



Footprints in Time: The Longitudinal Study of Indigenous Children

A guide for the uninitiated

Mick Dodson, Boyd Hunter and Matthew McKay

Indigenous disadvantage has many of its roots tied to experiences found within the context of early childhood. Policy recognises this as the overarching Overcoming Indigenous Disadvantage framework and includes positive child development as one component of the three priority areas (Steering Committee for the Review of Government Service Provision [SCRGSP], 2011). Of the other two priority areas, there is considerable focus on the importance of positive social environments in the local community and family. Given this emphasis, the historical lack of adequate research on the factors associated with positive development of Indigenous children, vis-à-vis other Australian children, constrain the ability of policy to achieve its stated goals. The very concept of child development means that it is a process that evolves, so the absence of longitudinal data on Indigenous children is one of the main reasons for poor evidentiary basis for policy.

Footprints in Time: The Longitudinal Study of Indigenous Children (LSIC) may be a crucial

landmark for the development of an effective policy to address Indigenous disadvantage early in the life cycle. Indeed, the main goal of LSIC is to determine what helps Aboriginal and Torres Strait Islander children “grow up strong” in their communities. After a preliminary development phase, the 2007 Federal Budget set aside funds to collect data for LSIC (Australian Government, 2007):

The Australian Government will conduct a targeted longitudinal survey twice a year from 2008 to 2011. The survey will collect data on Indigenous babies under 12 months and 4 to 5 year old Indigenous children from 1,650 families in 11 areas covering urban, regional and remote communities.

The main fieldwork for Wave 1 started on 21 April 2008 and finished on 23 February 2009. The eventual survey collected is only annual, which reflects the practical difficulties of conducting a unique and unprecedented survey such as LSIC (e.g., it took more than 6 months to collect Wave 1). The initial funding has been extended: Wave 4 interviews were completed late last year and now LSIC is in the field for

The main goal of LSIC is to determine what helps Aboriginal and Torres Strait Islander children “grow up strong” in their communities.

Wave 5, with additional waves in development. It is worth noting that LSIC’s sample size is substantial as it represents between 5 and 10% of Indigenous children of the appropriate ages.

It is time to take a step back and ask the question: How useful is the information provided? In order to answer this question it is necessary to reflect on the content of the survey, as well as how the survey data were collected and how that data were coded.

Rather than simply replicate information provided in the Data User Guide (Department of Families, Housing, Community Services and Indigenous Affairs [FaHCSIA], 2012a), this paper provides analysis and some personal reflections on the utility of LSIC data. It is informed by extensive experience of the authors in collaborating with Indigenous communities and conducting research firmly grounded in scholarly and peer-review processes. Two of the current authors were involved in the process of design of LSIC: Mick Dodson had been the chair of the LSIC Steering Committee since its inception in 2003, while Boyd Hunter was on the design sub-committee from 2003 and, more recently, was a member of the Steering Committee until 2011. Neither of the authors are specialists in child development but both have a long and extensive grounding in Indigenous research. For example, between August 1988 and October 1990, Dodson was Counsel assisting the Royal Commission into Aboriginal Deaths in Custody, which highlighted the need for data and research to analyse the embedded disadvantage that starts with early family life.¹

The paper is structured as follows. The next sections give a brief history of LSIC, which includes an extended rationale for the need for

such data and directly reflects on the survey design and methodology. This discussion leads to an analysis of the strengths and weaknesses of LSIC with reference to a few selected variables that may be useful in potential research. The penultimate section will attempt to identify some useful research questions that LSIC data may be used to address. The final section provides some concluding remarks that, among other things, reflect on growing research using LSIC data. There is clearly a high level of demand for good quality research that illuminates the processes driving positive development for Indigenous children. A key question we wish to address here is whether LSIC adequately meets the need of acquiring quality information to best inform policy on how to foster positive child development. The main contribution of this paper is to highlight what LSIC data can tell us, by providing some historical context about the survey design and collection.

Rationale for collecting longitudinal data about Indigenous children

The collection of longitudinal data is a relatively recent phenomenon in Australia. One of the first major longitudinal studies was the Household, Income and Labour Dynamics in Australia (HILDA) survey, which started collecting data in 2001 (Wooden & Watson, 2007). An impressive body of research based on HILDA is being collated, but the small Indigenous sub-sample means that it is not overly credible for research focusing on the first Australians. In any case, while HILDA allows analysis of adult outcomes and broader family dynamics, it is not particularly useful for understanding children.

The biennial *Growing Up in Australia: The Longitudinal Survey of Australian Children (LSAC)* does allow considerable analysis of the processes underlying child outcomes, but again the Indigenous sub-sample is probably too small to allow credible analysis of the developmental processes of Indigenous children. In 2003–04, LSAC collected data on a cohort of 5,000 children aged 0–1 years and a cohort of 5,000 children aged 4–5 years. Information collected from the study child, their parents (including both parents in separated families), their carers and teachers includes details of the children’s physical health and social, cognitive and emotional development, as well as their experiences in key environments such as the family, community, child care, preschool and school settings. LSAC was designed to



be representative of Australian children and contained more Indigenous respondents than HILDA, with 300 Indigenous mothers and 171 Indigenous fathers, the majority of whom resided in regional areas. Hunter (2008) argues that LSAC's Indigenous sub-sample from remote areas should be treated with caution, as Wave 1 data were inconsistent with representative cross-sectional Indigenous data from the 2002 National Aboriginal and Torres Strait Islander Social Survey (NATSISS). Furthermore, he argued that the relatively high attrition rates for mobile Indigenous families in subsequent waves would make it difficult to draw reliable inferences for any longitudinal analysis of Indigenous families from non-remote LSAC data.

The mobility issue is one manifestation of the rather distinctive (and diverse) social and cultural circumstances facing Indigenous peoples. Indigenous peoples are not only more likely than other Australians to live in remote areas, but the Indigenous cultures are so radically different from each other and from that of other Australians that it cannot be presumed that particular questions and answers will mean the same thing to various respondents (see various chapters in Arthur & Morphy, 2005).

Survey methods and contents also need to be relevant for Indigenous contexts to ensure that policy-makers have a solid foundation for their policy initiatives (Stewart, Lohar, & Higgins, 2011). The information on child functioning from LSIC is likely to be more informative, and culturally nuanced, than LSAC as many questions take into account the Indigenous context. For example, the instruments to measure vocabulary in LSIC take into account the possibility that Indigenous children will use an Indigenous language (see Australian Council for Educational Research [ACER], 2009, Table 10), whereas there is no facility to do so in LSAC. Furthermore, the heavy reliance of LSAC on telephone interviews and self-complete questionnaires means that it is more difficult to identify and redress cultural and other misapprehension of the question being asked.

The LSIC Steering Committee recommended that the ultimate focus of the study be on discrete communities as this was a more practical and cost-effective option for collecting detailed information on the local Indigenous context. Despite the manifest strengths of LSIC's approach, the fact that it is confined to 11 sites means that the resulting analysis cannot be generalised to the Indigenous population at large. It could be argued that the LSAC

Indigenous sub-sample has the advantage, in that it could be used for making tentative statements about the relevant populations in regional areas and metropolitan areas, but not remote areas (see Hunter, 2008).

To be fair, LSAC was not designed to provide a basis for the analysis of Indigenous child development. Nicholson, Sanson and the LSAC Research Consortium (2003) concluded that more intensive studies of subgroups were better conducted as separate studies. They argue that it would not constitute an efficient use of the LSAC sample to increase the Indigenous sub-sample to enhance the reliability of information provided for that section of the population.

By 2003, there were sound arguments for a specialised survey to begin to understand Indigenous child development. The 2003–04 Federal Budget provided the initial resources for the LSIC study. The first phase, from September 2003 to June 2004, involved extensive consultation with Indigenous peoples and communities about the study. The design and development of the study commenced in December 2005, with pilot testing continuing through 2006 and 2007. Wave 1 data have been available for researchers to analyse since late 2009, while data for Waves 2 and 3 were released in first half of 2011 and 2012.

Design of LSIC

The central issue for the design of LSIC was how to collect reliable data that captured the development pathways of Indigenous children and the full diversity of Indigenous-specific circumstances facing their parents and families.

LSIC is managed by FaHCSIA, and the LSIC Steering Committee has overseen the design, development and implementation of the study since 2003. Committee members are drawn from academic and community backgrounds, covering a wide range of disciplines such as health, early learning and child care. Sub-committees of the Steering Committee were formed to deal with particular issues as required. For example, the Design Sub-Committee provided expert advice on survey design and content.

Both the Steering Committee and Design Sub-Committee have a strong Indigenous representation and have an impressive commitment to community consultation. It is not only ethical to undertake considerable community consultation, but it is argued that consultation is essential for securing a high response rate to the initial wave and low attrition rates over subsequent waves. The reasoning was simple: the more information

A key question we wish to address here is whether LSIC adequately meets the need of acquiring quality information to best inform policy on how to foster positive child development.

Indigenous mobility is arguably complex and fundamentally different to that identified for other Australian groups.

that is available to the surveyed Indigenous families, and the greater the sense of local ownership of the study, the less resistance would be encountered to ongoing participation in LSIC.

The primary Human Research Ethics Committee (HREC) for the study is the Australian Government Department of Health and Ageing Departmental Ethics Committee. In addition, more decentralised regional ethics clearance and support were obtained for LSIC sites through state and territory HRECs or their equivalents (in accordance with the National Health and Medical Research Council and Australian Institute of Aboriginal and Torres Strait Islander Studies guidelines). The relevant departments of education and Catholic dioceses were also consulted to gain permission and support for preschool and school teachers to complete questionnaires about the children involved in the study. State and territory departments managing out-of-home care were also consulted. The agreement and approval to participate in the study was sought from communities and elders in these sites before research within the communities began.

There were two fundamental questions that needed to be resolved before credible information on Indigenous child development could be collected:

- What was the optimal sample size?; and
- How should the survey methodology accommodate Indigenous diversity?

In 2010, there were 16,100 births registered in Australia where at least one parent identified themselves as being of Aboriginal and Torres Strait Islander origin on the birth registration statement (Australian Bureau of Statistics [ABS], 2011). This number is rather small and precludes a large birth cohort survey (unless the budget was unconstrained), especially given the geographically dispersed nature of the underlying population. In the original deliberations about the LSIC sample at the Steering Committee and Design Sub-Committee, it was suggested that the ideal sample size would be at least 2,000 respondents per cohort—substantially fewer than LSAC, but sufficient to capture the manifest diversity within the Indigenous Australian community. Ultimately, this aspiration was not realistic within extant budget constraints as it became apparent that collecting Indigenous data was fundamentally different to collecting information from other Australians due to the unique challenges of collecting culturally sensitive and varied information in sometimes inaccessible locations and difficult circumstances. A plausible methodology was identified after substantial

debate and it became clearly evident that there was an additional cost in collecting information for Indigenous children.

The first challenge was how to find suitable families with children. For example, Medicare information on Indigenous status has only been collected on a voluntary basis since 2002—hence, Indigenous identification on administrative records is at best partial, meaning that the LSIC sample could not rely solely on such information to find suitable respondents. Accordingly, it was sometimes deemed necessary to supplement the sample using a “snowballing” methodology to identify additional potential respondents using known social networks. Snowball sampling is a form of non-probability sampling; another form of this sort of sampling is ad hoc quotas, where certain types of communities are sampled (see Box 1).

Given that the LSIC design deliberately focused on 11 sites, chosen in part to cover the range of socio-economic and community environments where Aboriginal and Torres Strait Islander children live, it is reasonable to construe LSIC as being based on implicit quotas. Non-probability sampling has been criticised on the basis that the resulting data and analysis is difficult to interpret because the respondents are selectively drawn from the population and hence data may be highly correlated and potentially biased (Magnani, Sabin, Sidel, &

Box 1: LSIC sampling

LSIC uses a non-random purposive sampling design that cannot be construed as a representative sample. Given that there are relatively few Indigenous children in the Australian population, potential respondents were identified using local social networks. LSIC sites were specifically chosen (non-randomly) to:

- represent the broad distribution of Aboriginal and Torres Strait Islander people around Australia;
- ensure approximately equal representation of urban, regional and remote areas, thus enabling some geographical comparison;
- contain a substantial Aboriginal and Torres Strait Islander population in the core and surrounding areas;
- include locations engaged in the pilot of the study where existing relationships could be built upon; and
- be located near relevant government offices, where the Indigenous interviewers could be based.

Heckathorn, 2005). In essence, this issue is just another aspect of the fact that the LSIC sample is not representative, a point that we will reflect on further below.

Another challenge for LSIC is the high rates of mobility among Indigenous Australians that affect the ability to follow up respondents in subsequent waves. Indigenous mobility is arguably complex and fundamentally different to that identified for other Australian groups (Taylor and Bell, 2004).

Geographically mobile populations are notoriously difficult to survey, especially in a cross-cultural longitudinal context (Martin & Taylor, 1996; Smith, 1992). The mobility-related issues documented in Hunter and Smith (2002) are likely to increase the costs of any longitudinal survey, especially those with an adequate sub-sample of Indigenous respondents. Indigenous mobility potentially affects (selective) individual response rates, subsequent relocation of respondents and sample attrition, and hence on data quality. Hunter and Smith suggest several strategies for minimising the cost of such a survey, including expanding operational definitions of households and combining a mix of qualitative and quantitative methodologies to maximise the chance that households can be found again (by enhancing the relationship between interviewer and respondent and hence arguably improving data quality).

Hunter and Smith (2002) identified several precedents for successful, relatively cost-effective longitudinal data collection where researchers worked closely with Indigenous research facilitators from each community. The argument was based on the need to engage the Indigenous community while maintaining the scholarly integrity of the survey methodology.

The community engagement strategy is integral to maximising participation and retention in a longitudinal survey of an Indigenous population that is arguably suffering survey fatigue. It was essential that Indigenous communities understood why LSIC was important and how the survey information was to be collected and analysed. The LSIC committees discussed how this might be best achieved and recommended that a series of reports and local profiles be provided to the communities involved, along with an extensive process of information dissemination at both a local and national level (e.g., workshops, seminars and conferences).

Initially, six full-time Indigenous Research Administrative Officers (RAOs) were employed and trained to manage the community engagement activities for the initial pilot

research in 2006–07 (including consent processes, data collection and dissemination of information in pilot communities). This process was deemed to be so successful that it was extended, enabling all interviews for the final LSIC survey to be conducted by RAOs. Even though RAOs tended to initially have less expertise or experience than most professional researchers or interviewers, their commitment was unquestionable and the formal skills were gained on the job. RAOs' connection with the local communities and culture clearly meant that they had vital advantages over other professional interviewers.

Readers may not realise how radical the community engagement process and employment of Indigenous interviewers was, but this unprecedented initiative (for a large-scale survey) has been responsible for the impressively high rates of ongoing participation in the survey that surprised many (including the authors of this paper). However, by going down the path of community engagement and RAOs, there were clearly additional costs involved: the direct cost of liaising with communities and preparing customised reports, and the training costs of building the research capabilities of Indigenous interviewers to work on the ground. While a nationally representative sample was desirable for the ability to generalise findings, and the community engagement strategy was deemed essential for getting credible information on Indigenous children over time. Adopting a methodology that solely relied on Indigenous RAOs may not have been possible for a nationally representative sample as it would have been almost impossible to employ, train and retain suitable RAOs throughout Australia (especially where the Indigenous community is relatively scattered). Community engagement was deemed to be both integral to the survey design and more affordable/practical than having a representative survey in a dispersed population, and hence the strategy involving Indigenous RAOs was formally endorsed.

After 2007, LSIC strategically focused solely on 11 sites and it was clear that the results would not be nationally representative. While it was a cost effective decision to focus resources on particular areas, LSIC arguably became more of a “proof of concept” at that stage rather than a provider of potentially definitive information on Indigenous child development. All research arising from LSIC data must be conditioned on the fact that the data was collected from discrete areas with a particular history. However, the focus on clearly identified areas had the distinctive advantage that it allowed a more intense community involvement. As the LSIC

It is not only ethical to undertake considerable community consultation, but it is argued that consultation is essential for securing a high response rate to the initial wave and low attrition rates over subsequent waves.

The Indigenous cultures are so radically different from each other and from that of other Australians that it cannot be presumed that particular questions and answers will mean the same thing to various respondents

Data User Guide indicates, the study would never have been possible without the support and trust of the Aboriginal and Torres Strait Islander families who opened their doors to the researchers and generously gave their time to talk openly about their lives (FaHCSIA, 2012).

Following the precedent set by LSAC, the LSIC Steering Committee recommended an accelerated, cohort cross-sequential design to allow an efficient use of the sample achieved, providing for coverage of two cohorts that would overlap after four years. LSIC identifies two cohorts of Indigenous children aged from 6 months to 2 years (Baby cohort) and from 3 years, 6 months to 5 years (Child cohort) in Wave 1 (FaHCSIA, 2012a). The design allows the data covering the first 9 or 10 years of Aboriginal and Torres Strait Islander children's lives to be collected in 6 years. The two-cohort design also facilitates the comparison of the cohorts when their ages overlap, in order to detect any changes due to different social conditions and policy initiatives.

The LSIC committees also spent considerable time exploring the possibility of including an explicitly qualitative research component in addition to the largely quantitative survey that LSIC eventually became. Indeed, extensive piloting of qualitative interviews and community engagement strategies was conducted from September 2004 to December 2005 in the Torres Strait and Northern Peninsula area, and in the ACT/Queanbeyan region. While stories were collected from some parents, they have not been transcribed, as resources were focused on getting the ground-breaking quantitative aspect of the study working. Among other things, qualitative data allows an in-depth analysis of context that, at worst, may be ignored or, at best, is modelled inadequately in quantitative research that is often motivated by (implicit) theoretical assumptions and rigid categories (enforced through coding assumptions) about what context matters. Both quantitative and qualitative analysis provide invaluable insights, but the intrinsically disparate goals and objectives of each mode of data collection involve distinct costs and hence it was not possible to fully incorporate all the qualitative components due to resource constraints. Those limitations notwithstanding, there are numerous open text questions available in LSIC data that allow qualitative insights to be attained.

Data collection method and response

The above section describes theoretical issues that were duly considered in the committee

phase of survey development; however, revisiting in detail how the data were collected and coded will highlight some of the potential issues that users of the data need to take into account.

Eligible families were approached and voluntary consent obtained. Prior to being interviewed, parents were provided with an introductory letter and a DVD describing the study and the consent process. A plain language statement to introduce the study was provided to parents at the interview. Interviewers went through each consent form with individual participants to ensure that parents could provide informed consent about their participation in the study.

Given that LSIC involved a substantial number of Indigenous children but was clustered in geographically discrete sites (see FaHCSIA, 2012a, map on p. 15), most interviewers were required to find a large proportion of the total number of Indigenous children on-site, usually in locations where the vast majority of the local population is predominantly non-Indigenous. Due to difficulties in sample recruitment related to small resident populations and sparse geographic spread of potential respondents, it was not possible to find sufficient numbers of children to meet the study's targets for some sites. Fortunately the number of eligible children in other sites was in excess of the required sample.

Content rationales were developed based on stakeholder and community consultations, as well as other research such as the Western Australian Aboriginal Child Health Survey (WAACHS) (Zubrick et al., 2004), and the NATSISS. These rationales were workshopped in November 2005 with members of the Steering Committee and other stakeholders, and then used to develop draft questionnaires and computer-assisted personal interview (CAPI) instruments.

Interviewers were instructed to select children for inclusion in the survey based on the agreement of their families to participate. If there was more than one in-scope child within the family, interviewers had been instructed to list more than one child during this initial stage, with the expectation that generally only one would actually be used for the main study. While the survey sites involved implicit quotas on where the surveys were to be conducted, no fixed quotas were imposed, either at the sample recruitment stage or the interviewing stage (Roy Morgan Research, 2009).

As indicated above, LSIC is not representative because of the focus on particular sites and the nature of the sample, specifically the

under-identification of Indigenous people on Medicare or Centrelink records and the need to use social networks to identify potential respondents not identified on administrative records. Self-selection bias may also arise from the fact that families needed to agree to participate in the survey.

For every child, each wave of data collection involved a number of questionnaires, including:

- Parent 1 Survey: household composition, pregnancy, diet and nutrition, health, emotional wellbeing, major life events, demographics, activities, community and neighborhood information;
- Study Child Survey: vocabulary, picture recognition, spatial development (i.e., copying shapes and letters), height and weight; and
- Parent 2/Father Survey: household composition, health, activities and demographics.

Parent 1 of the Study Child was the primary carer identified as the parent “who knew the child best”. Information was also collected from Parent 2, who was either Parent 1’s partner or another adult with a parental or carer relationship to the Study Child. Although the Parent 2 surveys were originally designed to be answered by grandmothers or aunts or other family members who had a caring role, there were few respondents who were not fathers (mostly biological fathers, but step-fathers were also common). Response rates for Parent 2 were quite low in Waves 1 and 2, so Wave 3 data were not collected from Parent 2. Responses from various stakeholders indicated that fathers should not be seen as secondary parents; as such, the interviews were redesigned to focus Parent 2 questions on fathers only from Wave 4.

Although it is the aim of the study to interview participants at 12-month intervals, this has not always been possible. The average intervening time between Waves 1 and 2 interviews was 10 months, which reflects some teething issues in the early stages with the process of organising interviews in this unique survey situation. In Wave 1, interviews with the Study Child went for between 10 to 20 minutes, interviews with Parent 1 were approximately 1 hour (ranging from 30 minutes to 3 hours), and interviews for Parent 2 were between 10 and 60 minutes. Despite this considerable respondent burden in terms of interview time, especially for Parent 1, most of these parents have been more than willing to return for subsequent waves. One suspects that the respondent who was interviewed for 3 hours (the maximum



time recorded) rather enjoyed the process or at least wanted to talk.

The latest release of data (Release 3) contains new information on schooling (including questions asked directly of the child), parental relationships, gambling activities, children’s physical abilities, identification with Indigenous and non-Indigenous groups, experience with and responses to racism, difficulties caused by money problems, homelessness, home maintenance, and community trust.

Each wave of data collection also included carer and teacher questionnaires, which have been released for the first time for Wave 3:

- Centre-based Carer Survey: group size and resources, carer demographics, child behaviour and centre type; and
- Teacher Survey: school type, class size and resources, teacher demographics, child behaviour.

These questionnaires include topics such as the program offered by the school or child care centre, parental involvement, the teacher or carer’s observations about the child, and the relationship between the child and the teacher or carer.

Once collected, the household data was cleaned to remove obvious inconsistencies. For example, where a member was described as “Aboriginal” in two waves and “neither” in the other, this was changed to “Aboriginal”. Where sex varied across waves it was corrected in line with the person’s name (where obvious). Issues such as these are crucial for users of the Confidentialised Unit Record File (CURF) of LSIC, which is the electronic form in which most analysts will access the data.

Brief overview of CURF

LSIC is an omnibus survey of child development that includes a wide range of data that may explain the outcomes measured. The strategy of pursuing diversity was essential so that the

The community engagement strategy is integral to maximising participation and retention in a longitudinal survey of an Indigenous population.

It was deemed to be important that the questions and answers be meaningful to Indigenous people and hence the questions are asked in the idiom of Aboriginal English where appropriate

results of the analysis were not excessively anticipated or the nature of the cultural context presumed. LSIC quantitative data is coded into categories that are broadly analogous with coded responses available in other surveys. However, many questions and coding of responses are LSIC-specific, particularly in the cultural domain and child development.

As is well known, Indigenous culture and language are specific to Aboriginal and Torres Strait Islander populations and hence there is no data against which it can be compared. The “Strong Souls” questions in LSIC were developed by the Menzies School of Health Research (specifically for its Aboriginal Birth Cohort Study). The original questions were intended to identify both the physical symptoms of parents and the frequency with which parents had been feeling depressed, anxious, angry or impulsive (Thomas, Cairney, Gunthorpe, Paradies, & Sayers, 2010). Note that LSIC asks only a sub-set of the questions used in the original study. There is no reason why such questions could not be asked for the general population, but we are not aware of any attempt to do so at this stage.

LSIC is fundamentally a study of child development with many of the indicators used having their roots in other relevant surveys such as LSAC. However, as an Indigenous survey, it was deemed to be important that the questions and answers are meaningful to Indigenous people and hence the questions are asked in the idiom of Aboriginal English where appropriate (RAOs can even use the local idiom to explain questions where appropriate). For example, developmental indicators such as “Who am I?”, Renfrew and Spatial Matrix Reasoning all had precedents in the existing literature, but were adapted to an Indigenous-specific context (e.g., ACER, 2009).² The Strength and Difficulties Questionnaire (SDQ) was asked in a very similar way but the validation process has shown that the resulting scales are not directly comparable with those used elsewhere (Zubrick, 2010).

One issue of the LSIC CURF is that some crucial socio-economic data that is usually used to control for family context is limited. The LSIC questionnaire rightly focuses on the child development outcomes rather than documenting all family circumstances that may explain development at the margin. Income is a core measure of resources available for family members, but income data on the CURF is heavily grouped because there were concerns that the resistance to answering this question in the pilot meant the finer grained income categories were not useful. For example, the small number of income groups provided on the initial LSIC waves cannot credibly be used to estimate “equivalent income” that controls for the resources needed by families of differing size and composition (Hunter, Kennedy, & Biddle, 2004). Furthermore, there was concern that finer income categories would create resistance, translating into higher attrition rates that the LSIC design process was seeking to avoid. Hopefully, finer grained income categories can be collected for future waves without compromising the integrity of the longitudinal nature of the study.

Sample description

Rather than attempt to introduce the gamut of LSIC data, this section reports the number of LSIC respondents on the confidentialised data from Waves 1 and 2 and conducts some basic analysis to illustrate some important issues for users.

The same families who were interviewed in Wave 1 were approached again for an interview in Wave 2. However, a proportion of families could not be interviewed again because they could not be located, had moved substantial distances, refused interviews, or could not be interviewed for other reasons. However, the reduction in the number of study children was partially offset by recruiting 88 additional children into the sample from the 11 existing sites. These children were from families who had either missed out on or refused to participate in Wave 1, but were available and willing to participate in Wave 2.

The numbers on the LSIC CURF may vary between versions/data releases as participants have the right to leave the study at any time and ask that their data be removed. The stability of the number of study children in the study means that this generally does not occur. However, to deal with this issue and to facilitate comparisons between waves, the household data has been reorganised so that each individual has a permanent household member number/position (from Release 1.2 onwards).

Table 1 Basic numbers of respondents to LSIC Waves 1, 2 and 3

	Wave 1	Wave 2	Wave 3
Parent 1	1,676	1,523	1,404
Parent 2	257	269	Not collected
Study child	1,485	1,472	1,394

Note: 1,435 Parent 1 respondents participated in both of the first two waves. The difference between Waves 1 and 2 is made up of 235 departures from the first wave, which was partially offset by 88 new entrants to Wave 2. In Wave 3, 92 respondents from Wave 1 participated even though they were not interviewed in Wave 2. The number of participants can vary from release to release because of irregularities across waves where a participant has requested to be removed from the study. The numbers of participants for Wave 1 listed in this table are those used in the statistical analysis reported in the text. FaHCSIA (2012b, p. 89) reports that 1,670 Wave 1 respondents were still participating at the later waves.

Arguably, the unique identifier makes working with LSIC data simpler than LSAC data, as the latter includes a rather complex system for identifying transitional members to the family, which is relational relative to the primary carer (each wave). The Wave 2 data included in Table 1 is the respondents for the latest release of Wave 2 CURF (Wave 3 *data* was available at the time of writing, but Wave 3 *CURF* was not). FaHCSIA (2012a) reports that the retention rates between waves were high—relative to other longitudinal surveys—at around 86%.

The vast majority of Wave 2 respondents indicated they had not changed address in the last 12 months (77.4%). While 1.6% did not provide a clear response to this question, only just over one-fifth (21%) indicated that they had changed address in the previous 12 months. Of those who indicated that they had moved (and where they had moved to), 64.7% said they had moved only within the local Indigenous area. We do not want to minimise the importance of mobility, as these families all have young families and the change of community context could be significant. Indeed, this is one of the key research questions for LSIC that needs to be explored.

The standard LSIC CURF does not include information on the 11 sites; however, the in-confidence data (that can only be accessed from a secure data facility approved by FaHCSIA) can include data on Indigenous areas for analysis (although researchers must ensure confidentiality of respondents and take care not to release information at that level publicly). Given the potential selectivity in response rates for the various sites, it is important to briefly analyse the geography of LSIC.

The main geographic information in the standard LSIC CURF is the Level of Relative Isolation (LORI), a measure of remoteness and local accessibility that was designed to take into account culturally specific characteristics (e.g., Indigenous language). One potential weakness of LSIC's use of LORI is that it is not widely used in research, with the important exception of the WAACHS survey, which provides the most comprehensive information on Indigenous child health (at least before LSIC was collected). However, LORI does have direct analogy with standard ABS remoteness classification and differences between measures of accessibility are generally not substantial (except perhaps for very remote areas).

Roy Morgan Research (2009) used ABS data from the 2006 Census to provide a rough analysis "sampling fraction" (i.e., the ratio of sample size to population size) that may have been achieved within each of the sites. ABS data

cover the full range of ages from 0 to 5, while the LSIC respondents are generally clustered in two cohorts. Analysis of these data reveal that some of the definitions of postal areas used to create site-equivalent information within the ABS data require significant adjustment. The postal areas used in the analysis, commonly known as postcodes, do not adequately cover the areas used by the RAOs to generate the LSIC sample and hence can at best be considered indicative. Furthermore this assumption will generate large variation in sampling fractions as the reference population will be measured with more error in remote areas where the correlation is low between the sampling area effectively used in LSIC and ABS postal areas.

In remote areas, postcodes are much larger than either the standard local ABS geographic areas or LSIC sampling areas. The notional sampling fraction varies from 3% to well over 100%. Given the difficulty in reconciling the LSIC geography with the Indigenous geography in the Census, it is not really surprising that the variation in sampling fractions was extremely large in remote areas. In contrast, the sampling fraction estimated in metropolitan areas tended to be more bounded within a narrow range as the notional sampling fraction was 10% plus or minus 2%. The variation in sampling fractions may either indicate the inadequacy of the concordance of the Census geography with LSIC study areas or selectivity in the sample.

In order to achieve some sense of the selectivity of the sample, we attempted to estimate regional averages for the Indigenous population in 143 Indigenous areas with LSIC respondents using both survey and 2006 Census data. We estimated regional averages for all the Census and LSIC data that was asked for in a broadly comparable manner in the two data sources, but there are two important differences that need to be noted before attempting to interpret any comparisons: differences in the population sub-group and timing of respective data collections.

Indigenous areas are designed to have enough Indigenous adults to make informative (reasonably accurate) regional estimates from Census data. However, it is not possible to compare LSIC data in Indigenous areas exactly to Census estimates as the ABS web-based facility for generating local statistics (called Table Builder) does not allow researchers to separately identify families and households with children less than 5 years of age. The best comparison available is for the proportion of 0–4 year olds who attend preschool in 2006 Census data that can be estimated in a broadly comparable manner to LSIC. Even then there

There are numerous open text questions available in LSIC data that allow qualitative insights to be attained

The great strength of the study is its sensitivity to cultural issues, especially in the design and implementation.

is a difference in the timing of the Census and LSIC of around 3 years, so the comparison of LSIC and Census estimates could only be broadly indicative of the representativeness of the sample. Another issue is that LSIC age distribution of 0–4 year olds is not designed to reflect that age group in the population, as it focuses on two age cohorts. Notwithstanding such issues, the LSIC sample is slightly less likely to attend preschool in that age group (3.8%); however, the difference is not statistically significant at the 5% level. This provides some direct evidence that LSIC sample is not selective with respect to the target population.

Other LSIC regional data also appears to be broadly consistent with population estimates once one takes into account likely behavioural differences in different population sub-groups in LSIC and Census data.³ The issue of the utility of geographic and community information in LSIC is explored elsewhere (Hunter & McKay, 2011). Irrespective of the findings of that research, one can be reasonably confident that there is a trade-off between the size of the area and the relationship with the community level analysis. If it is difficult to identify reliable local data comparable with the young Indigenous families in the Census, it will be difficult to do so for other potential data sources. One could aggregate the data to a higher level of geography, but that would lose the ability to claim that one is capturing the effect of community, which is often a localised phenomenon.

Selected strengths and weaknesses of LSIC

It is important not to become too obsessed with strengths and weaknesses of LSIC because, in all likelihood, they are probably highly correlated. The great strength of the study is its sensitivity to cultural issues (highlighted above), especially in the design and implementation. The main weaknesses are arguably the non-representativeness of the sample and the difficulty in making direct comparisons with the rest of the population. No population weights are provided in LSIC data and all analysis should be interpreted as being conditioned on the sample attained (i.e., the specific individual, geographic, historic and cultural conditions facing respondents). If these sites differ from other similar areas of Australia through historical circumstances or unmeasured regional characteristics, then the specific nature of these areas needs to be taken into account.⁴ Perhaps the most direct way to achieve this is to link suitable regional data, but if this is not available then geographically

disaggregated controls should be employed. The optimal strategy to do this is likely to involve accessing customised LSIC data in an approved secure data facility.

The second problematic issue is rooted partially in this conditionality, but also relates to the lack of comparability of questions asked in LSIC, LSAC and other surveys. For example, many of the child development outcomes and potential explanatory factors are measured in LSIC using questions specifically adapted for the Indigenous population. Sometimes the adaptations may be relatively minor but, even if one had similar outcomes in the non-Indigenous population, one still has to make the case that comparisons are valid. Whatever the conclusion about the comparability of data, it will be relatively difficult to make a strong case for policy action as the claims about the relative need of Indigenous and other Australian families will necessarily be heavily qualified. While some measure of comparability may have been sacrificed by adopting an Indigenous-specific approach, it should make the LSIC data more meaningful in the context of Indigenous culture.

Another important issue is the relatively small sample size, which will limit the statistical power of the resulting analysis. The overall LSIC sample is less than one-third of that available to LSAC researchers and hence, unless factors are associated with relatively large effects on child development, they will tend to be discounted by analysts as being not significant. At the very least, analysis will be confined to the major factors and researchers should not be too surprised if more subtle interactions anticipated in theory are not statistically significant.

Sample size is not an issue for qualitative data, the inclusion of which is, somewhat ironically, a strength of LSIC. While the proposed qualitative component of LSIC never eventuated, a range of qualitative data items are collected as part of the study in the form of free text responses to a small number of open-ended questions in the survey. References to particular places, individuals, employers, clans, family names, languages and rare circumstances are suppressed to ensure the confidentiality of respondents. The text responses incorporated in the datasets have been truncated to a fixed number of characters (around 30 characters on the CURF). The full responses, or rather confidentialised responses with a minimal truncation of text, can be viewed in the spreadsheets that are potentially available for approved purposes. Data users are permitted to directly quote free text responses on the basis that such usage poses no risk of the

respondent being rendered identifiable. Some of the questions that list text response on the CURF include: What has happened to the study child in past year?; What do you do to cope with stress?; Is there anything else you want to say about the community?; How do you deal with racist bullying?; and What would a good education be for the study child?

Text analytics software can be used to analyse the free-text data. For example, Leximancer uses emergent clustering algorithms to discover and extract concepts from the text to generate a thematic or semantic map (Smith & Humphreys, 2006). This automated content analysis derives concepts from an analysis of frequency, as well as the identification of co-location of phrases and words through clustering like concepts. Linguistics obviously involves a broader range of theories and contexts than can be taken into account in assumptions embedded in such algorithms, but analysis of open text responses using this sort of software can highlight particular issues of concern. Irrespective of the analytical tools used, free text data is an integral aspect of LSIC as a “proof of concept” in that it illustrates the diversity of responses and the assumptions required to code them into standard categories. Many quantitative researchers will take the pre-coded categories as given, but these assumptions need to be questioned in the context of LSIC, especially if the assigned categories do not capture the full diversity of Indigenous responses (which themselves might be quite different from non-Indigenous responses).

Moreover, the free-text data in LSIC may also be useful for generating hypotheses that could be tested. It is important not to be overly prescriptive as the direct responses of Indigenous people themselves can, and arguably should, inform the emphasis of research. Research questions also need to take in other considerations such as existing literature and theoretical perspective of each researcher, but if many Indigenous people nominate certain issues as paramount in child development then it needs to be researched.

Extant and potential uses of the data

Some interesting analysis has been published in international journals comparing Indigenous to non-Indigenous outcomes using LSAC (e.g., Leigh & Gong, 2009). Given the culturally specific nature of much of the LSIC data on child outcomes, it may not be possible to conduct exactly the same analysis using that data. LSIC does not lend itself to comparisons



to other Australian children, but is better placed to assist in identification of which Indigenous children are doing relatively well and the diversity of Indigenous outcomes.

LSIC has identified four key research questions, formulated under the guidance of the Steering Committee, which were designed to achieve this objective (FaHCSIA, 2012a):

- What do Aboriginal and Torres Strait Islander children need to have the best start in life to grow up strong?
- What helps Aboriginal and Torres Strait Islander children to stay on track or get them to become healthier, more positive and strong?
- How are Aboriginal and Torres Strait Islander children raised?
- What is the importance of family, extended family and community in the early years of life and when growing up?

These are “high level” questions that position the study in the minds of the community. Such questions may not actually be answered by any particular researcher, but may be more useful for explaining to respondents and communities the sorts of issues that animate researcher and policy interest. In this way it could be said to legitimise the survey and encourage participation by showing how the information may be used. Notwithstanding, the body of LSIC research could be taken as a whole, to provide some justification for some understanding of what “the best start in life” might be.

Objective measures of wellbeing almost always indicate that Indigenous Australians fare worse than the overall Australian population (Gray, 2012). Yet, on many of the subjective assessments of wellbeing, Indigenous Australians rate their wellbeing quite highly. According to parents’ ratings of the overall health status of their child, only 1 in 25 children had fair or poor health, with the majority in excellent or very good health (Shepherd & Zubrick, 2012). In contrast, more objective

The main weaknesses are arguably the non-representativeness of the sample and the difficulty in making direct comparisons with the rest of the population.

The commitment to community consultation and the training of RAOs were integral to the response rates achieved.

measures have Indigenous women as being twice as likely as non-Indigenous women to have a low birth weight baby, and Indigenous children as being substantially more likely to have been hospitalised relative to other Australian children (SCRGSP, 2011).

There is considerable variation in Indigenous child outcomes and this is likely to be the most productive avenue for future LSIC research—especially in documenting how Indigenous children in remote areas fare relative to those in more accessible areas. Shepherd and Zubrick (2012) use NATSISS data to show that living in remote areas can be associated with differing levels of prompts, facilitators and constraints to child development. The linkages between these prompts and the factors that drive them need to be researched.

Rowley et al. (2008) provide a positive characterisation of outstation and homeland living embodied by their analysis of the NT community Utopia. That study attributes the better health of Utopia residents to the culturally appropriate community-controlled Aboriginal Medical Service, and outstation living that generally includes a better diet and greater physical activity, in addition to an environment where people live more harmoniously with culture, family and land. Even if there is some health premium from living “on country”, one has to take into account other factors associated with child development, including social determinants of health that may not be so positive in remote outstations, such as unemployment, low income, overcrowding, lack of education and level of community harmony (at least in terms of perceived safety). LSIC data provides an opportunity to test for the relative importance of potential trade-offs in child development.

Concluding remarks

In order to operationalise an unprecedented data collection exercise such as LSIC, FaHCSIA has had to be extremely adaptive; in many ways LSIC has been experimental in nature. By making LSIC unique in terms of content and methodology, it may have limited the extent to which the data can be compared directly with other surveys. LSIC arguably provides qualitatively different information to that in LSAC, as it relied more heavily on telephone interviews and self-complete questionnaires, which cost substantially less than the interview methodology used in the Indigenous survey. LSIC’s full-time Indigenous RAOs are associated with a not insignificant cost, but they are invaluable in collecting valid information in Indigenous households, especially where

language, cultural and educational variation would otherwise undermine the quality of the data.

Some commentators, conservative or otherwise, may question whether it is appropriate for a government to be so heavily involved in the conduct of a survey such as LSIC. However, the experimental nature of the exercise has meant that writing contracts for private data collection companies that cover all contingencies that might arise in the operation of the survey would have been almost impossible. The commitment to community consultation and the training of RAOs were integral to the response rates achieved and both can also impose costs that are difficult to anticipate. For example, it was uncertain how many interviewers would leave LSIC between or even within waves, thus training costs are particularly uncertain. FaHCSIA has managed to minimise the turnover of RAOs and training costs by paying full-time government salaries and providing working conditions that are much more generous than the casual rates paid by other survey contractors.

In short, fundamental uncertainty about the way the survey was to be conducted and designed means that it would have been almost impossible to contract out the initial waves of the LSIC survey. Having successfully conducted several waves of LSIC, it may now become easier to write a contract for external operators, but there are some theoretical reasons why it might be difficult to specify terms of the contract that do not undermine the integrity of LSIC or some similar survey. Contractors theoretically will always have an incentive to cut training costs, as many of the skills attained are not transferable to other job situations (McConnell & Brue, 1992). If RAO wages were cut, this would raise the incentive for RAOs to look for another job in data collection and thus make it difficult for LSIC to build suitable relationships in the respective communities.

In addition to the secondary analysis of LSIC data as provided, there are also several less passive options for researchers. One possible mode of collaboration between researchers and the owners of the data (FaHCSIA) is to seek the consent of the LSIC respondent to be reinterviewed on a particular issue. This method is attractive to researchers in that it allows them to readily and cost-effectively identify young Indigenous families that can be difficult to identify in sufficient numbers. From FaHCSIA’s perspective, it is a way to add to LSIC’s public value and ensure that LSIC is actively used by researchers. At the same time, FaHCSIA can provide direct comments on the research

process and enhance the research capability of its staff. Last but not least, such collaboration can build research partnerships and expertise in the field, which is still nascent, to say the least. One such example of a collaborative model is the Paid Parental Leave Research conducted just after LSIC Wave 3 (organised by Maggie Walters).

While lead time in getting research published is relatively long, there are still exceptionally few entries in FaHCSIA's Longitudinal Surveys Electronic Research (FLoSse) repository that relates to Indigenous children. At the time of writing there were only three papers directly on Indigenous children in FLoSse. At least two of these relate to LSAC analysis, which, as this paper argues, is problematic for analysing development of Indigenous children.

There have been a number of promising LSIC-based research papers at conferences, but few have made it to the peer reviewed journals so far. One notable exception is a recent descriptive "data survey" by Mullan and Redmond (2012). At this stage, FaHCSIA staff members are still the major contributors of conference papers using LSIC, but the balance should shift towards scholarly research as more people become aware of the potential of the data. Hopefully, this article makes a contribution to enhancing understanding and confidence in the use of LSIC data for identifying positive pathways of the development of Indigenous children.

Endnotes

- 1 Professor Mick Dodson is a Yawuru man from the Broome area in Western Australia. He was the first Indigenous Australian to receive a law degree following studies at Monash University in Melbourne. He was Australia's first Aboriginal and Torres Strait Islander Social Justice Commissioner with the Human Rights and Equal Opportunity Commission (from April 1993 to January 1998). He is a former member of the United Nations Permanent Forum on Indigenous Issues and was until recently the Co-Chair of Reconciliation Australia. Professor Dodson has a distinguished academic career at the Australian National University, where he is Director of the National Centre for Indigenous Studies. In this role he engages with the research community and develops the next generation of Indigenous researchers.
- 2 An LSIC trial to assess the usefulness of *Who Am I?* was conducted in 2007. The instrument was found to be satisfactory for administration to Aboriginal and Torres Strait Islander children, although some modifications were made to it. In particular, trial sample results suggested that it would be wise to delete some of the items (Numbers, Letters, Words, Sentence).
- 3 The other Census data that was broadly comparable to the information collected in LSIC includes some information on income, housing and mobility. Household income is slightly lower in LSIC data—this is consistent with the young families having lower income, as there is a tendency for adults to

be outside the labour market (compared to other households in the local area). These compositional issues also explain why LSIC households are more likely than other local households to rent dwellings, although there is no significant difference in the proportion in private rental (i.e., sole parents are more likely to be in public housing). LSIC dwellings tend to have more bedrooms, which is consistent with the fact that they are more likely to have a greater number of people than Census data (i.e., families with young children). With respect to mobility, LSIC respondents are more likely than other local Indigenous people in the 2006 Census to have lived in the same address 1 year ago, but less likely to have lived at the same address 5 years ago. This last observation is consistent with a major event in LSIC households in the last few years; say, the birth of a new child.

- 4 The Australian Early Development Index (AEDI) is based on nation-wide data on the development of young children. Between 1 May and 31 July 2009, teachers provided information on five areas of early childhood development: physical health and wellbeing, social competence, emotional maturity, language and cognitive skills (school-based), and communications skills and general knowledge. It should be possible to link LSIC data with AEDI scores measured for all children in the local area. One historical limitation was the extent to which AEDI geography is compatible with LSIC geography on the CURF; however, both AEDI and LSIC data is held in geocoded form by the owners of the data so it may just be a matter of application of additional resources. Indeed, FaHCSIA can provide matched AEDI community level data now, but LSIC users have to make a specific request for it (and presumably have access to a secure data facility). Also, FaHCSIA has sought parental permission to link their child's AEDI data with LSIC data that may be available for a later release (in a separate CURF). Given that AEDI outcomes will partially reflect outcomes in local schools, this addition is likely to considerably enhance the geographic information in LSIC.

References

- Australian Bureau of Statistics. (2011). *The health and welfare of Australia's Aboriginal and Torres Strait Islander peoples, 2010* (Cat. No. 4704.0). Canberra: ABS.
- Arthur, W. S., & Morphy, F. (Eds.). (2005). *Macquarie atlas of Indigenous Australia*. Sydney: Macquarie Library Pty Ltd.
- Australian Council for Educational Research. (2009). *Footprints in Time: Who am I? and Renfrew Word Finding Vocabulary Test: Report on Wave 2 data* (LSIC Technical Report). Melbourne: ACER. Retrieved from <www.fahcsia.gov.au/sites/default/files/files/about-fahcsia/publication-articles/footprints/who_am_I.pdf>.
- Australian Government. (2007). *Budget 2007–08 Indigenous Affairs*. Canberra: Australian Government. Retrieved from <www.fahcsia.gov.au/sites/default/files/files/about-fahcsia/publication-articles/budget/07_indigenous_complete.pdf>.
- Department of Families, Housing, Community Services and Indigenous Affairs. (2012a). *Footprints in Time: The Longitudinal Study of Indigenous Children. An Australian Government initiative: Data user guide. Release 3.0*. Canberra: FaHCSIA. Retrieved from <fahcsia.gov.au/sites/default/files/files/about-fahcsia/publication-articles/footprints/who_am_I.pdf>.

There is considerable variation in Indigenous child outcomes, and future LSIC research should focus on documenting how Indigenous children in remote areas fare relative to those in more accessible areas.

LSIC data provides an opportunity to test for the relative importance of potential trade-offs in child development.

- fahcsia/publication-articles/footprints/data-user-guide-r3.pdf>.
- Department of Families, Housing, Community Services and Indigenous Affairs. (2012b). *Footprints in Time: The Longitudinal Study of Indigenous Children. Key summary report from Wave 3*. Canberra: FaHCSIA. Retrieved from <www.fahcsia.gov.au/our-responsibilities/indigenous-australians/publications-articles/families-children/footprints-in-time-the-longitudinal-study-of-indigenous-children-lsic/key-summary-report-from-wave-3>.
- Gray, M. (2012). A mile wide and an inch deep. In B. Hunter, & N. Biddle (Eds.), *Understanding Indigenous disadvantage: Survey data and social science analysis* (CAEPR Monograph No. 33). Canberra: ANU E-Press.
- Hunter, B. H. (2008). Benchmarking the Indigenous sub-sample of the Longitudinal Survey of Australian Children: Implications for Indigenous policy direction. *Australian Social Policy Journal*, 7, 61–84.
- Hunter, B. H., Kennedy, S., & Biddle, N. (2004). Indigenous and other Australian poverty: Revisiting the importance of equivalence scales. *Economic Record*, 80(251), 411–422.
- Hunter, B. H., & McKay, M. (2011). *Describing LSIC Indigenous communities: Some preliminary exploration of data quality of community variables in LSIC data*. Paper presented at the Australian Institute of Family Studies LSAC/LSIC Conference, Melbourne. Retrieved from <caepr.anu.edu.au/sites/default/files/StaffProfiles/stafffiles/Hunter%20and%20MacKay%20LSIC%20Conference%202011.pdf>.
- Hunter, B. H., & Smith, D. E. (2002). Surveying mobile populations: Lessons from recent longitudinal surveys of Indigenous Australians. *Australian Economic Review*, 35(3), 261–275.
- Leigh, A., & Gong, X. (2009). Estimating cognitive gaps between Indigenous and non-Indigenous Australians. *Education Economics*, 17(2), 239–261.
- Magnani, R., Sabin, K., Saidel, T., & Heckathorn, D. (2005). Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS*, 19(Supp. 2), S67–S72.
- Martin, D. F., & Taylor, J. (1996). Enumerating the Aboriginal population of remote Australia: Methodological and conceptual issues. *Journal of the Australian Population Association*, 13(1), 17–33.
- Mullan, K., & Redmond, G. (2012). A socio-economic profile of families in the first wave of the Longitudinal Study of Indigenous Children. *The Australian Economic Review*, 45(2), 232–245.
- McConnell, C. R., & Brue, S. (1992). *Contemporary labor economics* (3rd edn.). New York: McGraw-Hill.
- Nicholson, J. M., Sanson, A., & the LSAC Research Consortium. (2003). A new longitudinal study of the health and wellbeing of Australian children: How will it help? *Medical Journal of Australia*, 178(6), 282–284.
- Rowley, K., O’Dea, K., Anderson, I., McDermott, R., Sarawati, K., Tilmouth, R., et al. (2008). Lower than expected morbidity and mortality for an Australian Aboriginal population: 10-year follow-up in a decentralised community. *Medical Journal of Australia*, 188(5), 283–287.
- Roy Morgan Research (2009). *Draft sample report*. Melbourne: Roy Morgan Research.
- Shepherd, C., & Zubrick, S. R. (2012). What shapes the development of Indigenous children? In B. Hunter, & N. Biddle (Eds.), *Understanding Indigenous disadvantage: Survey data and social science analysis*. (CAEPR Research Monograph No. 33). Canberra: ANU E-Press.
- Smith, D. E. (1992). The cultural appropriateness of existing survey questions and concepts. In J. C. Altman (Ed.), *A national survey of Indigenous Australians: Options and implications* (CAEPR Research Monograph No. 3). Canberra: Centre for Aboriginal Economic Policy Research, Australian National University.
- Smith, A., & Humphreys, M. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods*, 38(2), 262–279.
- Steering Committee for the Review of Government Service Provision. (2011). *Overcoming Indigenous disadvantage: Key indicators 2011 report*. Melbourne: Productivity Commission.
- Stewart, J., Lohar, S., & Higgins, D. (2011). *Effective practices for service delivery coordination in Indigenous communities* (Closing the Gap Clearinghouse Resource Sheet No. 8). Canberra: Australian Institute of Health and Welfare.
- Taylor, J., & Bell, M. (2004). Continuity and change in Indigenous Australian population mobility. In J. Taylor, & M. Bell (Eds.), *Population mobility and Indigenous peoples in Australasia and North America*. London and New York: Routledge.
- Thomas, A., Cairney, S., Gunthorpe, W., Paradies, Y., & Sayers, S. (2010). Strong souls: Development and validation of a culturally appropriate tool for assessment of social and emotional wellbeing in Indigenous youth. *Australian and New Zealand Journal of Psychiatry*, 44(1), 40–48.
- Wooden, M., & Watson, N. (2007). The HILDA survey and its contribution to economic and social research (so far). *Economic Record*, 83, 208–231
- Zubrick, S. R. (2010). *Longitudinal Study of Indigenous Children: A measure of global social and emotional well being at Wave 1* (LSIC Technical Report). Perth: University of Western Australia and the Telethon Institute for Child Health Research. Retrieved from <www.fahcsia.gov.au/sa/indigenous/progserv/famchild/lsc/Documents/measure_global_social_emotional_wellbeing.pdf>.
- Zubrick, S. R., Lawrence, D. M., Silburn, S. R., Blair, E., Milroy, H., Wilkes, T., et al. (2004). *The Western Australian Aboriginal Child Health Survey: The health of Aboriginal Children & young people*. Perth: Telethon Institute for Child Health Research.

Mick Dodson is the Director of the National Centre for Indigenous Studies at the Australian National University (ANU). **Boyd Hunter** and **Matthew McKay** are, respectively, Senior Fellow and Research Assistant at the Centre for Aboriginal Economic Policy Research, the College of Arts and Social Sciences, also at the ANU.

Acknowledgements: We are indebted to the LSIC team at the Australian Government Department of Families and Housing, Community Services and Indigenous Affairs (FaHCSIA), and two anonymous referees for their comments on an earlier draft.

This paper uses unit record data from the Longitudinal Study of Indigenous Children (LSIC). LSIC was initiated and is funded and managed by FaHCSIA. The findings and views reported in this paper, however, are those of the authors and should not be attributed to FaHCSIA or the Indigenous people and their communities involved in the study.