The reliability and validity of the pupil behaviour questionnaire: a child classroom behaviour assessment tool

Matt Allwood, Kate Allen, Anna Price, Rachel Hayes, Vanessa Edwards, Susan Ball, Obioha C Ukoumunne & Tamsin Ford

To cite this article: Matt Allwood, Kate Allen, Anna Price, Rachel Hayes, Vanessa Edwards, Susan Ball, Obioha C Ukoumunne & Tamsin Ford (2018): The reliability and validity of the pupil behaviour questionnaire: a child classroom behaviour assessment tool, Emotional and Behavioural Difficulties, DOI: 10.1080/13632752.2018.1478945

To link to this article: https://doi.org/10.1080/13632752.2018.1478945

Published online: 23 May 2018.
The reliability and validity of the pupil behaviour questionnaire: a child classroom behaviour assessment tool

Matt Allwood, Kate Allen, Anna Price, Rachel Hayes, Vanessa Edwards, Susan Ball, Obioha C Ukoumunne and Tamsin Ford

ABSTRACT
This paper assesses the reliability and validity of the teacher-completed Pupil Behaviour Questionnaire (PBQ), by comparing it to the already extensively validated teacher-completed Strengths and Difficulties Questionnaire (SDQ). Participants included 2074 primary school children participating in a universal school-based trial and 41 vulnerable children who were taking part in a study exploring the impact of exclusion from school. Exploratory factor analysis results (first factor accounts for 80.8% of the variation in the items) and the high Cronbach’s alpha value of 0.85 indicate that the PBQ consists of one substantive factor/dimension. Strong correlations between the total PBQ score and the conduct subscale (Spearman’s correlation coefficient ($r_s$) = 0.67) and total difficulties score ($r_s = 0.59$) of the SDQ indicate convergent validity. This study suggests that the PBQ is a reliable measure, and provides some evidence of validity. Further work is needed to test the PBQ in an older, more diverse populations and to measure sensitivity to change.

KEYWORDS
Behaviour; school children; questionnaire; validity; reliability

Introduction

Disruptive childhood behaviour predicts poor educational, occupational, health, relationship and social outcomes, with associated increased costs to society (Scott 2007). Complaints about disruptive pupil behaviour in the classroom are common and are reported to be a major contributor to stress and burnout among teachers (Brouwers and Tomic 2000; Greene et al. 2002). In a recent survey, 42% of teachers reported that stress impacted on their ability to teach and 76% said that their stress levels had repercussions on their health (Teachers Assurance 2013). Only 77% of teachers described their pupils’ behaviour as good or very good in a recent survey completed by teachers on pupil behaviour Weaving, H. and H. Aston 2013, while the chief Inspector of Ofsted, Sir Michael Wilshaw, has expressed concern that ‘teachers were too often ignoring low-level disruption leading to a culture of casual acceptance of misbehaviour that is holding schools back’ Sky News (2013). Improvements in behaviour are now a major policy focus (Ofsted 2015).

There are many varying definitions of problematic behaviour in schools. Most of the discussion revolves around the perception of what is described as disruptive, a concept that includes both the frequency and severity of the behaviour Education Standards Analysis and Research Division, Department for Education 2012. Cameron (1998) has proposed grouping disruptive behaviour into five categories; behaviour that is aggressive, that challenges authority, or is disruptive either physically, socially or to the child themselves. This is in contrast to Watkins and Wagner (2000) who
describe low level disruption and ‘talking out of turn’ as the most frequent constellation of problematic behaviours that teachers report to be the trickiest to deal with. Valid and reliable measures of such pupil behaviour could be one way of clarifying the impact of different disruptive behaviours in schools.

It is important to quantify children’s behaviour for various purposes, which include: tracking a child’s progress over time, comparing whole class behaviour across a school, assessing behaviour at whole school level regionally or nationally, or to evaluate a behaviour management tool or intervention. Questionnaires provide a systematic approach that allows comparison across individuals, groups and times providing that the measure is reliable, valid and sensitive to the types of changes to be expected from common interventions. One of the first systematic attempts to measure children’s classroom behaviour was made by Wickman in the early 1920s (Haggerty 1925; Wickman 1928) but with the questionnaire including 26 items it was deemed too time consuming for teachers to complete about every child in their class (Rutter 1967).

In 1988 the Department of Education and Science commissioned an enquiry into discipline in schools in response to concerns about the difficulties facing teachers in England and Wales – the Elton Report (Department of Education and Science 1989). This enquiry consulted with 3,500 teachers who reported on their experiences in and around their school during the previous week. In 2009 the Scottish government’s study on the behaviour in schools in Scotland developed the survey used in the Elton Report, which included a teacher-reported measure of behaviour that was 26 items long and covered low level indiscipline, disengagement and serious indiscipline/violence (The University of Edinburgh 2009). Arguably the best teacher report measure of classroom behaviour is the revised Sutter-Eyberg questionnaire, which comprises 38 items that are rated on both Intensity and Problem scales and is validated for children aged 2–16 years (Eyberg and Pincus 1999). While the Sutter-Eyberg offers detail and depth of assessment for children with complex difficulties, the PBQ offers brevity for children with less severe difficulties or when teachers are required to complete measures on groups of children in their classroom.

In order that a questionnaire is feasible for completion by school teachers on every child in their class it needs to be short and easy to complete (Slade, Thornicroft, and Glover 1999). The aim of the current study was to assess the reliability and validity of a shorter version of the 26 items developed from the Elton Report, focusing only on the low level classroom disruption. The resulting Pupil Behaviour Questionnaire (PBQ) was developed for use in the Supporting Teachers And children in Schools (STARS) (Ford et al. 2012) trial in consultation with teachers and educational psychologists. We assessed the validity of the PBQ by comparing it to the already extensively validated teacher-completed Strengths and Difficulties Questionnaire (SDQ) (Goodman 1997) within a sample of children from mainstream schools. A second sample of children at risk of permanent exclusion was used to assess whether the PBQ could discriminate between a vulnerable sample and a community sample of children.

**Methods**

**Study samples/participants**

This study was a secondary analysis of data from two studies: the STARS trial (Ford et al. 2012), which evaluated the effectiveness of a teacher classroom management course and the SKIP (Supporting Kids, avoiding Problems) study that explored mental health, attainment and ability among children at risk of permanent exclusion from school (Parker et al. 2016). Both studies had ethical permission from the University of Exeter Medical School Research Ethics Committee. The STARS sample comprised the baseline data of a large cluster randomised controlled trial of teacher classroom management, which included 80 primary schools across Devon, Plymouth and Torbay (Figure 1). Included children were in Reception to Year 4 (aged 4 to 9 years). Data were collected
from three overlapping cohorts of schools at the beginning of the academic year in September 2012, 2013 and 2014.

The SKIP study (Parker et al. 2016) was a case-control study that assessed the mental health of children at risk of or who had experienced exclusion from primary school and the first year of secondary school. Children were identified by an educational or mental health practitioner and recruited from the Southwest of England between September 2011 and July 2013. Here we only present data from the cases, or those at risk of exclusion.

**Measures**

**Pupil behaviour questionnaire (PBQ)**

To be feasible for completion by primary school teachers on every child in their class, the PBQ was limited to six items from the questionnaire used in the Elton Report (Department of Education and

---

*Figure 1. Recruitment and outcomes shown per sample.*

1 Missing data represents children who had left the school (and thus, teachers were unable to provide data on).

2 Missing data represents teachers who did not respond to e-mails/letters about participation. In three cases, parents did not consent to involving the teacher in the study.
Science 1989) and, in consultation with teachers and educational psychologists, focused on the types of low level disruption reported to be common and most difficult to manage in primary schools (Watkins and Wagner 2000). The 6 items were; 1) talking out of turn, 2) interrupting other pupils, 3) making unnecessary noise, 4) verbal abuse towards other pupils, 5) physical aggression towards other pupils, and 6) cheeky or rude remarks to the teacher. Each of the six items addresses a negative behaviour, with responses on a 3 point scale, where 0, 1 and 2 correspond to the child never, occasionally and frequently exhibiting that behaviour, respectively. The total score for the PBQ is calculated as the sum of the six items and therefore has a possible range between 0 and 12, with higher scores indicating more disruptive behaviour. The PBQ is freely available to download (http://medicine.exeter.ac.uk/research/healthresearch/childhealth/child-mental-health/pbq/) provided the person downloading it is not charging for its use.

**Strengths and difficulties questionnaire (SDQ)**
The SDQ (Goodman 1997) is a well-validated measure of childhood mental health and behaviour. Unlike the PBQ, which is designed for the classroom setting, the SDQ is ordinarily conducted within a clinical setting. Baring this in mind, we would expect to see a medium association between individual children’s scores on the SDQ and PBQ. The SDQ is a 25-item questionnaire with responses to each item on a three point Likert scale; not true (0), somewhat true (1) and certainly true (2) for difficulties or reversed for strengths so that a high score indicates difficulty except for the pro-social scale where the opposite applies (Goodman 1997). Of the 25 items, 10 are thought of as strengths and 15 are thought of as difficulties. The questionnaire is split into five sub-scales: emotional, conduct, hyperactivity, peer problem and pro-social. The total score for each of these sub-scales ranges from 0 to 10, and is calculated by adding the scores of the items for that sub-scale. The sub-scale scores, excluding the pro-social sub-scale, are summed to form the total difficulties score which ranges from 0 to 40. The SDQ was completed by both teachers and parents in both studies.

The reliability and validity of the SDQ have been extensively tested. Goodman (2001) has shown that the internal consistency is satisfactory (Cronbach’s alpha (α) = 0.73) and after 4 to 6 months the retest stability of teacher ratings were satisfactory (mean 0.62) in a national sample of school aged children (Goodman 2001).

**Statistical analyses**
Participant demographic characteristics, PBQ and SDQ total and sub-scale scores were summarised using means and standard deviations for continuous variables and numbers and percentages for categorical variables. Analyses were carried out using Stata 13.1, R (3.2.0) and Mplus Version 6.

**Factor analysis**
An exploratory factor analysis was carried out on the 6 PBQ items to identify themes (factors) that underlie the measure (Bryman and Cramer 2005) using Mplus (Muthen and Muthen 2007). As the response set was comprised of only three levels the items were treated as ordinal categorical in the analysis, as opposed to continuous. Mean- and variance-adjusted (‘robust’) weighted least-square parameter estimates of the factor loadings were obtained based on analysis of polychoric correlations between the PBQ items. The analysis allowed for clustering within schools and we report within-cluster factor loadings.

**Internal consistency**
The internal consistency of the PBQ was assessed using Cronbach’s alpha, with values of 0.70 to 0.90 considered to demonstrate good reliability, whilst bearing in mind that a Cronbach’s alpha of greater than 0.90 could suggest that some items were redundant (Streiner 2003; Tavakol and Dennick 2011).
**Discriminant validity**
Using a random effects linear regression model, the mean PBQ total score in the community sample (STARS) was compared to the vulnerable children from the SKIP sample. This method takes account of the correlation between the responses related to children from the same class. We predicted that the children with difficulties in coping with school would have higher levels of disruptive behaviour due to the threatened or actual breakdown of their school placement.

**Construct validity**
Construct validity was assessed by examining convergent and divergent validity. A measure will be more closely related to another measure of a similar construct than a measure that assesses a different construct. We reported the Spearman’s correlation coefficient for the association between the PBQ with the different subscales of the SDQ as completed by teachers. We expected to see a medium association between these two constructs.

*Convergent validity* was quantified by the correlation between the PBQ total score and the behaviour subscale, hyperactivity subscale and the total difficulties score from the SDQ. These SDQ subscales measure disruptive behaviours that would be expected to positively correlate with the total score of the PBQ. *Divergent validity* was measured by comparing the total score from the PBQ with the pro-social subscale from the SDQ. This SDQ subscale measures positive behaviours which would be expected to negatively correlate with the PBQ focussing on disruptive behaviour.

**Gender and school effects**
Random effects linear regression models were fitted to examine the relationships between total PBQ score (outcome) and each of gender and school year (potential predictors). It would be expected that boys would score higher as they tend to exhibit more hyperactive and antisocial behaviour than girls (Sylva et al. 2012). We expected that low level disruptive behaviour would be more common in younger children, which would be reflected in a higher total PBQ (Ford et al. 2007).

**Attainment**
Using the STARS sample, we estimated the correlation between the total PBQ score and each of the literacy and numeracy attainment scores. Given the strong association between learning difficulties and disruptive behaviour, we expected a negative correlation between the PBQ and attainment (Ford et al. 2007).

**Results**

**Study participants**
The STARS study comprised 2,074 children in 80 different schools, taught by 80 different teachers who completed the PBQ and SDQ questionnaires on 2,074 of these children. The SKIP study provided a sample of 41 children aged 5–12 years from 22 schools in Devon. Of these 41 children, 30 (73%) had a PBQ completed by their current class teacher and 32 had a SDQ completed by their class teacher, as shown in Table 1 and Figure 1.

The distribution of children in the STARS sample was balanced in terms of gender and age. In contrast, the SKIP children, who were at risk of exclusion, were, as we had predicted, mainly boys (95%) and tended to be older. The STARS sample had more complete data from teachers who were the primary recruiting source, in the SKIP study contact with teachers was only made if the parent and child agreed. Item-level data for the PBQ shows scores were generally low for the STARS study and high for the SKIP study across all items, as shown in Figure 2.

In the factor analysis, the eigenvalue of the first factor was 4.845 (accounting for 80.8% of the total variation in the PBQ item scores) and those of the remaining five factors were all less than 0.5.
Table 1. Participant characteristics and response rates for the PBQ and SDQ by sample.

<table>
<thead>
<tr>
<th>DEMOGRAPHICS</th>
<th>Community sample</th>
<th>Exclusion sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male, n (%)</td>
<td>1,101 (53)</td>
<td>39 (95)</td>
</tr>
<tr>
<td>Age in years, mean (SD)</td>
<td>6.3 (1.3)</td>
<td>8.6 (2.1)</td>
</tr>
<tr>
<td>Year group, n (%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reception</td>
<td>270 (13)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Year 1</td>
<td>368 (18)</td>
<td>6 (15)</td>
</tr>
<tr>
<td>Year 2</td>
<td>410 (20)</td>
<td>4 (10)</td>
</tr>
<tr>
<td>Year 3</td>
<td>608 (29)</td>
<td>5 (13)</td>
</tr>
<tr>
<td>Year 4</td>
<td>418 (20)</td>
<td>5 (13)</td>
</tr>
<tr>
<td>Year 5</td>
<td>-</td>
<td>10 (25)</td>
</tr>
<tr>
<td>Year 6</td>
<td>-</td>
<td>5 (13)</td>
</tr>
<tr>
<td>Year 7</td>
<td>-</td>
<td>4 (10)</td>
</tr>
<tr>
<td>TEACHER COMPLETED PBQ</td>
<td>N = 2,074</td>
<td>N = 30</td>
</tr>
<tr>
<td>PBQ, mean (SD)</td>
<td>1.9 (2.4)</td>
<td>8.97 (2.4)</td>
</tr>
<tr>
<td>TEACHER COMPLETED SDQ</td>
<td>N = 2,074</td>
<td>N = 32</td>
</tr>
<tr>
<td>SDQ total difficulties score, mean (SD)</td>
<td>6.7 (5.9)</td>
<td>22.3 (5.1)</td>
</tr>
<tr>
<td>SDQ conduct sub scale score, mean (SD)</td>
<td>0.8 (1.5)</td>
<td>5.4 (1.8)</td>
</tr>
<tr>
<td>SDQ prosocial score, mean (SD)</td>
<td>7.5 (2.4)</td>
<td>3.9 (1.9)</td>
</tr>
</tbody>
</table>

<sup>a</sup>data only available for 40 children in the SKIP sample

Figure 2. Responses to the PBQ in both community (STARS) and high risk (SKIP) samples. This graph shows that behaviours measured by the PBQ were more frequently seen in the high risk sample than when compared to the community sample. This gives an indication that the PBQ can differentiate between a high risk population and the general population.

Notes: * 96.7% data completeness due to spoiled data
Q1.Talking out of turn (e.g. by making remarks, calling out, chattering
Q2.Interrupting other pupils (e.g. by distracting them from work)
Q3.Making unnecessary (non-verbal) noise (e.g. by scraping chairs, banging objects)
Q4.Verbal abuse towards other pupils (e.g. offensive or insulting remarks)
Q5.Physical aggression towards other pupils (e.g. offensive or insulting remarks)
Q6.Cheeky or rude remarks to the teacher

M. ALLWOOD ET AL.
This indicates that there is only one salient factor or substantive dimension that underlies the PBQ item scores and the items measure a single construct. The loadings for the first factor from the one factor solution indicate strong correlations between each of the PBQ items and the factor: 0.908 for talking out of turn, 0.952 for interrupting other pupils, 0.883 for making unnecessary noises, 0.899 for verbal abuse towards peers, 0.874 for physical aggression towards peers, and 0.828 for rudeness towards teachers. The PBQ demonstrated excellent internal consistency both overall (Cronbach’s alpha = 0.85) and for most subgroups defined by gender and age (Table 2). As expected, the SKIP sample demonstrated much higher levels of disruptive behaviour than the STARS sample; on average, total PBQ was 7 points higher in the SKIP sample (95% CI: 6.1 to 7.9, p < 0.0001). Moderate to strong positive correlations were found between the PBQ and each of the teacher-rated SDQ conduct, hyperactivity and total difficulties scores (p < 0.0001 for all), which demonstrates convergent validity, as shown in Table 3. Weak correlations were found between the PBQ and the teacher-rated SDQ emotional symptoms subscale and peer problems subscale and a moderate negative correlation was found with the pro-social subscale (p < 0.0001 for all), demonstrating divergent validity, as shown in Table 3.

**Relationships of PBQ score with gender and school year**

On average, total PBQ was estimated to be 1.3 points lower among females than males (95% CI for reduction: 1.1 to 1.5, p < 0.0001). There was little evidence of a relationship between total PBQ and school year (p = 0.9).

**Attainment and the PBQ**

There were small negative correlations found between the PBQ and both literacy (−0.20; 95% CI: −0.25 to −0.16; p < 0.0001) and numeracy (−0.13; 95% CI: −0.18 to −0.09; p < 0.0001) attainment scores.

**Discussion**

This is the first study to test the psychometric properties of this novel measure of low level disruption for primary school age pupils. Our initial findings suggest that it is both reliable and valid. The PBQ measures a single construct with excellent internal consistency and is correlated

<table>
<thead>
<tr>
<th>School year</th>
<th>Male N</th>
<th>Male α</th>
<th>Female N</th>
<th>Female α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reception</td>
<td>152</td>
<td>0.81</td>
<td>118</td>
<td>0.77</td>
</tr>
<tr>
<td>Year 1</td>
<td>184</td>
<td>0.85</td>
<td>184</td>
<td>0.81</td>
</tr>
<tr>
<td>Year 2</td>
<td>219</td>
<td>0.87</td>
<td>191</td>
<td>0.65</td>
</tr>
<tr>
<td>Year 3</td>
<td>319</td>
<td>0.86</td>
<td>289</td>
<td>0.79</td>
</tr>
<tr>
<td>Year 4</td>
<td>227</td>
<td>0.87</td>
<td>191</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables compared</th>
<th>Total PBQ vs teacher rated SDQ total</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.59</td>
<td>0.56 to 0.62</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total PBQ vs teacher rated conduct subscale</td>
<td>0.67</td>
<td>0.65 to 0.70</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total PBQ vs teacher rated pro-social subscale</td>
<td>−0.53</td>
<td>−0.56 to −0.50</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total PBQ vs teacher rated hyperactivity subscale</td>
<td>0.72</td>
<td>0.70 to 0.75</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total PBQ vs teacher rated peer problems subscale</td>
<td>0.19</td>
<td>0.14 to 0.23</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total PBQ vs teacher rated emotional symptoms subscale</td>
<td>0.01</td>
<td>−0.03 to 0.05</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
highly with teacher-completed SDQ scores in the expected directions. It is widely accepted that a scale with internal consistency estimates between 0.7 and 0.9 indicates that a test is a reliable measure (McCrae et al. 2011). It is encouraging that the children at risk of permanent exclusion from the SKIP sample scored significantly higher than those from the community sample, but also that they were not consistently scoring at the highest level on the PBQ. This suggests that there was not a ceiling effect with this scale, even among children whose behaviour was so challenging they were at risk of exclusion from mainstream school.

As expected, on average, girls had a lower average total PBQ score than boys, but no relationship was found between age and total score among our primary school sample. Differences might emerge if tested on a secondary school sample. The association found between disruptive behaviour reported on the PBQ and lower attainment scores replicates previous work (Ford et al. 2007).

Teachers completed both the PBQ and the teacher SDQ, both these measures aim to assess child behaviour however, differ in the sense that the former is designed for the classroom setting whereas the latter is intended for clinical settings as a diagnostic tool. Therefore, although we expected to see a medium correlation between the two, we did not expect this correlation to be strong. This is exactly what we found, with correlations ranging from strong to medium in the predicted directions. This is reassuring as it suggests that the PBQ has good construct validity.

Given our findings have demonstrated robust psychometric properties in this short-item measure, the PBQ has the potential to become a widely used method for quantifying the behaviour of primary school children in the classroom. The PBQ focuses on behaviours that commonly disrupt teaching.

Our findings are particularly pertinent currently given the Government’s recommendations for a greater emphasis on early interventions for children (Department of Health 2015) and the drive towards evidence-based interventions in schools and health care (Education Endowment Foundation 2015). A lack of effective support in a child’s early years has been identified as a potential cause of many of the costly and damaging social problems in society (Scott et al. 2001).

Many academics accept the use of Cronbach’s alpha to assess internal consistency as an acceptable measure of reliability but internal consistency does not necessarily equate to a measure that reliably reproduces the same results (McCrae et al. 2011). Future research is needed to assess the test-retest reliability of the PBQ as well as longitudinal studies assessing the stability of scores over time; epidemiological studies show that without intervention, disruptive behaviour is a highly persistent trait (Costello and Maughan 2015). In addition, to be useful as an outcome measure in practice, we need to test the PBQ’s sensitivity to change when evaluating evidence-based interventions. The appropriateness of the measure to children aged 10 and over and with severe behavioural difficulties also remains to be assessed.

This analysis benefitted from a comparison measure that was validated and frequently used in both practice and research and from the fact that the STARS sample offered access to a large unselected primary school sample. It is not, however, without limitations. As with all secondary data analyses, we were restricted by the constraints of the data that we had access to. We lacked data on children over the age of 9. The PBQ could be useful and feasible for older children, particularly those coming up to transition to secondary education or in the earlier years of secondary school, but this requires empirical testing. The longer Scottish questionnaire from which the PBQ was developed was aimed at both primary and secondary school aged children, but has never been validated (The University of Edinburgh 2009). Our findings suggest that it might be worth further testing of the PBQ with samples of secondary school aged children.

The STARS trial was based solely in the South West Peninsula of England (Ford et al. 2012), which has a lower than average ethnic diversity, compared to the rest of England and Wales (2.5% vs 14%) (Strategic Intelligence 2012). The findings, therefore, may not be generalizable outside the South West of England. This limitation can be addressed by comparing results from children from a wider ethnic background.
Further refinement of the measure could involve a qualitative study exploring the reasons why teachers made the responses they did and how these may be different for each teacher completing the measure.

In summary, our initial findings suggest that the PBQ is a promising, psychometrically valid measure of low level behavioural disturbance within primary school classrooms and could be of use in relation to the assessment of individual, class and school level difficulties. Future work needs to focus on testing the PBQ in different populations of children and on assessing how sensitive it is over time.

**Contributors**

The manuscript was prepared by MA with assistance from all co-authors. Data analysis were carried out by MA, SB, and OU. TF oversaw all aspects of the research project.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This project was funded by the National Institute for Health Research Public Health Research Programme (project number 10/3006/07) and the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care South West Peninsula. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR Public Health Research Programme, NIHR, NHS or the Department of Health and Social Care.

**Notes on contributors**

Matt Allwood is a Research Assistant in Child and Adolescent Mental Health. He is interested in child and adolescent mental health, wellbeing and psychometrics.

Kate Allen is a Postdoctoral Research Associate working across Child Health and Sport and Health Sciences and has over seven years experience conducting research with young children. She has an interest in child development, mental health and early intervention.

Anna Price is a Research Associate in Child Health who is also studying for a PhD. Her research focuses on improvements that can enhance young people’s experience of education and health services.

Rachel Hayes is a research fellow with over 8 years’ experience managing large trials of complex interventions to support the mental health of both children and adults. She has a background in developmental psychology and is interested in ways to support schools promote the mental health of children and young people.

Vanessa Edwards is a Trial Manager with 13 years experience of managing randomised controlled trials, mainly in the area of child mental health. She has extensive experience of user involvement in designing trials of complex interventions with the ultimate aim of making trials as acceptable and feasible as possible to participants.

Susan Ball is a medical statistician with more than ten years experience working on the design, analysis and reporting of research studies. Her work covers a range of areas, including maternal, fetal and child health.

Obioha Ukoumunne is a medical statistician who has played a key role in the design, conduct and analysis of many public health and health services research studies. His publications reflect contributions to the following major areas: statistical methodology for the design and analysis of data from cluster randomised trials; mental health research; and child health and wellbeing research.

Tamsin Ford is Professor of Child and Adolescent Psychiatry at the University of Exeter Medical School. She is a health services researcher and an epidemiologist, and her academic work focuses on the effectiveness of interventions and the efficiency of services in relation to the mental health of children and young people.
References


