Growing Up in Australia: **The Longitudinal Study of Australian Children**

Review of main educational program of 4-5 year olds

K cohort

In investigating the quality of the data for the child's educational program type (*cpc06a4)* at Wave 1, concerns were raised in regard to the consistency of responses to this item with other information from the face-to-face interview and the teacher questionnaire. To this extent it was decided to provide a corrected version of *cpc06a4* as well as the original version. The correction involved two processes:

1. If teacher data was present and contradicted the value given by the Parent 1, the value indicated by the teacher data was used instead; or

2. If no teacher data was present, a number of checks were performed on the consistency of the parent's response with other data given (e.g. number of hours in care, the age of the study child etc.). If a majority of cases with teacher data were corrected when they had the same the combination of the original response and number of inconsistencies, then those without teacher data were corrected to the majority value. For example, it was found that among those cases with two or more inconsistencies whose original response was 'Pre-year 1 in a school', more than 50 per cent of the teacher data where available indicated that the true response was 'Pre-school in a school'. This value was therefore assumed to be most likely for these cases in the absence of teacher data.

More information on this process is provided in Attachment 6 to the Data User Guide.

At Wave 3 respondents were asked to confirm the details of the educational program the child was in at the time of the Wave 2 interview two years prior and were then asked about the details of the child's educational programs from ages three to up to six years prior (working backwards until either the child wasn't in an educational program or was in pre-school/kindergarten).

Table 2 shows the information captured for each year. This section suggests improvements to the imputation based on this new data.

To determine how to best use this data, some determination has to be made as to its quality. As an initial check, the recall data was checked for reliability with itself. The data was considered unreliable if there was a greater gap in year level than the number of years between time points, or a lesser gap unless there was an indication that a year level was repeated. This check revealed 14.5 per cent of the cases were unreliable. The data collected at Wave 3 for these cases was not used to impute cpc06a4.

The data was then examined to quantify the number of inconsistencies with other data items from the Wave 1 questionnaire. The following circumstances were considered to be inconsistent:

- The child was in a 'pre-year 1 program' at school and was:
    - attending this program fewer than 5 days/week; or

o attending this program less than 30 hours/week; or

o younger than 55 months of age at Wave 1; or

o in 'Year 1' in Wave 2 unless indicated they had repeated at grade level.

- The child was attending a 'pre-school' (other than in day care) and was:

o attending this program for 30+ hours/week; or

o more than 62 months of age at Wave 1; or

o in 'Year 2' at Wave 2.

Table 1. Variables capturing previous years educational programs for the K cohort at Wave 3

Questions
1) What program did child attend the year before, that is in (3 years prior)?
Year 1 (Grade 1) → 3
Pre-year 1 program → 3
Pre-school/kindergarten program → 3
Long day care → 3
Home-schooled → 3
Other → 2
Child did not attend an educational program → End of recall items
2) Other specify → Previous year
3) Was that located in a school?
Yes →Previous year
No → epc59d?
4) Was it a…?
Preschool/kindergarten only centre → End of recall items
Preschool/kindergarten in a long day care centre → End of recall items
Mobile pre-school → End of recall items
Long day care centre → End of recall items
Other → End of recall items

Three different versions of this information were compared using these checks:

1. 'Original' – The original value entered from the Wave 1 Face-to-Face interview.

2. 'Teacher' – The original data corrected when it disagreed with data obtained from the Wave 1 teacher questionnaire.

3. 'Recall' – The information as recalled by the respondents at Wave 3 for four years prior.

Among those cases that had teacher data at Wave 1 and had reliable recall data at Wave 3[1], 80.6 per cent were found to have no inconsistencies when using the recall data. This compares with 64.7 per cent with no inconsistencies using the original data and 83.5 per cent when using the teacher data. So it would seem that the teacher data is still the most consistent indicator of the true value, however the recall data is also reasonably consistent.

In order to determine how to best use this data in the imputation, two different methods were tried. In the first, the recall data was substituted for the original data automatically. There was agreement between the value created using this scheme and the one using the teacher questionnaire data in 76.1 per cent of cases.

For the second approach to imputation, the recall data (when reliable) were used as an additional check to those listed above and imputations were made on the basis of the number of unlikely combinations of data.

Under the second scheme, the following corrections were made:

1. Children in 'Year 1' were automatically recoded to 'pre-year 1'.

2. Children in 'pre-year 1' with two or more inconsistencies were recoded to 'preschool in a school'.

3. Children attending a 'preschool in a school' with two or more inconsistencies were recoded to 'pre-year 1'.

4. Children attending a 'pre-school at a non-school centre' with two or more inconsistencies were recoded as being in a 'day care centre with a preschool program'.

For cases with Wave 1 teacher data, the data generated by these corrections matched the teacher data in 79.6 per cent of cases, better than using the recall data by itself, and better than the correction scheme used prior to the recall data becoming available (which matched in 72.8 per cent of cases). This approach has therefore been taken.

B cohort

Given the problems experienced for the K cohort at Wave 1, a different set of questions on educational programs was developed for the B cohort at Wave 3 (see Table 2). In Wave 1 for the K cohort, the data collected from the face-to-face interview on educational program differed from that collected in the teacher questionnaire in 28.7 per cent of cases. In Wave 3 for the B cohort there were differences in 12.9 per cent of cases.

However, when the consistency of the teacher data and the parent data was tested against other answers in the Wave 3 interview, it was found that neither version had many inconsistencies, however the teacher corrected version had slightly more (3.0 per cent v 2.7 per cent).

In the seven cases (so far) with inconsistencies when the teacher data was used, the teacher's response was 'Pre-Year 1 school program' while the parent's was 'Pre-school program in a school'. These cases may represent programs that don't fall neatly into either category, (e.g. classes at a Pre-Year 1 level that children attend part-time), although there is no consistency in terms of state of residence of the children, or the organisational basis of

---

[1] i.e. minus the 14.5 per cent mentioned above

the school (e.g. independent vs state vs Catholic). Whatever the situation is with these cases, there seems to be little reason to correct the parent data or teacher data when there is little indication of which is correct.

Outcomes:

1. *Teacher data still to be used to correct parent data when available in determining educational program at Wave 1 for the K cohort.*

2. *Recall data to be used as an extra consistency check within the existing process when imputing this information when teacher data is absent.*

3. *No imputation to be performed on Wave 3 B cohort educational program data.*

Table 2. Variables capturing current educational programs for the B cohort at Wave 3

Questions
1) (Thinking about the arrangement the child uses for the most hours per week) is this located in a school?
Yes → 1
No → 5
2) What class or program does child attend?
Year 1 (Grade 1) → 4
Pre-year 1 program → 4
Pre-school/kindergarten program → 4
Long day care → 4
Other e.g. multi age classes, early intervention → 3
3) Other specify → 4
4) Does child attend this program at…
A government school? → Further items
A Catholic school? → Further items
An independent or private school? → Further items
5) Which of the following best describes where child goes?
Preschool/kindergarten only centre → Further items
Preschool/kindergarten in a long day care centre → Further items
Mobile pre-school → Further items
Long day care centre → Further items
Other → 6
6) Other specify → Further items