

## Data supplement

| <b>Table DS1</b> Interobserver reliabilities for child and teacher observations over 5 min intervals |  |  |   |
|--|--|--|---|
|  | Interobserver reliabilities (ICC)                              |  |   |
|  | Baseline training reliabilities <sup>a</sup><br>Median (range) | Ongoing reliabilities <sup>b</sup><br>Median (range) | Post-intervention training reliabilities, <sup>a</sup> median (range) |
| Structured observations of child behaviour   |  |  |   |
| Aggressive/destructive behaviour   | 0.87 (0.79–0.94)   | 0.91 (0.78–0.95)                                     | 0.86 (0.80–0.95)  |
| Friendship skills  | 0.83 (0.79–0.94)   | 0.86 (0.70–1.00)                                     | 0.93 (0.80–0.98)  |
| Disruptive behaviour   | 0.89 (0.87–0.91)   | 0.80 (0.79–0.97)                                     | 0.96 (0.90–0.98)  |
| Rating scales of child behaviour   |  |  |   |
| Conduct problems   | 0.95 (0.90–0.97)   | 0.90 (0.84–0.99)                                     | 0.87 (0.79–0.93)  |
| Activity   | 0.86 (0.82–0.92)   | 0.67 (0.50–0.91)                                     | 0.74 (0.70–0.90)  |
| Task   | 0.89 (0.80–0.97)   | 0.87 (0.66–0.99)                                     | 0.93 (0.82–0.99)  |
| Rules  | 0.88 (0.82–0.92)   | 0.77 (0.60–0.92)                                     | 0.89 (0.80–0.92)  |

ICC, intraclass correlation coefficient  
a.  $n=20$  observations between trainer and each observer and minimum of 15 between each observer and other observers.  
b. 15% of all observations between senior research assistant and each observer and minimum of 15 between each observer and other observers during baseline assessments and minimum of 10 during post-test assessments.

| <b>Table DS2</b> Internal reliability and test–retest of parent and teacher report measures |                  |                   |                                |
|---|------------------|-------------------|--------------------------------|
|   | Cronbach's alpha |                   |                                |
|   | Baseline         | Post-intervention | Test–retest <sup>a</sup> (ICC) |
| Teacher report measures   |                  |                   |                                |
| Sutter–Eyberg Student Behaviour Inventory   | 0.96             | 0.97              | 0.92                           |
| Connor's Global Index   | 0.86             | 0.90              | 0.84                           |
| Preschool and Kindergarten Behaviour Scales (Social Skills Scale)                           | 0.92             | 0.93              | 0.85                           |
| Total behaviour difficulties (Strengths and Difficulties Questionnaire)                     | 0.78             | 0.80              | 0.80                           |
| Prosocial skills (Strengths and Difficulties Questionnaire)                                 | 0.80             | 0.81              | 0.78                           |
| Parent report measures  |                  |                   |                                |
| Eyberg Child Behaviour Inventory  | 0.80             | 0.83              | 0.90                           |
| Prosocial skills (Strengths and Difficulties Questionnaire)                                 | 0.65             | 0.74              | 0.70                           |
| Total behaviour difficulties (Strengths and Difficulties Questionnaire)                     | 0.67             | 0.67              | 0.85                           |
| Parent attitude to school   | –                | 0.74              | 0.78                           |

ICC, intraclass correlation coefficient  
a. Over 2 weeks:  $n=20$  for teacher report measures,  $n=18$  for parent report measures.