## Sample Selection Criteria

Initial selection took occurred on 30th March 2006. At this stage the 12 yr web tests were in mid-flow, so we had to select from the early finishers.

The main selection criterion was that both pairs should have finished the Maths and Reading tests (PIAT, Yes/No and GOAL). They need not necessarily have finished the entire test battery. At the time of selection, this was apparently 417 families.

The next stage was to remove exclusions. These were the standard exclusions used in TEDS analysis: medical exclusions, perinatal outliers, unknown sex/zygosity, families without 1st contact data. This brought the available sample down to 374 families.

Next families were divided into three SES groups, based on a variable "zases", which was our usual "ases" variable after $z$-standardisation on the entire 1 st contact sample. The cut-off points for the three groups were -1 and +1 , i.e. one standard deviation either side of the mean. The sizes of the three groups were 62 lowSES, 237 middle-SES and 75 high-SES.

Andy sent the sample list to Jane and Tricia for the final selection; and to Nicole a more detailed dataset including teacher NC data from the 10yr study. Jane and Tricia, then selected 3 lows, 3 highs and 9 middleSES families"

At this final stage, Tricia and Jane looked at selection in practical terms according to where the families lived, case notes, spin-offs etc. Tricia then sent Nicole our selected IDs for final confirmation before we contacted them. Did Nicole also check for their 10yr teacher scores?

## Time between web and in-person

The dataset contains a variable showing the time interval (in days) between completion of the Maths and Reading 12 yr web tests, and completion of the paper retests. The dates of completion of the web tests were taken from the main 12 yr dataset. The dates of the paper retests were looked up individually by Jane from our records. The output file shows the distribution of the time interval. The mean was 65 days (roughly 2 months), and the standard distribution was 22 days. The distribution is approximately uniform, varying between 20 and 94 days.

## Activities re-tested

## Reading tests

PIAT-R Reading Comprehension - administered from the start until discontinue rule met as per administration instructions

GOAL Formative Assessment in Literacy - All questions from Key stage 3 test booklet
Woodcock-Johnson III Tests of Achievement, Reading Fluency (Woodcock et al., 2001). - as per administration instructions
Author recognition list devised by Nicole - as web version

## Maths Paper and Pencil Battery

1. Administer the whole of the Booklet 12 (Non-calculator Section only). This adds up to 32 items. In the actual Booklet there are extra 18 items that require calculator. We do not use these items in the Web-battery, so we are not using them here. The children are required to do 50 items in 45 minutes. So, we expect them to do the 32 items in less than 30 minutes.
2. We will correlate performance on the whole of Booklet 12 (total score) with performance on the Web battery (total score). We can also only look at the items that are included in our Web-battery.
3. The only changes to the actual booklet are:

- get rid of the symbols indicating which maths category the question belongs to. This is because in our Web-battery the children have no such indication.
- Change question numbers

4. Administer an extra booklet, which consists of all the questions included in our Web-battery from Booklets $10,11,13$, and 14 . This includes 38 items:

- Category 1: 14 items: 9 easy ( 5 from Booklet $10 \& 4$ from Booklet 11) and 5 hard ( 1 from Booklet 13 \& 4 from Booklet 14).
- Category 2: 10 items: 6 easy ( 2 from Booklet $10 \& 4$ from Booklet 11) and 4 hard (3 from Booklet 13 \& 1 from Booklet 14).
- Category 3: 14 items: 10 easy ( 5 from Booklet $10 \& 5$ from Booklet 11) and 4 hard ( 3 from Booklet 13 \& 1 from Booklet 14).
- These 38 items should also take approximately 30 minutes, as the majority of the items are easier than expected grade level, and the total number of hard items is only 13.

5. We will also be able to look at correlation by category, when items are combined from both Booklets.

Total number of Items: 70. Estimated average time of completion: 1 hour.
Note that from Booklet 12 the category names and sometimes the meaning of categories change. So, the categories are assigned to the items (and sometimes items could belong to more than one category (see '?'). This does not matter, because we are most interested in the total score. Moreover, only items, where categories are unambiguous are included in the Web battery.

