

Productivity Commission Inquiry into the National Education Evidence Base

Submission on behalf of the Australian Research Council Centre of Excellence for Children and Families over the Life Course

Prepared by:

Professor Stephen Zubrick

Graduate School of Education, The University of Western Australia, Telethon Kids Institute.
Deputy Director, ARC Centre of Excellence for Children and Families over the Life Course (Life Course Centre)

Associate Professor David Lawrence

Graduate School of Education, The University of Western Australia, Telethon Kids Institute, ARC Life Course Centre

Ms Kirsten Hancock

Research Fellow, Telethon Kids Institute, and ARC Life Course Centre (UWA)

Dr Sarah Johnson

Research Fellow, Telethon Kids Institute, and ARC Life Course Centre (UWA)

Mr Francis Mitrou

Senior Research Fellow, Telethon Kids Institute, and ARC Life Course Centre (UWA)

Professor Janeen Baxter

Director, ARC Life Course Centre, Institute for Social Science Research, The University of Queensland

Summary / Overview

Improving Australia's national education evidence base would be a valuable resource for the continued development and evaluation of education policy in Australia. There are opportunities to make improvements in two broad areas — improving access to, and use of, existing data sets, and filling gaps that currently exist.

There is already a large amount of routinely collected education data in Australia that could be accessed more efficiently and used more effectively in order to improve outcomes for Australian students – including their educational outcomes, but also other aspects of health and wellbeing. Broadly, we find that the issues paper accompanying the inquiry has identified the major benefits, impediments and limitations of developing an education evidence base, and we commend the Productivity Commission for their coverage of the issues that have been considered so far. In this submission, we draw upon our own experiences accessing, analysing and interpreting education data sources to highlight what we think are the most significant limitations or concerns regarding the development of a national education evidence base.

Our experience:

This submission is prepared on behalf of researchers working within the Australian Research Council Centre of Excellence for Children and Families over the Life Course (Life Course Centre, or LCC). The LCC tackles the problem of deep and persistent disadvantage, characterised by the spread of social and economic poverty within families and across generations despite improvements in broader society. The potential to improve educational experiences and outcomes for children in disadvantaged families is integral in addressing deep and persistent disadvantage in Australia, and thus, we strongly support efforts to enhance education research in Australia. LCC government partners include the Australian Government Department of Education and the Victorian Government Department of Education and Training, and there are many LCC research fellows and affiliate researchers across the country currently conducting education research to fulfil the goals of the LCC. Our prior experience with education datasets, and our understanding of how better access to high quality data can shape and inform education policy in Australia underpin the comments we make in this submission.

The LCC researchers based in Western Australia have an extensive history of developing and working with population-level datasets. This has allowed us to examine trends and relationships for populations of children and their health and educational outcomes. Our work includes the development and administration of the *Western Australian Child Health Survey* (Zubrick et al., 1997), the *Western Australian Aboriginal Child Health Survey* (Zubrick et al., 2004), the *Second National Survey of Mental Health and Wellbeing of Children and Adolescents* (Young Minds Matter, Lawrence et al, 2015). Each project involved—and was enhanced by—extensive administrative linkage to health and education records. We have also worked with the Western

Australian Department of Education to use their administrative enrolment, attendance and achievement records to examine the attendance patterns of Western Australian students over time, and how these patterns correspond to achievement outcomes (Hancock et al., 2013).

More generally, the team has extensive experience in undertaking research using linked administrative data sets with Western Australia being an early leader in data linkage. This has given us the experience and knowledge to identify barriers to access, factors that aid linkage, data quality issues (e.g. family characteristics), and to identify further information that would be useful to collect to more broadly inform child development research. Beyond this, the LCC is also forging new partnerships with government departments to identify and make available administrative data for purposes of research and planning (DSS).

1.1 Governance and legal frameworks

In our view, the most significant obstacle to developing a National Education Evidence Base is a lack of governance and legal frameworks supporting the collection and sharing of data across multiple jurisdictions and government portfolios. Differences in data governance structures and privacy legislation across states, in addition to varying governance structures between the Government, Catholic and Independent sectors are associated with a raft of issues that make collating consistent data across the country very difficult. These include:

- Ownership and responsibility for collecting and managing data. Who owns data (e.g. NAPLAN records, attendance records), and who has the right to decide to share it with others?
- Different systems in place around the country (for example, for collecting routine administrative data, grading, reports) result in inconsistencies in the way data are collected, and therefore, comparability of data from different states and sectors. This poses significant barriers to harmonisation of data.
- There are also state and sectoral differences in starting age, finishing age, vocational pathways, attendance strategies, reporting and tertiary entrance assessments. These will necessarily occur when states are responsible for managing the education of students. While these variations may offer possibilities for counterfactual analysis and worthy comparisons, they potentially restrict the range of policy reach because of restrictions to generalisability.

Example – Linkage of Young Minds Matter survey data to NAPLAN

Young Minds Matter, the second Australian Child and Adolescent Survey of Mental Health and Wellbeing, surveyed over 6,300 families with children aged 4-17 years from around Australia. Funded by the Australian Government Department of Health, the survey set out to measure the prevalence and burden of mental disorders in Australian children and adolescents, and the use of

services, and unmet needs for services, in the health and education sectors. The survey found that mental disorders are common, persistent, and burdensome in children and adolescents, and have significant impacts on participation in education. Schools were also identified as major providers of support services for children and adolescents with mental disorders (Lawrence et al., 2015).

In Young Minds Matter (www.youngmindsmatter.org.au), survey participants (parents and children) were asked for their consent to link their survey information to their NAPLAN records. The purpose of this linkage is to examine how mental health problems in children and adolescents—in addition to the range of data collected on school experiences, family characteristics and service use—relates to educational outcomes. The survey also sought consent to access Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS) data. In comparison with accessing the MBS and PBS data, which was provided by a single data custodian, the process for obtaining approvals and ultimately obtaining NAPLAN data was long and costly. Even though survey participants had consented to linkage, project staff needed to contact each state and the different sectors in each state in order to access the data. The pathway for data access varied substantially for each of the Test Administration Authorities (TAAs) in each jurisdiction, with different procedures for application, varying access criteria and procedures.

This imparted uncertain and lengthy time horizons for third-party deliver of data. In addition, many of the application forms were generically designed to deal with the conduct of research in schools, not specifically for data already collected such as NAPLAN data. The inconsistency leads to differential assessments of research applications in terms of the criteria applied and the decisions made to support the request for data. From the perspective of the TAAs that are not independent of the State Education Departments, and organisations representing non-government schools, requests for NAPLAN data adds an extra burden to resources in terms of those processing requests to undertake research in schools.

A further complicating factor with applications for NAPLAN data arose from agreements between the TAAs and non-government schools about the use of NAPLAN data. In those States where the TAA is located within the State Education Department, permission for the release of NAPLAN testing results to third parties could only be given for public students. In these jurisdictions, NAPLAN test results from independent and Catholic schools could only be released to ACARA and therefore, we were required to gain additional permissions from Catholic and independent schools to gain access to this data. While such permissions for Catholic schools can be obtained from the representative body at the State level (Catholic Education Office or Commission), permission had to be sought from each individual independent school. The Independent Schools Associations did not have the authority to make decisions about access to NAPLAN data on behalf of its members. While all permissions were eventually gained, there were some initial refusals due to issues of resourcing and sensitivity about ownership of data, and it was necessary to negotiate access.

A central repository of identifiable NAPLAN data (for example, with ACARA) would have been a much more efficient process in linking this data. An important benefit would be greater consistency in the warehousing of NAPLAN data and in the criteria and protocols for providing access to third parties. The same rules for extraction and availability of data can be applied uniformly. At present, it seems that TAAs have varying capacity to maintain a database of linked NAPLAN test results for individual students, and varying capacity to support requests for data. From the perspective of data collection, it would mean a central repository of data already existed without the need for transfer of data from each State and Territory. It would also mean reduced administrative load in processing applications for data from researchers.

A further issue associated with access to NAPLAN data in Young Minds Matter was the requirement to agree that no analysis would be undertaken by school sector. While the survey collected information on whether the child or adolescent was attending a public, Catholic or independent school, this information will not be used in any analysis.

In terms of public acceptance of linking survey information with other datasets, we find there is generally wide acceptance among survey participants for linkage to occur. For example, 87% of parents and carers of 5-17 year-olds in Young Minds Matter gave their consent for obtaining their child's NAPLAN test results. With the Longitudinal Study of Australian Children (LSAC), over 95% of eligible parents in the pre-primary cohort gave their permission to link LSAC survey information with their study child's NAPLAN test results from the relevant testing authority, in addition to MBS and PBS data (Daraganova, Edwards & Siphthorp, 2013). Of the 5% where valid consent was not given, more than half were due to forms not being filled in correctly (e.g. box not ticked or signature not provided). The consent rate for the Western Australian Aboriginal Child Health Survey information to be linked to health and education administrative data including birth and hospital records was also very high at 96% (Zubrick et al., 2004).

1.2 Resources

To be used effectively, the national education evidence base needs to provide access to data that is high quality and that can be delivered and accessed in a timely manner. The resources required to achieve this need to be fully considered and recognised. These include:

- Resources necessary to support the states in collecting data, including the development, rollout and maintenance of IT systems, staff support and training in how to enter and access data.
- Resources to collate and clean data (e.g. systems and analyst staff in state education offices). Each jurisdiction requires a capacity to manage data collections, particularly if updates are required to a national system. Standardising and improving quality requires committed resources.

- Resources to share or provide access to data, and respond to data and data linkage requests. Again, this requires committed resources. In environments where budgets are restricted, allocating resources to managing and sharing data may be less of a priority.
 - o A major limitation of data linkage is the time taken to perform linkages, with permissions required from individual data custodians. Data linkage may not be a priority for particular custodians. The time it takes to produce data extracts is a major impediment to some research and it would be beneficial for both data custodians and researchers to create greater efficiencies or provide better financial support for these processes. This is evident even with an established system such as in WA, where lengthy delays to access data are common.

1.3 Analytic and research capacity

As noted on page 29 of the Issues Paper, ensuring sufficient analytic and research capacity to analyse large and complex datasets is integral to the success of the national education evidence base meeting the outlined goals and objectives. Any investment in a national education evidence base needs to include the requisite investment in developing and maintaining analytic capacity to maximise the value of the data collected.

Is there research capacity in Australia? There is arguably a large volume of education research already produced in Australia, including from government departments and institutes (e.g. Australian Institute of Family Studies, Australian Institute of Health and Welfare), independent institutes (e.g. Grattan Institute), and multiple departments across multiple universities, including education, but also disciplines such as economics, sociology, psychology, and epidemiology, and within research centres like the LCC. This research already draws upon an extensive range of data, many of which are listed in Table 1 in the Issues Paper. In addition to these resources, researchers also independently collect from schools and students for specific research projects. Currently, it is unclear how much of this research is based on population-level datasets (as opposed to small scale or qualitative research), or the extent to which education research informs policy or practice in Australia. Arguably, government commissioned research, or research conducted in partnership with government, is likely to have greater impacts on policy because such research directly addresses the information needs of government.

Any framework supporting the development of a national education evidence base should consider how research is initiated, funded and disseminated to ensure that researchers and policy makers linked in with data collected for the evidence base work in tandem to produce research that is relevant to the needs of government and policy makers. Part of this would involve funding research that explicitly meets the needs of policy makers, educators, families and students.

1.4 What data are desirable?

Notwithstanding the constraints and obstacles already identified (e.g. governance and resources) there are other avenues of data collection that would usefully inform about student learning outcomes.

More information on the early years

One of the largest education gaps occurs in the years between birth and when a child enters compulsory schooling. There is very little information routinely collected during this developmental period that usefully informs about child development. A reliance on administrative or population-level databases to track children's early education experiences would not capture a significant proportion of children whose families are not engaged in early childhood education, child care, or maternal health programs. For these children – potentially those at greatest risk of lower educational outcomes, collection of specific data is more valuable.

In this respect, a highly valuable data collection would be a standardised national data collection from parents upon enrolment in school (preschool). This data collection would ideally collect a snapshot of information from parents about the early education experiences of children (e.g. child care, playgroups, parenting programs, early education programs, along with standard family demographic information, number of siblings and so on) and a brief assessment of the child's school readiness.

The Australian Early Development Census (AEDC) is collected once every three years on the population of 5 year old children entering Year 1, and represents the first standardised collection of data of young children entering school. The instrument is designed as a community development tool to determine children's broad developmental standing with reference to local communities, and not as a measure that has been specifically designed to use as a measure of school readiness. While the Australian Government's commitment to the AEDC reflects our increased understanding of the critical importance of early childhood education, and the government's commitment to ongoing reform in this sector, the simple fact that the measure is used once every three years and not yearly keeps some of this enthusiasm in check.

Something that could help fill gaps in the AEDC methodology for low additional cost or effort would be the development of a nationally standardised short questionnaire for parents upon entry to compulsory schooling for their children. All school types across all jurisdictions already collect basic information from parents about themselves and their children upon initial school entry. A program to standardise these data items, and their collection via a secure remote repository, would provide the missing link between the limitations of the AEDC to predict outcomes at the individual child level and academic performance measurement represented by NAPLAN. Agreement on what represents the start of compulsory schooling would need to occur. A similar

system at first entry to secondary school would provide useful follow-up at another critical developmental period.

Currently, the best information available about the early educational experiences of children exists in the LSAC B-cohort data. As the youngest cohort of children in the LSAC are now 12–13 years old, the information collected on these children between 2004 and 2008 is now somewhat out of date. An updated survey with a new cohort of infants (perhaps with sample ‘refreshment’ every few years) would allow a more comprehensive evidence base of children’s early education and development. The value of such a survey would be enhanced through data linkage to birth and maternal health records, and other education and health records. Building in a capability in the data collection to routinely evaluate policy, trials and interventions would also significantly enhance the value of the data collected.

Other relevant data

The significant emphasis placed on NAPLAN as an indicator of student and school performance reflects government and community interest in improving literacy and numeracy domains. The high profile measurement of these domains, however, does bias the evidence base in favour of these outcomes. The national curriculum and national and jurisdictional education policies reflect a far greater diversity of domains that are expected to be influenced by Australia’s education system. Data collection across these additional domains is necessary to evaluate policy and program initiatives. More importantly, data collection in these domains is also likely to influence allocation of resources within the education sector. The collection of a more diverse range of wellbeing indicators would help to understand a broader array of factors that influence more ‘traditional’ student learning, such as measures of social and emotional wellbeing.

A record of what extra-curricular programs are being implemented that support the learning and social-emotional needs of students would be a worthwhile inclusion. For example, which schools are implementing MindMatters (www.mindmatters.edu.au) and Kids Matter (www.kidsmatter.edu.au), and other wellbeing and life skills programs, bullying initiatives and so on. This type of information may be useful to answer research questions about school resourcing/programming to assist student wellbeing, or to evaluate resourcing of non-teaching staff. One of the greatest areas of unmet need identified by Young Minds Matter was for life skills training (Lawrence et al., 2015).

The national Teaching Workforce Dataset could be expanded to reflect the contribution of non-teaching staff, such as counsellors, psychologists, nurses and chaplains. In conjunction with other education data, this information could usefully inform about the resources available to, or taken up by, schools and how these resources correlate with student learning outcomes.

Finally, linkage of individual or family records to the National Early Childhood Education and Care Workforce Census by geographic area would provide information on the availability and

characteristics of local childcare centres. This would enable research on the importance of the presence of local childcare facilities on children's development.

1.5 The development of a national education evidence base, alone, will not lead to improved educational outcomes

This statement relates to the question 1 on page 7 of the issues paper – *Do you agree that the objective of a national education evidence base should be to improve educational outcomes? Are there other objectives that should be included?*

A national education evidence base would serve as an integral component of lifting educational standards, but only through the potential to monitor progress or to enhance research and evaluation capacity. If the main objective of having a national education evidence base is to improve educational standards, the objective is destined to fail because no program logic or theory of action has been provided to connect the evidence base with improved educational outcomes – and this is likely to be beyond the scope of developing national education evidence base.

Administrative datasets are necessarily broad, and are therefore mainly useful in identifying broad problems and potential areas for intervention. For example, our report on student attendance clearly showed that students who were frequently absent had lower NAPLAN scores (Hancock et al, 2013), suggesting that one way of lifting NAPLAN results for vulnerable student populations is to improve their school attendance. Yet the attendance data revealed very little about *how* to change attendance patterns. It didn't tell us why students were missing from school, only whether their absences were authorised or unauthorised. Absences could be due to chronic illness, parent engagement, transport issues, or student-driven reasons like separation anxiety, among other reasons. Each of these factors would need a different intervention in order to improve attendance, and these interventions are not informed by broad databases. Data that provides more specific information about the *reasons* underpinning absence (e.g. specific absence codes for illness, family holidays etc) would be useful to understand which absences are more problematic than others (e.g. illness versus truancy) and the extent which to avoidable absences might be reduced, but simply collecting the information is unlikely to reduce chronic absence problems.

1.6 Caution should be exercised about how the evidence base is used

The Issues Paper notes that *“another issue relates to the ways in which data are used. Early consultations with stakeholders suggested that perceptions of the purposes for which data could be used (for example whether analysis of the data would be used as a basis for funding decisions) can affect the willingness (or otherwise) to provide, share and link data; the quality of data collection; and the publication of data and analysis”*

This is a key consideration, and we concur with these views.

Ideally, education policy should be informed by rigorous research and evaluation based on quality data, and a national education evidence base would in turn contribute to such research. However, policies that make data “high stakes”, for example by linking school funding or teacher pay to NAPLAN results or attendance data, may invite a risk of the data being compromised and therefore undermine the integrity of particular datasets. In the United States, where high-stakes testing has become pervasive after accountability policies were written into the *No Child Left Behind Act* of 2001, many states have experienced systematic corruption among educators to artificially increase test scores of students through a range of behaviours, including changing test answers of students. For example, 11 educators in Atlanta were convicted in April 2015 of racketeering charges relating to widespread cheating on standardised tests, with evidence of cheating in 44 of the 56 schools examined (Georgia Bureau of Investigations, 2011). Another example (among many) is in Columbus, Ohio, where school officials were found guilty of falsifying student attendance and grade records in order to improve their schools’ standing on state report cards. Along with receiving undeserved bonuses for ‘improving’ school performance, some students missed out on private school vouchers because their schools appeared to be achieving at a much higher level than was the case.

These examples from the United States provide a cautionary tale about how data might be inappropriately modified to ‘improve student outcomes’. These practices have implications for the quality of the data collected, an in turn compromise the usefulness of the data for the things it would actually be good for, such as monitoring student progress over time.

1.7 Multiple data sources and informants (ideally linked) are required to fully understand the contexts in which students learn and develop

The Issues Paper rightfully acknowledges the many influences on student progress, including student, family, school and political influences. The national education evidence base needs to include detailed information that reflects the contribution of each of these factors to learning outcomes.

Some of these influences are best captured by linking administrative records (e.g. student attendance) and other types of data are better captured through other means like surveys. The most valuable datasets are those that combine elements of both – for example, the linkage of the Longitudinal Study of Australian Children or Young Minds Matter data to NAPLAN records provides much needed student, family and community contexts to children’s educational experiences, that allow a better understanding of why some children excel and others fail to meet minimum standards.

Where population-level data can be collected this should continue, but these could be supplemented with well-designed surveys that capture other elements that are more difficult to collect, but that would provide important context to the population data (e.g. parent background items like engagement and involvement).

References

Columbus City Schools Office of Internal Audit (2012). Student Altered Attendance Records Review for the School Year 2010-2011: Special Review. Report available at <http://www.dispatch.com/content/downloads/2012/12/Student-Altered-Attendance-Special-Review.pdf>

Daraganova, G., Edwards, B., & Siphthorp, M. (2013). Using National Assessment Program-Literacy and Numeracy (NAPLAN) data in the Longitudinal Study of Australian Children. Melbourne: Australian Institute of Family Studies.

Georgia Bureau of Investigations (2011). Report to the Governor of Atlanta. Available at: <http://www.courthousenews.com/2011/07/27/APS1.pdf>

Lawrence, D., Johnson, S. E., Hafekost, J., Boterhoven de Haan, K., Sawyer, M., Ainley, J., & Zubrick, S. R. 2015. The mental health of children and adolescents. Report on the second Australian Child and Adolescent Survey of Mental Health and Wellbeing. Available at: <http://www.health.gov.au/internet/main/publishing.nsf/Content/mental-pubs-m-child2>

Hancock, K. J., Shepherd, C. C. J., Lawrence, D., & Zubrick, S. R. (2013). Student attendance and educational outcomes: Every day counts. Report prepared for the Australian Government Department of Education. Available at: http://telethonkids.org.au/media/472779/final_report_2013.pdf

Zubrick, S.R., Lawrence, D., Silburn, S.R., Blair, E., Milroy, H., Wilkes, T., Eades, S., D'Antoine, H., Read, A., Ishiguchi, P., Doyle, S. (2004). The Western Australian Aboriginal Child Health Survey: The Health of Aboriginal Children and Young People. Perth. Telethon Institute for Child Health Research.

Zubrick, S.R., Silburn, S.R., Gurrin, L., Teoh, H., Shepherd, C., Carlton, J., Lawrence, D. (1997). Western Australian Child Health Survey: Education, Health and Competence, Australian Bureau of Statistics and the TVW Telethon Institute for Child Health Research.