

Article Title:

'How I feel About My School': The construction and validation of a measure of wellbeing at school for primary school children

Corresponding author:

Ms Kate Allen, Department of Child Health, University of Exeter Medical School, Room 2.05, South Cloisters, St Luke's Campus, Exeter, EX1 2LU, UK E-mail: kate.allen@exeter.ac.uk

Co-authors names, affiliations and corresponding addresses:

Dr Ruth Marlow, Devon Autism and ADHD Service, Forde House, Park House, Harrier Way, Sowton Industrial Estate, Exeter, EX2 7HU, UK; R.Marlow@exeter.ac.uk

Ms Vanessa Edwards, Department of Child Health, University of Exeter Medical School, Room 2.05, South Cloisters, St Luke's Campus, Exeter, EX1 2LU, UK;
vanessa.edwards24@btinternet.com

Dr Claire Parker, Devon Autism and ADHD Service, Forde House, Park House, Harrier Way, Sowton Industrial Estate, Exeter, EX2 7HU, UK; C.H.Parker@exeter.ac.uk

Dr Lauren Rodgers, NIHR CLAHRC South West Peninsula (PenCLAHRC), University of Exeter Medical School, Room 1.09, South Cloisters, St Luke's Campus, Exeter, EX1 2LU, UK;
L.R.Rodgers@exeter.ac.uk

Dr Obioha C Ukoumunne, NIHR CLAHRC South West Peninsula (PenCLAHRC), University of Exeter Medical School, Room 1.05, South Cloisters, St Luke's Campus, Exeter, EX1 2LU, UK;
O.C.Ukoumunne@exeter.ac.uk

Dr Edward Chan Seem, Devon Partnership Trust, Wonford House, Dryden Road, Exeter, EX2
5AF, UK; echanseem@nhs.net

Dr Rachel Hayes, Department of Child Health, University of Exeter Medical School, Room
2.05, South Cloisters, St Luke's Campus, Exeter, EX1 2LU, UK; R.A.Hayes@exeter.ac.uk

Ms Anna Price, Department of Child Health, University of Exeter Medical School, Room 2.05,
South Cloisters, St Luke's Campus, Exeter, EX1 2LU, UK; A.Price@exeter.ac.uk

Professor Tamsin Ford, Institute of Health services research, University of Exeter Medical
School, Room 1.09, South Cloisters, St Luke's Campus, Exeter, EX1 2LU, UK;
T.J.Ford@exeter.ac.uk

Abstract

There is a growing focus on child wellbeing and happiness in schools, but we lack self-report measures for very young children. Three samples (N = 2345) were combined to assess the psychometric properties of the How I Feel About My School (HIFAMS) questionnaire, which was designed for children aged 4-8 years. Test re-test reliability was moderate (intraclass correlation coefficient = 0.62). HIFAMS assessed a single concept and had moderate internal consistency (Cronbach's alpha values from 0.62 to 0.67). There were low correlations between scores on the child-reported HIFAMS and parent- and teacher reports. Children at risk of exclusion had significantly lower HIFAMS scores than the community sample (mean difference = 2.4; 95% CI: 1.6 to 3.2; $p < 0.001$). Schools contributed only 4.5% of the variability in HIFAMS score; the remaining 95.5% reflecting pupil differences within schools. Girls' scores were 0.37 units (95% CI: 0.16 to 0.57; $p < 0.001$) higher than boys, while year group and deprivation did not predict HIFAMS score. HIFAMS is a promising measure that demonstrates moderate reliability and discriminates between groups even among very young children.

Keywords

Happiness at school, child mental health, assessment, child self-report measures, HIFAMS

Introduction

Wellbeing can be described as the presence of positive affect and the absence of negative affect (Diener, 1984, 1994) and comprises cognitive and affective components (Tennant et al., 2007). The growth of positive psychology has increased the focus on “wellbeing” as an important area of research (Diener, 2000; Veenhoven, 2000), although researchers have been surprisingly slow in developing studies that specifically address children’s wellbeing (Park & Peterson, 2006). Over the past 10-15 years, there has been an increasing policy focus in the UK on the wellbeing of children in school (Department for Children, 2007; Department for Education and Skills, 2004) combined with policies to promote children’s wellbeing (Public Health England, 2015). There is a similar international focus on the promotion of children’s wellbeing (Fauth, Renton, & Solomon, 2013; The Children's Society, 2012), hence the need for measures to assess well-being. One of the affective components of wellbeing can be usefully defined as *happiness*, which is one of the first emotions recognised by children (Harter, 1983) and can be expressed by children as young as four years old (Lewis & Michalson, 1983). However, most existing research focuses on what makes children happy, not how happy they are, and only considers children over the age of eight (see Chaplin, 2009; Park & Peterson, 2006).

There are four potential methods to assess happiness and wellbeing among young children: questionnaires completed by proxies or by the child, observations, interviews and physiological assessments. The latter is still not well established as a measure of happiness young children (see Holder & Callaway, 2010) and parents’, teachers’ and objective observers’ reports of children’s internal psychological states correlate poorly with each other and a child’s own self-assessment (Achenbach, McConaughy, & Howell, 1987; Collishaw, Goodman,

Ford, Rabe-Hesketh, & Pickles, 2009; Martin et al., 2002; Robinson, Emde, & Corely, 2001). This makes the validation of measures problematic but also stresses the need to include the child's voice. Furthermore, observations or interviews are costly in terms of researcher training and the time taken to complete them and maintain reliable administration and coding. There is, therefore, a need for simple measures that can be used directly with children without the need for sophisticated training to assess the success of interventions and policies in practice.

A number of existing wellbeing and happiness self-report measures have been successfully adapted and used with children (e.g., Holder & Coleman, 2008; Holder, Coleman, & Wallace, 2010; Holder & Klassen, 2010) and a few have been specifically designed for children and include measures of happiness at school (e.g., Huebner, 1994; McLellan & Steward, 2014). However, these measures have only been successfully tested and used with children over the age of seven years, while children as young as four years attend full-time education in the UK. Self-report measures for this younger age group have been shown to work best when employing scales that use cartoon-type faces (Chambers, Giesbrecht, Craig, Bennett, & Huntsman, 1999) or schematic drawings (Hunter, McDowell, Hennessy, & Cassey, 2000), with evidence suggesting children as young as three years perform best at recognising and labelling emotions represented as schematic drawings as opposed to photographs (MacDonald, Kirkpatrick, & Sullivan, 1996). For example the Pediatric Quality of Life Inventory Version 4.0 (PedsQL 4.0) is a measure of quality of life for children and adolescents aged 2 to 18 years that uses faces to represent emotions and feelings, and has consistently been shown to be reliable and valid (Stinson, Kavanagh, Yamada, Gill, & Stevens, 2006; Varni, Seid, & Kurtin, 2001). Younger children can, however, struggle to discern more subtle changes in degree of

sadness (Bieri, Reeve, Champion, Addicoat, & Ziegler, 1990), and some argue that scales with multiple response options are cognitively demanding for younger children (Borgers & Hox, 2000). For the young child self-report measure (ages 5-7), the PedsQL response scale was simplified to a three-point scale with response choices anchored to happy, ok and sad faces.

Due to the lack of other scales in this important area, we developed the “How I Feel About My School” (HIFAMS) questionnaire as a self-report measure of children’s happiness in school (see <http://medicine.exeter.ac.uk/hifams>). The questionnaire was developed for use with children in Foundation and Key Stage 1 (ages 4 to 8 years) and is freely available to download provided the person downloading is not charging for its use. The aim of the current paper was to assess the reliability and validity of the HIFAMS questionnaire.

Method

The HIFAMS measure

The HIFAMS questionnaire (see Appendix 1) comprises of seven questions that ask children how they feel about different school related situations: on the way to school; in the classroom; when doing work; in the playground; about other children; about their teacher; and about school in general. Children respond to each item on a three-point Likert scale, where emoticons are used to convey the responses: *sad* (0), *ok* (1) and *happy* (2). The total score is calculated as the sum of the individual items (possible score ranging from 0 to 14), with higher scores reflecting greater happiness.

The measure was developed as part of a larger research project *STARS* (Ford et al., 2012), by a steering group of teachers, parents, researchers, behavioural support specialists and healthcare professionals. Researchers led the development of the items in consultation with

the steering group about what may impact on wellbeing in school for young children, what were the best questions to ask and what was the best way to elicit reliable responses from children. In addition, the response-format was designed following consultation with three primary school councils (children aged 5-11 years) regarding what they thought the most easily understood response choices would be.

Measures used to assess construct validity of HIFAMS

HIFAMS parent and teacher versions. At the time of this study there were no similar measures for parents or teachers against which we could validate children's responses; therefore, parallel versions of the HIFAMS questionnaire were created for these informants (available from the authors on request). These ask parallel questions to HIFAMS using a five-point Likert scale response format: *very unhappy* (1), *unhappy* (2), *neither happy nor unhappy* (3), *happy* (4) and *very happy* (5). The teacher version excluded the items likely to be challenging for them to answer due to being outside their experience or subject to bias (i.e., how the child feels on the way to school and how the child feels when thinking about the teacher and the entire school).

Strengths and Difficulties Questionnaire. The Strengths and Difficulties Questionnaire (SDQ) (Goodman, 2001) is a reliable, well validated and widely used measure designed to assess the psychopathology of children aged between 3 and 16 years. The questionnaire comprises 25 positively and negatively worded items, which are rated on a three-point Likert-type scale; mostly as *not true* (0), *somewhat true* (1) and *certainly true* (2) but reverse scored for positive statements. The questionnaire is divided into five subscales: *behaviour*, *emotions*, *overactivity/concentration*, *peer relationships* and *prosocial behaviour*. The 20 items in the first four subscales are summed to create a *total difficulties score* (ranging from 0 to 40),

with higher scores indicating greater difficulties. A separate impact supplement asks the respondent whether the child has any difficulties (no, minor, definite or severe) and if so whether these have distressed the child or have impacted on the child's home life, peer relationships, classroom learning or leisure activities (not at all (0), a little (0), quite a lot (1) and a great deal (2)). The scale is scored out of 10 for parent report and 6 for teacher report. The internal consistency of the SDQ has been shown to be satisfactory (Cronbach's alpha (α) = 0.73) and the test-retest stability after four to six months is also reasonable (α = 0.62) (Goodman, 2001).

Participants

The psychometric properties of the HIFAMS questionnaire were assessed using three separate samples of school children aged between 4 and 12 years of age (see Figure 1 for outcomes and Table 1 for participant characteristics from each sample). The *Feasibility sample* (N = 268) was collected from three schools in Devon as part of the initial development and testing of the measure. One school was an inner city school, one was from a market town and the other was a rural village school. The children were aged between 4 and 8 years. The *STARS sample* (N = 2075) comprised the baseline data collected on children in a large cluster randomised controlled trial of teacher classroom management (see Ford et al., 2012 for further details on this sample), which included 80 primary schools across Devon, Plymouth and Torbay. The children in this sample were also between 4 and 8 years of age. Finally, the *SKIP sample* (N = 41) included data from a study exploring the mental health of pupils who were 'at risk' of exclusion or had been excluded from primary school or the first year of secondary school within a county in the South-West of England (see Parker et al., in press for further details). They were aged between 5 and 12.

[Figure 1: Recruitment and outcomes collected per sample]

Ethical permission was granted by the Peninsula College of Medicine and Dentistry Research Ethics Committee through separate applications in relation to all three samples.

Data collection

Figure 1 provides an overview of the measures collected from each sample.

Feasibility sample. All three schools approached, which were selected because of existing links to the researchers, agreed to participate and received £100 book tokens on completion. Parents were sent information about the study via their child's book bag. This information included an "opt-out"/refusal-of-consent form for parents to sign if they did not wish their child to be involved in the study. After the opt-out period of two weeks, researchers visited the schools to administer the HIFAMS questionnaire to the children. Children were asked to complete the HIFAMS questionnaire on two occasions (visits 1 and 2), 2 weeks apart, in February/March 2011. Children who joined the reception classes in September 2010 were included, while those who joined in January 2011 were excluded, as the schools were concerned that they would not have had time to settle in. Teacher HIFAMS questionnaires were given out at visit 1 and collected on visit 2. Instructions were given to complete the questionnaires without consulting the children. Parent HIFAMS questionnaires were sent to parents via their child's book bag and they were asked to complete them without consulting their child within two weeks and return them to school.

STARS sample. After gaining consent from head teachers and teachers, parental consent was obtained using a similar opt-out consent procedure to that described above (Ford et al., 2012). The HIFAMS questionnaire was completed with the children during

September/October of the school's first academic year involved in the study (2012-4).

During this time, teachers were asked to complete the SDQ for each participating child in their class using an online database. Teachers were supplied with half a day's supply cover to do this along with a £20 high street gift voucher. Parents were also asked to complete the SDQ on their child, which was sent home to parents in book bags at the beginning of the data collection period. Parents received a £5 gift voucher on completion of this questionnaire.

SKIP sample. Parents and children were recruited to the study between November 2011 and July 2013. Parents of children receiving a fixed term or permanent exclusion from school, or those who were deemed 'at risk' of exclusion by an educational or mental health practitioner (for example, Special Educational Needs Coordinator, Educational Psychologist, Behavioural Support Teacher) were provided with information about the study from the inclusion and reintegration team and invited to contact the researcher (Parker et al., in press). After giving consent, parents were asked to complete the SDQ. Children were visited at school/home by a researcher to help them complete the HIFAMS questionnaire alongside a larger battery of assessments.

Administering the HIFAMS

In the *Feasibility* and *STARS samples* the HIFAMS questionnaire was administered using the same procedure. Initially, the task was introduced to the whole class (or individual child in the *SKIP sample*) by a researcher with a standard script (see <http://medicine.exeter.ac.uk/hifams>). Set examples were given by the researcher during this introduction in an attempt to engage the children with the response scale being used. After explaining the task, children then completed the questionnaire. For key stage one (4 to

7 years of age), children were assisted on an individual basis by a researcher reading the items and repeating the examples given in the standard procedure if necessary. In this case, where possible, questionnaires were completed outside of the classroom. The questionnaire took between 5 and 10 minutes to complete with each child. For key stage two (7 to 8 years of age), children were asked to sit at their desks and work through the questionnaire independently. Where help was needed, children were encouraged to raise their hand and wait for one of the researchers to help them. In this instance, the questionnaires took around 30 minutes for the whole class to complete. In the *SKIP sample*, children completed their questionnaire alongside a researcher who was completing a separate battery of assessments.

Data Analysis

Data analysis was undertaken using Stata (version 13.1) and MPlus (Muthén & Muthén, 2007).

Participant characteristics, HIFAMS and SDQ responses were summarised using means and standard deviations, medians and interquartile ranges and ranges for continuous variables and frequencies and percentages for categorical variables.

The test-retest reliability of the HIFAMS items was quantified by reporting the weighted Kappa and percentage agreement between visit 1 and visit 2 responses in the *Feasibility sample*. The amount of agreement can be classified into almost perfect (>0.80) substantial (0.61-0.80) moderate (0.41-0.60) fair (0.21-0.40) and poor (<0.21), which is the terminology used in throughout this chapter (Landis & Koch, 1977). Test-retest reliability of the total HIFAMS score was assessed using the Bland-Altman (Bland & Altman, 1986) method and the

intraclass correlation coefficient. The Bland-Altman method summarises the variability of the change in the responses between visit 1 and visit 2 using the 95% range.

Factor analysis was used to explore the dimensional-structure of the HIFAMS measure (i.e., whether there are sub-dimensions) for the *Feasibility sample* (each of visits 1 and 2) and the *STARS sample*. The estimation method used for factor analysis recognised the ordinal nature of the item responses, using mean- and variance-adjusted (“robust”) weighted least squares parameter estimates. The factor analysis was based on the polychoric correlation coefficients between the items (Flora & Curran, 2004). Factor loadings greater than 0.4 were considered salient. Internal consistency across the HIFAMS items was quantified using Cronbach’s alpha (α).

The construct validity of the total HIFAMS score was further investigated by reporting the Spearman correlation coefficients among the three respondents (children, parents and teachers) for the *Feasibility sample* and reporting the correlation between the HIFAMS total score and the SDQ score in the *STARS sample*. Previous studies have shown that correlations from different informants reporting on psychopathology in children are approximately 0.3 (Achenbach et al., 1987). We anticipated that our analysis would also produce correlations of a similar size given that the measures were cross-informant and, in terms of HIFAMS – SDQ, with overlapping rather than equivalent constructs. We predicted that psychopathology measured by the SDQ total difficulties score and its impact (SDQ Impact subscale) would be negatively correlated with HIFAMS.

Random effects (“multilevel”) linear regression models were fitted using the *STARS sample* to examine the relationship of HIFAMS total score (outcome) with each of gender, year group and deprivation (predictors) allowing for variation across school clusters.

Results

Participant characteristics are summarised for each sample in Table 1. The *Feasibility* and *STARS samples* had a similar percentage of boys (49.6% and 53.1%, respectively) and mean ages (6.3 and 6.8 years, respectively). In contrast, the *SKIP sample* was predominately boys (95.3%) and older, as expected for a sample containing children that had been either excluded or were at risk of exclusion from school (Department for Education, 2014; Parker, 2014). Parent and teacher SDQ scores suggested higher levels of psychopathology among the *SKIP sample* compared to the *Feasibility* and *STARS samples*.

[Table 1: Participant characteristics by sample]

Table 2 summaries the response rates for each measure, informant and visit in each study. Completion rates were extremely high for most measures and over 70% for all except parents in the *Feasibility sample*. Teachers who responded provided complete data on the children in their classes, but one teacher's absence meant that no data could be collected on any of the children in that class in the *Feasibility sample*.

[Table 2: Response rates and median (IQR) scores for the HIFAMS and SDQ measures]

HIFAMS questionnaire descriptive data

Most children reported high levels of happiness, particularly in the *STARS* and *Feasibility samples*. The playground item had the greatest percentage of children reporting they were happy and the work item had the smallest percentage of children responding that they were happy (see Table 3), which mirrored the responses of teachers and parents. Parents and teachers consistently rated the children as happier than children rated themselves.

Similarly, both parents and teachers also consistently underestimated the child's reported unhappiness.

[Table 3: HIFAMS item responses from children, parents and teachers in the Feasibility sample]

Factor analysis

The factor loadings from the one-factor model are shown in Table 4 for the *Feasibility* and *STARS samples*. Factor analysis of the HIFAMS items indicated that there was only one underlying dimension in the measure as there was only one factor that had eigenvalue greater than 1 and the eigenvalue of the first factor was large relative to the second and all subsequent factors (Table 4). The playground item was the only one to have a factor loading below 0.4.

Internal consistency

Children's responses in the *Feasibility* (time 1), *Feasibility* (time 2) and *STARS samples* demonstrated moderate internal consistency (Cronbach's alpha values of 0.62, 0.67 and 0.62, respectively). Dropping the question about the playground in each case increased the alpha value by a small amount (to 0.65, 0.68 and 0.64, respectively), which suggests little advantage in doing this. There was a high level of consistency for the responses to questions for the parent and teacher HIFAMS with values of 0.81 and 0.82 respectively, which suggests that consistency differs across respondent type.

Test-retest reliability

There was moderate agreement on the child-reported total score between the two visits (intraclass correlation coefficient 0.62; 95% confidence interval ((CI): 0.54 to 0.69). The 95% limit of agreement was (-5.1 to 5.8) meaning that for most of the children the absolute change in total score was no greater than 5.8. Weighted kappa values for test-retest reliability for individual HIFAMS items ranged from 0.19 to 0.43 (see Table 4) and were poor (below 0.2) for how the child felt about the playground, the school and the teacher.

[Table 4: Test-retest reliability and HIFAMS factor loadings from factor analysis]

Correlation between adult HIFAMS responses and children's responses

There were low correlations between HIFAMS and the adult informants on the parallel measure in the *Feasibility sample* (Table 5), which is not an uncommon finding in this field. Correlations were in the expected directions and not markedly higher between parents and teachers, despite the greater internal consistency of the parallel measure that they completed.

[Table 5: Correlations between HIFAMS and adult informants on parallel measure]

Correlation between HIFAMS and SDQ responses

The negative correlations between the total HIFAMS scores with the SDQ total difficulties and Impact scores respectively reported by both informants in the *STARS sample* (see Table 6) were in the predicted directions, although very small.

[Table 6: Correlations between HIFAMS and teacher / parent completed SDQ]

Comparison of HIFAMS between unselected community (STARS) and at risk (SKIP) sample

The mean total HIFAMS score differed between the *STARS* and the *SKIP* samples with the *STARS* sample being happier at school than the *SKIP* sample (mean difference = 2.4; 95% CI: 1.6 to 3.2; $p < 0.001$) (see Table 2). This difference remained after adjusting for gender (mean difference = 2.2; 95% CI: 1.5 to 3.0; $p < 0.001$).

Year group, gender, deprivation and school effects (STARS sample)

On fitting a null random effects regression model (i.e., with no predictor variables) to the HIFAMS score, the intra-cluster (intra-school) correlation coefficient was estimated to be 0.049 which means that 4.9% of the variability in the HIFAMS measure was due to differences between schools and the remaining 95.1% due to differences between pupils within schools. Gender was a significant predictor of HIFAMS score, with girls having a mean HIFAMS score that was 0.37 units (95% CI: 0.16 to 0.57; $p < 0.001$) higher than boys, but this only explained a negligible amount of variation. Neither year group ($p = 0.15$) nor deprivation ($p = 0.85$) were significant predictors of HIFAMS score.

Discussion

Our findings suggest that HIFAMS assesses a single concept and has moderate internal consistency in the measurement of happiness in school among very young children for whom there are a lack of school-focused measures of happiness. Given that test-retest reliability was moderate, while the test-retest reliability for individual items was poor, the HIFAMS total score may be more useful than individual items when assessing change in both research and practice. This is not entirely unexpected given the greater affective volatility of younger children and their lower levels of emotional regulation. We found that HIFAMS

scores differed as predicted between children at risk of exclusion from school compared to whole school samples that would be expected to have more positive experiences of school.

Happiness is experienced internally and therefore validity is difficult to assess. Previous studies demonstrate poor inter-informant agreement on measures of psychopathology in children (Achenbach et al., 1987), so the low correlations obtained between HIFAMS and adult completed measures was not unexpected. The correlations, although very small, particularly for the single items of the HIFAMS, were in the predicted directions, with the exception of teacher reported prosocial score in relation to happiness in the playground and parent reported impact on leisure activities. The latter are likely to include out of school as well as school related pastimes, and strong prosocial skills may not necessarily make playtimes go smoothly in all schools. One explanation for these low correlations is that parents and teachers are reporting on observed behaviour, whereas children are reporting on internal states. The low correlation between children's and adults' responses in this study, therefore, should not be interpreted at face value as absence of validity. In contrast, they could be taken as a strong argument to include the child's voice when assessing the outcome of interventions at individual, classroom, school and regional level. This suggestion is emphasised by our findings that parents and teachers consistently overestimated levels of happiness and underestimated levels of unhappiness, and that the parent-teacher reports were also poorly correlated.

When assessing validity it is preferable to use a measure that considers the same construct (Prince, Stewart, Ford, & Hotopf, 2003). Although the SDQ is a well-validated measure of child psychopathology, it is not a measure of child happiness per se. Unfortunately, at the time of the study there was a lack of alternative well-validated measures assessing young

children's happiness for us to use and we were conscious not to place too many demands on young children in the school context. Future research could try to compare HIFAMS with new measures of happiness that have since been published (e.g. McLellan & Steward, 2014). However, these new measures are not validated for children under the age of six.

Although it has been difficult to demonstrate validity of HIFAMS, our finding that HIFAMS scores varied significantly between unselected community samples and children at risk of the breakdown of their school placement provides some evidence that HIFAMS can discriminate between groups of children with different experiences. We found that children who had either been excluded, or were at risk of exclusion, from school reported significantly lower levels of happiness using HIFAMS as compared to our unselected sample of children. This would be predicted given that the former group were struggling to cope at school (Parker, 2014). In addition to this, the high proportion of children who reported happiness in the playground and the lower proportion of children who were happy when working in the classroom is in keeping with common experience and suggests face validity of the measure. Validity could be further studied by examining differences during the school year and in response to school events that could be predicted to affect happiness, such as exams, fetes, or residential courses.

As anticipated the youngest children required support to complete HIFAMS, and, even though completion with each child took less than 10 minutes, this requirement may limit its application. The moderate internal consistency and test-retest reliability may be explained by a poorer understanding of the questions or how to respond, and/or a poorer ability to understand and regulate emotions, with fewer strategies to cope with adversity. It may be that there is a lower age limit to which this measure is appropriate, but studies involving

larger samples of each age group are required to test this empirically. It would also be useful to test the measure in older children; we would predict that reliability would increase with age, as internal consistency was higher among parents and teachers. The small increase in internal consistency that could be obtained by dropping the item about being in the playground is offset by the importance of including a question about unstructured, less supervised time to important issues that are highly likely to influence children's happiness at school, such as bullying, that might be missed were it omitted.

In addition to exploring the reliability and validity of HIFAMS, we found an interesting difference in happiness at school between genders and among schools. Approximately 5% of variance in HIFAMS score was attributed to the school. This suggests that the influence of school environment on children's happiness is smaller than that of individual characteristics of the child. However, school environment may be more tractable than temperament or home environment.

We have demonstrated that HIFAMS shows promising reliability and validity, but our findings need to be interpreted in the context of several methodological issues. The study benefitted from three separate samples of children but all three samples were recruited from schools in Devon. The findings, therefore, may not be generalisable to children from outside the West of England, and in particular, to children of Black and Minority Ethnicity groups who were very poorly represented in the current analyses. With the *Feasibility* and *STARS samples*, parent questionnaires were left in book bags, and parents who responded are likely to differ from those who did not respond or opted out. In general though, parent response rates were high (Cook, Dickinson, & Eccles, 2009) for what was essentially a postal survey. In addition, parents completing the HIFAMS questionnaire about their child might

have completed the questionnaire with their child rather than independently, although the low levels of correlation in responses suggests that did not occur often, if at all.

Preliminary work was carried out with children in schools around the layout, wording and response formats for the questionnaire as far as time and resources permitted. Future research should explore what influences their happiness in school, as well as the thoughts behind their answers to the questionnaire, which could inform the further development of the current or other measures (Tennant et al. 2007). The inclusion of additional items and adaptation of existing ones might increase the validity and reliability of the measure as well as ensuring the measure truly reflects factors important to the happiness of children in school.

Seligman *et al.*, (2009), highlight that 'happy and settled' children are likely to perform better, thus wellbeing is likely to impact greatly on academic attainment. Bearing in mind the moderate reliability and sensitivity of the measure there are several suggested uses for the HIFAMS moving towards the future. Firstly, HIFAMS could be used to assess the wellbeing of children at class or school level, either as a one off measurement or to assess the impact of events or interventions. Secondly, whilst collecting these data a small number of children disclosed major difficulties with bullying to the researchers that had not been identified by their teacher. This suggests that HIFAMS could be used as a tool with which to explore school experience with children at an individual level, although we cannot currently make any judgement about what scores in any particular range might indicate.

The current study suggests that HIFAMS is a promising measure that assesses happiness in the school setting, demonstrates moderate reliability and discriminated between groups of children expected to have differing experiences of school even in very young children. It

requires additional testing in larger samples with more varied socio-demographic backgrounds in order to better understand how it functions.

Ethics

The *STARS*, *feasibility* and *SKIP* studies had institutional approval from the University of Exeter Medical School.

Funding acknowledgement

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*(2), 213-232. doi: 10.1037/0033-2909.101.2.213
- Bieri, D., Reeve, R. A., Champion, G. D., Addicoat, L., & Ziegler, J. B. (1990). The Faces Pain Scale for the self-assessment of the severity of pain experienced by children: development, initial validation, and preliminary investigation for ratio scale properties. *Pain*, *41*(2), 139-150.
- Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *i*, 307-310.
- Borgers, N., & Hox, J. J. (2000). *Reliability of responses in questionnaire research with children*. Paper presented at the Fifth International Conference on Logic and Methodology, Cologne, Germany.
- Chambers, C. T., Giesbrecht, K., Craig, K. D., Bennett, S. M., & Huntsman, E. (1999). A comparison of faces scales for the measurement of pediatric pain: children's and parents' ratings. *Pain*, *83*(1), 25-35. doi: [http://dx.doi.org/10.1016/S0304-3959\(99\)00086-X](http://dx.doi.org/10.1016/S0304-3959(99)00086-X)
- Chaplin, L. (2009). Please May I Have a Bike? Better Yet, May I Have a Hug? An Examination of Children's and Adolescents' Happiness. *Journal of Happiness Studies*, *10*(5), 541-562. doi: 10.1007/s10902-008-9108-3
- Collishaw, S., Goodman, R., Ford, T., Rabe-Hesketh, S., & Pickles, A. (2009). How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? *Journal of Child Psychology and Psychiatry and allied disciplines*, *50*(5), 571-580. doi: 10.1111/j.1469-7610.2008.02026.x
- Cook, J. V., Dickinson, H. O., & Eccles, M. P. (2009). Response rates in postal surveys of healthcare professionals between 1996 and 2005: an observational study. *BMC Health Serv Res*, *9*, 160. doi: 10.1186/1472-6963-9-160

Department for Children, Schools and Families. (2007). *The children's plan: building brighter futures*.

UK: The Stationary Office.

Department for Education. (2014). Permanent and Fixed Period Exclusions in England: 2012 to 2013.

London.

Department for Education and Skills. (2004). *Every child matters: change for children*. Nottingham,

UK: Department for Education and Skills Publications Retrieved from

<http://webarchive.nationalarchives.gov.uk/20130401151715/https://www.education.gov.uk/publications/eOrderingDownload/DfES10812004.pdf>.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, 95(3), 542-575.

Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social Indicators Research*, 31(2), 103-157. doi: 10.1007/BF01207052

Diener, E. (2000). Subjective well-being. The science of happiness and a proposal for a national index. *Am Psychol*, 55(1), 34-43.

Fauth, B., Renton, Z., & Solomon, E. (2013). Tackling child poverty and promoting children's well-being: lessons from abroad. London: National Children's Bureau.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychol Methods*, 9(4), 466-491. doi: 10.1037/1082-989x.9.4.466

Ford, T., Edwards, V., Sharkey, S., Ukoumunne, O., Byford, S., Norwich, B., & Logan, S. (2012).

Supporting teachers and children in schools: the effectiveness and cost-effectiveness of the incredible years teacher classroom management programme in primary school children: a cluster randomised controlled trial, with parallel economic and process evaluations. *BMC Public Health*, 12(1), 1-16. doi: 10.1186/1471-2458-12-719

Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(11), 1337-1345. doi:

<http://dx.doi.org/10.1097/00004583-200111000-00015>

- Harter, S. (1983). Developmental perspectives on the self-system. In E. M. Hetherington (Ed.), *Socialization, personality, and social development* (Vol. 4, pp. 275-385). New York: Wiley.
- Holder, M. D., & Callaway, R. J. (2010). Happiness in children: a review of the scientific literature. In A. Mäkinen & P. Hájek (Eds.), *Psychology of happiness* (pp. 51-70). New York: Nova Science Publishers, Inc.
- Holder, M. D., & Coleman, B. (2008). The contribution of temperament, popularity, and physical appearance to children's happiness. *Journal of Happiness Studies*, *9*(2), 279-302. doi: 10.1007/s10902-007-9052-7
- Holder, M. D., Coleman, B., & Wallace, J. M. (2010). Spirituality, Religiousness, and Happiness in Children Aged 8–12 Years. *Journal of Happiness Studies*, *11*(2), 131-150. doi: 10.1007/s10902-008-9126-1
- Holder, M. D., & Klassen, A. (2010). Temperament and Happiness in Children. *Journal of Happiness Studies*, *11*(4), 419-439. doi: 10.1007/s10902-009-9149-2
- Huebner, E. S. (1994). Preliminary development and validation of a multidimensional life satisfaction scale for children. *Psychological Assessment*, *6*(2), 149-158. doi: 10.1037/1040-3590.6.2.149
- Hunter, M., McDowell, L., Hennessy, R., & Cassey, J. (2000). An evaluation of the Faces Pain Scale with young children. *J Pain Symptom Manage*, *20*(2), 122-129.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159-174. doi: 10.2307/2529310
- Lewis, M., & Michalson, L. (1983). *Children's emotions and moods: developmental theory and measurement*. New York: Plenum Publishing Company Limited.
- MacDonald, P. M., Kirkpatrick, S. W., & Sullivan, L. A. (1996). Schematic drawings of facial expressions for emotion recognition and interpretation by preschool-aged children. *Genetic, Social and General Psychology Monographs*, *122*(4), 373-388.
- Martin, L. R., Friedman, H. S., Tucker, J. S., Tomlinson-Keasey, C., Criqui, M. H., & Schwartz, J. E. (2002). A Life Course Perspective on Childhood Cheerfulness and its Relation to Mortality

Risk. *Personality and Social Psychology Bulletin*, 28(9), 1155-1165. doi:

10.1177/01461672022812001

McLellan, R., & Steward, S. (2014). Measuring children and young people's wellbeing in the school context. *Cambridge Journal of Education*, 1-26. doi: 10.1080/0305764X.2014.889659

Muthén, L. K., & Muthén, B. O. (2007). *Mplus User's Guide* (Fifth Edition ed.). Los Angeles, CA: Muthén & Muthén.

Park, N., & Peterson, C. (2006). Character Strengths and Happiness among Young Children: Content Analysis of Parental Descriptions. *Journal of Happiness Studies*, 7(3), 323-341. doi: 10.1007/s10902-005-3648-6

Parker, C. (2014). *An exploration of the relationship between exclusion from school and children's mental health and well-being: a mixed methods approach*. (Doctor of Philosophy in Psychology), Universities of Exeter and Plymouth, University of Exeter.

Parker, C., Marlow, R., Kastner, M., May, F., Mitrofan, O., Henley, W., . . . Ford, T. (in press). SKIP 'Supporting Kids, avoiding Problems' study- case control, descriptive paper. *Journal of Children's Services*.

Prince, M., Stewart, R., Ford, T., & Hotopf, M. (Eds.). (2003). *Practical Psychiatric Epidemiology*. Oxford: Oxford University Press.

Public Health England. (2015). *Promoting children and young people's emotional health and wellbeing: A whole school and college approach*. London: Public Health England Publications.

Robinson, J. L., Emde, R. N., & Corely, R. P. (2001). Heterogeneity of joy in infancy. In R. N. Emde & J. K. Hewitt (Eds.), *Infancy to early childhood: genetic and environmental influences on developmental change* (pp. 163-177). New York: Oxford University Press.

Stinson, J. N., Kavanagh, T., Yamada, J., Gill, N., & Stevens, B. (2006). Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity

measures for use in clinical trials in children and adolescents. *Pain*, 125(1-2), 143-157. doi:
10.1016/j.pain.2006.05.006

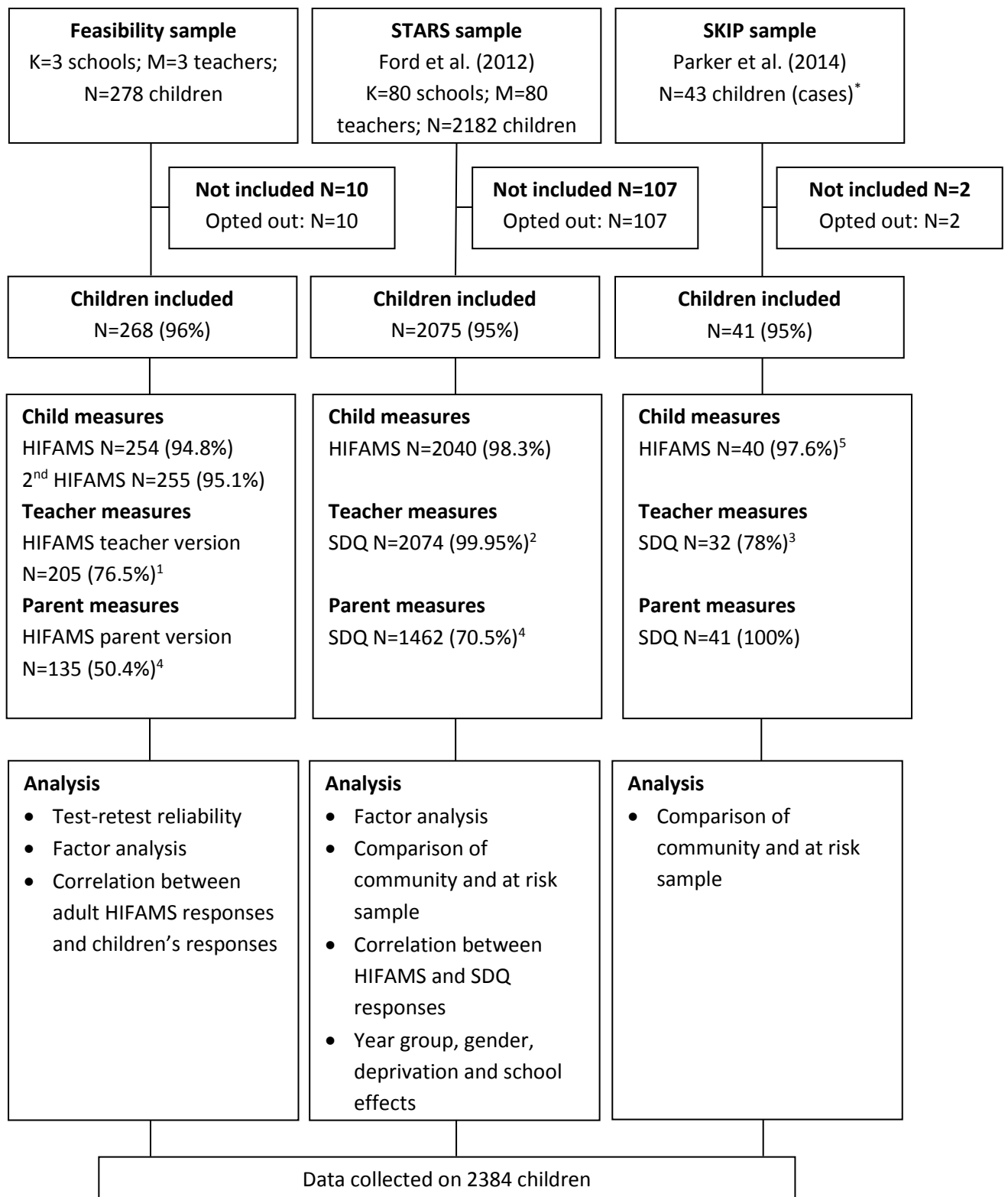
Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., . . . Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health and Quality of Life Outcomes*, 5(1), 63.

The Children's Society. (2012). The good childhood report 2012: a summary of our findings (pp. 1-64): The Children's Society.

Varni, J. W., Seid, M., & Kurtin, P. S. (2001). PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Med Care*, 39(8), 800-812.

Veenhoven, R. (2000). The Four Qualities of Life. *Journal of Happiness Studies*, 1(1), 1-39. doi:
10.1023/A:1010072010360

Figure 1: Recruitment and outcomes collected per sample



¹ Missing data represents entire classes where information was missing due to staff sickness

² Missing data represents children who had left the school (and thus, teachers were unable to provide data on)

³ 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.

⁴ Missing data represents parents who failed to respond to e-mails/letters asking them to complete questionnaires

⁵ Missing data represents children who were unable to complete HIFAMS

Table 1: Participant characteristics by sample

	Feasibility sample N = 268 pupils (3 schools)	STARS sample N = 2075 pupils (80 schools)	SKIP sample N = 41 pupils (40 schools)
Female, %	50.4	46.9	4.9
Age, mean (SD)	6.3 (0.8)	6.8 (1.3)	9.0 (2.0)
Year Group			
Reception, %	19.8	13.0	2.4
Year 1, %	38.8	17.7	14.6
Year 2, %	36.6	19.8	9.8
Year 3, %	4.9	29.4	12.2
Year 4, %	0	20.1	12.2
Year 5, %	0	0	24.4
Year 6, %	0	0	12.2
Year 7, %	0	0	9.8
in lowest 20% ¹ of SES ² , %	. ³	16.3 ⁴	7.3
Number of schools	3	80	40

¹ lowest quintile of index of multiple deprivation

² Social Economic Status (SES) at child level

³ data not available for *Feasibility sample*

⁴ data only available for parents who returned parent questionnaire n=1456

Table 2: Response rates and median (IQR) scores for the HIFAMS and SDQ measures

	Informant/Measure	Response n (%)	Total score Mean (SD)	Total score Median (IQR)	Total score Range
Feasibility sample N=268	Child HIFAMS visit 1	254 (94.8)	11.1 (2.4)	11 (10 to 13)	0 to 14
	Child HIFAMS visit 2	255 (95.1)	10.7 (2.6)	11 (9 to 13)	1 to 14
	Parent HIFAMS	135 (50.4)	28.9 (3.3)	28 (27 to 31)	19 to 35
	Teacher HIFAMS	205 (76.5)	15.9 (2.4)	16 (14 to 18)	9 to 20
STARS sample N=2075	Child HIFAMS	2040 (98.3)	11.0 (2.4)	11 (10 to 13)	0 to 14
	Parent SDQ	1462 (70.5)	7.0 (5.6)	6 (3 to 9.5)	0 to 32
	Teacher SDQ	2074 (99.95)	6.7 (5.9)	5 (2 to 10)	0 to 31
SKIP sample N=41	Child HIFAMS	40 (97.6)	8.3 (3.1)	8 (7 to 10.5)	0 to 14
	Parent SDQ	41 (100)	22.3 (6.9)	22 (18 to 27)	6 to 35
	Teacher SDQ	32 (78)	22.3 (5.1)	23 (20 to 26)	6 to 29

Table 3: HIFAMS item responses from children, parents and teachers in the *Feasibility sample*

	HIFAMS Question						
	Way to School %	Classroom %	Work %	Playground %	Other Children %	Teacher %	School %
Child HIFAMS first visit (Feasibility sample, N=254)							
Sad	6.7	7.5	12.6	4.7	8.3	6.3	7.9
OK	30.7	30.3	32.7	8.3	31.9	22.1	28.7
Happy	62.6	62.2	54.7	87	59.8	71.7	63.4
Child HIFAMS second visit (Feasibility sample, N=254)							
Sad	7.5	7.5	18.4	5.5	5.9	9.4	7.1
OK	38.8	31	31.8	12.5	38	23.9	27.8
Happy	53.7	61.6	49.8	82	56.1	66.7	65.1
Parent (Feasibility sample N=135)							
Very unhappy	0	0	0	0	0	0.7	0
Unhappy	2.2	0.7	1.5	1.5	3	0	2.2
Neither happy nor unhappy	12.4	10.2	16.1	10.9	18.5	10.9	13.1
Happy	55.5	62.8	46.7	56.2	57.8	42.3	62.8
Very happy	29.9	26.3	35.8	31.4	20.7	46	21.9
Teacher							

(Feasibility sample, N=205)							
Very unhappy	-	0	0.5	0	0	-	-
Unhappy	-	1.5	4.4	1.5	2	-	-
Neither happy nor unhappy	-	18	24.9	20	22	-	-
Happy	-	62.9	44.9	49.8	50.2	-	-
Very happy	-	17.6	25.4	28.8	25.9	-	-
Child HIFAMS							
(STARS sample, N=2053)							
Sad	7.9	5.8	9.8	6.1	7.4	3.5	11.3
OK	32.6	29.6	36.8	15.6	31.2	21.4	29.7
Happy	59.5	64.6	53.4	78.4	61.5	75.1	59.1
Child HIFAMS							
(SKIP sample, N=40)							
Sad	10.3	18.0	25.6	15.4	20.5	20.5	28.2
OK	43.6	56.4	56.4	20.5	33.3	46.2	43.6
Happy	46.6	25.6	18.0	64.1	46.2	33.3	28.2
Total	100	100	100	100	100	100	100

Table 4: Test-retest reliability and HIFAMS factor loadings from factor analysis

Item	% agreement between time 1 and time 2 (N=244)	Weighted kappa and 95% CI	Feasibility visit 1 Factor loading (N=254)	Feasibility visit 2 Factor loading (N=255)	STARS Factor loading (N=2040)
Way to school	65.6	0.41 (0.29 to 0.52)	0.55	0.72	0.58
Classroom	63.9	0.31 (0.18 to 0.43)	0.53	0.48	0.58
Work	60.3	0.43 (0.30 to 0.54)	0.51	0.58	0.46
Playground	76.2	0.19 (0.01 to 0.37)	0.13	0.25	0.26
Other Children	64.8	0.40 (0.26 to 0.50)	0.60	0.58	0.50
Teacher	68.0	0.35 (0.23 to 0.49)	0.55	0.60	0.52
School	61.9	0.28 (0.14 to 0.42)	0.71	0.66	0.74

NB Values above 0.4 were interpreted as salient loadings

Table 5: Correlations between HIFAMS and adult informants on parallel measure

Variables Compared	Sample size	r_s	95% CI	p	r_s	95% CI	p
HIFAMS Child vs Parent Visit 1	135	-	-	-	0.21	0.04 to 0.38	0.02
HIFAMS Child vs Parent Visit 2	135	-	-	-	0.19	0.01 to 0.35	0.03
HIFAMS Child vs Teacher Visit 1	205	-	-	-	0.06	0.08 to 0.20	0.42
HIFAMS Child vs Teacher Visit 2	205	-	-	-	0.00	-0.11 to 0.17	0.95
HIFAMS Teacher vs Parent	135	0.20	0.01 to 0.38	0.04	-	-	-

Table 6: Correlations between HIFAMS and teacher / parent completed SDQ

Variables compared	r_s	95% CI	p
<i>Correlations between total scores</i>			
Total HIFAMS vs teacher rated SDQ total	-0.17	-0.21 to -0.13	0.00001
Total HIFAMS vs parent rated SDQ total	-0.11	-0.16 to -0.06	0.00001
Total HIFAMS vs teacher rated SDQ impact score	-0.12	-0.16 to -0.08	0.00001
Total HIFAMS vs parent rated SDQ impact score	-0.10	-0.15 to -0.05	0.0002
<i>Correlations between individual items (see footnotes for key)</i>			
HIFAMS Q2 ¹ vs teacher rated impact on classroom learning	-0.09	-0.13 to -0.04	0.0001
HIFAMS Q2 ¹ vs parent rated impact on classroom learning	-0.07	-0.13 to -0.02	0.0047
HIFAMS Q4 ² vs teacher rated prosocial subscale score	0.03	-0.01 to 0.08	0.1158
HIFAMS Q4 ² vs parent rated prosocial subscale score	0.05	0.00 to 0.10	0.0429
HIFAMS Q4 ² vs parent rated impact on leisure activities	-0.04	-0.09 to 0.01	0.1105
HIFAMS Q5 ³ vs teacher rated prosocial subscale score	0.09	0.04 to 0.13	0.0001
HIFAMS Q5 ³ vs parent rated prosocial subscale score	0.05	0.00 to 0.10	0.0486
HIFAMS Q5 ³ vs teacher rated impact on peer relationships	-0.07	-0.12 to -0.03	0.0011

¹ HIFAMS Q2 - 'When I am in the classroom I feel ...'






















² HIFAMS Q4 - 'When I am in the playground I feel ...'

³ HIFAMS Q5 - 'When I think about the other children at school I feel ...'

Name.....

How I feel about my school

Please put a circle around the face that shows how you feel

	Happy	OK	Sad
On my way to school I feel ...			
When I am in the classroom I feel ...			
When I am doing my work I feel ...			
When I am in the playground I feel ...			
When I think about the other children at school I feel ...			
When I think about my teacher I feel ...			
When I think about school I feel ...			

Thank you very much for doing this

