grade level were similar for families completing retrospective questionnaires compared to questionnaires concurrent with NAPLAN grades. Declining response rates across multiple waves of data collection are not uncommon in longitudinal research (Gustavson et al., 2012), and since, for example, all 2018 participants are at least at their third follow-up occasion, it is to be expected that declining proportions of families will remain engaged with the study over time.

Due to the recruitment plan, the first round of questionnaires collected in 2013 contained a proportion of retrospectively asked questions, along with a proportion of concurrently asked questions (see Table 3). Some of these questions related to relatively stable constructs, such as twins' preschool attendance, first language, and whether or not twins shared a classroom at each grade level.

https://doi.org/10.1017/thg.2020.49 Published online by Cambridge University Press

Other questions related specifically to behaviors of twins during each NAPLAN test year, such as twin enjoyment of reading and mathematics, sleep behavior, diet, and time spent on extracurricular activities. Retrospective questionnaire responses have been shown to be a potential cause of bias (Bowling, 2005); however, the retrospective questionnaire format attempted to overcome some of this by presenting the questions in reverse chronological order for those participants with retrospective questions. For example, in the questionnaire for students in Grade 9 2011, parents were asked to 'Think about the twins around the time of the Grade 9 NAPLAN' for the first set of questions; then 'Think about the twins around the time of the Grade 7 NAPLAN in 2009' when questions were repeated. Preliminary exploration of retrospective and concurrent reporting shows higher correlations between time points when parents answered multiple questionnaire waves retrospectively, as opposed to responses given concurrent to grade level, suggesting some bias in retrospectively reported data. However, MZ and DZ correlations were similar regardless of reporting time, thus differences between twins can still be observed in both retrospective and concurrently reported data. Table 3 shows the number of twin pairs for whom questionnaire data are held at each grade level and the proportion that is retrospective.

Representativeness of the Sample

The highest level of education of the mother and father of the twins was recorded using a nine-point scale. Table 4 shows parent education levels and the percentage of responses at each level. In this sample, mean education level for mothers was 4.92 (*SD* 1.88), and for fathers, 4.54 (*SD* 1.98). A higher proportion of mothers (81.6%) had completed postschool qualifications than fathers (75.5%). Fathers (35.7%) were more likely than mothers (28.2%) to have completed a trade diploma or certificate. However, a higher proportion of mothers (53.3%) than fathers (39.7%) had completed a 3-year university degree or above.

Australian Bureau of Statistics (ABS) data from the 2012 census indicate that of the Australian population aged between 25 and 54 years, 62% of females and 64% of males hold postschool qualifications (ABS, 2019). In the same age bracket, 34% of females and 29% of males have attained a 3-year university degree or above. These figures indicate that the parents of the participants in this study have higher levels of educational attainment than the Australian population, with 19.6% more mothers and 11.5% more fathers holding postschool qualifications compared to the general population of similarly aged women and men.

Parents were also asked to report their current occupations, and 99% of mothers and 97% of fathers provided this information. Occupations were subsequently coded using the International Socio-Economic Index (ISEI) of occupational status (Ganzeboom, 2010), which ranks occupational prestige on a standardized scale of 10–90. Table 4 shows means and standard deviations of occupational rating for mothers and fathers, and percentages of parents with occupations in each quartile. Australian data from the Programme for International Student Assessment tests (Lokan et al., 2001), which includes the higher coded occupational status of students' mother or father as a measure of SES, indicate mean ISEI for parents of Australian students is 52, lower than that reported by highest coded parent in this study ($M = 59.5$, $SD = 13.3$). A small percentage of fathers indicated they were unemployed (1.5%), and 336 mothers (12.2%) indicated their occupation as 'full-time mother' or 'stay-at-home mother'. These responses were coded at the lowest two unused categories of the ISEI, '8' and '9', respectively. Excluding the 'full-time mother' category, occupational status and education level were moderately correlated for both mothers ($r = 0.53$) and fathers ($r = 0.55$), which aligns with data used in the development of the ISEI (Ganzeboom, 2010).

Ancestry of the mother and father of the twins was recorded and subsequently coded into ancestry categories using the Australian Standard Classification of Cultural and Ethnic Groups (ABS, 2016). Parents were predominantly of European ancestries with 96% of mothers and 95% of fathers listing some European ancestry. Breakdown of ancestry by parent can be seen in Table 4. A comparison of these figures with national Australian data indicate that the sample had a higher proportion of participants identifying Australian ancestry than the general population (33.5%), and a lower proportion of both Asian descent participants (5.6% of Australians identify Chinese ancestry alone), and Indigenous Australian participants (2.8% of the population; ABS, 2017).

Table 4. Demographic characteristics of mothers and fathers

	Mother	Father
Education completed (%)		
1 — Some high school	3.6	7.6
2 — School certificate (Grade 10)	6.6	7.1
3 — Higher School Certificate (Grade 12)	8.2	9.8
4 — TAFE or Trade (certificate/diploma)	28.2	35.7
5 — 3-year university degree	14.3	10.9
6 — 4-year university degree	13.2	10.1
7 — Some postgraduate study	16.3	8.8
8 — Master's degree	7.6	7.3
9 — Doctoral degree	1.9	2.6
N	2753	2712
Mean (*SD*)	4.92 (1.88)	4.54 (1.98)
International Socio-Economic Index of (ISEI[a]) occupational status (%)		
First quartile[b]	4.2	7.0
Second quartile	37.6	33.9
Third quartile	34.1	45.6
Fouth quartile	11.8	11.4
Stay-at-home parent	12.2	0.6
Unemployed	0.1	1.5
N	2748	2679
Mean (*SD*)	53.01 (13.98)	54.01 (15.72)
Ancestry		
Australian	49.0	47.0
UK	35.0	34.0
Other European	11.7	14.1
Asian	2.2	1.9
Indigenous Australian	0.7	1.0
Other	1.4	2.0
N	2712	2684

Note: [a]Ganzeboom (2010); [b]Quartiles are defined by this project as 1 = unskilled or semi-skilled laboring, retail, agricultural occupations (codes 10–30); 2 = Trades or skilled laboring, retail or entry-level administration occupations (codes 31–50); 3 = associate professionals, teachers, managerial occupations (codes 51–70); 4 = highly skilled professional occupations (codes 71–90).

Upon entry to the study, most twins (82%) lived with both their biological mother and father, while 11% lived with a single mother, and a further 4% lived with their biological mother and a nonbiological father. Australian census data for 2012–2013 shows that of families with children aged 0–17 years, 81% were couple families and 19% were single-parent families, predominantly single mothers (ABS, 2015). For families whose youngest child was aged 5–9 years, 72% lived with both biological parents. This rate reduced as the youngest child aged so that at age 15–17 years only 60% resided with both biological parents (ABS, 2015). The sample recruited for this research thus has a lower proportion of single-mother families, and a higher proportion of intact families with both biological parents resident with children than the general population. Census data show that families in Australia have an average of 1.8 children (ABS, 2017). Obviously having twins means that twin families have slightly more children than average. Families in this sample also had an average of 1.6 siblings, with 43% of families including one sibling, 27% of twins with no siblings and the remaining 30% with two or more siblings.

Parents who did not return any follow-up questionnaires tended to have slightly lower education levels (mothers, $M = 4.69$, $SD = 0.09$; fathers, $M = 4.12$, $SD = 0.04$) than those who responded (mothers, $M = 5.06$, $SD = 0.04$; fathers, $M = 4.62$, $SD = 0.04$). Likewise, nonresponding parents had slightly lower occupational prestige on the ISEI scale (mothers, $M = 50.71$, $SD = 13.94$; fathers $M = 50.15$, $SD = 16.15$) than responding parents (mothers. $M = 53.49$, $SD = 13.94$; fathers, $M = 54.81$, $SD = 15.51$). On other demographics, nonresponding families were similar to responding families, including whether mothers identified as ‘stay-at-home mothers’ (12% in each group), number of siblings, and whether twins were identical or fraternal.

NAPLAN: Literacy and Numeracy Standardized Tests

The key measurement of educational achievement in ADSAT is scores on the NAPLAN tests. The Australian Federal Government introduced the NAPLAN in 2008 as a nationwide standardized testing program designed to assess the literacy and numeracy capabilities of students in Grades 3, 5, 7 and 9. The tests are carried out in May of each year, approximately 3 weeks into the second school term; 94%–97% of students in Grades 3, 5 and 7, and 91%–94% of students in Grade 9 participate in the tests each year (Australian Curriculum, Assessment and Reporting Authority ACARA, [2017]). NAPLAN tests are aligned with the Australian National Curriculum and are designed to assess basic skills of students expected at each grade level (ACARA, 2016).

NAPLAN tests include one numeracy component, and four literacy components: reading, writing, spelling and grammar. In Grades 7 and 9, the numeracy component incorporates two subtests, with one allowing the use of calculators. Reading, spelling, grammar and numeracy tests are a combination of multiple-choice items and short responses, which are scored correct or incorrect. In the writing test, students must respond to a prompt and write a persuasive or narrative text, which is subsequently graded using prescriptive marking criteria. All writing tasks are double-marked. Example tests and past papers for the 2008–2011 tests are publicly available (ACARA, 2016).

Each NAPLAN test score is translated to a standardized scale score ranging from 1 to 1000 and equating to an achievement band between 1 and 10. An expected minimum achievement band is set for each grade level, for example, the national minimum standard for all Grade 3 test domains is band 2, increasing to band 4 for Grade 5, band 5 for Grade 7 and band 6 for Grade 9 (ACARA, 2017). A procedure of equating between year cohorts and across grade levels is carried out each year to ensure that scores are comparable between different cohorts and across the four testing grades (ACARA, 2014). For example, achieving at band 6 in Grade 3 reading in 2014 has the same meaning as achieving at band 6 in Grade 3 reading in 2017; and achieving at band 7 in Grade 5 numeracy has the same meaning as achieving at band 7 in Grade 7 numeracy. In this way, student achievement can be tracked across time, and different cohorts at each grade level can be compared.

After being granted consent by state and territory education departments, project staff requested NAPLAN scale scores in each of the five domains. Six of the seven jurisdictions provided NAPLAN data with a match rate of 91% for Grade 3 participants reducing to 88% for Grade 9 participants. Data were not available for 40 pairs in the Northern Territory. Additionally, 71 families (3%) failed to provide school- and grade-level information, even after follow-up from the researchers; therefore, it was not possible to collect NAPLAN data for these participants.

Inevitably, there is an amount of missing NAPLAN data due to students withdrawing or being absent for the tests, families moving interstate or overseas, or students changing schools. The main determinant of missing NAPLAN data, however, is the recruitment procedure. Because the project recruited students who had completed NAPLAN tests at any grade beginning in 2008, the year NAPLAN testing began, there are students who only sat NAPLAN in one or two grades (i.e., they were in Grade 9 when NAPLAN was introduced in 2008 or they were in Grade 3 at the study's most recent recruitment in 2017). Participants with data at Grade 3, who continue to participate longitudinally, will eventually have four time points of NAPLAN data (see Table 1). Table 5 shows the total number of individuals for whom NAPLAN data are currently held at each cross-sectional grade level. These numbers will increase slightly over the coming years as remaining participants move toward completing Grade 9.

Table 5 also shows means and standard deviations by NAPLAN grade and domain for participants with data at any grade from 2008 to 2018, compared with publicly available national data (ACARA, 2020). While the mean scores in each domain are higher for this sample than the national data, the standard deviations are comparable. It should be noted that we were not able to gain access to data from participants in the Northern Territory, the jurisdiction that consistently reports the lowest mean performance in NAPLAN tests. Higher mean performance might also be expected because of the opt-in nature of study recruitment. Socioeconomically advantaged participants are more often recruited into survey-based research (Gustavson et al., 2012) and there is a clear association between NAPLAN performance and SES in these data, in population statistics, and in other projects that collect NAPLAN results (e.g., Marks, 2016). Interestingly, NAPLAN means and standard deviations in this sample are also similar to those collected in the Longitudinal Study of Australian Children, which recruited two representative cohorts of approximately 5000 children in 2004 (Australian Government Department of Social Services, 2019). With comparable standard deviations and mean scores (on average) 0.38 of a standard deviation higher in the ADSAT sample than in the national data, it follows there is less sampling in ADSAT from the national distribution of those who have lower scores. While important to acknowledge this limitation in the sample distribution, the ADSAT data do have considerable variation among the lower

https://doi.org/10.1017/thg.2020.49 Published online by Cambridge University Press

Table 5. NAPLAN means and standard deviations by grade and domain for ADSAT participants compared with national data 2008–2018

	ADSAT		National
	N[a]	*M* (*SD*)	*M* (*SD*)
Reading			
Grade 9	4247	616.2 (66.2)	579.4 (66.1)
Grade 7	4779	577.2 (68.9)	542.3 (68.0)
Grade 5	4872	529.3 (78.6)	496.8 (75.9)
Grade 3	4362	452.5 (87.2)	419.5 (85.4)
Numeracy			
Grade 9	4213	622.1 (71.2)	587.6 (70.0)
Grade 7	4770	575.7 (72.0)	545.6 (71.6)
Grade 5	4868	513.8 (71.2)	488.6 (68.7)
Grade 3	4356	425.3 (74.2)	399.6 (71.8)
Writing			
Grade 9	4248	587.9 (86.1)	556.4 (83.8)
Grade 7	4780	544.8 (71.0)	519.9 (74.8)
Grade 5	4874	498.0 (64.3)	477.6 (68.2)
Grade 3	4362	433.8 (62.0)	414.0 (67.9)
Spelling			
Grade 9	4253	601.5 (68.1)	580.3 (72.6)
Grade 7	4789	562.3 (66.8)	543.9 (72.0)
Grade 5	4880	507.9 (68.7)	493.0 (73.0)
Grade 3	4366	427.4 (77.1)	409.9 (83.2)
Grammar and punctuation			
Grade 9	4253	606.8 (76.0)	573.2 (70.1)
Grade 7	4789	572.6 (77.7)	538.9 (73.4)
Grade 5	4881	531.3 (83.9)	500.1 (79.0)
Grade 3	4361	457.5 (93.3)	425.4 (89.9)

Note: [a]Cross-sectional numbers of individuals (i.e., twins, siblings, triplets) with NAPLAN data 2008–2018.

scores, with 22% of ADSAT data being sampled from the lowest third of national scores.

Summary of Key Findings

In line with the initial aim of the project, Grasby et al. (2016) explored the relative influence of genes and the environment on NAPLAN scores at each of the grade levels and reported similar findings to those in international samples (e.g., Davis et al., 2009; Hart et al., 2009). Specifically, heritability of achievement in all five NAPLAN domains was moderately high, ranging between 0.39 and 0.79, although shared environment contributed to a small but significant portion of variance in most domains, between 0.02 and 0.19. Consistently, high genetic correlations between domains were reported, lending further support to the generalist genes hypothesis put forward by Kovas and Plomin (2007). Nonetheless, Grasby et al. also reported a portion of genetic influence on numeracy domains that was independent of genetic influence on literacy domains. High shared environment correlations were also reported among most of the domains at each grade level, leading to the conclusion that shared environmental factors influencing performance in different domains are fairly constant — presumably due in part to consistency across schools, teachers and the curriculum in Australia.

In a follow-up study, Grasby and Coventry (2016) investigated genetic and environmental influences on stability and growth in literacy and numeracy, finding that performance in NAPLAN tests is highly stable across time, and that genes mediate most of this stability. For reading, variation in growth was predominantly influenced by genetic factors; however, in the other literacy domains, variation in growth was principally influenced by the shared environment. Variation in growth in numeracy differed between girls and boys, with girls' growth influenced by both genes and the shared environment while boys' growth was significantly influenced solely by shared environmental factors.

Two further studies arising from this project have employed gene-by-environment interaction designs to explore the moderating effects of measured shared environments. Grasby, Coventry et al. (2019) investigated whether SES moderated the heritability of academic achievement. In contrast to results from the USA (see Tucker-Drob & Bates, 2016), their results demonstrated little evidence that SES is a moderator of the heritability of academic achievement in this Australian sample. Similarly, Gould et al. (2018) showed that neither SES nor a measure of home disorder moderated the heritability of ADHD symptoms. It should be noted that while these results provide a counterpoint to international research, both studies should be considered preliminary: both were underpowered to detect small moderation effects and both will be replicated when larger sample sizes are available.

Another two studies have used behavior genetic methods to investigate environmental features widely considered to impact on educational attainment: preschool attendance and classroom effects. Using a novel quantile regression twin design, Little et al. (under review) investigated whether the preschool attendance moderated the heritability of reading, writing and numeracy skills across the quantiles of ability. Preschool attendance did not moderate the heritability of these skills, and contrary to widely held beliefs that preschool attendance leads to higher academic achievement, preschool attendance was not associated with later achievement in NAPLAN. This is the first application of quantile regression in an examination of gene-by-environment interactions across the distribution of ability, and the manuscript details the procedure for future researchers interested in exploring gene–environment interplay. These results provide a counterpoint to existing research, which suggests that preschool attendance causally affects NAPLAN achievement (Warren & Haisken-DeNew, 2013).

In a follow-up to a finding of the ILTS (Byrne et al., 2010), Grasby, Little et al. (2019) explored the impact of classroom assignment on the variance of NAPLAN scores. Again, contrary to the widely held idea that classroom factors, usually conceptualized as teacher effects, contribute a substantial amount of variance in school performance (e.g., Hattie, 2008), this study revealed small and mostly nonsignificant portions of variance in NAPLAN achievement that could be attributed to classroom factors.

Finally, in an exploration of the nonshared environment, a qualitative study followed up families of identical twins notably and consistently discordant for achievement in reading, writing or numeracy (Larsen et al., 2019). Interviews with the parents of such twins (approximately 5% of the total sample) revealed three categories of explanations for extreme and ongoing discordance in academic achievement between genetically identical pairs. These included discordance in biomedical conditions between twins, discordant school experiences and personality differences. While these results are preliminary in terms of identifying causal factors at work in differentiating identical twins, they do mirror the results of the one existing study using a similar design (Asbury et al., 2016).

Future Plans

ADSAT will continue to follow up existing cohorts of twins until the final cohort completes Grade 9 in 2023. Once questionnaire data and NAPLAN results have been obtained for these cohorts, the database will comprise 10 sequential cohorts with data at 4 biennial time points, in addition to 6 cohorts with data at between 1 and 3 instances (see Table 1). Plans are underway for further follow-up with twins who have graduated from secondary school in order to obtain data on school completion, final academic grades and early adult life information. Alongside the potential for further behavior genetic studies, the data collected in this project provide an opportunity for a range of phenotypic investigations of school achievement in Australia. Indeed, some work has already begun to this end with an investigation of the associations between dietary patterns and NAPLAN achievement (Burrows et al., 2017), and a study exploring whether delaying the school entry of children was associated with higher NAPLAN achievement in all four grades (Larsen et al., in press).

Another future aim of this project is to investigate the mechanisms through which disadvantaged environments are contributing to poor achievement for Australian students. The overall objective is to use geospatial analyses to identify potential environmental 'levers' that can be adjusted to improve educational outcomes. To achieve this objective, twins addresses have been geo-located, and will be matched with publicly available census data, along with other geocoded features of home and school neighborhoods (e.g., proximity to resources) to identify potential protective or risk factors within the shared environment that are associated with school achievement.

A database of the size and scope such as the one reported here provides a wealth of opportunities for future investigations into the predictors of academic achievement from Grades 3 to 9 and beyond. There are no similar existing twin datasets in Australia, thus these data have the capacity both to test findings made in other educational jurisdictions and to explore features specific to the Australian schooling environment. Despite the reservations about the NAPLAN testing program articulated by much public commentary (see Bahr & Pendergast, 2018, as one example), the standardized tests provide broad comparative data on the progress of Australian school students in literacy and numeracy domains. Indeed, Grasby et al. (2015) reported high genetic correlations between NAPLAN reading tests and 'gold standard' literacy tests often used in twin projects of academic achievement, evidence that goes some way toward confirming that NAPLAN tests appropriately assess the skills they are intended to assess. Combining NAPLAN results data at four school grades with the range of questionnaire data collected at multiple time points presents unique opportunities to both test existing research findings and provide new insights into genetic and environmental contributors to academic skills.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/thg.2020.49.

Acknowledgments. We would like to acknowledge the twins and their families who have donated their time to contribute to this project over many years.

Funding. This research was supported by two Australian Research Council Discovery Project Grants: DP 120102414 (2012–2014) and DP 150102441 (2015–2018). Access to the sample was facilitated by Twins Research Australia, a national resource supported by a Centre of Research Excellence Grant (ID: 1079102), from the National Health and Medical Research Council.

References

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

Asbury, K., Moran, N., & Plomin, R. (2016). Nonshared environmental influences on academic achievement at age 16: A qualitative hypothesis-generating monozygotic-twin differences study. *AERA Open*, *2*, 112.

Asbury, K., & Wai, J. (2019). Viewing education policy through a genetic lens. *Journal of School Choice.* doi: 10.1080/15582159.2019.1705008

Australian Bureau of Statistics (ABS). (2015). Family characteristics and transitions, Australia, 2012–2013. Retrieved from https://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/E6A9286119FA0A85CA25699000255C89?opendocument

Australian Bureau of Statistics (ABS). (2016). Australian Standard Classification of Cultural and Ethnic Groups (ASCCEG), 2016. Retrieved from https://www.abs.gov.au/AUSSTATS/abs@.Nsf/7d12b0f6763c78caca257061001cc588/964d7ded6ce6f93aca2578ed0021374b!OpenDocument

Australian Bureau of Statistics (ABS). (2017). *2016 Census QuickStats*. Retrieved from https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/quickstat/036

Australian Bureau of Statistics (ABS). (2019). Education and Work Australia, May 2018. Retrieved from https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/6227.0

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2014). National Assessment Program — Literacy and Numeracy 2013: Technical report. Retrieved from https://www.nap.edu.au/_resources/NAPLAN_2013_technical_report.pdf

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2017). National Assessment Program Literacy and Numeracy. Achievement in Reading, Writing, Language Conventions and Numeracy: National report for 2017. Retrieved from https://www.nap.edu.au/docs/default-source/default-document-library/naplan-national-report-2017_final_04dec2017.pdf?sfvrsn=0

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2016). National Assessment Program: The tests. Retrieved from https://www.nap.edu.au/naplan/the-tests

Australian Curriculum, Assessment and Reporting Authority (ACARA). (2020). National Assessment Program: Results. Retrieved from https://reports.acara.edu.au/Home/Results

Australian Government: Department of Social Services. (2019). Growing up in Australia: The Longitudinal Study of Australian Children (LSAC). Retrieved from https://www.dss.gov.au/about-the-department/publications-articles/research-publications/longitudinal-data-initiatives/footprints-in-time-the-longitudinal-study-of-indigenous-children-lsic/growing-up-in-australia-the-longitudinal-study-of-australian-children-lsac

Bahr, N. & Pendergast, D. (2018). Let's abandon NAPLAN — We can do better. *The Conversation*. Retrieved from https://theconversation.com/lets-abandon-naplan-we-can-do-better-95363

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

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, *27*, 281–291.

Buckingham, J., Beaman, R., & Wheldall, K. (2014). Why poor children are more likely to become poor readers: The early years. *Educational Review*, *66*, 428–446.

Burrows, T., Goldman, S., Olson, R. K., Byrne, B., & Coventry, W. L. (2017). Associations between selected dietary behaviours and academic achievement: A study of Australian school aged children. *Appetite*, *116*, 372–380.

Byrne, B., Coventry, W. L., Olson, R. K., Wadsworth, S. J., Samuelsson, S., Petrill, S.A., & Corley, R. (2010). 'Teacher effects' in early literacy development: Evidence from a study of twins. *Journal of Educational Psychology*, *102*, 32–42.

Byrne, B., Olson, R. K., & Samuelsson, S. (2019) Behavior-genetic studies of academic performance in school students: A commentary for professionals in psychology and education. In D. Kilpatrick, R. M. Joshi & R. K. Wagner (Eds.), *Reading development and difficulties: Bridging the gap between research and practice* (pp. 213–232). Basel, Switzerland: Springer International Publishing.

Byrne, B., Wadsworth, S. J., Boehme, K., Talk, A. C., Coventry, W. L., Olson, R. K., & Corley, R. (2013). Multivariate genetic analysis of learning and early reading development. *Scientific Studies of Reading*, *17*, 224–242.

Davis, O. S. P., Haworth, C. M. A., & Plomin, R. (2009). Learning abilities and disabilities: Generalist genes in early adolescence. *Cognitive Neuropsychiatry*, *14*, 312–331.

Daucourt, M. C., Erbeli, F., Little, C. W., Haughbrook, R., & Hart, S. A. (2020). A meta-analytical review of the genetic and environmental correlations between reading and attention-deficit/hyperactivity disorder symptoms and reading and math. *Scientific Studies of Reading*, *24*, 23–56.

Ganzeboom, H. B. G. (2010, May). *A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with data from the ISSP 2002–2007*. Paper presented at the Annual Conference of International Social Survey Programme, Lisbon.

Gould, K. L., Coventry, W. L., Olson, R. K., & Byrne, B. (2018). Gene-environment interactions in ADHD: The roles of SES and chaos. *Journal of Abnormal Child Psychology*, *46*, 251–263.

Grasby, K. L., Byrne, B., & Olson, R. K. (2015). Validity of large-scale reading tests: A phenotypic and behaviour–genetic analysis. *Australian Journal of Education*, *59*, 5–21.

Grasby, K. L., & Coventry, W. L. (2016). Longitudinal stability and growth in literacy and numeracy in Australian school students. *Behavior Genetics*, *46*, 649–664.

Grasby, K. L., Coventry, W. L., Byrne, B., & Olson, R. K. (2019). Little evidence that socioeconomic status modifies heritability of literacy and numeracy in Australia. *Child Development*, *90*, 623–637.

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*, *46*, 627–648.

Grasby, K. L., Little, C. W., Byrne, B., Coventry, W. L., Olson, R. K., Larsen, S.A., & Samuelsson, S. (2019). Estimating classroom-level influences on literacy and numeracy: A twin study. *Journal of Educational Psychology*. doi: 10.1037/edu0000418

Gustavson, K., von Soest, T., Karevold, E., & Røysamb, E. (2012). Attrition and generalizability in longitudinal studies: Findings from a 15-year population-based study and a Monte Carlo simulation study. *BMC Public Health*, *12*, 918.

Harlaar, N., Dale, P. S., & Plomin, R. (2007). From learning to read to reading to learn: Substantial and stable genetic influence. *Child Development*, *78*, 116–131.

Harlaar, N., Cutting, L., Deater-Deckard, K., DeThorne, L. S., Justice, L. M., Schatschneider, C., & Petrill, S. A. (2010). Predicting individual differences in reading comprehension: A twin study. *Annals of Dyslexia*, *60*, 265–288.

Harlaar, N., Kovas, Y., Dale, P. S., Petrill, S. A., & Plomin, R. (2012). Mathematics is differentially related to reading comprehension and word decoding: Evidence from a genetically-sensitive design. *Journal of Educational Psychology*, *104,* 622–635.

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*, 388–402.

Hattie, J. (2008). *Visible learning: A synthesis of meta-analyses relating to achievement.* New York: Routledge.

Haughbrook, R., Hart, S. A., Schatschneider, C., & Taylor, J. (2017). Genetic and environmental influences on early literacy skills across school grade contexts. *Developmental Science*, *20*, e12434.

Kovas, Y., Malykh, S., & Gaysina, D. (Eds.). (2016). *Behavioural genetics for education*. New York: Palgrave Macmillan

Kovas, Y., & Plomin, R. (2007). Learning abilities and disabilities: Generalist genes, specialist environments. *Current Directions in Psychological Science*, *16*, 284–288.

Larsen, S. A., Byrne, B., Little, C. W., Coventry, W. L., Ho, C. S., Olson, R. K., & Stevenson, A. (2019). Identical genes, unique environments: A qualitative exploration of persistent monozygotic-twin discordance in literacy and numeracy. *Frontiers in Education*, *4*. doi: 10.3389/feduc.2019.00021

Larsen, S. A., Little, C. W., & Coventry, W. L. (in press). Exploring the associations between delayed school entry and achievement in primary and secondary school. *Child Development.*

Little, C. W., Haughbrook, R., & Hart, S. A. (2017). Cross-study differences in the etiology of reading comprehension: A meta-analytical review of twin studies. *Behavior Genetics*, *47*, 52–76.

Little, C. W., Larsen S. A., Coventry, W. L., Byrne, B., & Olson, R. (under review). *Exploring the influence of early childhood education and care on the etiology of achievement.*

Logan, J. A. R., Hart, S. A., Cutting, L., Deater-Deckard, K., Schatschneider, C., & Petrill, S. (2013). Reading development in young children: Genetic and environmental influences. *Child Development*, *84*, 2131–2144.

Lokan, J., Greenwood, L., & Cresswell, J. (2001). *15-up and counting, reading, writing, reasoning: How literate are Australia's students? The PISA 2000 Survey of Students' Reading, Mathematical and Scientific Literacy Skills.* Australian Council for Educational Research. Retrieved from https://research.acer.edu.au/indigenous_education/7/

Lykken, D. T., Bouchard, T. J., Jr., McGue, M., & Tellegen, A. (1990). The minnesota twin family registry: Some initial findings. *Acta Geneticae Medicae Gemellologiae*, *39*, 35–70.

Marks, G. N. (2015a.). Do Catholic and Independent schools 'add-value' to students' Tertiary Entrance Performance? Evidence from longitudinal population data. *Australian Journal of Education*, *59*, 133–157.

Marks, G. N. (2015b). School sector differences in student achievement in Australian primary and secondary schools: A longitudinal analysis. *Journal of School Choice*, *9*, 219–238.

Marks, G. N. (2016). The relative effects of socio-economic, demographic, non-cognitive and cognitive influences on student achievement in Australia. *Learning and Individual Differences*, *49*, 1–10.

Organisation for Economic Co-Operation and Development (OECD). (2009). *Comparative child well-being across the OECD: Doing better for children.* Retrieved from https://www.oecd.org/social/family/43570328.pdf

Perry, L. B., & McConney, A. (2013). School socioeconomic status and student outcomes in reading and mathematics: A comparison of Australia and Canada. *Australian Journal of Education*, *57*, 124–140.

Plomin, R., & Daniels, D. (2011). Why are children in the same family so different from one another? *International Journal of Epidemiology*, *40*, 563–582.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science*, *11*, 3–23.

Samuelsson, S., Byrne, B., Olson, R. K., Hulslander, J., Wadsworth, S., Corley, R., DeFries, J. C. (2008). Response to early literacy instruction in the United States, Australia, and Scandinavia: A behavioral-genetic analysis. *Learning and Individual Differences*, *18*, 289–295.

Smith, C., Parr, N., & Muhidin, S. (2019). Mapping schools' NAPLAN results: A spatial inequality of school outcomes in Australia. *Geographical Research*, *57*, 133–150.

Thomas, M. S. C., Kovas, Y., Meaburn, E. L., & Tolmie, A. (2015). What can the study of genetics offer to educators? *Mind, Brain, and Education*, *9*, 72–80.

Tikhodeyev, O. N., & Shcherbakova O. V. (2019). The problem of non-shared environment in behavioral genetics. *Behavior Genetics*, *49*, 259–269.

Tucker-Drob, E. M., & Bates, T. C. (2016). Large cross-national differences in gene × socioeconomic status interaction on intelligence. *Psychological Science*, *27*, 138–149.

Twins Research Australia. (n.d.). *Twins Research Australia: Who we* are. Retrieved from https://www.twins.org.au/about-us/who-we-are

Warren, D., & Haisken-Denew, J. (2013). *Early bird catches the worm: The causal impact of pre-school participation and teacher qualifications on Year 3 National NAPLAN Cognitive Tests* (*IDEAS Working Paper Series from RePEc*). Retrieved from https://ideas.repec.org/p/iae/iaewps/wp2013n34.html

Wright, M. J., & Martin, N. G. (2004). Brisbane adolescent twin study: Outline of study methods and research projects. *Australian Journal of Psychology*, *56*, 65–78.

https://doi.org/10.1017/thg.2020.49 Published online by Cambridge University Press