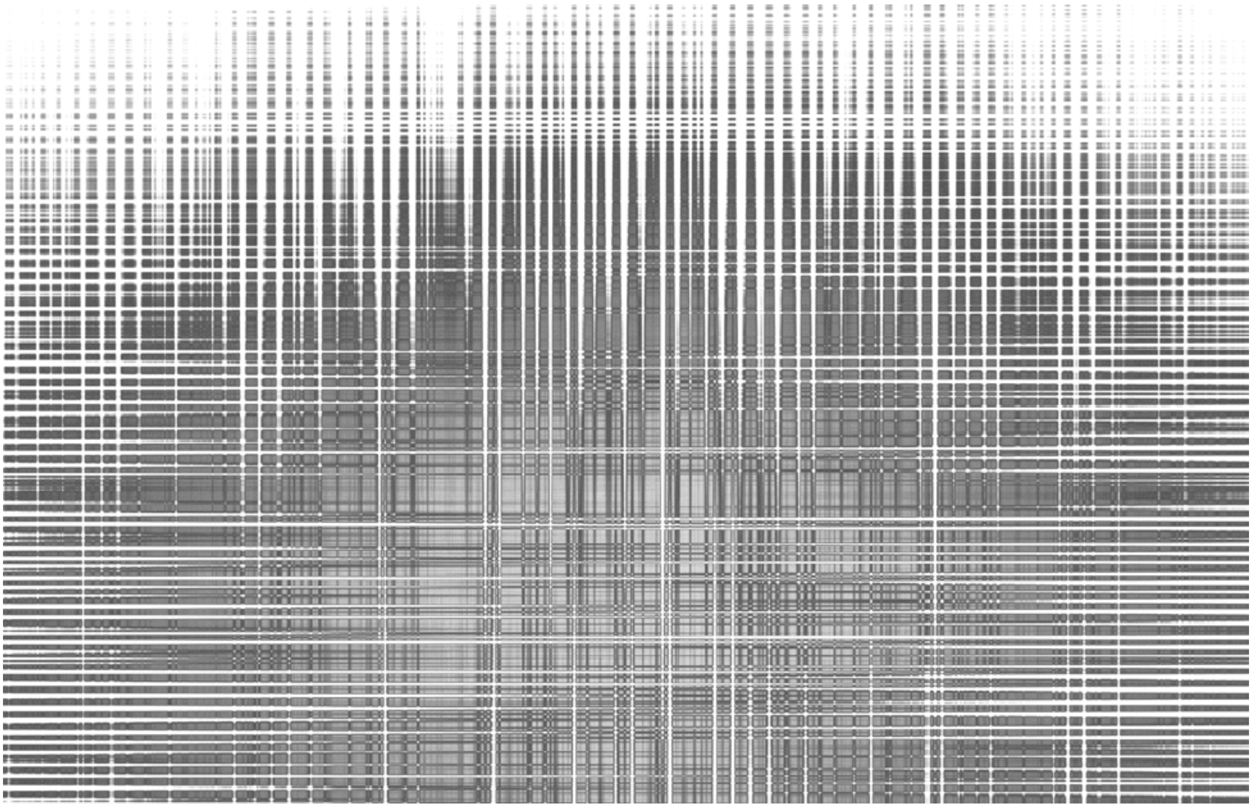
Australian National University

CENTRE FOR SOCIAL RESEARCH & METHODS

Social Research Centre

A SUBSIDIARY OF THE AUSTRALIAN NATIONAL UNVERSITY



Longitudinal Study of Indigenous Children (LSIC) technical report: education

N Biddle, B Edwards, R Lovett, P Radoll, K Sollis and K Thurber

CSRM & SRC METHODS PAPER

NO. 5/2019

Series note

The ANU Centre for Social Research & Methods (CSRM) was established in 2015 to provide national leadership in the study of Australian society. CSRM has a strategic focus on:

- development of social research methods
- analysis of social issues and policy
- training in social science methods
- providing access to social scientific data.

CSRM publications are produced to enable widespread discussion and comment, and are available for free download from the CSRM website (http://csrm.cass.anu.edu.au/research/publications).

CSRM is located within the Research School of Social Sciences in the College of Arts & Social Sciences at the Australian National University (ANU). The centre is a joint initiative between the Social Research Centre and the ANU. Its expertise includes quantitative, qualitative and experimental research methodologies; public opinion and behaviour measurement; survey design; data collection and analysis; data archiving and management; and professional education in social research methods.

As with all CSRM publications, the views expressed in this Methods Paper are those of the authors and do not reflect any official CSRM position.

Professor Matthew Gray
Director, ANU Centre for Social Research & Methods
Research School of Social Sciences
College of Arts & Social Sciences
The Australian National University
November 2019

Methods Paper No. 5/2019

ISSN 2209-184X
ISBN 978-1-925715-18-7

An electronic publication downloaded from http://csrm.cass.anu.edu.au/research/publications.

For a complete list of CSRM Methods Papers, see http://csrm.cass.anu.edu.au/research/publications/methods-research-papers.

ANU Centre for Social Research & Methods
Research School of Social Sciences
The Australian National University

Longitudinal Study of Indigenous Children (LSIC) technical report: education

N Biddle, B Edwards, R Lovett, P Radoll, K Sollis and K Thurber

Nicholas Biddle is Professor of Economics and Public Policy, Associate Director of the ANU Centre for Social Research & Methods, and Fellow of the Tax and Transfer Policy Institute.

Ben Edwards is Associate Professor in Child and Youth Development and Longitudinal Studies at the ANU Centre for Social Research & Methods.

Raymond Lovett is a Senior Research Fellow with the Epidemiology for Policy and Practice group at the National Centre for Epidemiology and Population Health.

Peter Radoll is Professor of Information Technology and Dean, Aboriginal and Torres Strait Islander Leadership and Strategy at the University of Canberra; and Director of the Ngunnawal Centre.

Katherine Sollis is a Research Officer and PhD candidate at the ANU Centre for Social Research & Methods.

Katerine Thurber is an NHMRC Early Career Fellow (2019–23) with the Aboriginal and Torres Strait Islander Health Program at the National Centre for Epidemiology and Population Health, ANU Research School of Population Health.

Abstract

This report evaluates the education measures in the Longitudinal Study of Indigenous Children (LSIC). Education measures in the LSIC were found to be internally valid and perform as expected. The LSIC is a robust dataset that, if used carefully, can improve our understanding of the development of Indigenous children, and help design good public policy. For analysts, we recommend using the data with confidence, while remaining aware that some variables perform better than others and that models using the education measures (especially those specific to the LSIC) tend to have low explanatory power. We also recommend taking advantage of the longitudinal data rather than the cross-sectional data. For reviewers of papers based on LSIC data, we recommend taking into account the unique circumstances of the survey, and that models will be estimated with low precision and with variables that differ from those collected in other datasets. Finally, for policy makers, we recommend making decisions using longitudinal research and considering funding a top-up sample.

Acknowledgments

We highly commend the contributions of the Indigenous children, their families, the interviewers and field staff, the Indigenous researchers and other stakeholders in the ongoing success of the LSIC. We would like to thank the Australian Government Department of Social Services for the opportunity to review the education measures, and for the collaborative approach they took to the project. All opinions are those of the authors, and should be attributed as such.

Acronyms and abbreviations

AEDC	Australian Early Development Census
ANU	Australian National University
CFI	comparative fit index
CSRM	ANU Centre for Social Research & Methods
DSS	Australian Government Department of Social Services
LSIC	Londitudinal Study of Indigenous Children
NAPLAN	National Assessment Program – Literacy and Numeracy
NHMRC	National Health and Medical Research Council
P1	parent 1
PSSM	Psychological Sense of School Membership
RMSEA	root-mean-square error of approximation
SDQ	Strengths and Difficulties Questionnaire
SLAQ	School Liking and Avoidance Questionnaire
SRMR	standardised root mean residual
STRS	Pianta Student-Teacher Relationship Scale
TLI	Tucker–Lewis index

Contents

Tables and figures

Executive summary

Globally, Australia's Longitudinal Study of Indigenous Children (LSIC) is the only longitudinal child cohort study on the developmental outcomes of Indigenous children. The study surveys Aboriginal and Torres Strait Islander Australian children aged either 6–18 months (B cohort) or 3.5–5 years (K cohort) when the study began in 2008. This report evaluates the education measures in the LSIC and contains recommendations for data collectors and analysts, researchers, and policy makers.

Education measures in the LSIC are internally valid and perform as expected. The LSIC is a robust dataset that, if used carefully, can improve our understanding of the development of Indigenous children, and help design good public policy. We highly commend the contributions of the Indigenous children, their families, the interviewers and field staff, the Indigenous researchers, and other stakeholders to the ongoing success of LSIC.

Some measures perform better than others. In particular, the school climate measures and student-rated teacher relationship showed high correlation. However, measures of academic self-concept were not related to outcomes. Most academic measures showed only small correlations with student outcomes, except the Student–Teacher Relationship Scale, which showed the strongest and most consistent relationships with outcomes. The Student–Teacher Relationship Scale is one of the few education measures consistently collected over the waves, and may be particularly important in understanding Indigenous children's learning.

The education measures had similar patterns of associations for children living in remote areas, but this was not evident for school affective disengagement, as measured with the School Liking and Avoidance Questionnaire (SLAQ), for children in remote areas. For children in remote areas, associations between SLAQ and other education measures were lower and few were statistically significant, including those for child outcomes. The pattern of correlations was similar for the B and K cohorts, and suggests that affective disengagement reported by study children has a qualitatively different meaning for children in remote areas. Data users should exercise caution if using SLAQ for research that includes children from remote areas.

One limitation of the LSIC identified in this paper is that few variables appear to explain the variation in the education measures. When undertaking multivariate analysis of the LSIC, researchers may find constructing models with high explanatory power difficult. We recommend including data items (directly or through data linkage) that can be used to understand variation in education measures. For data collectors, we recommend asking fewer questions but asking them consistently, and being careful and explicit when attempting to balance specificity and generalisability.

For analysts, we recommend using the data with confidence, while remaining aware that some variables perform better than others and that models using the education measures (especially those specific to the LSIC) tend to have low explanatory power. We also recommend taking advantage of the longitudinal data rather than the the cross-sectional data. We showed a few variables that were significantly associated with change over time in the National Assessment Program – Literary and Numeracy (NAPLAN) (in particular, housing circumstances), and a consistent association between parent-reported health and school attendance. Worse health at a given time was associated with a lower probability of attending school every day in the previous week.

For reviewers of papers based on LSIC data, we recommend taking into account the unique circumstances of the survey and that models will be estimated with low precision and with variables that differ from those collected in other datasets. Finally, for policy makers, we recommend making decisions using longitudinal research and considering funding a top-up sample. We also suggest that a dedicated analytical hub is established in a research institution to increase the visibility and use of the data.

1 Overview of project and the data

This report evaluates the education measures in the Longitudinal Study of Indigenous Children (LSIC). Globally, Australia's LSIC is the only longitudinal child cohort study on the developmental outcomes of Indigenous children. The study surveys Aboriginal and Torres Strait Islander Australian children aged either 6–18 months (B cohort) or 3.5–5 years (K cohort) when the study began in 2008.

The original objectives of the LSIC (also known as Footprints in Time) were to generate high-quality quantitative and qualitative data to provide a better insight into how a child's early years affect their development. According to the (then) Australian Government Department of Families, Community Services, Housing and Indigenous Affairs, Footprints in Time was designed to answer four key research questions. These continue to guide the ongoing development and analysis of the LSIC:

- What do Aboriginal and Torres Strait Islander children need to get the best start in life and grow up strong?
- What helps Aboriginal and Torres Strait Islander children to stay on track or encourages them to become healthier, more positive and stronger?
- How are Aboriginal and Torres Strait Islander children raised?
- What is the importance of family, extended family and community for both young children and as they grow up?

The focus on the specific needs and circumstances of the Indigenous population has gained considerable support within the community for the aims and goals of the LSIC. For example, the study has achieved high sample retention given the mobile and hard-to-reach population, as discussed below.

1.1 Structure of the data and the interviews

The unit of analysis in the LSIC is the study child. However, data in the LSIC are collected from four informants (Table 1). Parent 1 (P1) is the main parent or carer interviewed about the study child. Interviews were conducted with a median interview length of 31 minutes in wave 9 (according to the data user guide). This median length has declined from 48 minutes in wave 8 and about an hour in wave 2.

In wave 8, 84.5% of P1 interviews were conducted with the study child's mother. In addition, 5.5% of P1 interviews were conducted with the child's grandmother and 3.7% with the study child's father.

The data collected directly from the study child have become increasingly important as the children have grown up. In wave 9, the median interview length for the study child in the B cohort was 15 minutes, whereas for the K cohort the median length was 33 minutes. Most P1s consented to interview of the study child in the most recent waves; these interviews were in addition to direct testing or measurement of the study child.

The third informant's identity has changed over time. In initial waves, this informant was parent 2 (P2), who was either P1's partner or another adult with a parent or carer relationship to the study child. For most respondents, this was the biological father, but people with other caring roles were also interviewed. In waves 1 and 2, P2s were primarily fathers and response rates were relatively low. Hence, wave 4 focused on 'Dads' ('either fathers or men performing a father-like role in a Study Child's life'). The low response rates continued throughout the data collection: only 10.8% of study children had a P2/Dad who completed a face-to-face interview in wave 8; a further 5.5% self-completed the interview.

The fourth informant (introduced when the child entered the formal school system) was the study child's teacher. In wave 8, only 40.9% of the children had their teacher complete an interview. Informants completed the interview either on paper, with the data entered by the Department of Social Services (DSS) (33.8% of the sample) or online (7.1% of the sample). However, over the course of the study, most children have had at least one of their teachers complete a survey. Analysis by the DSS provided for this report indicates that, in the first 10 years of the study, teacher surveys were collected for more than 80% of the study children; for only 45% of the children, teachers have responded three or more times. Thus, longitudinal analysis of teacher responses is difficult, but some cumulative cross-sectional analysis can be done.

1.2 Sample retention

Up to and including wave 8, 865 of a total of 1255 respondents had completed all eight waves (54.6% of the original sample). Of these, 280 were from the K cohort (56% of the original K cohort sample) and 405 were from the B cohort (53.6% of the original B cohort sample).

At present, the most recent wave of LSIC data available for researchers is wave 9. These data were collected in 2016 when the B cohort was aged 9.5–11 years, and the K cohort 11.5–13 years. In wave 9, 1268 interviews were conducted. Of these, 1117 were conducted with those who participated in wave 8 (retention rate 89.0%) and 151 were conducted with those missing from the previous wave.

Up to and including wave 9, 647 of a total of 1268 respondents had completed all nine waves (51.0% of the original sample). Of these, 264 were from the K cohort (51.4% of original K cohort sample) and 383 were from the B cohort (50.8% of original sample).

The number of children lost from the sample means selective attrition is a concern (an issue we consider in this paper). Nevertheless, our retention rate compares favourably with retention rates of Indigenous cohorts in other longitudinal surveys. However, it has come at a cost. A trade-off was made between community support and representativeness on the one hand and the need to control for geographic clustering of individuals and the associated effect on standard errors and inference on the other (Hewitt 2012).

1.3 Sample distribution

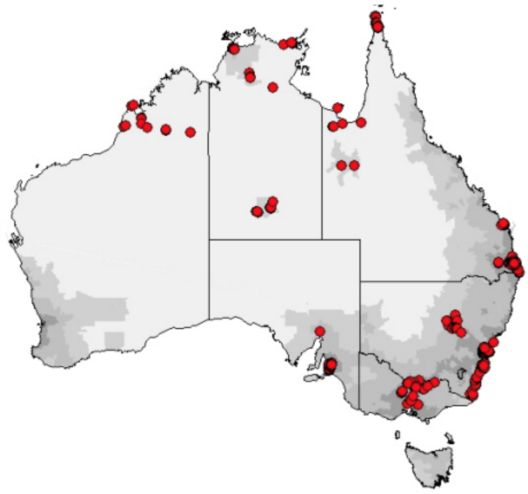
We will discuss the issue of sample retention regarding education measures later in this report, but first will consider the broad geographic

Table 1 Informants for data collection in the waves of the LSIC

Wave	P1	P2/Dads	Study child	Teacher or carer
1	1671	257	1469	44
2	1523	268	1472	163
3	1404	–	1394	326
4	1283	213	1269	442
5	1258	180	1244	473
6	1239	–	1241	543
7	1253	222	1244	549
8	1255	215	1240	517
9	1268	175	1247	583
10	1270	110	1254	631

– = no informants for that wave; LSIC = Longitudinal Study of Indigenous Children; P1 = parent 1; P2 = parent 2 distribution of respondents. The original sample of the LSIC was highly geographically clustered, not a representative sample from a randomly selected set of areas across Australia. The 11 identified sites spanned five states (not Tasmania) and the Northern Territory. Figure 1 shows the distribution of the original P1 interviews.

Figure 1 Geographic distribution of parent 1 interviews, wave 1



The distribution of the LSIC has changed over time, as study children continue to be surveyed even if they move out of the original study areas. Further, participant attrition varied with the area in which they lived. For example, we can examine the representativeness of the LSIC sample by geographic area in wave 8 – that is, the proportion of the sample who live in the five remoteness classifications usually used for analysis and policy making about the Indigenous population. In Table 2, the LSIC sample is separated into those who completed all LSIC waves (the complete responders), those who missed at least one of the waves (incomplete responders) and data for the entire wave 8 sample.

The final column of Table 2 illustrates the proportion of children aged 7–11 years who identified as Aboriginal and/or Torres Strait Islander according to the 2016 Census.[1] As population estimates are only available for major cities, regional areas (combining inner regional and outer regional areas) and remote areas (combining remote and very remote areas), the table shows a three-category remoteness classification. However, the LSIC separates data from regional and remote areas into the component parts.

In this analysis, the complete responders in LSIC more closely resemble the census-based estimates of the geographic distribution of the Indigenous population than do the incomplete responders. However, the total wave 8 sample is less likely to live in major cities than the 2016 estimates, and more likely to live in a remote areas (Table 2).

The LSIC sample may not represent the total Indigenous population (based on the 2016 census) for three reasons. Biddle (2011) and Thurber et al. (2015) discuss two of these reasons (baseline representativeness and nonrandom attrition). However, an emerging issue is identification change among the Indigenous population. Specifically, an increasing number of people (relative to the baseline Indigenous population) who did not previously identify or were not identified as Indigenous now identify as Indigenous.

Table 2 Geographic distribution of LSIC sample (wave 8) and census-based population estimates

Geographic area	LSIC sample (%) Complete responders (%)	Incomplete responders (%)	Total (%)	2016 Census-based estimates of Indigenous children aged 7–11 years (%)
Major cities	33.1	18.1	26.3	36.6
Regional areas	47.9	39.5	44.1	46.7
Remote areas	19.0	42.5	29.6	17.0

LSIC = Longitudinal Study of Indigenous Children

Campbell et al. (2018) analysed the level and determinants of changes in Indigenous identification between 2006 and 2011. The analysis showed that the newly identified people tend to have very different outcomes from the previously identified people, and are concentrated in urban areas and in relatively advantaged families. More recent estimates from the 2011–16 Australian Census longitudinal dataset suggest that net identification change in this period was between 15% and 20% of the baseline Indigenous population. Because the LSIC sample fails to capture these newly identified Indigenous children, it probably becomes less representative over time, apart from the issues of nonrandom attrition and baseline representativeness. We return to this issue in our concluding comments.

1.4 Data linkage

The LSIC also collects information via data linkage to supplement the interview data collected. These data are important for the analysis of education as both an outcome and a determinant.

The Australian Early Development Census (AEDC) is a 'nationwide data collection of early childhood development at the time children commence their first year of full-time school'.[2] The data is currently available at the aggregate level and linked to the children in wave 2. However, we have sought consent to link data at the individual level for future waves. The current documentation does not indicate the consent rate.

Information is also available from the MySchool data collection, including but not limited to data from the National Assessment Program – Literacy and Numeracy (NAPLAN). Unlike the AEDC (which occurs every 3 years), NAPLAN occurs every year: 'Students in Years 3, 5, 7 and 9 are tested on the fundamental literacy and numeracy skills that every child needs to succeed in school and beyond. NAPLAN is a national, consistent measure to determine whether or not students are meeting important educational outcomes.'[3]

Further, unlike the AEDC, information on NAPLAN is available at both the school and individual study child levels. P1s have been asked for consent to use NAPLAN data from wave 4, with consent rates high and increasing over time. For example, in 2011 when consent was first asked, 106 (7.4%) respondents refused to consent. By 2015, this figure had declined to 51 (3.4%) P1s. Linkage consent is lower for those missing from some of the waves than for those included in all waves. Specifically, consent rates in 2015 were 95.6% for incomplete responders and 98.0% for complete responders. We will return to the external validity of individual NAPLAN data for the LSIC later in this paper.

1.5 Structure of the report

Having introduced the data here, in Section 2 we discuss existing research that uses LSIC education measures to identify how data items have previously been used, particularly within research conducted by or in partnership with Aboriginal or Torres Strait Islander researchers or organisations. In Section 3 we outline our methodological approach and assumptions, including a review and synthesis of the existing evidence on the validity of the education measures for the Aboriginal and Torres Strait Islander population. Where evidence for this population is not available, we use evidence from other populations. For example, the Longitudinal Study of Australian Children contains many similar education measures.

Sections 4–6 present the detailed data analysis for the project. In Section 4 we use cross-sectional bivariate and univariate analysis to test the consistency of the education measures using Cronbach's alpha; undertake exploratory factor analysis and confirmatory factor analysis to identify how the items cluster together; and test the correlations against related and unrelated outcomes to test construct validity. Most of this analysis focuses on wave 9 of the LSIC.

Often, researchers are less interested in the levels of the education measures than in what predicts the measures. In Section 5, therefore, we present multivariate analysis of the relationships among demographic, family, community, socioeconomic and geographic variables to test convergent and discriminant validity (i.e. education as a dependent variable), and the relationship between education and key developmental, physical and mental health measures to test predictive validity (i.e. education as the independent variable).

In the final set of analyses (Section 6), we test the longitudinal consistency of LSIC measures and the observed longitudinal relationship with other variables. Specifically, we formally test the measurement invariance of education constructs over time and undertake longitudinal multivariate analysis.

In Section 7, we provide recommendations for three (sometimes overlapping) groups of individuals: those involved in data collection, those analysing the LSIC, and those using analysis of the LSIC for policy decisions. Section 8 provides a summary and concluding comments.

2 Literature review and existing knowledge on validity of the education data

To validate the education measures in the LSIC, we must understand how they are currently being used, as well as any comments or criticisms of the measures in the existing literature. It is always a challenge to identify a comprehensive list of research outputs that use a particular dataset, because not all are publicly available or identified in searchable databases. Nonetheless, we believe the summary below identifies most of the main papers used or cited by other researchers and policy makers.

2.1 Identification of articles

We conducted the literature review through a Google Scholar search with the search terms LSIC, Education, and Aboriginal OR Torres OR Indigenous. We restricted the search results to those published since 2010 to ensure only papers that contained analysis were reviewed and to avoid papers that discussed methodology for a future LSIC. This search was undertaken early in the project, but was repeated on 5 April 2019 to capture any recent publications, or those recently available through Google Scholar. We also checked specific education databases such as ERIC and Proquest.

From the initial list, publications written exclusively by the DSS, formerly the Department of Families, Housing and Community Services, were excluded, because this study examined data use by researchers external to the data custodian. Articles co-authored with DSS staff were included. We also excluded systematic reviews of LSIC analysis (although these reviews were assessed to check for any additional papers that may have been missed).

An initial review of the articles' title and abstract identified those relevant to this project. Next, we checked for papers that cited articles found in the initial scan of the literature, resulting in a total Google Scholar sample of 16 articles. Finally, we similarly searched FLoSse, DSS's online repository of research, using its main longitudinal datasets. With the same exclusion criteria and the keyword 'education' within LSIC-based research, we did not identify any additional articles.

The list below contains the 16 identified articles. Three articles from one book are grouped together.

- Arcos Holzinger LA & Biddle N (2015). *The relationship between early childhood education and care (ECEC) and the outcomes of Indigenous children: evidence from the Longitudinal Study of Indigenous Children (LSIC)*, Working Paper 103, Centre for Aboriginal Economic Policy Research, Australian National University, Canberra.
- Armstrong S, Buckley S, Lonsdale M, Milgate G, Kneebone LB, Cook L & Skelton F (2012). *Starting school: a strengths-based approach towards Aboriginal and Torres Strait Islander children*, https://research.acer.edu.au/indigenous_education/27.
- Colquhoun S & Dockery AM (2012). *The link between Indigenous culture and wellbeing: qualitative evidence for Australian Aboriginal peoples*, Discussion Paper Series 2012/01, Centre for Labour Market Research, Curtin University, Perth.
- Walter M, Martin KL & Bodkin-Andrews G (eds) (2017). *Indigenous children growing up strong: a longitudinal study of aboriginal and Torres Strait Islander families*, Palgrave Macmillan, London.
  - Bodkin-Andrews G, Whittaker A, Cooper E, Parada RH, Denson N & Bansel P (2017). Moving beyond essentialism: Aboriginal parental perceptions of school bullying and school engagement. In: *Indigenous children growing up strong*, 153–178.

   - Trudgett M, Page S, Bodkin-Andrews G, Franklin C & Whittaker A (2017). Another brick in the wall? Parent perceptions of school educational experiences of Indigenous Australian children. In: *Indigenous children growing up strong*, 233–258.
   - Anderson I, Lyons JG, Luke JN & Reich HS (2017). Health determinants and educational outcomes for Indigenous children. In: *Indigenous children growing up strong*, 259–285.
- Martin KL (2017). It's special and it's specific: understanding the early childhood education experiences and expectations of young Indigenous Australian children and their parents. *Australian Educational Researcher* 44(1):89–105.
- Hewitt B & Walter M (2014). Preschool participation among Indigenous children in Australia. *Family Matters* 95:41–50.
- Prout Quicke S & Biddle N (2017). School (non-) attendance and 'mobile cultures': theoretical and empirical insights from Indigenous Australia. *Race Ethnicity and Education* 20(1):57–71.
- Biddle N (2011). *An exploratory analysis of the Longitudinal Survey of Indigenous Children*, Working Paper 77, Centre for Aboriginal Economic Policy Research, Australian National University, Canberra.
- Gilroy J & Emerson E (2016). Australian Indigenous children with low cognitive ability: Family and cultural participation. *Research in Developmental Disabilities* 56:117–127.
- Blunden S, Magee C, Attard K, Clarkson L, Caputi P & Skinner T (2018). Sleep schedules and school performance in Indigenous Australian children. *Sleep Health* 4(2):135–140.
- Dunstan L, Hewitt B & Tomaszewski W (2017). Indigenous children's affective engagement with school: the influence of socio-structural, subjective and relational factors. *Australian Journal of Education* 61(3):250–269.
- Azpitarte F, Chigavazira A, Kalb G, Farrant BM, Perales F & Zubrick SR (2018). Childcare use and its role in Indigenous child development: evidence from the Longitudinal Study of Indigenous Children in Australia. *Economic Record* 95:1–33.
- Peacock H & Prehn J (2017). *Successful Indigenous student transition to secondary education,* presentations from the AIATSIS National Indigenous Research Conference, 21–23 March 2017, Canberra, Australia, 1.
- Marmor A & Harley D (2018). What promotes social and emotional wellbeing in Aboriginal and Torres Strait Islander children? Lessons in measurement from the Longitudinal Study of Indigenous Children. *Family Matters* 100:4–18.

2.2 Use of education data from the LSIC in the literature

The LSIC is an underutilised resource for research and policy development related to Indigenous child education outcomes. Nevertheless, the identified research provides substantive insights and indicates how the LSIC can be used for educational research. The identified publications are probably not exhaustive, and likely miss some Honours, Masters and PhD theses. However, several themes emerge from the cited articles.

Previous analysis of the education data from the LSIC tends to be cross-sectional rather than longitudinal. This bias is diminishing in later papers. For example, Blunden et al. (2018) investigated whether sleep schedule predicted change in NAPLAN results between year 3 and year 5. Arcos Holzinger and Biddle (2015) and Azpitarte et al. (2018) used participation in early childhood education and care to predict outcomes after children begin full-time schooling. Most researchers, however, used only one wave of LSIC data and none of the studies we found used panel data techniques such as fixed or random effects modelling to control for unobserved heterogeneity.

Researchers from a **range of disciplines** have analysed the LSIC education data. Our review identified analyses by economists, public health researchers, education researchers, psychologists and sociologists; probably other disciplines have also used the data.

Qualitative use of the LSIC data is limited but important. For example, Dockery (2010) examined responses to two open-ended questions and found that 'Aboriginal parents place great importance upon education, but also upon their child maintaining and learning about aspects of their culture for identity development, upon the positive experience of the traditional culture and the significance of support from the community to which they belong.' In addition, several authors have used attitudinal data that, while quantitative in application, attempt to capture qualitative concepts. For example, Trudgett et al. (2017) investigated 'Parent 1's perception that the teacher of the study child was sensitive to the needs of Indigenous families' as a key explanatory variable. However, we were unable to identify any truly mixed methods studies that incorporate insights from the qualitative analysis into longitudinal analysis with quantitative methods.

In terms of analytical techniques, studies used a mix of descriptive and multivariate analysis. Recent papers use techniques that attempt to draw out causal inference, including propensity score matching and instrumental variables techniques. **Multivariate analyses mostly use education as an outcome rather than an explanatory variable.** The main exception is using participation in early childhood education to predict future cognitive outcomes. For example, Anderson et al. (2017) used data from wave 6 of the LSIC and 'found both individual and family-level health determinants have an adverse effect on educational attainment among children.'

In terms of education outcomes, the **most common outcomes are participation measures**, especially school attendance as an education outcome. For example, Peacock and Prehn (2017) cross-sectionally examined the relationship between parental engagement in school and school attendance. Several papers have used outcome measures that span education outcomes and child development, especially papers on policy effectiveness (e.g. Arcos Holzinger & Biddle 2015, Azpitarte et al. 2018). Only Blunden et al. (2018) used the linked outcome data as outcome measures.

Our literature review identified an important and apparently **increasing use of the LSIC by researchers who identify as Indigenous**, including the Palgrave Macmillan publication edited by Professors Walter, Martin and Bodkin-Andrews (2017). Three chapters identified in this book specifically concern education outcomes. However, many of the other chapters touched on education more broadly. In addition, several articles by the above authors and other Indigenous authors feature in this review.

A final point is the use of **Indigenous-specific measures, particularly as explanatory variables**. A notable body of research examines the experience of racism and bullying in school (and the wider society) as a determinant of child outcomes. Bodkin-Andrews et al. (2017) investigated bullying (as perceived by parents) and its relationship with a range of school-based outcomes. Two additional articles not included in our list – one was not directly related to education and one was unpublished at the time – are worth mentioning. Shepherd et al. (2017) looked longitudinally at the relationship between racism and health, and Biddle and Priest (2019) looked longitudinally at the relationship between exposure to racism and education outcomes. Similar research with Indigenous-specific measures would be impossible with the Indigenous sample in the Longitudinal Study of Australian Children or with administrative datasets.

3 Methodological approach and assumptions

We take a highly data-driven approach to understand and validate the education measures in the LSIC. Our analysis is supported by education (and related) theories, but we avoid testing any theoretical models of education outcomes for the Indigenous population. Here we outline our broad methodological approach. We discuss the specific methodologies in more detail in the relevant sections.

Section 4 presents the tests for various elements of construct validity and also for internal consistency. Much of the section is devoted to testing factorial validity – that is, whether a coherent factor structure exists for the scales used to assess various aspects of education measures. In the LSIC, some of these scales come from standardised measures or are replications of data items from other datasets to give us a theoretical idea of what to expect. However, the origins of the LSIC scales are often unknown, or they were developed specifically for the dataset and the particular cohort. In both instances, we are looking for the presence of several characteristics:

- a coherent factor structure from an exploratory factor analysis in Stata
- factors identified through exploratory factor analysis, adequate scale reliability or internal consistency (Cronbach's alpha)
- a factor structure that fits our expected factor structure or has 'good model fit' when tested using confirmatory factor analysis in Mplus 8.1.

In addition, Section 4 shows the extent to which the different scales demonstrate convergent and discriminant validity through correlation matrices of the education measures. Convergent validity is the extent to which a scale correlates with another similar measure. Discriminant validity is the extent to which a scale does not correlate (i.e. has a low correlation) with another dissimilar measure.

In Section 5, we test the extent to which the education items can be explained by demographic, geographic and socioeconomic outcomes. We assume that a useful education measure varies in significant and predictable ways throughout the population of interest. For this test, we use a regression approach to measure the association between an outcome of interest (the dependent variable) and another variable (the independent or explanatory variable) that we expect to be associated with the first variable. All other independent variables are held constant. We analysed three types of dependent variables with appropriate analytical methods:

- Continuous variables (i.e. variables that can take on multiple values across the distribution) are estimated with a linear model, using ordinary least squares estimation.
- Binary variables (i.e. variables that can take on two possible values) are estimated with the probit model, using maximum likelihood estimation.
- Categorical variables with a logical rank or categorisation are estimated with the ordered probit model, using maximum likelihood estimation.

We present results from the regression analysis as coefficients. The values of the coefficients in the different models do not directly compare, but the direction (positive or negative) and the statistical significance do compare. For each model we estimate, we also include the adjusted or pseudo R-squared, which measures the amount of variation in the dependent variable that is explained by the model estimates. This measure varies from 0 (none of the variation explained) to 1 (all of the variation explained).

In the final section of empirical results, we test the longitudinal validity of the data. For the regression analysis, we use two different types of models. In the simplest model, we estimate the change through time (between time t and $t + \Delta$) in one of the dependent variables as a function of a set of

independent variables at time *t*. Because there is only one observation per individual, we estimate this using ordinary least squares and a linear model. For multiple waves of observations per individual, we estimate a random effects model, where the outcome at time *t* is a function of the explanatory variable at time *t*, a time dummy variable, and an individual-level error term.
The coefficients are estimated using maximum likelihood estimation.

4 Cross-sectional, bivariate and univariate analysis

The LSIC contains scales that measure affective school engagement (School Liking and Avoidance Questionnaire [SLAQ]), academic self-concept in school subjects (academic self-concept in maths and reading), students' perceptions of the school environment (e.g. school climate, teacher relationship, Psychological Sense of School Membership [PSSM] scale) and teachers' ratings of their relationship with students (Pianta Student–Teacher Relationship Scale [STRS]). Table 3 indicates when these scales are included in LSIC surveys for the B and K cohorts. The findings are important when considering optimal candidates for longitudinal analyses. Data users should note some key points:

- Wave 9 includes more education scales than any other wave (particularly for the K cohort).
- For the B cohort, SLAQ was the only child-reported measure consistently measured for more than one wave (waves 6–9). The teacher-reported STRS was the only other measure that enabled longitudinal analyses (waves 4–9).
- For the K cohort, SLAQ was the only child-reported measure at every wave from wave 3 onwards, except for the academic self-concept, which was measured at waves 6, 7 and 9.
- Two other child self-report scales included in more than one wave of data were a measure of the teacher relationship (waves 6–9) and a measure of perceptions of school climate (waves 8–9).
- The teacher-reported STRS was included in waves 4–9.

This section also explains the reliability and construct validity of the scales. Given the limited research using these educational measures, and that some measures are newly developed or are abridged versions of existing scales, this section aims to:

1. establish coherent factor structure from an exploratory factor analysis in Stata
2. for factors identified through exploratory factor analysis, test for internal consistency (Cronbach's alpha)
3. test the expected factor structure through confirmatory factor analysis in Mplus 8.1
4. test whether these scales show convergent and discriminant validity.

For all exploratory factor analyses we used principal component analysis with a quartimax rotation using Stata 15.1. Bartlett's test of sphericity indicated the correlations were high enough for the exploratory analyses and the KMO Measure of Sampling Adequacy indicated that data were also used (Tabachnick & Fidell 2007). We used an orthogonal rotation (quartimax rotation) because the factor loadings are generally easier to interpret and report (Tabachnick & Fidell 2007), and because we expected that in most instances the first factor (Thompson 2004) would explain sufficient variation. The number of factors was determined by eigenvalues above 1 and the scree plot (Thompson 2004).

To test for internal consistency, we used Cronbach's alpha (Cronbach 1951). Nunnally (1978) considered that alpha coefficients of 0.70 and above are acceptable, but others disagree because this criterion is not based on any empirical research (Cho & Kim 2015).[4] Therefore, data users should be mindful of the 0.70 alpha convention. For important research, an alpha of 0.70 or higher may be needed. However, much of the research undertaken with LSIC data is exploratory, and thus alpha measures of 0.60 to 0.69 are practical.

Table 3 **Educational scales in LSIC by cohort and wave for factor analysis**

Age of B cohort (years)	0–2	1–3	2–4	3–5	4–6	5–7	6–8	7–9	8–10
Age of K cohort (years)	3–5	4–6	5–7	6–8	7–9	8–10	9–11	10–12	11–13
	Wave								
Construct	1	2	3	4	5	6	7	8	9
Academic self-concept						(K)	(K)		✓
Psychological Sense of School Membership Scale									(K)
School climate								(K)	(K)
School Liking and Avoidance Questionnaire			(K)	(K)	(K)	✓	✓	✓	✓
Teacher relationship						(K)	(K)		✓
Pianta Student–Teacher Relationship Scale (teacher report)				✓	✓	✓	✓	✓	✓

LSIC = Longitudinal Study of Indigenous Children

Note: Approximate age ranges of study children at each wave are presented in years. Questions are asked in relation to all children, unless indicated as B cohort (B), K cohort (K) or B and K cohort (✓).

As a further test of factorial validity after the exploratory factor analysis, we estimated confirmatory analysis. For the confirmatory factor analyses we used all available data.[5] Confirmatory factor analysis enables the specification of particular items loading on a latent factor (hence the name confirmatory). The technique provides information about the extent to which the theorised factor structure fits the empirical data in the form of model fit indices (for further details, see Zubrick et al. 2014). We evaluated model fit according to the guidelines provided by Browne and Cudeck (1993) and Hu and Bentler (1998), using the model chi-square (χ^2), the comparative fit index (CFI), the Tucker–Lewis index (TLI), the standardised root mean residual (SRMR) and the root-mean-square error of approximation (RMSEA). A nonsignificant χ^2 indicates a good model fit. The CFI ranges from 0 to 1.00, with a cutoff of 0.95 or higher indicating a well-fitting model and 0.90 indicating an adequate fit. A TLI value close to 1 indicates a good fit; RMSEA values below 0.05 indicate a good model fit, and values between 0.06 and 0.08 indicate an adequate fit. For the SRMR, a value less than 0.05 is considered a good fit and below 0.10 an adequate fit. Following Zubrick et al. (2014), confirmatory factor analysis models were judged:

- good – model meets specified criteria for SRMR, TLI and CFI
- acceptable – model meets SRMR criteria and least one TLI or CFI criterion
- not acceptable – model fails to meet SRMR or model meets SRMR but no TLI or CFI criteria.

In the literature, authors debate about the most appropriate indices. We include the RMSEA for completeness but do not use it to make final judgements on model fit.

In this section, we also test the construct validity of the measures. Construct validity is the degree to which the measure captures the essential features of the concept or construct that it is intended to measure (Cronbach & Meehl 1955). Convergent and discriminant validity are elements of construct validity. Convergent validity occurs when two different measures of the same construct provide similar results or are highly correlated (Kvien et al. 1998). Discriminant validity occurs when one measure of a construct can be differentiated from a different measure or construct (Kvien et al. 1998). For example, when a measure of negative mood strongly correlates with a measure of anxiety, this is an example of convergent validity. One way to assess convergent and discriminant validity is to test the correlations between like (convergent) and unlike (discriminant) measures in a correlation matrix.

4.1 School Liking and Avoidance Questionnaire

Affective engagement in schooling has commonly been measured by SLAQ (Ladd & Price 1987). Affective engagement or school liking predicts greater cooperation and independence in classroom activities throughout the first year of school (Ladd et al. 2000). Further, children who are more comfortable at school are better equipped to invest effort when confronted by novel and challenging tasks. Therefore, affective engagement promotes productive classroom behaviours and subsequent academic achievement (Ladd & Dinella 2009).

First, to understand the underlying factor structure of the SLAQ items, we conducted exploratory factor analysis for each individual wave and cohort from waves 7 to 9 (Table 4).

We used principal component analysis with a quartimax rotation using Stata 15.1. All exploratory factor analyses indicated a one-factor structure. Table 4 shows factor loadings for each wave and by cohort. For the B cohort, the factor loadings for waves 7 and 8 were highly consistent, suggesting that the factor can be interpreted as 'school liking'. In wave 9, an additional avoidance item was added to the survey (Do you try and find ways of getting out of going to school?). Consequently, the direction of the sign of the factor loadings reversed (from negative to positive and vice versa). This result suggests that for this wave the factor can be interpreted as a measure of school avoidance. Measures of internal consistency of factors for waves 7–9 for the B cohort ranged from 0.64 to 0.73, which is adequate for research purposes (for the lack of empirical agreement, see Nunnally 1978, Cho & Kim 2014).

Table 4 Exploratory factor analyses of SLAQ and internal consistency, by wave and cohort

Variable	Wave		
	7	8	9
B cohort			
Is school fun?[a] (csc5a)	0.71	0.71	−0.69
When you get up in the morning, do you feel happy about going to school? (csc6b)	0.79	0.79	−0.72
Do you try and find ways of getting out of going to school? (csc18b)	na	na	0.63
Do you ask to stay home from school? (csc9b)	−0.53	−0.67	0.74
Do you wish you didn't have to go to school? (csc7b)	−0.75	−0.74	0.72
Cronbach's alpha	0.64	0.69	0.73
N	689	687	733
K cohort			
When you get up in the morning, do you feel happy about going to school? (csc6b)	−0.67	−0.70	−0.55
Do you try and find ways of getting out of going to school? (csc18b)	0.72	0.79	0.72
Do you ask to stay home from school? (csc9b)	0.78	0.69	0.78
Do you wish you didn't have to go to school? (csc7b)	0.76	0.77	0.80
Cronbach's alpha	0.71	0.71	0.67
N	486	482	482

na = not applicable; SLAQ = School Liking and Avoidance Questionnaire

a The response format for the SLAQ data items are (1) Always, (2) Most of the time, (3) Fair bit, (4) Little bit, (5) Not much and (6) Never.

For the K cohort, a more consistent pattern emerged across waves 7–9. The factor loadings suggest that they measure school avoidance, with the only item measuring school liking negatively related to the factor. Cronbach's alphas for waves 7 to 9 were marginally better than for the B cohort and adequate for research purposes.

Given the nature of the SLAQ factor loadings (a single factor), we estimated one-factor congeneric models for each wave by cohort. Table 5 shows the model fit indices for waves 7–9 for the B and K cohorts for SLAQ. For the B cohort, all of the fit indices except the RMSEA showed at least an acceptable fit for the congeneric model in waves 7–9. For wave 9, the B cohort one-factor congeneric model showed good fit for every index except the SRMR.

The K cohort showed mixed results for the model fit indices. In waves 7 and 8, but not in wave 9, the mix of indices showed an overall acceptable level of fit according to the criteria of Zubrick et al. (2014).

Table 5 Confirmatory factor analyses of SLAQ, by wave and cohort

Variable	Wave		
	7	8	9
B cohort			
SRMR (<0.05 good fit, <0.10 acceptable)	**0.03**	**0.03**	**0.007**
RMSEA (<0.05 good fit, <0.08 acceptable)	0.09	0.11	**0.00**
CFI (>0.95 good fit, 0.90–0.95 acceptable)	**0.98**	**0.98**	**1.00**
TLI (>0.95 good fit, 0.90–0.95 acceptable)	**0.94**	**0.94**	**1.00**
Chi-square	14.46	18.25	1.80
df	2	2	3
P	0.0007	0.0001	0.62
N	716	742	744
Overall rating	Good	Good	Good
Estimator	ML	WLSMV	ML
K cohort			
SRMR (<0.05 good fit, <0.10 acceptable fit)	**0.03**	**0.046**	**0.05**
RMSEA (<0.05 good fit, <0.08 acceptable)	0.17	0.28	0.31
CFI (>0.95 good fit, 0.90–0.95 acceptable)	**0.95**	**0.91**	0.88
TLI (>0.95 good fit, 0.90–0.95 acceptable)	0.85	0.73	0.63
Chi-square	32.34	77.25	93.85
df	2	2	2
P	0.0000	0.0000	0.0000
N	514	490	495
Overall rating	Acceptable	Acceptable	Not acceptable
Estimator	WLSMV	WLSMV	WLSMV

CFI = comparative fit index; ML = maximum likelihood; RMSEA = root-mean-square error of approximation; SLAQ = School Liking and Avoidance Questionnaire; SRMR = standardised root mean residual; TLI = Tucker–Lewis index; WLSMV = weighted least square mean and variance

Note: Bold numbers indicate good fit.

4.2 Academic self-concept

The Marsh Self-Description Questionnaire is a widely used, validated instrument to measure self-concept in children (Marsh & MacDonald Holmes 1990). Baumeister (1999) defined self-concept as 'the individual's belief about himself or herself, including the person's attributes and who and what the self is'. As used in the LSIC, the questionnaire comprises scales to measure both academic (reading, writing, mathematics and general school) and nonacademic (physical ability, physical appearance, and peer/parent relations) self-concept. A large body of research shows that academic self-concept is strongly associated with academic outcomes (Marsh & MacDonald Holmes 1990, Ferla et al. 2009) and school satisfaction (Briones & Tabernero 2012).

Studies have found that the Marsh Self-Description Questionnaire is a valid instrument to measure self-concept among Indigenous children (Bodkin-Andrews et al. 2010ab). For example, Yeung et al. (2013) compared levels of academic self-concept and school satisfaction in Indigenous and non-Indigenous Australian children. They found that, in both urban and rural settings, non-Indigenous children had higher self-concept, greater enjoyment and participation at school, as well as better self-ratings of their schoolwork. Bodkin-Andrews et al. (2012) examined the relationship between Indigenous identification, academic self-concept and academic disengagement. They found that Indigenous children are significantly more likely to be academically disengaged. The lower academic self-concept of Indigenous children largely accounts for this disengagement.

In the LSIC, a subset of items from the Marsh Self-Description Questionnaire were asked, with a focus on maths and reading for both cohorts and an additional focus on writing for the K cohort. The response format for the data items was (1) Yes: always, (2) Yes: most of the time, (3) Sometimes: fair bit, (4) Sometimes: little bit, (5) No: not much and (6) No: never.

We performed this analysis on wave 9. We used principal component analysis with a quartimax rotation using Stata 15.1. As an indication of factor structure we used eigenvalues above 1 and the scree plot. For the B cohort, two eigenvalues were greater than 1 (2.20, 2.02) (Table 6). The factor loadings suggest a clear two-factor structure, with a maths factor and a reading factor. The Cronbach's alphas were acceptable (Nunnally 1978).

The same factor analytic approach was adopted for wave 9 academic self-concept items for the K cohort (Table 7). There were three eigenvalues greater than 1 (2.55, 2.03, 1.91), which suggests a three-factor solution. The factor loadings align with the hypothesised factor structure of one-factor for the maths, reading and writing concepts. However, 'I like reading' cross-loaded on the reading and writing factors (higher loading on reading factor). This suggests that the reading and writing factors are not as 'pure' as the maths factor. Internal consistency was good according to the Cronbach's alphas (Nunnally 1978).

Table 6 Factor loadings and Cronbach's alphas for academic self-concept, B cohort, wave 9

Variable	Maths	Reading
I am good at reading (original, Marsh[a] 1990) (icsc35a)		0.84
I learn fast in reading (adapted, Marsh 1990) (icsc34a)		0.80
I like reading (Marsh 1990) (icsc33a)		0.79
I am good at maths (adapted, Marsh 1990) (icsc39a)	0.87	
I learn fast in maths (adapted, Marsh 1990) (icsc38a)	0.83	
I like maths (adapted, Marsh 1990) (icsc37a)	0.82	
Cronbach's alpha	0.81	0.76

a Marsh Self-Description Questionnaire

Notes:

1. *N* = 738.
2. Factor loadings under 0.36 are suppressed.

Table 7 Factor loadings and Cronbach's alphas for academic self-concept, K cohort, wave 9

Variable	Maths	Reading	Writing
I am good at reading (original, Marsh[a] 1990) (icsc35a)		0.89	
I learn fast in reading (adapted, Marsh 1990) (icsc34a)		0.86	
I like reading (Marsh 1990) (icsc33a)		0.58	0.49
I am good at maths (adapted, Marsh 1990) (icsc39a)	0.90		
I learn fast in maths (adapted, Marsh 1990) (icsc38a)	0.91		
I like maths (adapted, Marsh 1990) (icsc37a)	0.89		
I like writing (adapted, Marsh 1990) (icsc51)			0.91
I am good at writing (adapted, Marsh 1990) (icsc52)			0.85
Cronbach's alpha	0.89	0.80	0.83

a Marsh Self-Description Questionnaire

Note: Factor loadings under 0.36 are suppressed.

Next, we estimated one-factor congeneric models for both cohorts to provide estimates of the factor structure as specified in the exploratory factor analysis (Table 8). For the B cohort model, fit indices suggested that the residuals from two sets of items from each of the two factors should correlate with one another. We allowed them to correlate because they shared the same item stems. The first set was 'I like maths' (icsc37a) and 'I like reading' (icsc33a), and the second set was 'I learn fast in maths' (icsc38a) and 'I learn fast in reading' (icsc34a). Figure 2 depicts the final model. The model fit indices for this final model were very good, with all indices indicating a good fit.

For the K cohort, we tried to fit a model with maths, reading and writing academic self-concepts. However, this model had poor fit, probably because of the cross-loading of 'I like reading' on the writing latent factor. Therefore, given that only wave 9 captured writing self-concept and that data users wanting to conduct longitudinal analyses would benefit from information about model fit of the maths and reading self-concepts for the K cohort, we re-ran the confirmatory factor analysis for these two sets of items. The models had good fit (Zubrick et al. 2014).

For the B cohort, the items 'I am good at maths' (icsc39a) and 'I am good at reading' (icsc35a) were the highest-loading items (Figure 2); the lowest loading items for the two factors reflected 'liking'. The K cohort (Figure 3) showed the same pattern of item loadings, which suggests this construct has some stable temporal features, at least for two age cohorts. Section 6 presents more thorough analyses of factorial invariance over time.

Table 8 **Confirmatory factor analyses for academic self-concept in wave 9 by type and cohort**

Variable	B cohort	K cohort with writing	K cohort without writing
SRMR (<0.05 good fit, <0.10 acceptable fit)	**0.017**	0.05	**0.032**
RMSEA (<0.05 good fit, <0.08 acceptable)	**0.04**	0.13	**0.077**
CFI (>0.95 good fit, 0.90–0.95 acceptable)	**0.995**	0.93	**0.99**
TLI (>0.95 good fit, 0.90–0.95 acceptable)	**0.988**	0.89	**0.97**
Chi-square	13.40	157.40	23.81
df	6	17	6
P	0.0371	0.0000	0.0006
N	745	495	495
Overall rating	Good	Not acceptable	Good
Estimator	ML	ML	ML

CFI = comparative fit index; df = degrees of freedom; ML = maximum likelihood; RMSEA = root-mean-square error of approximation; SRMR = standardised root mean residual; TLI = Tucker–Lewis index

Note: Bold numbers indicate good fit.

Figure 2 **Standardised factor loadings for maths and reading academic self-concepts, B cohort, wave 9**

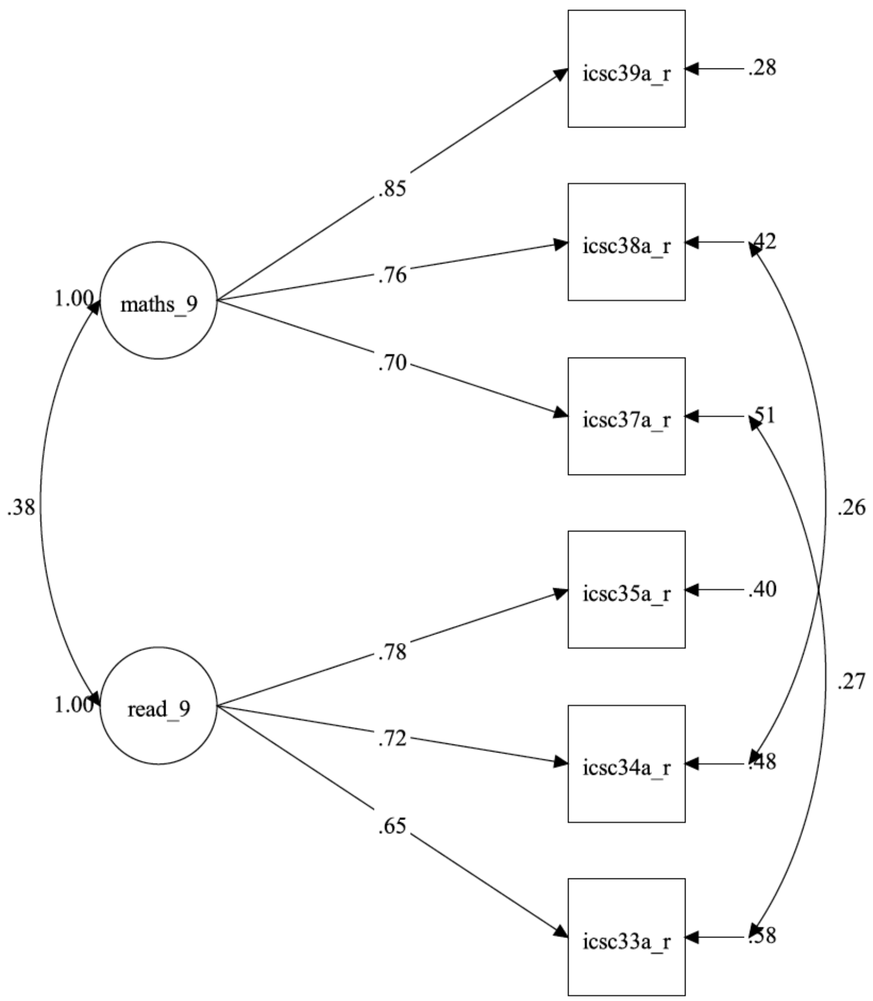
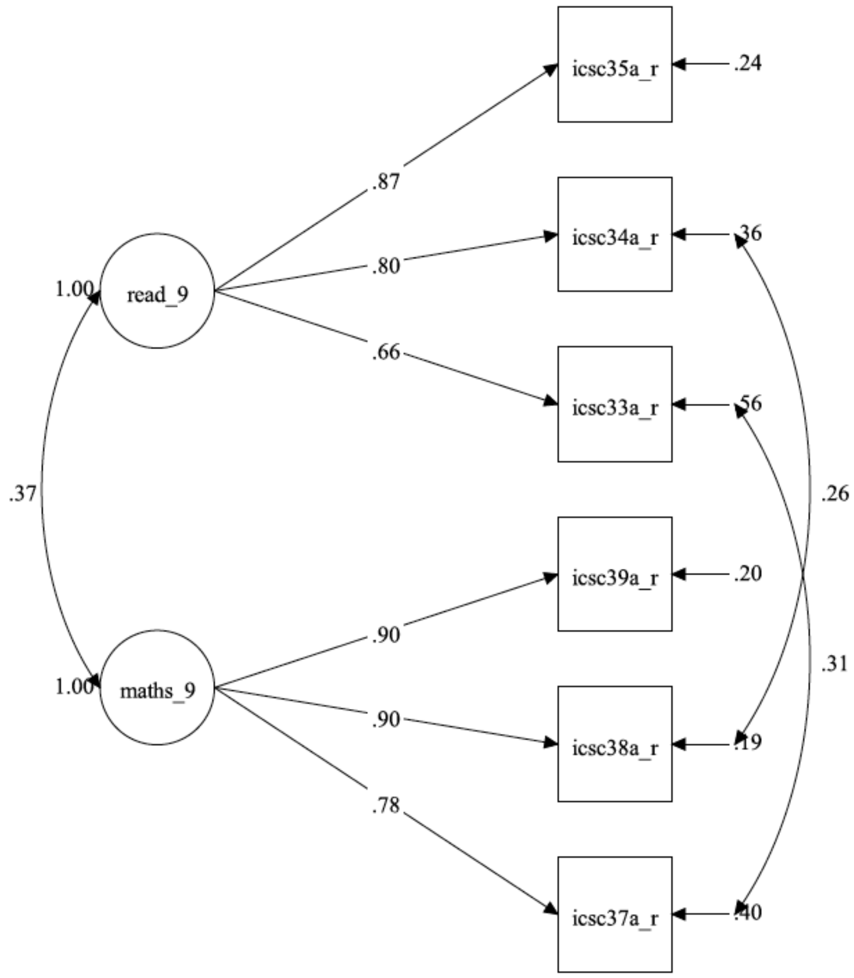


Figure 3 Standardised factor loadings for maths and reading academic self-concepts, K cohort, wave 9



4.3 Psychological Sense of School Membership scale

The PSSM scale, a widely used measure of school belonging, was developed 'to assess the adolescent's perceived belonging or psychological membership in the school environment' (Goodenow 1993:79). School belonging is associated with academic engagement, academic achievement, motivation, mental health and problem behaviour (see Slaten et al. 2016 for a recent review).

The PSSM scale was conceptualised as unidimensional. However, various studies have identified two or three factors; some studies needed to exclude items to obtain a clear factor structure (Abubakar et al. 2016). A recent cross-national study suggested a one-factor structure (Abubakar et al. 2016), but lack of empirical agreement still plagues this measure.

Considering the uncertainties around the scale factor structure of the original PSSM scale, and that the LSIC includes only 5 of the original 18 items, we were unclear how many factors to expect. The initial factor analysis suggested a two-factor solution with two eigenvalues greater than 1 (2.23, 1.02). However, after quartimax rotation the factor solution showed only one item loads on factor 2: 'I feel very different from most other students here'. We therefore conducted a second run of the exploratory factor analysis excluding this item. This finding reflects previous research that conducted factor analysis on the full version of the PSSM, which showed that negatively worded items loaded on their own factor (Ye & Wallace 2014). After removing this item, only one eigenvalue was greater than 1 (2.25) and all the items loaded on one factor (Table 9). The Cronbach's alpha (0.73) suggests that the factor has an acceptable level of internal consistency (Nunnally 1978).

We then ran a confirmatory factor analysis in Mplus to test whether a single factor fitted the data. The model fit indices suggested a good fit according to the criteria of Zubrick et al. (2014), and all individual items showed good fit (Table 10). Factor loadings were moderately high (Figure 4). However, the factor loadings on the item 'There's at least one teacher or other adult in this school I can talk to if I have a problem' suggests that this item was less reflective of the PSSM factor than the other factors that reflected 'respect' and 'belonging'. Nonetheless, the measurement properties of this factor are good.

Table 9 Factor loadings and Cronbach's alpha for the PSSM scale, K cohort, wave 9

Variable	Loading
I feel proud of belonging to my school (icsc65_1)	0.79
I am treated with as much respect as other students (icsc65_2)	0.81
The teachers here respect me (icsc65_4)	0.83
There's at least one teacher or other adult in this school I can talk to if I have a problem (icsc65_5)	0.52
Cronbach's alpha	0.73

PSSM = Psychological Sense of School Membership

Note: The response format for the data items is (1) Completely true, (2) Somewhat true, (3) Neither, (4) Not very true, (5) Not at all true.

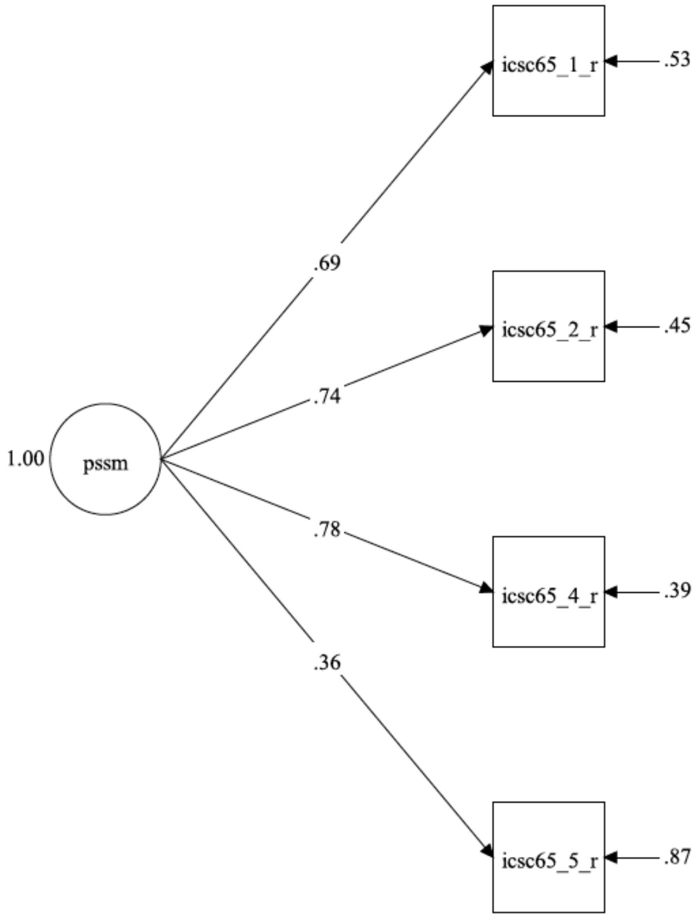
Table 10 Confirmatory factor analyses for the PSSM scale, wave 9

Index	Value
SRMR (<0.05 good fit, <0.10 acceptable fit)	**0.003**
RMSEA (<0.05 good fit, <0.08 acceptable)	**0.00**
CFI (>0.95 good fit, 0.90–0.95 acceptable)	**1.00**
TLI (>0.95 good fit, 0.90–0.95 acceptable)	**1.00**
Chi-square	0.121
df	2
P	0.94
N	493
Overall rating	Good
Estimator	ML

CFI = comparative fit index; df = degrees of freedom; ML = maximum likelihood; PSSM = Psychological Sense of School Membership; RMSEA = root-mean-square error of approximation; SRMR = standardised root mean residual; TLI = Tucker–Lewis index

Note: Bold numbers indicate good fit.

Figure 4 Standardised factor loadings for the PSSM, K cohort, wave 9



PSSM = Psychological Sense of School Membership

4.4 School climate

Cohen et al. (2009:182) suggested that school climate 'refers to the quality and character of school life. It is based on patterns of people's experiences of school life and reflects norms, goals, values, interpersonal relationships, teaching and learning practices, and organizational structures'. School climate scales are often measured in four major domains: physical and emotional safety; teaching and learning such as teaching quality, professional development and leadership; relationships such as respect for diversity, school community and collaboration and connectedness; and structural aspects such as cleanliness, space and extracurricular offerings (National School Climate Center 2014). In waves 8 and 9 of the LSIC, the K cohort were asked five items that mostly captured two aspects of school climate, safety and relationships. These two aspects are most commonly covered by school climate measures (Ramelow et al. 2015). However, this scale was adapted for use in the LSIC; hence, making any hypotheses about the specific factor structure is difficult.

We performed exploratory factor analysis using principal component analysis with a quartimax rotation in Stata 15.1 to test for factors in the school climate items. The analysis suggested a one-factor solution with only one eigenvalue above 1 (3.32). When interpreting a factor analysis of a potential scale with no source, the highest-loading items indicate the construct that the factor is measuring (Table 11). In this example, the two highest-loading items were 'My school has safe places' and 'My school is good for me'. However, all items loaded highly, which is reflected in the high Cronbach's alpha (0.87).

After identifying a one-factor solution, we formally tested this through confirmatory factor analysis in Mplus 8.1 (Table 12). Model fit indices suggested an acceptable model fit according to Zubrick et al. (2014), with the SRMR and the CFI indicating good fit but the TLI only acceptable. The RMSEA fit was not acceptable but we deemed the factor acceptable for research purposes given the results for the other indices.

The standardised factor loadings from the confirmatory factor analysis show a similar pattern to that of the exploratory factor analysis, with high loadings for all items (Figure 5). 'My school has safe places' and 'My school is good for me' were the highest-loading items.

Table 11 Factor loadings and Cronbach's alpha for school climate, K cohort, wave 9

Item	Factor loading
My school is good for me (icsc46_1)	0.83
My school has safe places (icsc46_2)	0.85
My school has people I trust (My school has good people) (icsc46_3)	0.81
My school has people who help each other (icsc46_4)	0.80
My school helps me learn (icsc46_50)	0.78
Cronbach's alpha	0.87

Notes:

1. *N* = 489.
2. The response format for the data items are (1) Yes: always, (2) Yes: most of the time, (3) Sometimes: fair bit, (4) Sometimes: little bit, (5) No: not much and (6) No: never.

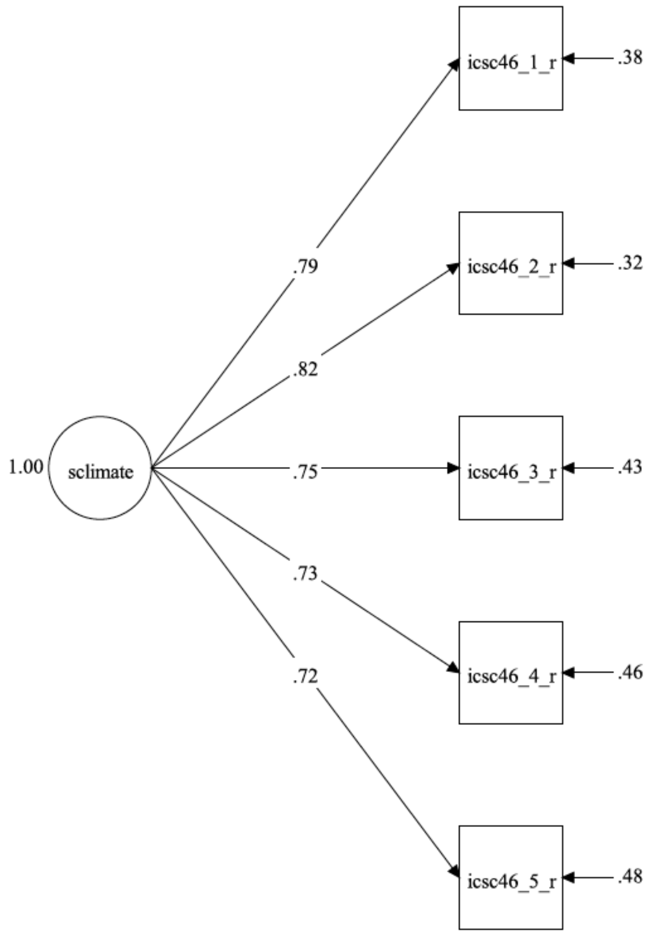
Table 12 Confirmatory factor analyses for school climate in wave 9

Index	Value
SRMR (<0.05 good fit, <0.10 acceptable fit)	**0.031**
RMSEA (<0.05 good fit, <0.08 acceptable)	0.14
CFI (>0.95 good fit, 0.90–0.95 acceptable)	**0.96**
TLI (>0.95 good fit, 0.90–0.95 acceptable)	**0.92**
Chi-square	50.906
df	5
P	0.00
N	494
Overall rating	Acceptable
Estimator	ML

CFI = comparative fit index; df = degrees of freedom; ML = maximum likelihood; RMSEA = root-mean-square error of approximation; SRMR = standardised root mean residual; TLI = Tucker–Lewis index

Note: Bold numbers represent good fit.

Figure 5 **Standardised factor loadings for school climate, K cohort, wave 9**



4.5 Teacher relationship

Good student–teacher relationships are associated with lower rates of disruptive behaviours, suspensions and dropouts, and with better school attendance, psychological engagement and academic achievement (Quin 2017). Longitudinal studies also show that student–teacher relationships tend to decline alongside attendance levels and behavioural and psychological engagement of students (Quin 2017).

The Teacher Relationship scale was developed in collaboration with Associate Professor Gawaian Bodkin-Andrews from the University of Technology Sydney. Items were derived from the Seeding Success project (Craven et al. 2013) (Table 13).

In wave 9 of the LSIC, B- and K-cohort children were asked four items capturing the student–teacher relationship. For both cohorts, we initially ran a principal component analysis with quartimax rotation (exploratory factor analysis) to determine whether the items had a factor structure (Table 14). Results from the B cohort suggested a one-factor solution (eigenvalue 2.22). Factor loadings suggested that all the items asked of the B cohort load similarly. Items about teachers being 'fair' and 'listens to me' had the highest-loadings. K-cohort results also suggested a one-factor solution (eigenvalue 2.51). The highest-loading items were 'listens to me' and 'care about me and want me to do well at school'. Cronbach's alphas for both cohorts were acceptable (Nunnally 1978).

We formally tested the one-factor solution for these sets of items through confirmatory factor analysis in Mplus 8.1. Results from these analyses suggested good model fit according to the Zubrick et al. (2014) criterion, and individually for the SRMR, the CFI and the TLI (Table 15). For the B cohort, the RMSEA also indicated a good fit.

Table 13 Items in Teacher Relationship scale and source

Variable	Source
My teacher listens to me (csc24)	A variant of this item (My teacher listens to what I have to say) was used within the factor on teacher rapport within the Seeding Success project
My teachers make sure my class is a fun place to be (csc28)	
My teachers care about me and want me to do well at school (csc25)	A variant of this item (My teacher cares about me) was used within the factor on teacher rapport within the Seeding Success project
My teachers are fair to me (csc26)	This item was used within the factor on teacher rapport within the Seeding Success project
My teachers understand how I talk (csc53)	This item was used within the factor on learning fun within the Seeding Success project

Note: The response format for the data items in the Teacher Relationship scale are (1) Yes: always, (2) Yes: little bit, (3) Sometimes: more yes, (4) Sometimes: more no, (5) No: not much and (6) No: never.

Table 14 Factor loadings from a principal component analysis and Cronbach's alpha, wave 9, by cohort

Variable	B cohort	K cohort
My teacher listens to me (icsc24a)	0.74	0.81
My teachers make sure my class is a fun place to be (icsc28a)	0.72	0.79
My teachers care about me and want me to do well at school (icsc25a)	0.72	0.83
My teachers are fair to me (icsc26a)	0.79	na
My teachers understand how I talk (icsc53)	na	0.75
Cronbach's alpha	0.73	0.79
N	740	481

na = not applicable

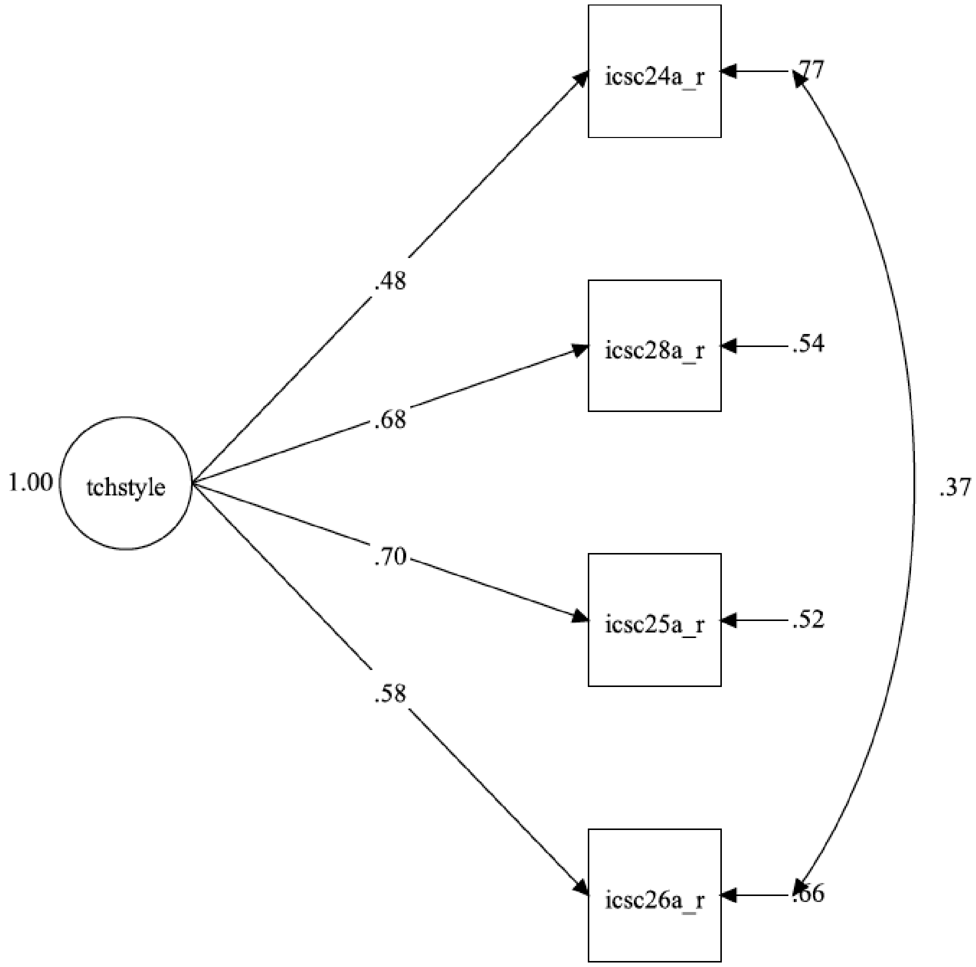
Table 15 Confirmatory factor analyses of teacher relationship, wave 9, by cohort

Variable	B cohort	K cohort
SRMR (<0.05 good fit, <0.10 acceptable fit)	**0.006**	**0.02**
RMSEA (<0.05 good fit, <0.08 acceptable)	**0.023**	0.096
CFI (>0.95 good fit, 0.90–0.95 acceptable)	**0.999**	**0.99**
TLI (>0.95 good fit, 0.90–0.95 acceptable)	**0.996**	**0.96**
Chi-square	1.397	11.073
df	1	2
P	0.2373	0.0039
N	743	495
Overall rating	Good	Good
Estimator	ML	ML

CFI = comparative fit index; df = degrees of freedom; ML = maximum likelihood; RMSEA = root-mean-square error of approximation; SRMR = standardised root mean residual; TLI = Tucker–Lewis index

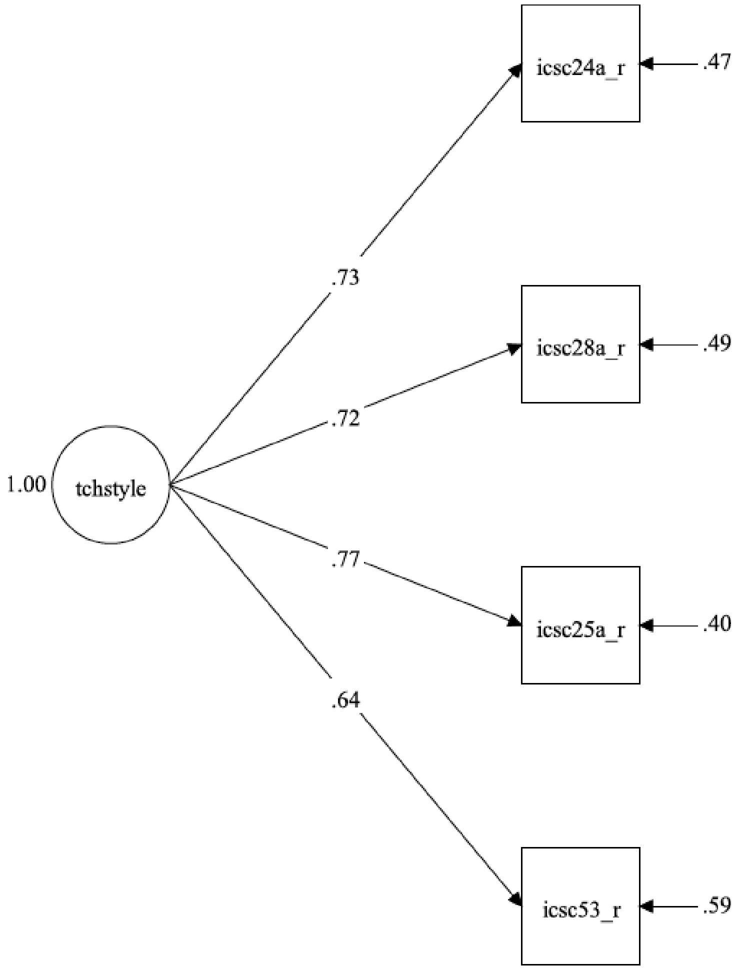
Note: Bold numbers indicate good fit.

Figure 6 **Standardised factor loadings for teacher style, B cohort, wave 9**



For the B cohort, standardised factor loadings suggested that the factor was most indicative of two items: 'Teachers care about me and want me to do well at school' and 'Teachers make sure my class is a fun place to be' (Figure 6). Similarly for the K cohort, the item 'Teachers care about me and want me to do well at school' correlated most strongly with the factor overall, followed by 'My teachers listen to me' and 'making the class a fun place to be' (Figure 7).

Figure 7 Standardised factor loadings for teacher style, K cohort, wave 9



4.6 Pianta Student–Teacher Relationship scale

Section 4.5 discussed the importance of the student–teacher relationship. In the LSIC, teacher responses regarding relationship quality with the study child were measured by the 15-item STRS-short form (STRS-SF; Pianta 1992). The scale was designed to assess teacher's feelings and beliefs about the student's behaviour toward them. The short-form has two subscales: conflict (7 items) and closeness (8 items). Conflict measures the degree to which the teacher perceives his or her relationship with a student as negative and high in conflict. Higher scores on this scale suggest that the teacher perceives the student as angry and unpredictable, resulting in the teacher feeling emotionally drained and believing they are ineffective. High-conflict relationships as measured by the STRS lead to increased rates of school absence and affect academic achievement (DiLalla et al. 2004, Hamre & Pianta 2007). Closeness measures the degree to which a teacher experiences affection, warmth and open communication with a particular student. Greater closeness as measured by the STRS is associated with fewer behavioural problems, greater social skills and more positive feelings about school (Pianta & Stuhlman 2004, Buyse et al. 2008).

Initially, we conducted exploratory factor analyses to test the underlying factor structure. We used principal component analysis with a quartimax rotation using Stata 15.1. We used an orthogonal rotation (quartimax rotation) because the factor loadings are generally easier to interpret and report (Tabachnick & Fidell 2007). To determine the number of factors, we used eigenvalues above 1 and the scree plot.

For the B cohort waves 6–8, the eigenvalues indicated a two-factor solution[6] (Table 16). All factor loadings aligned with the hypothesised factor structure with a subscale for conflict and another for closeness. The exception was item 4 'This child is uncomfortable with physical affection or touch from me', which negatively loaded on closeness for waves 6–8. Item 4 has been removed from wave 9 onward because it is no longer appropriate for children as they reach high school and because of negative responses from interviewees on the wording of the question. Those using waves 6–8 of the data should also consider excluding this item from their analysis.

Both subscales had very good internal consistency in all three waves, with Cronbach's alphas from 0.81 to 0.92 (note that we calculated Cronbach's alphas for closeness removing item 4). For the B cohort at wave 9, the initial exploratory factor analysis suggested a three-factor solution (eigenvalues: 5.12, 3.15, 1.22); however, the third factor was largely driven by item 4. When the exploratory analysis was re-run, the eigenvalues (5.24, 3.12) suggested a two-factor solution. The factor loadings were consistent with conflict and closeness factors, and the Cronbach's alphas were very good.

For the K cohort, we estimated exploratory factor analyses for waves 6–8. For waves 6 and 7, eigenvalues[7] and scree plots suggested a two-factor solution. The factor loadings were consistent with closeness and conflict factors (Table 17), with item 4 again not loading positively on either factor. For wave 8, the initial factor analysis suggested a three-factor solution (eigenvalues: 5.05, 3.64, 1.22); however, this third factor was largely driven by item 4. We removed this item and the eigenvalues (5.83, 2.99) were consistent with a two-factor solution. Again, the factor loadings were consistent with conflict and closeness. The internal consistency of all the factors was very good for waves 6–8, from 0.85 to 0.92 (Nunnally 1978).

Table 16 Exploratory factor analyses of the STRS-SF in waves 6–9, B cohort

	Wave 6		Wave 7		Wave 8		Wave 9	
Item	Conflict	Closeness	Conflict	Closeness	Conflict	Closeness	Conflict	Closeness
1. I share an affectionate, warm relationship with child (dcc33_1)	−0.31	0.49		0.71		0.67	−0.30	0.52
2. This child and I always seem to be struggling with each other (dcc33_2)	0.76		0.67		0.77		0.74	
3. If upset, this child will seek comfort from me (dcc33_3)	−0.30	0.66		0.76		0.70	0–.31	0.70
4. This child is uncomfortable with physical affection or touch from me (dcc33_4)		−0.43		−0.36		−0.32	-	-
5. This child values his/her relationship with me (dcc33_5)		0.68		0.63		0.69	−0.32	0.65
6. When I praise this child, he/she beams with pride (dcc33_6)		0.59		0.57		0.57		0.52
7. This child spontaneously shares information about himself/herself (dcc33_7)		0.75		0.67		0.73		0.73
8. This child easily becomes angry with me (dcc33_8)	0.80		0.78		0.79		0.87	
9. It is easy to be in tune with what this child is feeling (dcc33_9)	−0.39	0.65		0.60		0.68		0.64
10. This child remains angry or resistant after being disciplined (dcc33_10)	0.78		0.78		0.86		0.81	
11. Dealing with this child drains my energy (dcc33_11)	0.82		0.80		0.82		0.83	
12. When this child is in a bad mood, I know we're in for a long and difficult day (dcc33_12)	0.83		0.86		0.87		0.88	
13. This child's feelings towards me can be unpredictable or can change suddenly (dcc33_13)	0.87		0.84		0.86		0.89	
14. This child is manipulative with me (dcc33_14)	0.77		0.75		0.77		0.81	
15. This child openly shares his/her feelings and experiences with me (dcc33_15)		0.69		0.71		0.77		0.81
Cronbach's alpha	0.82	0.89a	0.81	0.90a	0.82	0.92a	0.81	0.93[a]

STRS-SF = Student–Teacher Relationship Scale-short form

a Cronbach's alpha calculated without item 4

Notes:

1. The response format for these data items is (1) Definitely does not apply, (2) Not really, (3) Neutral/Not sure, (4) Applies somewhat, and (5) Definitely applies.
2. Blank cells are loadings <0.30.

Table 17 Exploratory factor analyses of the STRS-SF in waves 6–8, K cohort

	Wave 6		Wave 7		Wave 8	
Item	Conflict	Closeness	Conflict	Closeness	Conflict	Closeness
1. I share an affectionate, warm relationship with child (dcc33_1)		0.63	−0.30	0.75		0.68
2. This child and I always seem to be struggling with each other (dcc33_2)	0.72		0.74		0.72	−0.33
3. If upset, this child will seek comfort from me (dcc33_3)		0.72		0.73		0.78
4. This child is uncomfortable with physical affection or touch from me (dcc33_4)		−0.40				-
5. This child values his/her relationship with me (dcc33_5)		0.73		0.78		0.76
6. When I praise this child, he/she beams with pride (dcc33_6)		0.61		0.69		0.62
7. This child spontaneously shares information about himself/herself (dcc33_7)		0.71		0.78		0.78
8. This child easily becomes angry with me (dcc33_8)	0.79		0.84		0.77	
9. It is easy to be in tune with what this child is feeling (dcc33_9)		0.68		0.65		0.63
10. This child remains angry or resistant after being disciplined (dcc33_10)	0.80		0.82		0.78	
11. Dealing with this child drains my energy (dcc33_11)	0.82		0.85		0.88	
12. When this child is in a bad mood, I know we're in for a long and difficult day (dcc33_12)	0.89		0.82		0.92	
13. This child's feelings towards me can be unpredictable or can change suddenly (dcc33_13)	0.86		0.83		0.88	
14. This child is manipulative with me (dcc33_14)	0.82		0.79		0.84	
15. This child openly shares his/her feelings and experiences with me (dcc33_15)		0.83		0.79		0.79
Cronbach's alpha	0.86	0.92	0.86	0.92	0.85	0.92

STRS-SF = Student–Teacher Relationship Scale-short form

Notes:

1. The response format for these data items is (1) Definitely does not apply, (2) Not really, (3) Neutral/Not sure, (4) Applies somewhat and (5) Definitely applies.
2. Blank cells are loadings <0.30.

We then estimated a confirmatory factor analysis using Mplus 8.2 for the B cohort (Table 18). We specified a two-factor structure consistent with the intended item loadings of the original STRS-SF.

In wave 6, while the model fit was adequate χ^2 (75) = 352.51, $P < 0.001$; (RMSEA = 0.10; CFI = 0.95; TLI = 0.95; SRMR = 0.08), the modification indices suggested that item 4 ‘This child is uncomfortable with physical affection or touch from me’ should also load on closeness ($\chi^{2\Delta}$ = 99.46). This cross-loading provides further evidence that this item is not a ‘pure’ measure of conflict or closeness. Therefore, we decided to re-run the model excluding this item (see above). The model fit improved after the re-run. Using Zubrick et al. (2014) criteria, the model fit indices suggest a good model fit.

In wave 7, we had similar issues with item 4. Again, the modification indices suggested that item 4 should load on closeness ($\chi^{2\Delta}$ = 72.32), although model fit was adequate (χ^2 (89) = 305.28, $P < 0.001$; RMSEA = 0.09; CFI = 0.96; TLI = 0.96; SRMR = 0.08). Given the difficulties with this particular item, we dropped it from analyses and re-ran the model. After this modification, the model fit was good according to our criteria.

In wave 8, we encountered similar issues with item 4 to those in previous waves and cross-loadings (χ^2 (89) = 318.55, $P < 0.001$; RMSEA = 0.09; CFI = 0.97; TLI = 0.96; SRMR = 0.07), and again removed this item. Model fit improved ($\chi^{2\Delta}$ = 89.64) and the fit indices indicated good model fit.

In wave 9, we had similar issues with item 4. The modification indices suggested that item 4 should load on closeness ($\chi^{2\Delta}$ = 81.82), although model fit was adequate ($\chi^{2\Delta}$ (89) = 471.78, $P < 0.001$; RMSEA = 0.11; CFI = 0.96; TLI = 0.95; SRMR = 0.08). After item 4 was removed, the model fit was good using the Zubrick et al. (2014) criteria.

The consistency in model fit for the final model across the three waves is encouraging. However, to establish the validity of the measure for longitudinal analyses requires further information about factorial invariance over time.

All the standardised factor loadings were above 0.82 for conflict. For closeness, they were 0.68 or above for wave 6 (Figure 8). For conflict, the highest-loading item was item 13 (This child’s feelings toward me can be unpredictable or can change suddenly) and for closeness, item 5 (This child values his/her relationship with me). Conflict and closeness were negatively correlated.

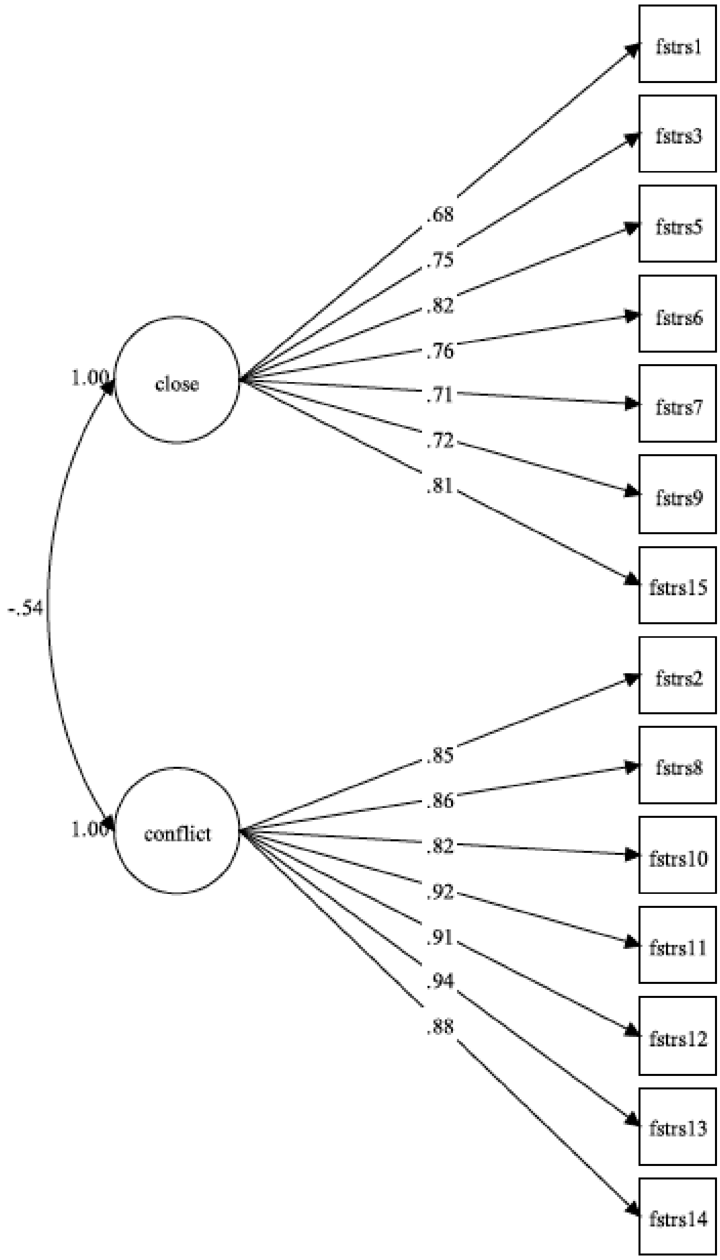
Table 18 Confirmatory factor analyses of the STRS-SF by wave, B cohort

	Wave			
Variable	**6**	**7**	**8**	**9**
SRMR (<0.05 good fit, <0.10 acceptable fit)	**0.07**	**0.066**	**0.06**	**0.08**
RMSEA (<0.05 good fit, <0.08 acceptable)	0.086	**0.076**	**0.077**	0.11
CFI (>0.95 good fit, 0.90 to 0.95 acceptable)	**0.97**	**0.98**	**0.98**	**0.97**
TLI (>0.95 good fit, 0.90 to 0.95 acceptable)	**0.97**	**0.97**	**0.98**	**0.96**
Chi-square	236.55	218.90	211.01	380.52
df	76	76	76	76
P	0.0001	0.0001	0.0001	0.0001
N	283	326	302	353
Overall rating	Good	Good	Good	Good
Estimator	WLSMV	WLSMV	WLSMV	WLSMV

CFI = comparative fit index; df = degrees of freedom; RMSEA = root-mean-square error of approximation; SRMR = standardised root mean residual; STRS-SF = Student–Teacher Relationship Scale-short form; TLI = Tucker–Lewis index; WLSMV = weighted least square mean variance

Note: Bold numbers indicate good fit.

Figure 8 **Factor loadings on closeness and conflict for the STRS-SF, B cohort, wave 6**

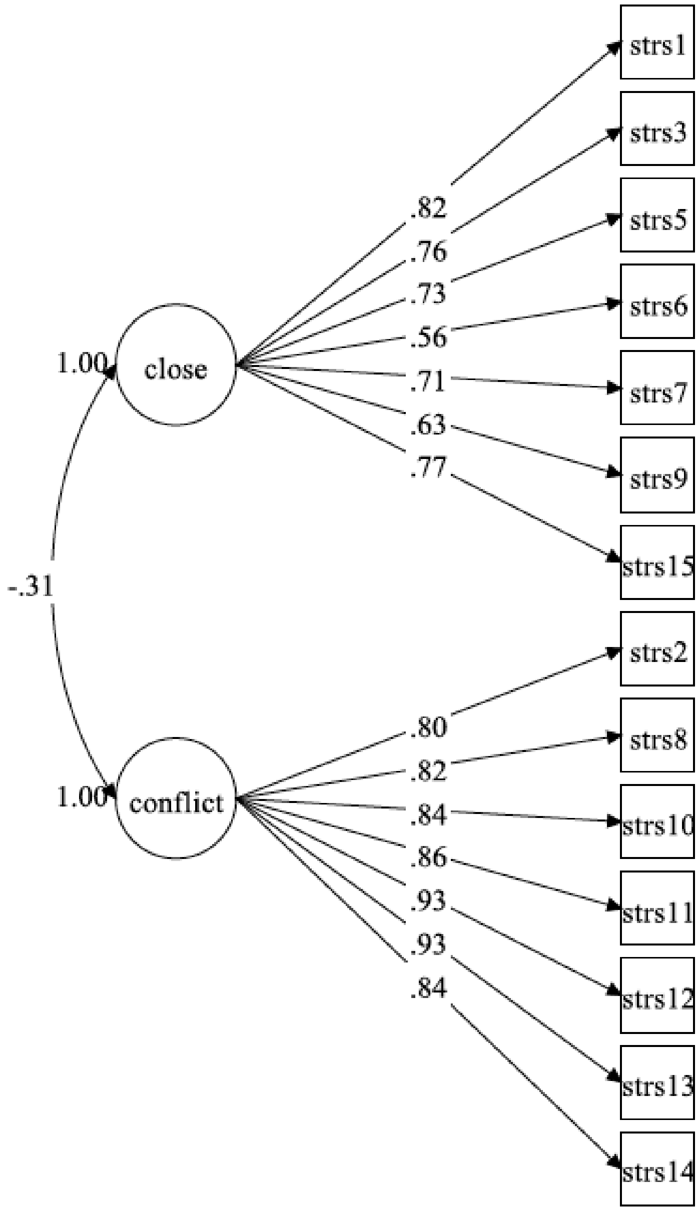


STRS-SF = Student–Teacher Relationship Scale-short form

Conflict and closeness were negatively correlated in wave 7 (Figure 9). All the standardised factor loadings were above 0.80 for conflict, with the highest-loading items 'This child's feelings toward me can be unpredictable or can change suddenly' and 'When this child is in a bad mood, I know we're in for a long and difficult day'. For closeness, standardised factor loadings ranged from 0.56 to 0.82. The highest factor loading was 'I share an affectionate, warm relationship with this child', and the lowest 'When I praise this child, he/she beams with pride'.

For wave 8, all the standardised factor loadings were above 0.83 for conflict and 0.63 or above for closeness (Figure 10). For conflict, the highest-loading item was 'This child's feelings toward me can be unpredictable or can change suddenly', and for closeness was 'It is easy to be in tune with what this child is feeling'. Closeness and conflict were negatively correlated (r = −0.27).

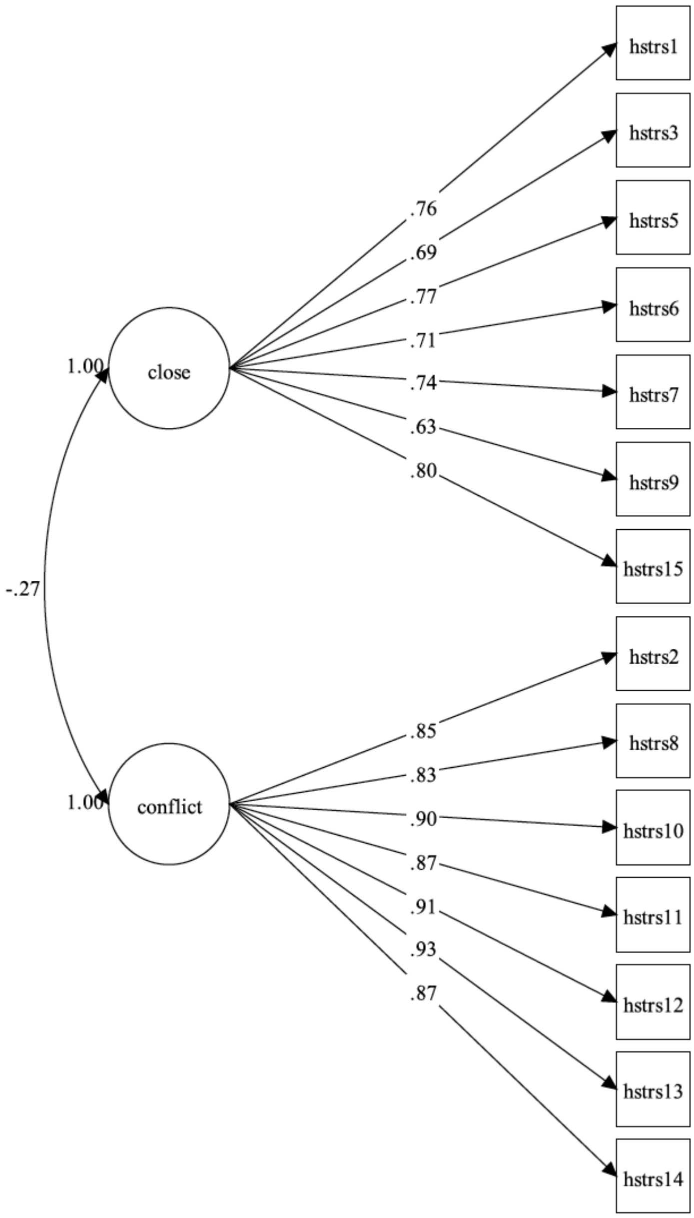
Figure 9 **Factor loadings on closeness and conflict for the STRS-SF, B cohort, wave 7**



STRS-SF = Student–Teacher Relationship Scale-short form

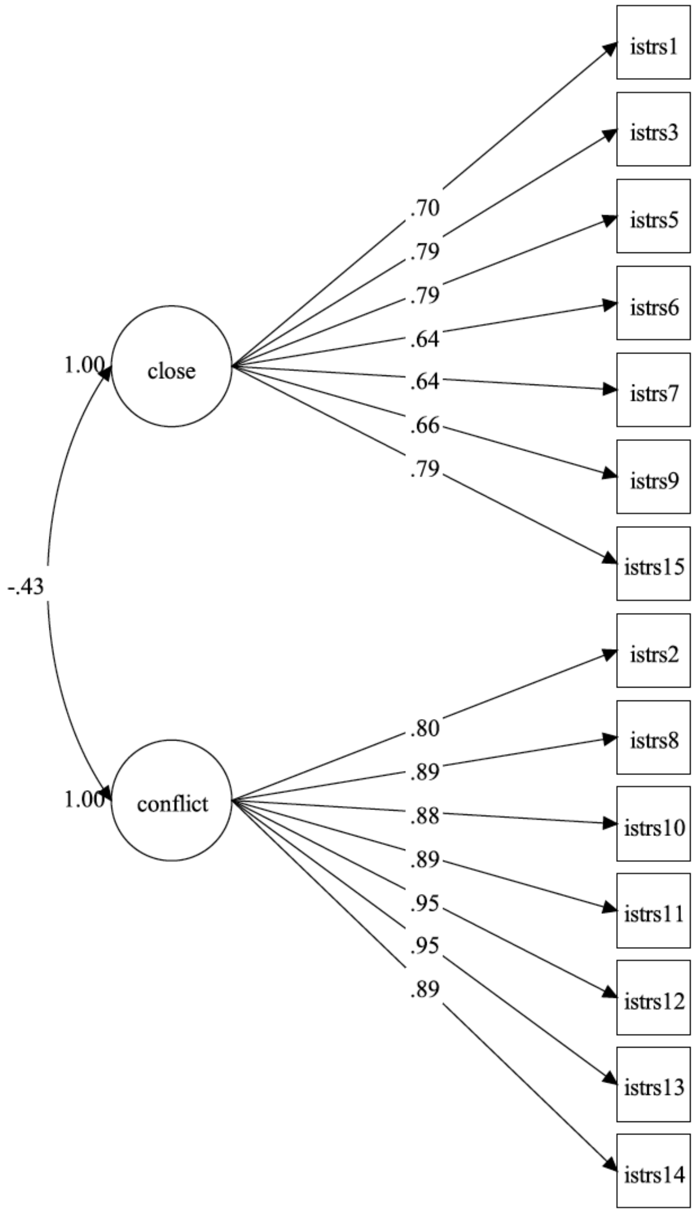
For wave 9, the factor loadings showed similar patterns to those of waves 6–8. All the standardised factor loadings were above 0.80 for conflict (Figure 11). The highest-loading items were ‘This child’s feelings toward me can be unpredictable or can change suddenly’ and ‘When this child is in a bad mood, I know we’re in for a long and difficult day’. As in previous waves, standardised factor loadings for closeness were lower than for conflict. Factor loadings for closeness ranged from 0.64 to 0.79. The highest-loading items were ‘If upset, this child will seek comfort from me’, ‘The child values his/her relationship with me’ and ‘This child openly shares his/her feelings and experiences with me’.

Figure 10 Factor loadings on closeness and conflict for for the STRS-SF, B cohort, wave 8



STRS-SF = Student–Teacher Relationship Scale-short form

Figure 11 **Factor loadings on closeness and conflict for the STRS-SF, B cohort, wave 9**



STRS-SF = Student–Teacher Relationship Scale-short form

4.7 Construct validity

As outlined in Section 4.1, construct validity is the degree to which the measure captures the essential features of the concept it is intended to measure (Cronbach & Meehl 1955). Convergent validity occurs when two different measures of the same construct are highly correlated. Low correlations between unlike constructs suggest discriminant validity (Kvien et al. 1998). One way to assess convergent and discriminant validity is to test the correlations between like (convergent) and unlike (discriminant) measures in a correlation matrix. In practice, assessment of convergent and discriminant validity relies on subjective interpretations of correlation matrices and the degree to which the theoretical constructs are similar (and correlated) or different (and uncorrelated). These assessments can also be confounded or obscured by shared method variance, which is the phenomenon where two measures of a different construct may have a higher correlation because they are assessed by the same respondent (e.g. study child or parent or teacher), or method (direct assessment).

Higher intercorrelations may also occur when the respondent is considering the same target (teacher or school or themselves).

The education scales applied in the LSIC ask the study children to rate three different targets. First, they rate themselves through SLAQ and the maths and reading academic self-concepts. Second, they rate the teacher through the Teacher Relationship Scale. Third, they rate the school through the PSSM and school climate scales. Therefore, we expect higher intercorrelations for measures that share the same target than for measures with different targets.

In the correlation matrices below, we have included measures of the extent to which teachers rated the study child as a fluent reader, the PAT Maths scale score, and the teacher-rated Strengths and Difficulties Questionnaire (SDQ) total score. We expect education scales more theoretically relevant to these outcomes to correlate more highly with these measures (e.g. reading academic self-concept with teacher-rated reading fluency) and less highly with unlike measures. In the correlation matrices, higher scores on the education scales suggest worse outcomes. For example, a higher score on SLAQ indicates higher school avoidance and lower levels of school liking. In addition, measures of teacher-rated conflict and cohesion only appear for the B cohort because we were unable to estimate confirmatory factor analyses for the K cohort.

Given the large differences among Aboriginal and Torres Strait Islander populations, conceptualisations of education probably vary among groups, particularly those living in remote areas of Australia (Maher 2010). Conceptualisations may also vary with particular educational constructs. Therefore, we present correlations for both those living in remote areas and the overall sample.

Table 19 shows correlations for the K cohort in wave 9. For SLAQ, a measure of affective disengagement from school, the highest positive correlations were for measures of poorer ratings of school climate (PSSM, school climate), a poorer teacher relationship, and poorer reading academic self-concept. As expected, poorer teacher ratings of reading fluency were associated with more affective disengagement (a negative relationship with SLAQ). According to the SDQ, more problems were associated with more school avoidance.

The maths academic self-concept correlated best with the other academic self-concept measure – reading (supporting its convergent validity). The maths academic self-concept was not correlated with teacher ratings of reading, which supports its discriminant validity. This self-concept also correlated with measures of school climate and teacher relationship ($r_s = 0.24–0.28$). Its nonsignificant relationship with the PAT Maths is surprising and contrasts with findings of previous longitudinal research with Indigenous students in Sydney. Craven et al. (2013) found an association between maths self-concept and maths achievement. However, they found that maths self-concept does not predict maths achievement; rather, maths achievement predicts maths self-concept.

For reading academic self-concept, correlations with the PSSM, school climate and teacher relationships were unexpectedly higher than with maths academic self-concept. However, higher reading academic self-concept (i.e. lower scores) correlated with higher teacher ratings of reading fluency (supporting convergent validity) but not with maths achievement (discriminant validity). The association of higher levels of reading academic self-concept and teacher ratings of reading fluency is consistent with findings from longitudinal studies that suggest reading self-concept is associated with reading achievement (Craven et al. 2013). A stronger relationship existed between behavioural problems (determined through the SDQ total scores) and poor reading academic self-concept than between behavioural problems and reading fluency. This finding may be due to measurement error because the reading rating of fluency is based on a single item, whereas the SDQ total score is based on 25 items; more items in a construct can reduce measurement error. Moreover, the SDQ total score has the subscales of hyperactivity ($r = 0.15$, $P < 0.0001$) and conduct problems ($r = 0.14$, $P < 0.0001$), which may interact with children's reading self-concept. Consistent with these findings, other research suggests that lower levels of reading self-concept are associated with a greater desire to truant and lower classroom participation. Both of these concepts relate to hyperactivity (i.e. classroom participation) and conduct problems (i.e. truancy) (Craven et al. 2013).

The PSSM scale correlated most highly with school climate (*r* = 0.71). The PSSM scale also correlated highly with study children's ratings of their relationship with teachers (*r* = 0.63), which is consistent with the view that the PSSM captures the school environment more generally. The high correlations (r_s = 0.63) between the Teacher Relationship scale and measures of school climate support its convergent validity. In contrast, ratings of school avoidance and maths and reading academic self-concepts correlated only moderately with the PSSM scale (*r* = 0.24–0.38). Given the high correlation between school climate and the PSSM, correlations with other education scales and child outcomes (reading fluency, maths achievement and SDQ total problems) were very similar.

The student-rated measure of Teacher Relationship had highest associations with measures of school climate (PSSM, school climate), followed by reading academic self-concept, school avoidance and maths academic self-concept. The Teacher Relationship measure did not significantly correlate with any child outcomes (reading fluency, maths achievement or SDQ total problems).

For children in remote areas, the pattern of associations between education measures was broadly similar to the estimates for the whole sample on most measures, except for school avoidance in SLAQ (Table 20). A very different picture emerged for school avoidance in remote areas, with only maths and reading academic self-concept having significant associations. None of the other education measures or school climate were significantly associated with affective disengagement. For remote areas, school avoidance appears unrelated to school climate measures (PSSM, school climate) and the teacher relationship. Therefore, other family and community factors may drive disengagement from school (see Daraganova et al. 2014).

Next, we turn to the construct validity of B cohort measures in wave 9. For the B cohort, we estimated scores for the Student–Teacher Relationship scale, which consists of two subscales: closeness and conflict. For both these subscales, higher scores indicate more closeness and more conflict.

Higher levels of school affective disengagement (SLAQ) were associated with poorer teacher relationship as rated by study children, and

Table 19 Correlations between education scales, teacher ratings of reading, and SDQ and PAT Maths scale scores, K cohort

Variable	1	2	3	4	5	6	7	8
1. SLAQ								
2. Maths academic self-concept	0.17***							
3. Reading academic self-concept	0.26***	0.34***						
4. PSSM scale	0.25***	0.28***	0.38***					
5. School climate	0.29***	0.28***	0.41***	0.71***				
6. Teacher relationship	0.27***	0.24***	0.38***	0.63***	0.63***			
7. Reads fluently (teacher)	−0.21**	−0.02	−0.16*	0.00	−0.07	−0.10		
8. PAT Maths	−0.08	−0.05	−0.03	−0.07	−0.05	−0.06	0.30***	
9. SDQ total (teacher)	0.14**	0.11*	0.18***	0.10*	0.14**	0.14	−0.38***	−0.01

*** = *P* < 0.01; ** = *P* < 0.05; * = *P* < 0.1; PSSM = Psychological Sense of School Membership; SDQ = Strengths and Difficulties Questionnaire; SLAQ = School Liking and Avoidance Questionnaire

Table 20 Correlations between education scales, teacher ratings of reading, and SDQ and PAT Maths scale scores for children in remote areas only, K cohort

Variable	1	2	3	4	5	6	7	8
1. SLAQ								
2. Maths academic self-concept	0.20***							
3. Reading academic self-concept	0.28***	0.57***						
4. PSSM scale	0.07	0.43***	0.41***					
5. School climate	0.05	0.45***	0.35***	0.58***				
6. Teacher relationship	0.13	0.44***	0.36***	0.43***	0.48***			
7. Reads fluently (teacher)	−0.09	−0.16	0.17	−0.02	−0.09	−0.15		
8. PAT Maths	−0.09	0.13	0.10	0.09	0.06	0.09	0.23**	
9. SDQ total (teacher)	0.16	0.18*	0.07	−0.04	0.08	0.03	−0.36***	0.02

*** = P < 0.01; ** = P < 0.05; * = P < 0.1; PSSM = Psychological Sense of School Membership; SDQ = Strengths and Difficulties Questionnaire; SLAQ = School Liking and Avoidance Questionnaire

greater levels of conflict with teachers (STRS-conflict) (Table 21). Lower levels of maths and reading academic self-concepts were associated with higher levels of school affective disengagement. As expected, students with high levels of affective disengagement from school also had more behaviour problems (as measured by SDQ total problems).

Similar to the K cohort, the reading academic self-concept of the B cohort most strongly related to the maths academic self-concept (evidence of convergent validity). Higher levels of reading self-concept were also associated with higher ratings of the study children's Teacher Relationship but not with teacher-rated closeness and conflict. High reading self-concept (lower scores) was also associated with better reading fluency as rated by teachers (convergent validity). Low reading self-concept was associated with behavioural problems (SDQ total problems), consistent with findings from the K cohort and previous literature. In support of its discriminant validity, reading self-concept was not associated with PAT Maths scores.

Again similar to the K cohort, for the B cohort maths self-concept was associated with reading self-concept but not with PAT Maths scores. This finding also aligns with findings from a longitudinal study of Indigenous children in Sydney (Craven et al. 2013). Although maths self-concept was associated with study children's ratings of their relationship with teachers, it was not associated with teachers' ratings of their level of closeness or conflict with study children.

As expected, a more positive teacher relationship as rated by the study child was associated with the teacher reporting feeling closer (STRS-closeness) and having less conflict with the student (STRS-conflict). Better teacher relationships were also associated with higher academic self-concepts (maths and reading), higher maths achievement and lower levels of behavioural problems.

Higher scores of closeness of teachers (STRS-closeness) with students were associated with better student-reported teacher relationships and lower levels of teacher-reported conflict (STRS-conflict) but not with academic self-concept. Higher levels of student–teacher closeness were associated with greater reading fluency and lower levels of behavioural problems (SDQ total problems). Conversely, higher levels of student-teacher conflict were associated with lower reading fluency, lower maths achievement and more behavioural problems. Student–teacher conflict was more strongly associated with these child outcomes than with any other education measure.

Table 21 Correlations between education scales, teacher ratings of reading, and SDQ and PAT Maths scale scores, B cohort, wave 9

Variable	1	2	3	4	5	6	7	8
1. SLAQ								
2. Maths academic self-concept	0.20***							
3. Reading academic self-concept	0.28***	0.34***						
4. Teacher relationship	0.32***	0.32***	0.34***					
5. STRS-closeness	−0.07	−0.02	−0.07	−0.16**				
6. STRS-conflict	0.17**	−0.00	0.10	0.15**	−0.31***			
7. Reads fluently (teacher)	−0.07	−0.06	−0.26***	−0.15**	0.23***	−0.28***		
8. PAT Maths	−0.07	−0.02	−0.07	−0.07*	0.11	−0.20***	0.40***	
9. SDQ total	0.08*	0.01	0.09*	0.10**	−0.29***	0.76***	−0.40***	−0.01

*** = *P* < 0.01; ** = *P* < 0.05; * = *P* < 0.1; SDQ = Strengths and Difficulties Questionnaire; SLAQ = School Liking and Avoidance Questionnaire; STRS = Student–Teacher Relationship Scale

Again, we examine children living in remote areas of Australia separately given that the social and cultural environments can differ markedly from those in other areas. The correlations were consistent with findings for the K cohort (Table 22). Associations among B cohort education measures were largely similar in remote and nonremote areas with two exceptions:

- For children living in remote areas, school affective disengagement (SLAQ) and academic self-concepts (maths and reading) were not associated.
- For children living in remote areas, higher teacher-rated closeness was associated with higher maths self-concept, which differs from the full sample result.

Several associations between education measures and child outcomes (reading fluency, maths achievement and behavioural problems) differed between the full sample and remote areas:

- Reading self-concept was not associated with reading fluency or behavioural problems in remote areas.
- Study children's ratings of their relationship with their teacher was not associated with behavioural problems in remote areas.
- Teacher-rated conflict with the study child was not associated with reading fluency and maths achievement in remote areas, although it was in the full sample.

In general, our findings suggest that the education measures displayed good convergent and discriminant validity. In particular, the school climate measures (PSSM and school climate) and student-rated teacher relationship correlated strongly. Higher levels of school affective disengagement (SLAQ) were associated with a poorer school climate and student-rated teacher relationship, and more conflict with teachers and more behavioural problems (SDQ). We were only able to estimate factor scores for the STRS for the B cohort, but it also correlated with student ratings of teacher relationships and academic self-concept. In addition, maths and reading academic self concepts were correlated with school climate measures and student-rated teacher relationships (but not teachers' ratings). Except for a statistically significant relationship in the B cohort of reading fluency with reading self-concept, measures of academic self-concept were not related to outcomes (reading fluency, PAT Maths and behavioural problems). Mostly, these student outcomes showed only small correlations with academic measures, except

Table 22 Correlations between education scales, teacher ratings of reading, and SDQ and PAT Maths scale scores for children living in remote areas only, B cohort, wave 9

Variable	1	2	3	4	5	6	7	8
1. SLAQ								
2. Maths academic self-concept	0.12							
3. Reading academic self-concept	0.12	0.48***						
4. Teacher relationship	0.17*	0.40***	0.34***					
5. STRS-closeness	−0.07	−0.24*	−0.09	−0.21				
6. STRS-conflict	0.05	0.15	−0.12	0.08	−0.27*			
7. Reads fluently (teacher)	−0.14	−0.03	−0.12	−0.16**	0.42***	−0.13		
8. PAT Maths	−0.04	−0.04	−0.05	−0.05*	0.15	−0.02	0.37**	
9. SDQ total	0.07	0.01	0.01	0.02	−0.42***	0.70***	−0.39***	0.00

*** = $P < 0.01$; ** = $P < 0.05$; * = $P < 0.1$; SDQ = Strengths and Difficulties Questionnaire; SLAQ = School Liking and Avoidance Questionnaire; STRS = Student–Teacher Relationship Scale

for the STRS, which had the strongest and most consistent relationships with outcomes. The STRS is one of the few education measures consistently collected over the waves, and thus may be particularly important in understanding Indigenous children's learning.

The education measures displayed similar patterns of associations for children living in remote areas, except for school affective disengagement (SLAQ). For children in remote areas, associations between SLAQ and other education measures were lower and few were statistically significant, including those for child outcomes (SDQ, PAT Maths and reading fluency). In addition, the patterns of correlations were similar in the B and K cohorts,[8] which suggests that child-reported affective disengagement has a qualitatively different meaning for children in remote areas. Hence, data users should exercise caution when using SLAQ for research that includes children from remote areas.

5 Cross-sectional, multivariate analysis

The previous section examined the extent to which the data items within the education scales were related to each other in the way we expected, based on how they are constructed, theory and existing evidence. There is much interest as to what factors these (and other) education measures are correlated with. In this section we present multivariate analysis of the relationship between demographic, family, community, socioeconomic and geographic variables. Section 6 will present how the education measures change over time.

Ideally, a number of explanatory variables should have a statistically significant relationship with the education outcome variables. In addition, the sign and size of the associations should vary across the dependent variables, and associations that differ from those found in the literature should be explainable

5.1 Explanatory variables

Given the relatively small sample size, we use a parsimonious model, which is kept consistent across the education variables used for analysis. A total of 16 explanatory variables were included: 13 binary, 2 categorical and 1 continuous (Table 23). Depending on the availability of data

Table 23 Variables used in multivariate modelling, wave 8

Variable	Type	Mean/ proportion	Minimum	Maximum
Female	Binary	0.507	0	1
Age	Continuous	9.286	7	12
Level of relative isolation – low	Binary	0.517	0	1
Level of relative isolation – medium	Binary	0.135	0	1
Level of relative isolation – high	Binary	0.087	0	1
Decile of Indigenous relative socioeconomic outcomes	Categorical	5.637	1	10
Parent 1 employed	Binary	0.435	0	1
Partner of parent 1 employed	Binary	0.404	0	1
Parent 1 does not have a partner	Binary	0.399	0	1
Family has relatively low income (under $600 per week)	Binary	0.256	0	1
Family has relatively high income ($1000 or more per week)	Binary	0.369	0	1
Living in a house owned or being purchased by residents	Binary	0.232	0	1
Living in a house rented through the private rental market	Binary	0.208	0	1
School grade at time of survey	Categorical	3.631	1	8
Moved school since last survey	Binary	0.102	0	1
Parent 1 has completed year 12 or has a post-school qualification	Binary	0.533	0	1

items, some of the models analysed in this section use explanatory variables from different waves of the LSIC.

For the binary independent variables, there is an implicit omitted category against which the coefficients below should be compared. For all the estimates, this omitted or reference category has these characteristics: male; lives in an area with no relative isolation; has a P1 who is not employed and has not completed year 12 or a post-school qualification; has a P1 with a partner who is not employed; lives in a middle-income family with a community or government rental; and did not move in the 12 months before the survey.

5.2 Analysis of attendance measures

We begin our analysis of the convergent and discriminant validity of the LSIC education measures by looking at the relationship between school attendance (dependent variable) and demographic, geographic and family variables (independent variables). We examine two dependent variables:

- study child attended school every day last week (ace13) – a binary variable, analysed with the probit model
- absences per term (ace78) – a categorical variable with six categories; larger values represent more absences, analysed with the ordered probit model.

Table 24 presents the coefficient estimates, standard errors and significance levels.

Even using a relatively low significance cut-off (10%), few explanatory variables were significantly associated with school attendance. The only variable with a consistent association was P1 employment, which was associated with better attendance on both measures of school attendance.

Females were more likely to have attended school every day in the last week than males, but measuring school attendance over a longer time abolished the difference. Those in more isolated areas had a somewhat higher chance of full attendance, although the standard errors were reasonably large. Home ownership was negatively associated with absences over the term. P1 education was unexpectedly associated with more absences. However, both of these associations were significant only for the longer period at the 10% significance level.

5.3 Analysis of education scales

In this subsection, we extend the analysis of the factors associated with education outcomes by including a selection of the education scales as dependent variables with the same set of explanatory variables as in Table 24. In Section 4, we introduced and validated the different scales. The scales we included in our multivariate analysis are listed below, with results of the analysis presented in Tables 25 and 26.

Beginning with Table 25, we examine the factors associated with four measures related to teachers and schools. The first two measures (teacher closeness and teacher conflict) are estimated for the combined cohorts, whereas school liking is estimated separately for the B and K cohorts. The final column gives the factors associated with the PSSM scale, available for the K cohort only. For all measures except the teacher closeness scale, higher values represent poorer outcomes.

Apart from sex, few explanatory variables were consistently associated with the teacher and school measures. For all variables, girls' outcomes were better than boys' outcomes. For the school closeness measure, age was positively associated, whereas school grade was negatively associated (conditional on age). The opposite occurred for the school liking measure for the B cohort. Income and housing status were associated with school liking for the K cohort, with those in low-income houses having lower factor scores (i.e. greater school liking).

In Table 26, we examine maths and reading self-concepts for the K and B cohorts. Sex and geographic isolation were the only variables that correlated strongly with the self-concept measures; however, these associations were inconsistent across the different measures. Girls tended to have higher self-concept about reading, whereas boys had higher self-concept about maths (K cohort only). Children in more isolated areas tended to have higher self-concepts, but the association was only significant at the 5% level for the K cohort.

Table 24 Factors associated with school attendance

Explanatory variable	No absences last week Binary, cross-sectional			Absences per term Categorical, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig
Female	0.200	0.087	**	−0.096	0.069	
Age	−0.117	0.084		0.076	0.067	
Level of relative isolation – low	0.029	0.110		0.037	0.086	
Level of relative isolation – medium	−0.354	0.188	*	−0.216	0.154	
Level of relative isolation – high	−0.289	0.222		−0.099	0.182	
Decile of Indigenous relative socioeconomic outcomes	0.008	0.025		−0.003	0.019	
Parent 1 employed	0.241	0.102	**	−0.193	0.081	**
Partner of parent 1 employed	−0.080	0.140		−0.043	0.110	
Parent 1 does not have a partner	−0.238	0.124	*	−0.029	0.099	
Family has relatively low income (under $600 per week)	0.078	0.117		0.100	0.094	
Family has relatively high income ($1000 or more per week)	0.031	0.118		−0.006	0.094	
Living in a house owned or being purchased by residents	0.083	0.132		−0.172	0.104	*
Living in a house rented through the private rental market	−0.035	0.119		−0.038	0.095	
School grade at time of survey	0.132	0.084		−0.082	0.067	
Moved school since last survey	−0.030	0.144		−0.072	0.115	
Parent 1 has completed year 12 or has a post-school qualification	−0.110	0.094		0.141	0.075	*
Constant (probit)	1.106	0.558				
Cut-value 1				−0.500	0.444	
Cut-value 2				0.603	0.444	
Cut-value 3				1.227	0.445	
Cut-value 4				1.549	0.446	
Cut-value 5				2.010	0.449	
Number of observations	971			952		
Pseudo R-squared	0.0249			0.0091		

** = $P < 0.05$; * = $P < 0.1$; SE = standard error; Sig = significance level

Table 25 Factors associated with wave 8 factors scores – school and teacher measures

Explanatory variable	Teacher closeness			Teacher conflict			School liking (SLAQ) – K cohort			School liking (SLAQ) – B cohort			School membership (PSSM)		
	Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig
Female	0.201	0.069	***	−0.456	0.082	***	−0.546	0.295	*	−1.407	0.324	***	−0.544	0.182	***
Age	0.160	0.068	**	−0.066	0.081		−0.582	0.380		0.949	0.407	**	−0.093	0.237	
Level of relative isolation – low	0.087	0.085		−0.001	0.104		0.066	0.362		−0.706	0.405	*	−0.195	0.226	
Level of relative isolation – medium	0.096	0.161		0.153	0.194		0.550	0.664		0.083	0.733		−0.446	0.416	
Level of relative isolation – high	−0.170	0.208		−0.132	0.238		−0.288	0.786		−0.363	0.858		−0.440	0.484	
Decile of Indigenous relative socioeconomic outcomes	0.005	0.021		−0.017	0.024		−0.012	0.089		−0.085	0.092		−0.083	0.055	
Parent 1 employed	0.043	0.083		−0.055	0.096		−0.217	0.338		−0.038	0.381		−0.063	0.208	
Partner of parent 1 employed	0.077	0.112		−0.233	0.136	*	0.178	0.492		0.805	0.512		−0.305	0.306	
Parent 1 does not have a partner	0.077	0.103		−0.014	0.119		0.267	0.445		0.564	0.452		−0.344	0.276	
Family has relatively low income (under $600 per week)	0.173	0.098	*	−0.139	0.117		−0.836	0.420	**	−0.201	0.433		0.372	0.258	
Family has relatively high income ($1000 or more per week)	−0.027	0.093		0.023	0.114		0.102	0.391		−0.776	0.450	*	0.022	0.243	

continued

Table 25 ***continued***

Explanatory variable	Teacher closeness			Teacher conflict			School liking (SLAQ) – K cohort			School liking (SLAQ) – B cohort			School membership (PSSM)		
	Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig
Living in a house owned or being purchased by residents	0.100	0.105		0.028	0.124		−0.880	0.418	**	0.306	0.497		0.115	0.260	
Living in a house rented through the private rental market	−0.025	0.096		−0.008	0.114		−0.211	0.411		0.343	0.451		0.188	0.257	
School grade at time of survey	−0.263	0.070	***	0.065	0.082		0.311	0.307		−0.795	0.328	**	0.313	0.190	
Moved school since last survey	0.032	0.124		0.037	0.153		0.406	0.501		−0.036	0.545		−0.181	0.309	
Parent 1 has completed year 12 or has a post-school qualification	0.112	0.075		−0.029	0.090		−0.563	0.318	*	−0.442	0.354		0.301	0.198	
Constant	3.197	0.440	***	2.539	0.527	***	−2.814	3.659		−8.959	3.080	***	4.061	2.300	*
Number of observations	372			345			335			522			331		
Pseudo *R*-squared	0.0731			0.0784			0.0147			0.0389			0.0163		

*** = $P < 0.01$; ** = $P < 0.05$; * = $P < 0.1$; PSSM = Psychological Sense of School Membership scale; SE = standard error; Sig = significance level; SLAQ = School Liking and Avoidance Questionnaire

Table 26 Factors associated with wave 8 factors scores – school and teacher measures

Explanatory variable	Math self concept – K cohort			Reading self concept – K cohort			Math self concept – K cohort			Reading self concept – K cohort		
	Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig
Female	−0.124	0.388		−0.890	0.309	***	0.600	0.261	**	−0.614	0.234	***
Age	1.011	0.501	**	0.190	0.401		0.333	0.329		0.521	0.293	*
Level of relative isolation – low	−1.288	0.477	***	−0.131	0.381		−0.131	0.327		−0.289	0.292	
Level of relative isolation – medium	−2.868	0.876	***	−1.153	0.700	*	−1.071	0.593	*	−0.878	0.528	*
Level of relative isolation – high	−3.438	1.038	***	−1.754	0.829	**	−0.111	0.687		−0.063	0.612	
Decile of Indigenous relative socioeconomic outcomes	0.018	0.117		−0.004	0.094		0.022	0.074		−0.067	0.066	
Parent 1 employed	0.503	0.444		0.340	0.353		0.228	0.309		−0.258	0.276	
Partner of parent 1 employed	0.036	0.648		−0.259	0.517		0.681	0.414	*	−0.330	0.372	
Parent 1 does not have a partner	0.115	0.583		−0.287	0.466		0.317	0.363		−0.031	0.325	
Family has relatively low income (under $600 per week)	0.165	0.555		0.363	0.441		0.181	0.349		−0.135	0.312	
Family has relatively high income ($1000 or more per week)	0.255	0.516		−0.184	0.411		−0.141	0.366		0.508	0.327	

continued

Table 26 continued

Explanatory variable	Math self concept – K cohort Continuous, cross-sectional			Reading self concept – K cohort Continuous, cross-sectional			Math self concept – K cohort Continuous, cross-sectional			Reading self concept – K cohort Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig
Living in a house owned or being purchased by residents	−0.986	0.548	*	−0.028	0.438		−0.729	0.402	*	−0.207	0.360	
Living in a house rented through the private rental market	−0.873	0.543		−0.028	0.433		−0.298	0.364		−0.076	0.325	
School grade at time of survey	0.057	0.403		0.205	0.323		0.131	0.264		−0.213	0.236	
Moved school since last survey	−0.052	0.669		−0.393	0.529		−0.300	0.439		0.030	0.391	
Parent 1 has completed year 12 or has a post-school qualification	−0.152	0.418		−0.245	0.334		0.338	0.287		−0.041	0.256	
Constant	−3.549	4.833		3.072	3.860		1.179	2.491		1.991	2.218	
Number of observations	338			338			527			523		
Pseudo R-squared	0.0464			0.0101			0.0074			0.0043		

*** = $P < 0.01$; ** = $P < 0.05$; * = $P < 0.1$; SE = standard error; Sig = significance level

5.4 Analysis of NAPLAN

In this section, we investigate the factors associated with the linked NAPLAN results. Specifically, we consider those who undertook the year 3 or the year 5 NAPLAN exams in either 2014 or 2015. In Table 27, we look at the average score across the five components of NAPLAN, with the first set of columns for those who undertook the year 3 tests and the second set of columns for those who undertook the year 5 tests. In Tables 28 and 29, we respecively examine the individual year 3 and year 5 tests.

As the NAPLAN scores are continuous variables, we analysed the factors associated with them using ordinary least squares regression.

Despite the relatively small sample size (due to many students not in scope for NAPLAN), some factors associated with both year 3 and year 5 NAPLAN results were not associated with the LSIC-specific education measures. This finding may result from the greater variation across the sample of the NAPLAN scores than the variation in most of the LSIC-specific measures, largely due to the number of individual questions used to construct.

Sex had a very strong association with the NAPLAN scores: average NAPLAN scores for Indigenous females were higher than those of their male counterparts. Those in relatively isolated areas tended to have lower NAPLAN scores, though the large standard error of the estimate for the most isolated areas meant that the coefficient was not statistically significant. For year 3 students, those who lived in the most advantaged areas tended to have the highest scores (Table 28). P1 employment was also significant for this group. In contrast, for year 5 students the tenure type of the child's house had a strong association; school mobility also had a very strong association.

Most factors associated with the individual NAPLAN items were relatively consistent (Table 29). However, the socioeconomic characteristics of the area in which a person lives was associated with year 5 spelling but not with the other year 5 tests; generally, the factors associated with one test type tended to be associated with the others. This finding implies that a researcher would make very similar conclusions about the determinants of NAPLAN outcomes if they used either the average of all five tests or the averages of each item. We confirmed this assumption with a factor analysis of the five individual test scores; the results heavily favoured a one-factor solution.

Table 27 Factors associated with linked NAPLAN scores – averages for years 3 and 5

Explanatory variable	Year 3 Continuous, cross-sectional			Year 5 Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig
Female	49.444	10.762	***	22.998	7.209	***
Age	–1.805	10.388		17.008	9.641	*
Level of relative isolation – low	–32.485	13.227	**	–31.278	8.768	***
Level of relative isolation – medium	–49.983	19.979	**	–67.903	16.232	***
Level of relative isolation – high	–25.200	28.095		–16.532	19.435	
Decile of Indigenous relative socioeconomic outcomes	6.672	2.912	**	2.775	2.133	
Parent 1 employed	25.467	12.678	**	5.395	8.349	
Partner of parent 1 employed	–19.810	18.296		4.573	12.027	
Parent 1 does not have a partner	–7.192	15.812		0.575	10.890	
Family has relatively low income (under $600 per week)	–15.931	14.571		–9.168	10.039	
Family has relatively high income ($1000 or more per week)	12.036	15.163		–9.534	9.625	
Living in a house owned or being purchased by residents	39.275	16.714	**	44.815	9.926	***
Living in a house rented through the private rental market	21.199	14.753		34.456	9.907	***
School grade at time of survey	6.822	19.993		–16.402	8.256	**
Moved school since last survey	–17.278	17.684		–31.222	11.930	***
Parent 1 has completed year 12 or has a post-school qualification	–4.045	12.453		12.943	7.579	*
Constant	283.504	81.614	***	303.739	96.684	***
Number of observations	237			288		
Pseudo *R*-squared	0.2301			0.2904		

*** = $P < 0.01$; ** = $P < 0.05$; * = $P < 0.1$; SE = standard error; Sig = significance level

Table 28 Factors associated with linked NAPLAN scores – individual year 3 items

Explanatory variable	Reading			Writing			Spelling			Grammar			Numeracy		
	Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig
Female	50.671	11.022	***	51.122	10.725	***	50.410	11.106	***	51.731	10.857	***	48.395	10.878	***
Age	0.819	10.615		–1.120	10.352		–8.887	10.696		–4.752	10.457		1.987	10.476	
Level of relative isolation – low	–31.931	13.505	**	–31.303	13.182	**	–28.383	13.607	**	–35.510	13.303	***	–33.497	13.328	**
Level of relative isolation – medium	–45.022	20.427	**	–53.200	19.910	***	–47.495	20.582	**	–55.409	20.121	***	–55.455	20.159	***
Level of relative isolation – high	–26.191	28.691		–22.607	27.999		–25.203	28.909		–27.959	28.262		–29.038	28.314	
Decile of Indigenous relative socioeconomic outcomes	6.171	2.995	**	6.613	2.902	**	6.738	3.018	**	5.885	2.951	**	5.867	2.956	**
Parent 1 employed	33.239	12.954	**	21.250	12.635	*	26.319	13.052	**	25.420	12.760	**	24.345	12.784	*
Partner of parent 1 employed	–22.181	18.680		–15.856	18.234		–15.474	18.822		–22.435	18.401		–20.589	18.435	
Parent 1 does not have a partner	–11.313	16.141		–3.279	15.758		–10.197	16.264		–6.152	15.900		–6.365	15.929	
Family has relatively low income (under $600 per week)	–13.385	14.873		–19.522	14.522		–12.902	14.986		–19.183	14.651		–14.401	14.678	
Family has relatively high income ($1000 or more per week)	12.774	15.479		10.089	15.111		8.817	15.597		16.737	15.248		13.430	15.276	

continued

Table 28 continued

Explanatory variable	Reading			Writing			Spelling			Grammar			Numeracy		
	Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig
Living in a house owned or being purchased by residents	37.569	17.060	**	42.246	16.657	**	35.287	17.190	**	40.158	16.805	**	40.157	16.837	**
Living in a house rented through the private rental market	23.804	15.129		19.724	14.702		19.802	15.244		27.514	14.903	*	23.393	14.931	
School grade at time of survey	5.734	20.714		16.014	19.924		12.276	20.871		4.664	20.404		15.349	20.442	
Moved school since last survey	–20.277	18.059		–17.138	17.624		–7.505	18.196		–21.998	17.789		–22.616	17.822	
Parent 1 has completed year 12 or has a post-school qualification	–5.614	12.725		–3.440	12.410		–4.565	12.822		–1.498	12.535		–1.671	12.558	
Constant	265.644	83.478	***	249.373	81.336	***	322.649	84.112	***	317.252	82.230	***	232.369	82.383	***
Number of observations	236			237			236			236			236		
Pseudo *R*-squared	0.2216			0.2386			0.2026			0.2521			0.2325		

*** = $P < 0.01$; ** = $P < 0.05$; * = $P < 0.1$; SE = standard error; Sig = significance level

Table 29 Factors associated with linked NAPLAN scores – individual year 5 items

Explanatory variable	Reading			Writing			Spelling			Grammar			Numeracy		
	Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig
Female	20.187	7.989	**	19.159	7.888	**	24.189	8.172	***	26.122	8.536	***	21.588	7.615	***
Age	16.780	10.697		24.123	10.584	**	22.030	10.997	**	13.933	11.486		11.479	10.230	
Level of relative isolation – low	–35.071	9.683	***	–28.220	9.592	***	–27.085	9.936	***	–38.535	10.377	***	–28.238	9.215	***
Level of relative isolation – medium	–87.831	17.904	***	–68.643	17.689	***	–36.691	18.418	**	–79.117	19.237	***	–64.352	17.394	***
Level of relative isolation – high	–25.689	21.454		–33.210	21.182		–1.869	22.005		–1.641	22.983		–25.324	20.439	
Decile of Indigenous relative socioeconomic outcomes	0.766	2.361		1.790	2.325		6.137	2.423	**	2.509	2.531		1.901	2.248	
Parent 1 employed	1.935	9.223		2.761	9.119		11.612	9.436		3.474	9.855		8.384	8.837	
Partner of parent 1 employed	2.962	13.496		4.159	13.112		–0.910	13.701		9.458	14.310		5.368	12.711	
Parent 1 does not have a partner	–4.595	12.245		1.195	11.897		5.513	12.458		0.089	13.012		–3.290	11.569	
Family has relatively low income (under $600 per week)	–8.617	11.154		–17.454	11.002		–7.940	11.484		–4.771	11.995		–8.962	10.661	
Family has relatively high income ($1000 or more per week)	–12.696	10.681		–13.346	10.537		–1.075	10.876		–6.132	11.359		–14.298	10.180	

continued

Table 29 ***continued***

Explanatory variable	Reading			Writing			Spelling			Grammar			Numeracy		
	Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional			Continuous, cross-sectional		
	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig	Coefficient	SE	Sig
Living in a house owned or being purchased by residents	45.293	10.942	***	49.007	10.855	***	39.781	11.224	***	48.092	11.722	***	38.704	10.428	***
Living in a house rented through the private rental market	36.182	10.986	***	32.795	10.841	***	37.143	11.203	***	34.288	11.701	***	29.949	10.419	***
School grade at time of survey	–13.491	9.158		–33.168	9.073	***	–1.261	9.454		–17.776	9.875	*	–14.261	8.822	
Moved school since last survey	–31.206	13.166	**	–37.774	12.998	***	–29.921	13.467	**	–31.535	14.065	**	–26.123	12.499	**
Parent 1 has completed year 12 or has a post-school qualification	11.581	8.415		15.959	8.284	*	8.803	8.594		12.923	8.976		17.629	7.977	**
Constant	314.711	107.171	***	319.062	106.200	***	144.826	109.911		344.243	114.795	***	358.801	102.003	***
Number of observations	286			285			284			284			283		
Pseudo *R*-squared	0.2456			0.2818			0.2329			0.2510			0.2306		

*** = $P < 0.01$; ** = $P < 0.05$; * = $P < 0.1$; SE = standard error; Sig = significance level

6 Longitudinal validity and reliability over time

As outlined in Zubrick et al. (2014), one way to assess reliability of a construct over time is to examine the correlations from wave to wave. The criteria to assess the strength of a correlation between the same construct over time are:

- ≥0.4 is a strong relationship
- 0.30–0.39 is a moderate relationship
- <0.30 is a weak relationship.

In addition to examining the size of correlations as an indication of reliability of the construct over time, we also tested for measurement invariance over time (Little 2013). In confirmatory factor analysis, this refers to testing whether the factor structure of a given construct is stable over time. There are several steps to testing measurement invariance. Zubrick et al. (2014:24) provide a short description but do not test for factorial invariance and note that it is a 'complex method'. In this section, we test for configural invariance, the first step in measurement invariance testing. In this step, the pattern of factor loadings is tested to see whether they are equal.

As noted in Section 4.1, SLAQ and the STRS are the only scales that have been measured three or more times in consecutive waves. Therefore, we tested the reliability over time and configural invariance for these scales only.

6.1 School Liking and Avoidance Questionnaire

Testing for configural invariance is also done through the model fit indices used in Section 4. In Table 30, we see that the SLAQ model fit was good for both the B and K cohorts at waves 7 and 8, with all indices meeting the standard for good fit. For the B cohort, the correlation between waves 7 and 8 of SLAQ was 0.48 (Figure 12), indicative of a strong relationship and good construct reliability over time. For the K cohort, the relationship was even stronger (r = 0.68) (Figure 13). These findings provide evidence of strong reliability over time for both cohorts between waves 7 and 8.

6.2 Pianta Student–Teacher Relationship Scale

The STRS has two factors, closeness and conflict. First, we tested whether the factor structure of closeness was consistent over waves 6, 7, 8 and 9. We estimated a confirmatory factor analysis using weighted least square mean variance (WLSMV) because we specified the items as categorical and took account of missing variables. We allowed correlation of latent variables at waves and of each individual item with the corresponding item at the subsequent wave (item 1 at waves 7, 8 and 9). The model fit was good (χ^2 (165) = 337.43, P < 0.001; RMSEA = 0.04; CFI = 0.97; TLI = 0.96; SRMR = 0.07), which supports configural invariance of closeness. The modification indices suggested that model fit could be improved by correlating items 7 ('spontaneously shares information') and 15 ('openly shares feelings and experiences') at each wave. Because the model fit was already good, we parsimoniously chose not to correlate the error terms of these items.

The standardised factor loadings on closeness across the three waves ranged from 0.61 to 0.83 (Figure 14). There was evidence of moderate reliability over time between waves 7 and 8 (r = 0.35) and weak reliability between waves 8 and 9 (r = 0.28). However, for waves 8 and 9, the correlation between the two latent variables showed only weak reliability (r = 0.28).

We then tested whether the factor structure of conflict was consistent over waves 7, 8 and 9. We estimated a confirmatory factor analysis using WLSMV because we specified the items

Table 30 Confirmatory factor analyses of SLAQ by waves and cohort

Variable	Waves 7–8
B cohort	
SRMR (<0.05 good fit, <0.10 acceptable fit)	**0.02**
RMSEA (<0.05 good fit, <0.08 acceptable)	**0.03**
CFI (>0.95 good fit, 0.90–0.95 acceptable)	**0.99**
TLI (>0.95 good fit, 0.90–0.95 acceptable)	**0.98**
Chi-square	23.451
df	13
P	0.0366
N	824
Overall rating	Good
Estimator	ML
K cohort	
SRMR (<0.05 good fit, <0.10 acceptable fit)	**0.02**
RMSEA (<0.05 good fit, <0.08 acceptable)	**0.017**
CFI (>0.95 good fit, 0.90–0.95 acceptable)	**0.998**
TLI (>0.95 good fit, 0.90–0.95 acceptable)	**0.994**
Chi-square	12.928
df	11
P	0.298
N	575
Overall rating	Good
Estimator	ML

CFI = comparative fit index; df = degrees of freedom; ML = maximum likelihood; RMSEA = root-mean-square error of approximation; SLAQ = School Liking and Avoidance Questionnaire; SRMR = standardised root mean residual; TLI = Tucker–Lewis index

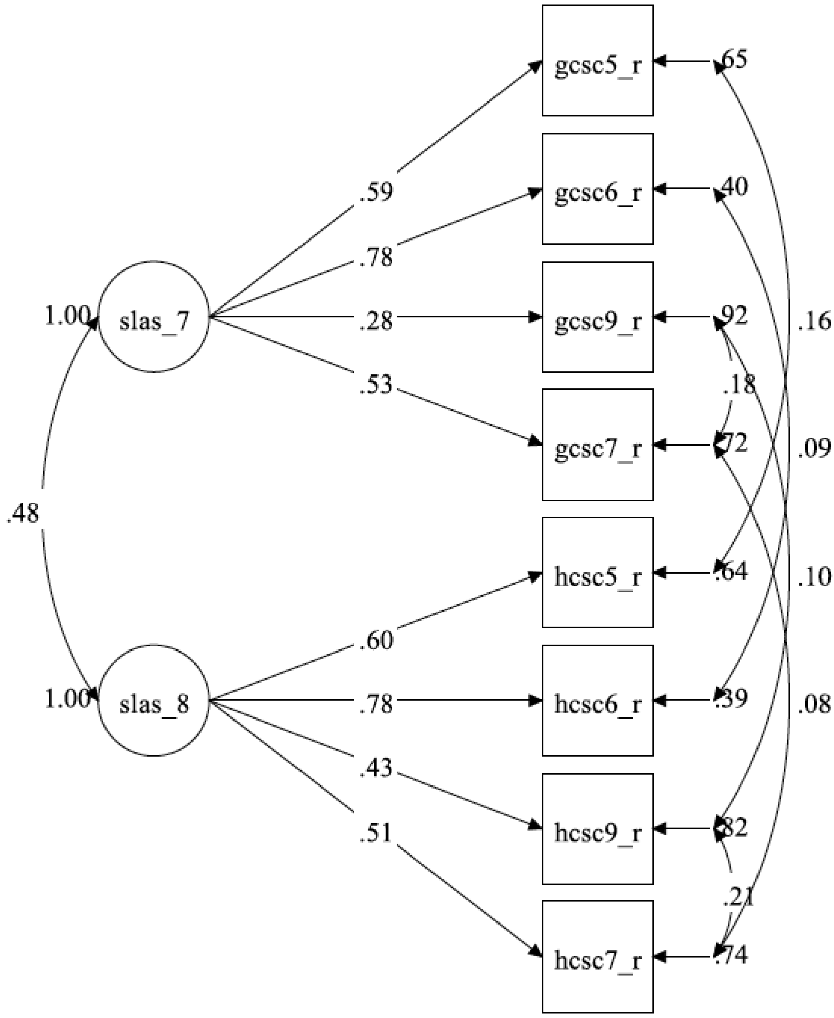
Note: Bold numbers indicate good fit.

as categorical and took account of missing variables. We allowed correlation of latent variables at waves and of each individual item with the corresponding item at the subsequent wave (item 1 at waves 7, 8 and 9). The model fit was excellent (χ^2 (165) = 232.09, $P < 0.001$; RMSEA = 0.03; CFI = 0.99; TLI = 0.99; SRMR = 0.05).

The standardised factor loadings on conflict across the three waves ranged from 0.77 to 0.95 (Figure 15). There was evidence of strong reliability over time for waves 7 to 8 and 8 to 9 (waves 7 and 8, *r* = 0.70; waves 8 and 9, *r* = 0.62), which support a strong relationship and reliability over time. The correlation between waves 7 and 9 was similar (*r* = 0.68).

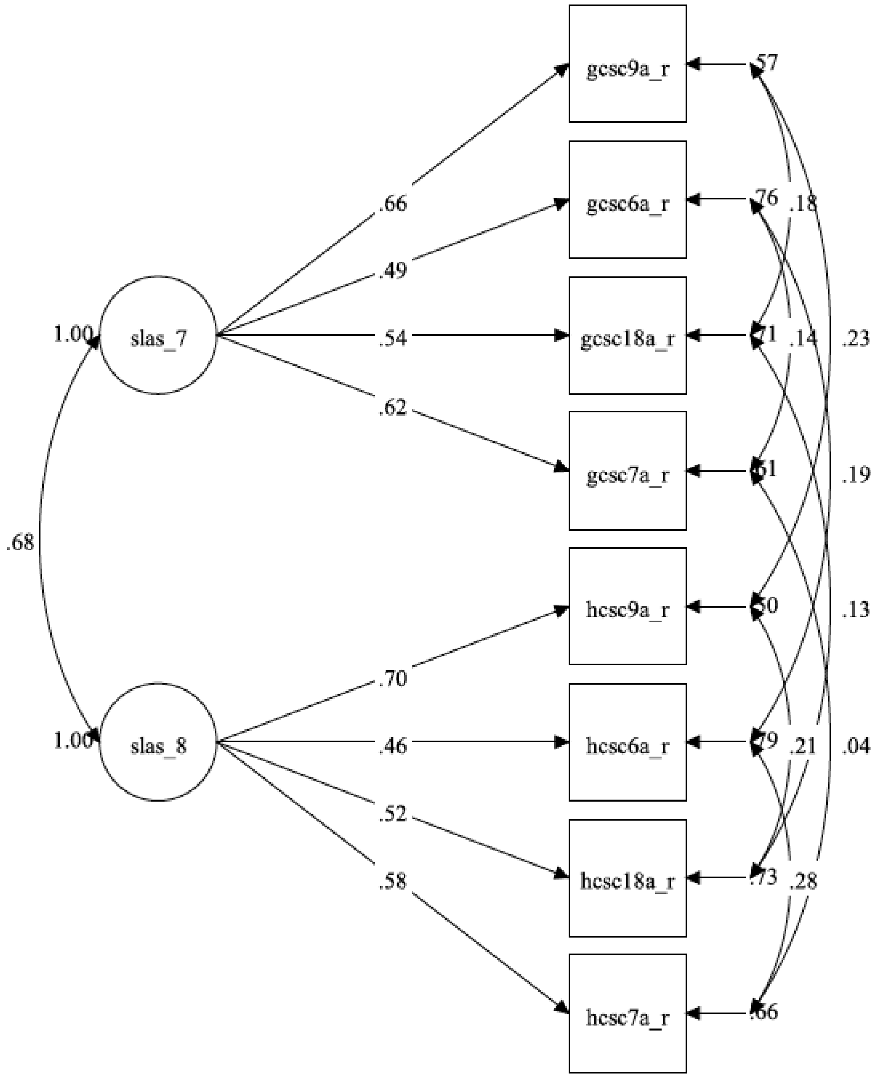
The results of the longitudinal confirmatory factor analyses of closeness and conflict both suggest configural invariance and, for conflict, strong reliability over time.

Figure 12 Configural invariance for SLAQ, B cohort, waves 7–8



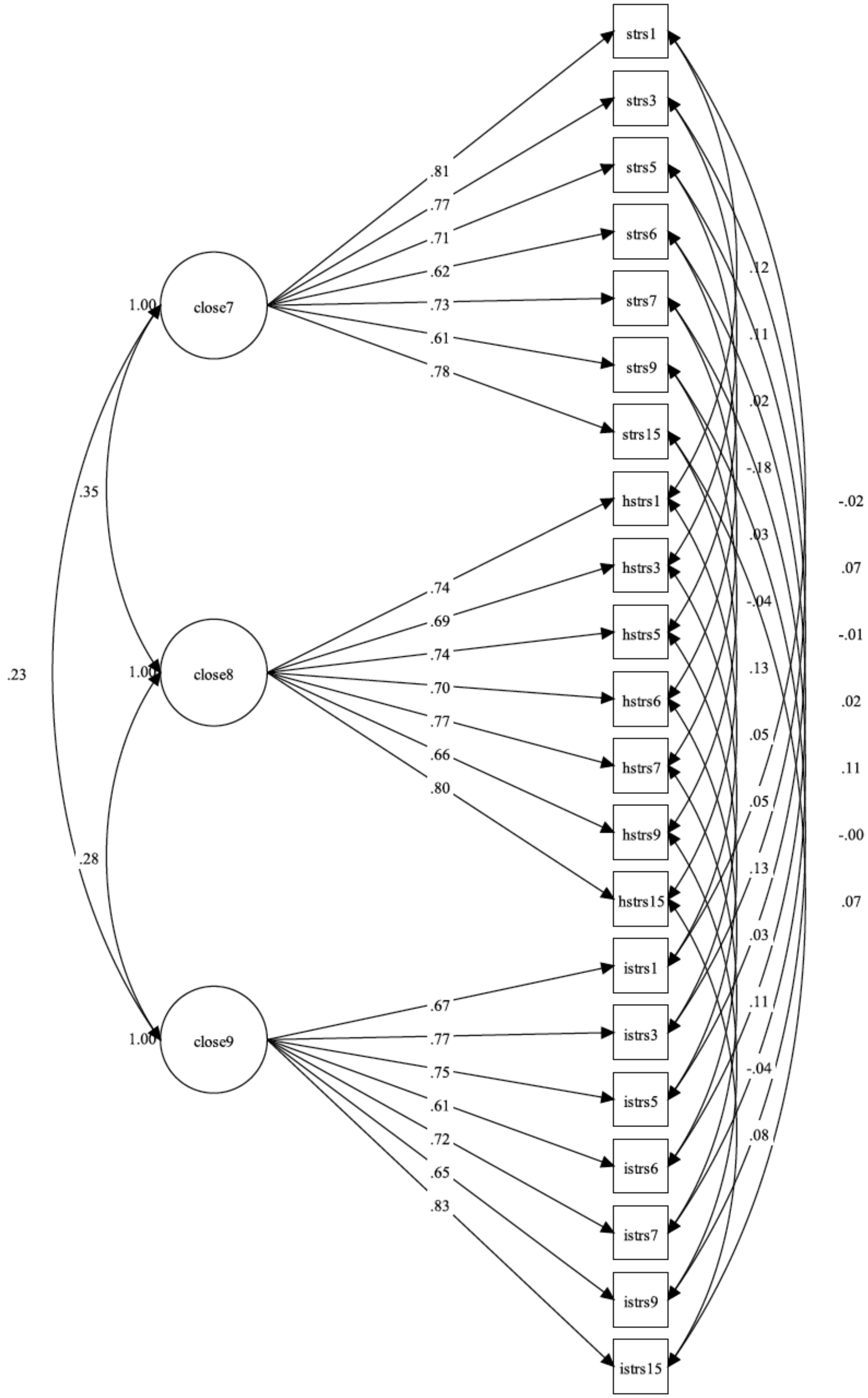
SLAQ = School Liking and Avoidance Questionnaire

Figure 13 Configural invariance for SLAQ, K cohort, waves 7–8



SLAQ = School Liking and Avoidance Questionnaire

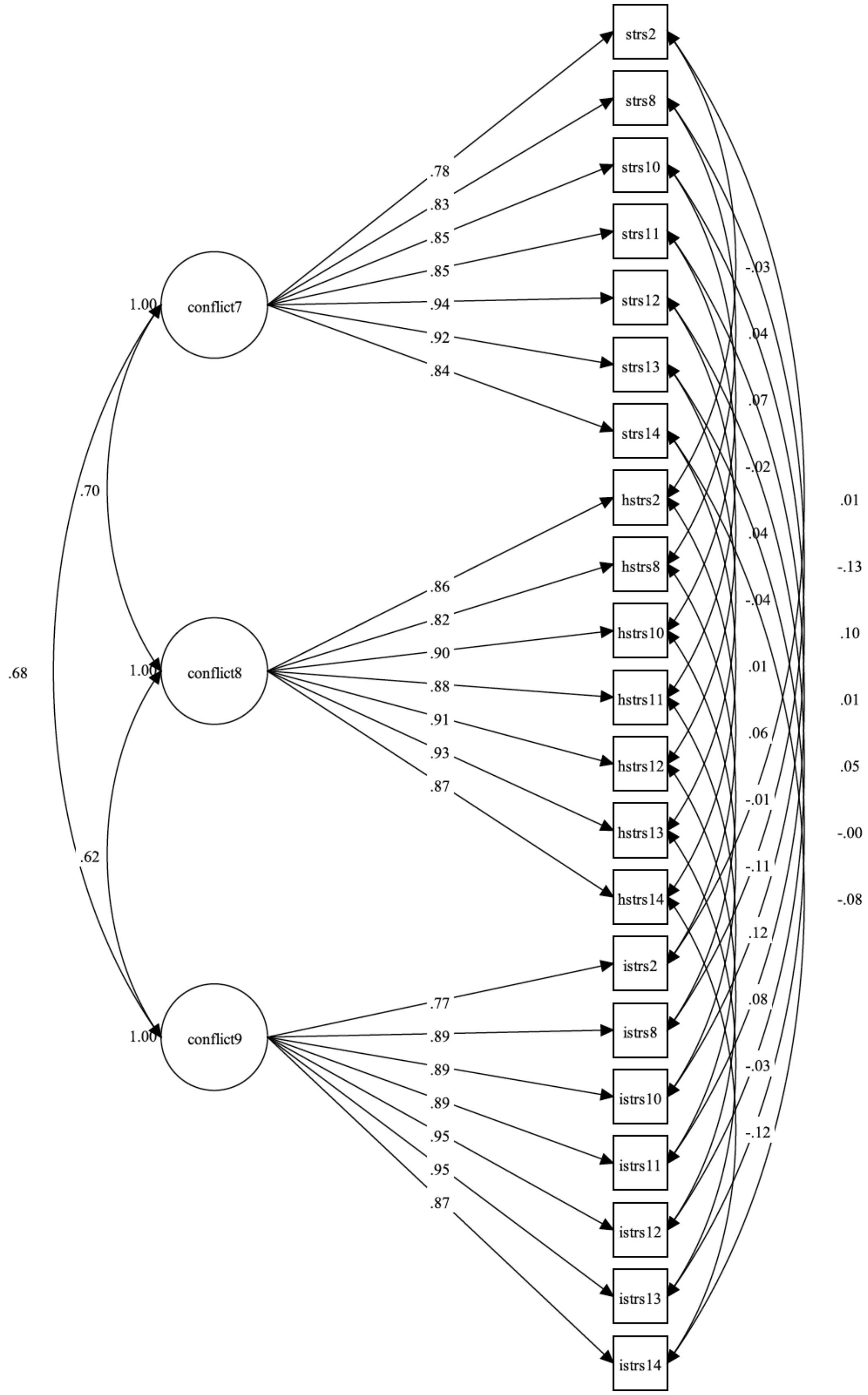
Figure 14 Factor loadings on closeness for the STRS, B cohort, waves 7–9



STRS = Student–Teacher Relationship Scale

Note: N = 617.

Figure 15 **Factor loadings on conflict for the STRS, B-cohort, waves 7–9**



STRS = Student–Teacher Relationship Scale

Note: N = 617.

6.3 Predictors of change in education measures

We began our analysis of predictors of change through time by looking at changes in NAPLAN results for those in the sample for whom year 3 and year 5 results were available. We modelled this change as a function of the wave 6 characteristics of the individual for the B and K cohorts combined, as well as the NAPLAN test scores in year 3 (to capture reversion to the mean) (Table 31).

In similar models using the full NAPLAN sample (i.e. not only the LSIC-linked sample), NAPLAN score in time t is negatively associated with NAPLAN scores in time $t + 2$ (Biddle & Edwards 2018). This finding is probably due to a combination of random variation in test scores across years (i.e. someone with a randomly high or low score in time t is likely to decrease or increase their relative score between time t and $t + 2$) and natural floors and ceilings on NAPLAN scores (i.e. those who do well/poorly in one year have less scope to increase/decrease their scores over the next two years).

Table 31 Factors associated with change in linked NAPLAN scores – averages for year 3 to year 5, combined B and K cohorts

Explanatory variable	Year 3 to year 5 Continuous, longitudinal		
	Coefficient	SE	Sig
Year 3 NAPLAN	–0.257	0.053	***
Female	7.084	6.262	
Age	4.636	8.302	
Level of relative isolation – low	–16.581	7.228	**
Level of relative isolation – medium	–9.934	14.436	
Level of relative isolation – high	45.295	19.280	**
Decile of Indigenous relative socioeconomic outcomes	0.007	1.852	
Parent 1 employed	–2.197	6.785	
Partner of parent 1 employed	–14.113	10.854	
Parent 1 does not have a partner	–10.287	10.492	
Family has relatively low income (under $600 per week)	13.049	8.975	
Family has relatively high income ($1000 or more per week)	14.863	8.013	*
Living in a house owned or being purchased by residents	24.302	8.041	***
Living in a house rented through the private rental market	18.836	7.884	**
School grade at time of survey	–3.134	6.663	
Moved school since last survey	–11.988	10.644	
Parent 1 has completed year 12 or has a post-school qualification	1.971	6.795	
Constant	132.426	71.674	*
Number of observations	206		
Pseudo R-squared	0.1574		

*** = $P < 0.01$; ** = $P < 0.05$; * = $P < 0.1$; SE = standard error; Sig = significance level

Even after controlling for these baseline scores, a few variables were significantly associated with change through time. In particular, housing circumstances and living outside public housing were significantly associated with an increase in NAPLAN scores between times t and $t + 2$.

Multiple waves of observations per individual need a different estimation technique. We estimated the relationship between the global health measure (ahc1) and a set of education dependent variables, while controlling for unobserved, time-invariant, individual-level characteristics (Table 32). As discussed previously, few education measures are measured consistently over long periods. One dependent variable we could estimate was school or preschool attendance in the previous week.

Table 32 shows a consistent association between parental-reported health and school attendance. Apart from those reported by their P1 to have poor health (a very small sample size), worse health at a particular point in time was associated with a lower probability of attending school every day in the previous week.

Table 32 Random effects modelling of the relationship between health and education measures

Explanatory variable	School attendance		
	Binary, longitudinal		
	Coefficient	SE	Sig
Health very good	−0.091	0.044	**
Health good	−0.210	0.053	***
Health fair	−0.515	0.121	***
Health poor	0.312	0.587	
Wave	−0.089	0.013	***
Constant	1.317	0.091	***
Number of individuals	1545		
Number of observations	5830		
Number of waves	5		

*** = $P < 0.01$; ** = $P < 0.05$; SE = standard error; Sig = significance level

Note: The omitted category for the global health measure is 'Excellent health'.

7 Recommendations

In this section, we outline some of the key recommendations that follow from the analysis presented in sections 4, 5 and 6. We outline separate recommendations for data collection, users of the data, and policy makers who are making decisions based on the LSIC.

7.1 Recommendations for data collection

Clearly, the DSS makes considerable effort to ensure that the LSIC is relevant to the Indigenous population and that Indigenous researchers contribute to the design of the questions and sample retention strategy. These factors should remain a high priority component of the LSIC when considering our recommendations. However, often trade-offs must be made in terms of sample representation and question consistency.

Our first recommendation for data collectors is to seriously consider boosting the sample of the LSIC to account for both the growth in the Indigenous population through identification change and sample attrition. The current LSIC is not representative of the Australian Indigenous population. Although complete representation is never possible, those Indigenous Australians or families who have begun to identify as Indigenous since the first sample recruitment (either because of new knowledge or increased comfort with identifying to a government data collector) are missing from the dataset. Boosting the sample will allow future longitudinal analysis of this new cohort. Further, asking some retrospective questions and linking to NAPLAN and AEDC data will enable historical longitudinal analysis.

In this project, it was difficult to determine the items that corresponded to each scale, and which scales were standardised measures as opposed to those constructed for the LSIC. For identified standardised measures, the source of the scale was often unclear. Based on correspondence for this project, it appears that SLAQ (which has been abridged), the academic self-concept (abridged), the PSSM (abridged) and the Pianta STRS are the only standardised measures. However, only one of these four measures has the complete set of items.

This lack of documentation and lack of standardisation present barriers to researchers using the data, because successful paper publication can depend on reviewers' confidence in the validity and reliability of measures. Notwithstanding the validation of scales provided in this report, the research community is unlikely to use new measures that have no items from pre-existing scales, despite their cultural appropriateness for LSIC participants.

Finally, the reporting of measures across the waves is inconsistent. Hence, longitudinal statistical analyses cannot be done with many of these measures. Currently, in addition to the attendance measures and the linked NAPLAN results, SLAQ and the STRS appear to be the only measures that afford true longitudinal analyses. Some limited longitudinal analyses are possible with academic self-concept and school climate for the K cohort, but the absence of these and other variables in the B cohort presents many missed opportunities. Therefore, we make these recommendations:

- Provide adequate documentation about the source of scales and scale items, and their reliability and validity (partly remedied by this report).
- Use standardised measures where possible. When an Indigenous-specific reason precludes this, the explicit trade-offs in terms of comparability and publishability should be presented.
- Include standardised measures consistently and in every wave.

This last recommendation may involve asking fewer items but asking them more frequently.

Although this will diminish the use of the LSIC as a cross-sectional survey, we contend this is not where the LSIC's potential strengths lie, or what its priority should be. The National Aboriginal and Torres Strait Islander Social Survey is the preferred vehicle for Indigenous-specific estimates at particular time points. To be a truly useful longitudinal survey, the balance should shift towards the use of consistent questions, especially because the sample ages, concepts and frames are relevant for more waves.

7.2 Recommendations for analysts

We make the following recommendations for analysts.

7.2.1 Reliability

Internal consistency. The internal consistency of all the education scales was in the acceptable or good range for the K cohort. For the B cohort, the internal consistency was also acceptable or good for all scales except for waves 7 and 8 of SLAQ, when children were aged 6–8 years and 7–9 years, where it approximated acceptability.

Factorial validity and validity over time. For both cohorts, all confirmatory factor analyses showed good or acceptable model fit for research purposes, except for wave 9 of SLAQ in the K cohort, where the model fit was not acceptable. When we examined the validity of the factor structure over time (waves 7 to 8), SLAQ showed good model fit for both the B and K cohorts.

Reliability over time. For SLAQ and the STRS, reliability over time was good (Zubrick et al. 2014).

Validity. Data users should note several key messages about the convergent and discriminant validity of the education scales:

- SLAQ shows evidence of convergent validity with measures of academic self-concept and teacher and school climate. There is limited evidence of its correlation with child outcomes (SDQ total problems, reading fluency – K cohort only). Moreover, associations with other education measures for children living in remote areas are limited.
- While reading academic self-concept shows good evidence of convergent and discriminant validity, maths academic self-concept seems unrelated to maths performance.
- For the K cohort, the PSSM and school climate scales show good convergence with one another, and some relationship with behavioural problems.
- The measure of teacher relationship seems to show convergent validity and some relationship with reading fluency, maths achievement and behavioural problems.
- The teacher-rated measure of student–teacher relationship shows convergent validity with teacher relationship and the strongest associations with reading fluency, maths achievement and behavioural problems.

Overall, the education scales in wave 9 of the LSIC show promising signs of convergent validity, with the exception of the maths academic self-concept scale, which lacks correlation with maths scores.

7.2.2 Overall recommendations based on the psychometric characteristics of education scales

Given the young and heterogeneous Indigenous population of children interviewed in the LSIC, the psychometric quality of scales is very good. Analysts are encouraged to be mindful of when the data from education scales are collected, but should have confidence that the psychometric properties of the education scales are sound, with only a few exceptions.

Analysts should also be aware of some specific variables that did not load highly on the factor constructs. Some of these variables have been removed from later waves of the survey but others are still included (and may be useful as individual data items). Analysts should take care when using these as part of a battery of questions.

We also recommend that data users exercise caution about making comparisons between findings from the LSIC and findings from other datasets. Not only is the Indigenous population unique in its geographic distribution, historical experience and exposure to policy interventions, but also the LSIC sample does not necessarily represent the Indigenous Australian population. This lack of representativeness has become more pronounced as the survey continues, due in part to nonrandom sample attrition, but also because of the change in the population of interest (primarily from identification change). Therefore, all conclusions using LSIC data should be conditional on the sample.

A fair proportion of researchers will first encounter the LSIC as reviewers of papers or other publications. A standard question reviewers ask when deciding if a paper should be published is whether the measures used in the study have been validated in other contexts. This reasonable question should be asked of analysts of the LSIC. However, we recommend that reviewers consider that the LSIC was designed in collaboration with Indigenous communities and researchers, and that the data is designed to answer questions relevant to the Indigenous population. We recommend to data collectors that they make the trade-off between sample-specific and validated data items explicit. Similarly, we recommend that reviewers remember that data items on the LSIC may not be fully externally validated, but may be more relevant to the communities involved than measures designed in a different context.

Users of the LSIC should also note that multivariate analysis using the data often leads to a large proportion of insignificant variables, especially in longitudinal analysis but also in cross-sectional analysis. The relatively small sample size and the significant measurement error around the dependent and independent variables account for this factor. This implies use of parsimonious models that focus on the main explanatory variables of interest and include a limited number of control variables. Readers or reviewers of research papers using the LSIC should keep this in mind and expect imprecision from estimated relationships.

Finally, we recommend that the research community looks for opportunities to further validate LSIC data items. We have validated many measures, but we focused on cross-sectional validation, validation for the total sample, and validation of education-specific measures. Opportunities exist for further longitudinal validation. For example, researchers could validate items for particular subpopulations (e.g. sex, geography, family circumstance) and measures more indirectly related to education.

7.3 Recommendations for policy makers

One of the overall aims of the LSIC is to provide an evidence base for important policy debates. Policy makers urgently need to become more familiar with, and make greater use of, the LSIC, to consider the findings of researchers who have used the LSIC dataset, and to provide a proper funding environment for the LSIC to continue.

When using the LSIC for education or related policy deliberations, policy makers should place greater weight on findings that use the longitudinal nature of the LSIC rather than those making inferences on cross-sectional prevalence. The LSIC is not well designed to identify the proportion of Indigenous children or youth who have, for example, a certain level of academic self-concept, school attendance, literacy or numeracy. In particular, it cannot compare these measures with findings in other Australian children or youth, or Indigenous populations in other countries. Rather, the LSIC is designed to understand the extent to which certain outcomes or experiences predict growth or other forms of change through time in other outcomes. We recommend that policy makers fund and consider this type of research in their decisions.

With this in mind, we also recommend that policy makers note the difficulty in publishing research using LSIC data. The LSIC clearly decided to prioritise Indigenous-specific measures. Thus academic journals may be less inclined to publish research using the LSIC than research using externally validated measures. We believe most of the education measures in the LSIC are internally valid. However, this is sometimes a difficult argument to make with reviewers. We recommend that policy makers who assess research output when making funding decisions consider this when evaluating research track records. Researchers using the LSIC should definitely be encouraged to publish academically, but the difficulty of such publication should be noted.

Policy makers can also directly contribute to the research from the LSIC through supporting the dataset. Early-career researchers will more likely invest in understanding the complexity of the dataset if they know the LSIC will continue and how the data are collected. Funding uncertainty will hamper usage of the LSIC. Therefore, we recommend that policy makers clearly signal whether they are open to funding a top-up of the LSIC sample to make it more representative. We expect more people will use the LSIC if it better reflects the contemporary distribution of the Indigenous population.

The LSIC differs from other Australian longitudinal datasets in that it is embedded within a policy agency rather than a research centre. This characteristic has benefits because decisions on questionnaires and methodologies can draw on the priorities of the current government. However, it also brings disadvantages. The researchers within the National Centre for Longitudinal Data (NCLD) are highly skilled and familiar with the intricacies of the dataset. Nevertheless, they do not have the same incentives as external researchers to generate research findings from the dataset, and even less incentive to publish their findings in peer-reviewed academic journals. This factor limits the wider research community's exposure to findings from the LSIC, as evidenced by the small number of papers that have used the education measures.

We strongly recommend that policy makers fund a research organisation to undertake ongoing analysis of the LSIC. This organisation should research questions relevant to the Indigenous community and the broader policy community, and promote the use of the LSIC in the broader research community. The data collection and study design roles of the current Steering Committee and NCLD team should continue alongside the new organisation's focus on increasing use of the data.

8 Summary and concluding comments

This report evaluated the education measures in the LSIC. The study surveys Aboriginal and Torres Strait Islander Australian children who were aged either 6–18 months (B cohort) or 3.5–5 years (K cohort) when the study began in 2008. Analysis presented in this report examined data from the first nine waves. We also made some longitudinal comparisons between waves to test whether relationships are consistent through time.

We found that education measures in the LSIC are mostly internally valid and perform as we expect based on the existing literature and what we know about the measures from other populations. Some measures perform better than others, and some items load more strongly on individual factors. However, our general conclusion is that the LSIC is a useful and robust dataset for answering education-related aspects of the study's aims:

- What do Aboriginal and Torres Strait Islander children need to get the best start in life and grow up strong?
- What helps Aboriginal and Torres Strait Islander children to stay on track or encourages them to become healthier, more positive and stronger?
- How are Aboriginal and Torres Strait Islander children raised?
- What is the importance of family, extended family and community for both young children and as they grow up?

We also outlined specific recommendations for people involved in the collection and analysis of LSIC data and for policy makers who use or make funding decisions about the dataset.

For data collectors, we recommended asking fewer questions but asking them more consistently, and to continue exercising great care and being explicit when balancing specificity and generalisability.

For analysts we recommended using the data with confidence, but being aware that some variables perform better than others and that models using the education measures (especially those specific to the LSIC) tend to have low explanatory power. We also recommended to exploit the longitudinal nature of the dataset rather than focusing on specific waves.

For reviewers of papers based on the LSIC (a subset of data users), we recommended taking into account the unique circumstances of the survey and that models will be estimated with less precision and with variables that may differ from those collected on other datasets.

Finally, for policy makers we recommended making decisions using longitudinal research and to consider funding a top-up sample. We further recommended establishing a dedicated analytical hub in a research institution to increase the visibility and use of the data.

We can learn much from the LSIC about the changing education outcomes of Aboriginal and Torres Strait Islander children, about what predicts positive outcomes for these children, and about the ways in which policy and community can improve the lives of Indigenous Australians. The education data from the LSIC should be analysed with care, but with confidence.

Appendix Factor loadings for the Student–Teacher Relationship Scale-short form

Table A1 STRS-SF, wave 8, K cohort

Item	Wave 8		
	Conflict	Closeness	Other
1. I share an affectionate, warm relationship with child (dcc33_1)		0.68	
2. This child and I always seem to be struggling with each other (dcc33_2)	0.73	−0.34	
3. If upset, this child will seek comfort from me (dcc33_3)		0.75	
4. This child is uncomfortable with physical affection or touch from me (dcc33_4)			0.80
5. This child values his/her relationship with me (dcc33_5)		0.78	
6. When I praise this child, he/she beams with pride (dcc33_6)		0.59	
7. This child spontaneously shares information about himself/herself (dcc33_7)		0.77	
8. This child easily becomes angry with me (dcc33_8)	0.74		0.36
9. It is easy to be in tune with what this child is feeling (dcc33_9)		0.54	−0.53
10. This child remains angry or resistant after being disciplined (dcc33_10)	0.77		
11. Dealing with this child drains my energy (dcc33_11)	0.87		
12. When this child is in a bad mood, I know we're in for a long and difficult day (dcc33_12)	0.94		
13. This child's feelings towards me can be unpredictable or can change suddenly (dcc33_13)	0.88		
14. This child is manipulative with me (dcc33_14)	0.82		
15. This child openly shares his/her feelings and experiences with me (dcc33_15)		0.76	

Table A2 STRS-SF, wave 9, B cohort

Item	Conflict	Closeness	Other?
1. I share an affectionate, warm relationship with child (dcc33_1)		0.34	0.73
2. This child and I always seem to be struggling with each other (dcc33_2)	0.72		
3. If upset, this child will seek comfort from me (dcc33_3)		0.63	0.36
4. This child is uncomfortable with physical affection or touch from me (dcc33_4)		–0.53	0.48
5. This child values his/her relationship with me (dcc33_5)		0.56	0.49
6. When I praise this child, he/she beams with pride (dcc33_6)		0.54	
7. This child spontaneously shares information about himself/herself (dcc33_7)		0.76	
8. This child easily becomes angry with me (dcc33_8)	0.86		
9. It is easy to be in tune with what this child is feeling (dcc33_9)		0.68	
10. This child remains angry or resistant after being disciplined (dcc33_10)	0.81		
11. Dealing with this child drains my energy (dcc33_11)	0.82		
12. When this child is in a bad mood, I know we're in for a long and difficult day (dcc33_12)	0.89		
13. This child's feelings towards me can be unpredictable or can change suddenly (dcc33_13)	0.88		
14. This child is manipulative with me (dcc33_14)	0.78		
15. This child openly shares his/her feelings and experiences with me (dcc33_15)		0.82	

Table A3 STRS-SF, wave 9, K cohort

Item	Conflict	Closeness	Other?
1. I share an affectionate, warm relationship with child (dcc33_1)		0.34	0.73
2. This child and I always seem to be struggling with each other (dcc33_2)	0.72		
3. If upset, this child will seek comfort from me (dcc33_3)		0.63	0.36
4. This child is uncomfortable with physical affection or touch from me (dcc33_4)		−0.53	0.48
5. This child values his/her relationship with me (dcc33_5)		0.56	0.49
6. When I praise this child, he/she beams with pride (dcc33_6)		0.54	
7. This child spontaneously shares information about himself/herself (dcc33_7)		0.76	
8. This child easily becomes angry with me (dcc33_8)	0.86		
9. It is easy to be in tune with what this child is feeling (dcc33_9)		0.68	
10. This child remains angry or resistant after being disciplined (dcc33_10)	0.81		
11. Dealing with this child drains my energy (dcc33_11)	0.82		
12. When this child is in a bad mood, I know we're in for a long and difficult day (dcc33_12)	0.89		
13. This child's feelings towards me can be unpredictable or can change suddenly (dcc33_13)	0.88		
14. This child is manipulative with me (dcc33_14)	0.78		
15. This child openly shares his/her feelings and experiences with me (dcc33_15)		0.82	

Notes

1. Source: https://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3238.0.55.001Main+Features1June%202016?OpenDocument
2. https://www.aedc.gov.au/about-the-aedc
3. https://www.acara.edu.au/assessment/naplan
4. Kline (2016) also warns against very high values for Cronbach's alpha (0.90). The coefficient alpha is mathematically related to the number of items in the test; Kline made this comment regarding test construction and item redundancy rather than fundamental issues with high alpha values.
5. Patterns of missing values can be provided on request.
6. Wave 6 eigenvalues: 5.13, 3.24; wave 7 eigenvalues: 4.47, 3.30; wave 8 eigenvalues: 4.95, 3.48
7. Wave 6 eigenvalues: 4.77, 3.75; wave 7 eigenvalues: 4.90, 4.06
8. We tested whether levels of affective disengagement were significantly different between children living in remote and nonremote areas. Independent t-tests suggested that there was no significant difference for the B cohort ($t(731) = 0.64$, $P > 0.05$), but K-cohort children in remote areas had lower levels of affective disengagement ($t(480) = -2.52$, $P < 0.05$).

References

Abubakar A, van de Vijver F J R, Alonso-Arbiol I, Suryani AO, Pandia WS, Handani P, Arasa J, Mazrui L & Murugumi M (2016). Assessing sense of school belonging across cultural contexts using the PSSM: measurement and functional invariance. *Journal of Psychoeducational Assessment* 34(4):380–388.

Anderson I, Lyons JG, Luke JN & Reich HS (2017). Health determinants and educational outcomes for Indigenous children. In: Walters M, Martin KL & Bodkin-Andrews G (eds), *Indigenous children growing up strong*, Palgrave Macmillan, London, 259–285.

Arcos Holzinger L & Biddle N (2015). *The relationship between early childhood education and care (ECEC) and the outcomes of Indigenous children*, Working Paper 103, Centre for Aboriginal Economic Policy Research, Australian National University, Canberra.

Azpitarte F, Chigavazira A, Kalb G, Farrant BM, Perales F & Zubrick SR (2018). Childcare use and its role in Indigenous child development: evidence from the Longitudinal Study of Indigenous Children in Australia. *Economic Record* 95(308):1–33.

Baumeister RF (ed.) (1999). *The self in social psychology*, Psychology Press, Philadelphia.

Biddle N (2011). *An exploratory analysis of the Longitudinal Survey of Indigenous Children*, Working Paper 77, Centre for Aboriginal Economic Policy Research, Australian National University, Canberra.

——— & Edwards B (2018). *The characteristics and potential effects of the schools that Indigenous Australians attend*, Working Paper 119, Centre for Aboriginal Economic Policy Research, Australian National University, Canberra.

——— & Priest N (2019). *The importance of reconciliation in education*, CSRM Working Paper 1, Centre for Social Research and Methods, Australian National University, Canberra.

Blunden S, Magee C, Attard K, Clarkson L, Caputi P & Skinner TJSH (2018). Sleep schedules and school performance in Indigenous Australian children. *Sleep Health* 4(2):135–140.

Bodkin-Andrews GH, O'Rourke V & Craven RG (2010a). The utility of general self-esteem and domain-specific self-concepts: their influence on Indigenous and non-Indigenous students' educational outcomes. *Australian Journal of Education* 54:277–306.

———, Ha MT, Craven RG & Yeung AS (2010b). Factorial invariance testing and latent mean differences for the Self-Description Questionnaire II (short version) with Indigenous and non-Indigenous Australian secondary school students. *International Journal of Testing* 10:47–79.

———, O'Rourke V, Dillon A, Craven RG, & Yeung AS (2012). Engaging the disengaged?: a longitudinal analysis of the relations between Indigenous and non-Indigenous Australian students' academic self-concept and disengagement. *Journal of Cognitive Education and Psychology* 11(2):179–195.

———, Whittaker A, Cooper E, Parada RH, Denson N & Bansel P (2017). Moving beyond essentialism: Aboriginal parental perceptions of school bullying and school engagement. In: Walters M, Martin KL & Bodkin-Andrews G (eds), *Indigenous children growing up strong*, Palgrave Macmillan, London, 153–178.

Briones E & Tabernero C (2012). Social, cognitive and demographic factors related to adolescents' intrinsic satisfaction with school. *Social Psychology of Education* 15(2):219–232.

Browne M W & Cudeck R (1993). Alternative ways of assessing model fit. In: Bollen KA & Long JS (eds), *Testing structural equation models*, Sage, Newbury Park, California, 136–162.

Buyse E, Verschueren K, Doumen S, Van Damme J & Maes F (2008). Classroom problem behavior and teacher-child relationships in kindergarten: the moderating role of classroom climate. *Journal of School Psychology* 46:367–391.

Campbell P, Biddle N & Paradies Y (2018). Indigenous identification and transitions in Australia: exploring new findings from a linked micro-dataset. *Population* 73:771–796.

Cho E & Kim S (2015). Cronbach's coefficient alpha: well known but poorly understood. *Organizational Research Methods* 18(2):207–230.

Cohen J, McCabe L, Michelli NM & Pickeral T (2009). School climate: research, policy, practice, and teacher education. *Teachers College Record* 111:180–213.

Craven R, Yeung A, Munns G, Bodkin-Andrew G, Denson N & O'Rourke V (2013). *Seeding success for Aboriginal primary students*, Centre for Positive Psychology and Education, University of Western Sydney & New South Wales Department of Education and Communities, Sydney.

Cronbach LJ (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16:297–334.

——— & Meehl PE (1955). Construct validity in psychological tests. *Psychological Bulletin* 52(4):281–302, doi:10.1037/h0040957.

Daraganova G, Mullan K & Edwards B (2014). *Attendance in primary school: factors and consequences*, Occasional Paper 51, Australian Government Department of Social Services, Canberra.

DiLalla LF, Marcus JL & Wright-Phillips M (2004). Longitudinal effects of preschool behavioral styles on early adolescent school performance. *Journal of School Psychology* 42:385–401, doi:10.1016/j.jsp.2004.05.002.

Dockery AM (2010). Culture and wellbeing: the case of Indigenous Australians. *Social Indicators Research* 99(2):315–332.

Ferla J, Valcke M & Cai Y (2009). Academic self-efficacy and academic self-concept: reconsidering structural relationships. *Learning and Individual Differences* 19(4):499–505.

Goodenow C (1993). The psychological sense of school membership among adolescents: scale development and educational correlates. *Psychology in the Schools* 30:79–90.

Hamre BK & Pianta RC (2007). Learning opportunities in preschool and early elementary classrooms. In: Pianta R, Cox M & Snow K (eds), *School readiness and the transition to kindergarten in the era of accountability*, Brookes Publishing, Baltimore, 49–84.

Hewitt B (2012). *The Longitudinal Study of Indigenous Children: implications of the study design for analysis and results*, Institute for Social Science Research, University of Queensland, Brisbane.

Hu L & Bentler PM (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychological Methods* 3(4):424–453.

Kline P (2016). *A handbook of test construction: introduction to psychometric design*, Methuen, London.

Kvien TK, Kaasa S & Smedstad LM (1998). Performance of the Norweigen SF-36 health survey in patients with rheumatoid arthritis. II. A comparison of the SF-36 with disease-specific measures. *Journal of Clinical Epidemiology* 51:953–959.

Ladd GW & Dinella LM (2009). Continuity and change in early school engagement: predictive of children's achievement trajectories from first to eighth grade? *Journal of Educational Psychology* 101(1):190–206.

——— & Price JM (1987). Predicting children's social and school adjustment following the transition from preschool to kindergarten. *Child Development* 58:1168–1189.

——— , Buhs ES & Seid M (2000). Children's initial sentiments about kindergarten: is school liking an antecedent of early classroom participation and achievement? *Merrill-Palmer Quarterly* 46(2):255–279.

Little TD (2013). *Longitudinal structural equation modeling*, Guilford Press, New York.

Maher M (2010). Indigenous teacher education initiative: shared conceptualisation leading to social justice and social capital in remote Australian Aboriginal communities. *International Journal of Interdisciplinary Social Sciences* 5(1):357–366.

Marsh HW (1990). A multidimensional, hierarchical self-concept: theoretical and empirical justification. *Educational Psychology Review* 2:77–171.

——— & MacDonald Holmes IW (1990). Multidimensional self-concepts: construct validation of responses by children. *American Educational Research Journal* 27(1):89–117.

National School Climate Center (2014). *School climate*, www.schoolclimate.org.

Nunnally, JC (1978). *Psychometric theory*, 2nd edn, McGraw-Hill, New York.

Peacock H & Prehn J (2017). *Successful Indigenous student transition to secondary education*, presentations from the AIATSIS National Indigenous Research Conference, 21–23 March 2017, Canberra, Australia, 1.

Pianta R (1992). *Student–Teacher Relationship Scale*, short form, University of Virginia, https://curry.virginia.edu/sites/default/files/uploads/resourceLibrary/STRS-SF.pdf.

——— & Stuhlman MW (2004). Teacher-child relationship and children's success in the first years of school. *School Psychology Review* 33:444–458.

Quin D (2017). Longitudinal and contextual associations between teacher–student relationships and student engagement: a systematic review. *Review of Educational Research* 87(2):345–387.

Ramelow D, Currie D, & Felder-Puig R (2015). The assessment of school climate: review and appraisal of published student-report measures. *Journal of Psychoeducational Assessment* 33:731–743.

Shepherd CC, Li J, Cooper MN, Hopkins KD & Farrant BM (2017). The impact of racial discrimination on the health of Australian Indigenous children aged 5–10 years: analysis of national longitudinal data. *International Journal for Equity in Health* 16(1):116.

Slaten CD, Ferguson JK, Allen K, Brodrick D & Waters L (2016). School belonging: a review of the history, current trends, and future directions. *Educational and Developmental Psychologist* 33(1):1–15.

Tabachnick BG & Fidell LS (2007). *Using multivariate statistics*, 5th edn, Pearson, Sydney.

Thompson B (2004). *Exploratory and confirmatory factor analysis: understanding concepts and applications*, American Psychological Association, Washington, DC.

Thurber KA, Banks E, Banwell C & LSIC team (2015). Cohort profile: Footprints in Time, the Australian Longitudinal Study of Indigenous children. *International Journal of Epidemiology* 44(3):789-800.

Trudgett M, Page S, Bodkin-Andrews G, Franklin C & Whittaker A (2017). Another brick in the wall? Parent perceptions of school educational experiences of Indigenous Australian children. In: Walters M, Martin KL & Bodkin-Andrews G (eds) *Indigenous children growing up strong*, Palgrave Macmillan, London, 233–258.

Walters M, Martin KL & Bodkin-Andrews G (eds) (2017). *Indigenous children growing up strong: a longitudinal study of Aboriginal and Torres Strait Islander families*, Palgrave Macmillan, London.

Ye F & Wallace TL (2014). Psychological Sense of School Membership scale method effects associated with negatively worded items. *Journal of Psychoeducational Assessment* 32:202–215.

Yeung AS, Craven RG & Ali J (2013). Self-concepts and educational outcomes of Indigenous Australian students in urban and rural school settings. *School Psychology International* 34(4):405–427.

Zubrick SR, Lucas N, Westrupp EM & Nicholson JM (2014). *Parenting measures in the Longitudinal Study of Australian Children: construct validity and measurement quality, waves 1 to 4*, Australian Government Department of Social Services, Canberra.

CENTRE FOR SOCIAL
RESEARCH & METHODS

+61 2 6125 1279
csrm.comms@anu.edu.au

The Australian National University
Canberra ACT 2601 Australia

www.anu.edu.au