

A new measure of unhealthy school environments and its implications for critical assessments of health promotion in schools

Nikki Shackleton PhD
Research Fellow
Centre of Methods and Policy Application in the Social Sciences
University of Auckland
Fale – Office Building - Bldg 273
Level 2
20 Wynward Street
AUCKLAND 1010
New Zealand
Tel: +64 9 923 5109
Email: nicholashackleton@hotmail.com

Adam Fletcher PhD
Reader in Health and Social Science
School of Social Sciences
Cardiff University
Room 2.03, 1-3 Museum Place
Cardiff CF10 3BD
UK
Tel: +44 (0)29 2087 9679
Email: fletcher@cardiff.ac.uk

Farah Jamal PhD
Research Officer
Department of Social Science
UCL Institute of Education
18 Woburn Square
London WC1H 0NR
UK
Tel: +44 (0)20 7612 6573
Email: f.jamal@ioe.ac.uk

Wolfgang Markham PhD
Associate Professor

Warwick Medical School- Statistics and Epidemiology
University of Warwick
Coventry CV4 7AL
UK
Tel: +44(0)24 7657 4129
Email: Wolfgang.Markham@warwick.ac.uk

Paul Aveyard PhD
Professor of Behavioural Medicine
Department of Primary Care Health Sciences
University of Oxford
New Radcliffe House
Radcliffe Observatory Quarter
Woodstock Road
Oxford OX2 6GG
Tel: +44 (0)1865 617860
Email: Paul.aveyard@phc.ox.ac.uk

Anne Mathiot MSc
Trial Manager
University College London Institute of Child Health
30 Guildford Street
London WC1N 1EH
UK
Tel: +44 (0)207 905 2772
Email: a.mathiot@ucl.ac.uk

Elizabeth Allen PhD
Senior Lecturer
Department of Medical Statistics
Room G35
London School of Hygiene and Tropical Medicine
Keppel Street
London WC1E 7HT
UK
Tel: 020 79272943
Email: elizabeth.allen@lshtm.ac.uk

Russell Viner PhD
Professor in Adolescent Health

University College London Institute of Child Health
30 Guildford Street
London WC1N 1EH
UK
Tel: +44 (0)20 7242 9789
Email: r.viner@ucl.ac.uk

Chris Bonell PhD*
Professor of Public Health Sociology
Department of Social and Environmental Health Research
London School of Hygiene and Tropical Medicine
15-17 Tavistock Place
London WC1H 9SH
UK
Tel. +44 (0)20 7612 7918
Email chris.bonell@lshtm.ac.uk

Word count: 4791

* corresponding author

Acknowledgements

We would like to thank the other co-investigators working on the INCLUSIVE trial as well as the students and staff of the participating schools. This project is funded from a grant by the National Institute for Health Research Public Health Research Programme (grant 12/153/60). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the National Institute for Health Research Public Health Research Programme or the Department of Health. *The work was undertaken with the support of The Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research Centre of Excellence. Joint funding (MR/KO232331/1) from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the Welsh Government and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged.*

Author contributions

CB co-directed the study and led the planning and drafting of this paper. NS designed and implemented statistical analyses with support from EA and both contributed to drafting the paper. FJ and AF contributed to planning and drafting the paper, and FJ collected staff data. WM and PA advised on construction of the VAE measure and contributed to planning and drafting. AM led data collection and contributed to the methods sections. RV co-directed the study and contributed to planning the paper.

Disclosure statement

The authors declare that they have no conflict of interest.

A new measure of unhealthy school environments and its implications for critical assessments of health promotion in schools

Abstract

243 words

The theory of human functioning and school organisation informed by Basil Bernstein's sociology of education suggests that to gain the commitment and promote the health of students, particularly those from disadvantaged backgrounds, schools require radical transformations eroding various 'boundaries': between and among staff and students; between students' academic learning and broader social development and welfare; and between schools and their local communities. Existing research examining this theory has reported associations between school-level proxy measures of student commitment and lower rates of student smoking, drinking alcohol, use of drugs and violence. But this research has not directly assessed whether reduced school boundaries explain this. We piloted a new scale derived from teacher reports to measure unhealthy school boundaries and examined its inter-item reliability and its criterion validity in terms of associations with various measures of school commitment and smoking. Data on boundaries came from 101 teachers across 40 schools. Data on student commitment and smoking came from 6667 students. We assessed reliability by examining correlations between scale-items and criterion validity in terms of associations with student-reported commitment and smoking. Inter-item reliability was sub-optimal but better within the sub-scales about boundaries between academic/broader learning and schools/local communities. The scale had good criterion validity, strongly associated with reduced student-reported school commitment and increased student-reported smoking. We reflect on the implications of these findings in terms of critical perspectives on health promotion in schools and the strengths and limitations of quantitative research in examining health behaviours as opposed to practices.

Keywords

Schools, smoking, multi-level models, environment

A new measure of unhealthy school environments and its implications for critical assessments of health promotion in schools

Introduction

Over the last three decades there has been increasing interest in how the social environment influences health and health inequalities (Marmot, 2010). This includes quantitative research on how the school social environment affects smoking, drug use and other ‘health behaviours’. Most empirical research on this topic has been atheoretical and the few studies that do use theory, do so using theoretical frameworks concerning how students can successfully connect with or commit to school, rather than how schools can successfully engage students (Bonell, Fletcher, et al., 2013). The key exception is a body of empirical research (Aveyard et al., 2004; Bisset, Markham, & Aveyard, 2007; W.A. Markham et al., 2008; W.A. Markham, Young, Sweeting, West, & Aveyard, 2012; Tobler, Komro, Dabroski, Aveyard, & Markham, 2011) and theoretical work (W. A. Markham & Aveyard, 2003) focused on the theory of human functioning and school organisation. Drawing on Basil Bernstein’s critical sociology of the reproduction of educational inequalities (Bernstein, 1975), as well as Martha Nussbaum’s ethical-political analyses of human needs (Nussbaum, 1990) and Paolo Freire’s pedagogy of the oppressed (Freire, 1989), this theory considers the institutional processes via which schools might succeed or fail in engaging students and the implications of this for students’ health.

Drawing on Bernstein, the theory suggests that schools which successfully gain students’ commitment to school (especially working class students) do so by eroding ‘boundaries’: between senior leaders and other staff (so that power is more distributed); between staff and students (so that relationships are collaborative rather than authoritarian); among students (so developing a broad sense of solidarity and community); between academic and wider learning; and between schools and their local communities. Informed by Bernstein’s work, these boundaries can best be understood as asymmetries of power or resource allocation: for example, disparities of power between staff and students and disparities of time and resource allocation between academic and wider learning. This use of the term is quite different to that used in the public administration literature where it typically refers to discontinuities between organisations and sectors with regard to social networks, information

or ideology (Bacharach, Bamberger, & McKinney, 2000; Williams, 2002). The theory of human functioning and school organisation asserts that students are more likely to commit to schools which erode these boundaries because the school then becomes more relevant to the student's own experiences, cultural values and aspirations. Informed by Nussbaum and Freire, the theory goes on to suggest that students who do commit to school are then generally more likely to avoid actions such as smoking which might damage their health.

The original articulation of the theory suggested this occurred because students who are committed to school are more likely to develop the psychological and cultural ('practical reasoning') and social ('affiliation') resources to make and implement decisions which prioritise their own health and wellbeing (W. A. Markham & Aveyard, 2003). Later papers also acknowledged that students who were uncommitted to school might also engage in risks not merely because of deficits in practical reasoning or social affiliation but through rationale choices albeit in constrained circumstances (W.A. Markham, 2015). Qualitative research including by some of the present authors has explored how students might engage in smoking and drug use as practices of cognitive escape, self-medication, rebellion or the formation of protective social relationships with peers, particularly when they feel they have few other options for expression or protection (A. Fletcher & Bonell, 2008; A. Fletcher, Bonell, & Rhodes, 2009; A. Fletcher, Bonell, Sorhaindo, & Rhodes, 2009; Jamal et al., 2013). This research recognises that smoking and drug use are best considered as practices not merely behaviours, with social meanings specific to certain times and place and which can be transformed and not merely enacted by those engaging in them; resonating with recent conceptual work published in *Critical Public Health* (Blue, Shove, Carmona, & Kelly, 2016).

Although the theory of human functioning and school organisation has been examined in, and is generally well supported by, a raft of empirical studies (Aveyard et al., 2004; Bisset et al., 2007; W.A. Markham et al., 2008; W.A. Markham et al., 2012; Tobler et al., 2011), these studies are limited by their reliance on school-level proxy measures of student commitment to school (termed 'valued added education', indicative of a school having better student attainment and attendance than would be predicted by students' social profile) and their not directly examining the school organisational processes that lie at the centre of the theory of human functioning and school organisation. We therefore sought to develop a new quantitative scale to measure unhealthy school-level boundaries, as depicted in the theory of human functioning and school organisation, between and among

staff and students, between academic learning and broader social development, and between a school and its community.

We are not proposing that our measure of unhealthy school boundaries be used as a ‘metric’ to monitor or manage school performance. Rather, the measure is intended to be used as a means for understanding the processes via which schools might influence student health behaviours. But this point notwithstanding, we believe that the new measure is quite different from the sorts of ‘performance metrics’ that are currently used to monitor and control schools - for example in the UK and USA, on which some of us have previously written very critically (Bonell, Fletcher, Sorhaindo, Wells, & McKee, 2012; A. Fletcher, Bonell, & Rhodes, 2009). Current metrics focus overwhelmingly on narrow measures of academic attainment in high-stakes tests and public examinations. These encourage many if not most schools to narrow their provision, for example: teaching to the test; neglecting non-examined subjects; neglecting students’ broader personal development and enjoyment of school life; and neglecting students’ health and wellbeing. Rather than contributing towards this neo-liberal agenda of school performance management, our research is intended to critically examine processes such as the erecting of rigid boundaries between academic learning and broader activities, and the separation of students based on academic attainment.

We sought to test the inter-item reliability of this scale and its criterion validity in terms of associations with existing measures: routine measures of value-added education; student-reported commitment to school; and student-reported smoking. We predicted that the items would strongly correlate with each other and that school-level boundaries as measured by the overall scale would be associated with reduced student-reported commitment to school and increased student-reported smoking. Our discussion then critically assesses the limitations of such work and its implications for public health interventions in schools.

Method

Development of the new measure

Our scale involves several multi-item subscales examining the unhealthy boundaries described in the theory of human functioning and school organisation (W. A. Markham & Aveyard, 2003). We decided that information

on boundaries should be collected from teachers because they are well placed to report on school organisation and because, when examining whether a measure of school boundaries correlates with student health behaviours, this avoids the problem of same-source bias (Diez-Roux, 2007). The choice of scale items was informed by our understanding of the theory of human functioning and school organisation, as well as by a recent systematic review of qualitative research on the school environment and student health (Bonell, Jamal, et al., 2013). Basing our scale on theory and evidence and experience was intended to maximise the content validity of the scale so that it might reflect all aspects of school boundaries. Where possible we drew on existing questionnaire items (ALSPAC, 2002; Day et al., 2007; Moore, Littlecott, Fletcher, Hewitt, & Murphy, 2016) but re-grouped these to reflect our organisation of sub-scales around the theorised boundaries (text box 1). Final choice of the included items was agreed by a subset of the authors (CB, WM, PA, AF) to maximise face validity.

Participants and data collection

We collected data for the unhealthy boundaries scale from staff across 40 secondary schools in south-east England (table 1). These were participating in a cluster randomised controlled trial of a restorative practice intervention to reduce bullying and aggressive behaviour. Trial recruitment targeted state secondary schools within one hour's train journey from central London judged by the national school inspectors as 'requires improvement'/'satisfactory' or higher rating. Private schools or pupil referral units (catering exclusively for students expelled from mainstream schools) or those with learning disabilities were excluded. Because of resource constraints, we drew on a small convenience sample of three teachers per school: one member of the school's senior leadership team and two other teachers, including classroom teachers, middle leaders (e.g. heads of subject area) and/or teachers coordinating pastoral care. Data on smoking and sense of belonging and commitment to academic values were derived from student self-completion questionnaires. Survey data targeted all students at the end of year 7 (age 11/12 years).

[Table 1 near here.]

Data collection from staff and students occurred at trial baseline (2014) before implementation of the intervention. We contacted the liaison teacher at each of the schools to identify and approach three staff

members to participate in the survey. Structured telephone interviews were used to collect data, with the interviewer recording answers to multiple choice questions directly on to a database. This method was chosen over postal questionnaires in order to increase the response rate. Student questionnaire surveys were also collected at trial baseline before allocation or intervention. Details of the sample size calculation, recruitment and data collection methods for the student survey are described elsewhere [reference redacted].

Data analysis

Data were analysed using Stata version 13 (StataCorp, 2013). The data were treated as ordered categorical variables in all analyses. First, we examined response rates for each item and each subscale. We also examined the distribution of response for each item.

Reliability

We then assessed inter-item scale reliability using Cronbach's alpha (Cronbach, 1951) and ordinal alpha for the overall scale and sub-scales (Gadermann, Guhn, & Zumbo, 2012). Items were recoded such that for all items a higher score indicated increased boundaries. Inter-item correlations and item-rest correlations were also used to assess the scales and the individual items. Item level inter-item correlations provide a measure of how much the average inter-item correlation would improve with the removal of an item. Item-test correlations show the correlation between the individual items and the scale as a whole and item-rest correlations show the correlation between the item and the scale if it was created without this item.

Confirmatory factor analysis was then used to test how well the data fitted two a priori models: one in which all items measured a single construct of overall school boundaries, and one in which there was evidence of five specific boundary domains as listed above (allowing for covariance between the five subscales). Goodness-of-fit statistics and factor loadings were used to assess the fit of the model, with an acceptable fit being indicated by fit indices of chi squared/degrees of freedom ($df \leq 4$), root mean square error of approximation ($RMSEA \leq 0.08$) and confirmatory fit index ($CFI \geq 0.90$), and a good fit by fit indices approximating chi squared/ $df \leq 2$, $RMSEA \leq 0.05$, $CFI \geq 0.95$ (Hooper, Coughlan, & Mullen, 2008). Standard errors were clustered to take into account the structure of the data using Mplus version 7 (Muthén & Muthén, 1998-2012) with the weighted least-square-mean-

variance-adjusted (WLSMV) estimator to estimate the model and obtain model fit statistics. If, as was the case, confirmatory factor analysis established that the data did not fit the model well, we would undertake exploratory factor analysis to reconfigure the subscales to try to produce more reliable, but still substantively meaningful, sub-scales. Visual inspection of correlation coefficients, Bartlett's tests of sphericity, and the Kaiser-Mayer-Olkin (KMO) test would be used to determine whether the data were appropriate for factor analysis (Bartlett, 1937; Kaiser, 1974). If so, scree plots (Cattell, 1966) would then be used to assess how many factors were appropriate to extract within subscales and for all items. Uniqueness values were also to be used to assess the proportion of the variance of the variable not accounted for by the factor structure. To explore the extent to which teachers within the same school were consistent in their reporting of that school we partitioned the variance in the sub-scales and the full scale into the within school variance and the between school variance.

Validity

Criterion validity of the boundaries scale was then tested by examining whether our measure of unhealthy school boundaries was associated with reduced 'value-added education' (the current 'gold standard' proxy measure of schools hypothesised as achieving high levels of student commitment) as well as student-reported measures of school commitment plus smoking, the latter being the health outcome most commonly associated with value-added education in empirical studies (Aveyard et al., 2004; Bisset et al., 2007; W.A. Markham et al., 2008; W.A. Markham et al., 2012; Tobler et al., 2011).

Assessment of value-added education drew on routine data on school-level attainment and attendance rates. Attainment rates were measured for each school by the proportion of year 11 students (aged 15–16) passing at least five General Certificate of Secondary Education (GCSE) examinations with A*–C grades (5A*–C). Attendance rates were measured per school as the proportion of all half-days students attended overall. Five-year averages (2009–2013) were used to improve reliability. Two logistic regression models were then created using school-level 5A*–C and absence rates as outcomes with the following variables as exposures: proportion of white students; proportion of females; income deprivation affecting children index (IDACI), a measure of neighbourhood deprivation indicating the proportion of children age under 16 years in a local area that live in low-income households (Department for Education, 2015); proportion of students eligible for free school meals, a measure of welfare entitlement; proportion of students who speak English as an additional language; and the

proportion of students reporting high affluence on the family affluence scale (Currie et al., 2008). Free school meals, IDACI, English as an additional language and the proportion of female students were taken from government websites. FAS and proportion of White students were taken from the student survey. The standardized residuals from these two logistic regression models represent the difference between the observed attainment and attendance rates and what would be expected based upon the student socio-demographic profile of each school. Examination of the standardised residuals produced from these two logistic regression models showed that schools achieving better than expected examination results also achieved better than expected attendance rates ($r=-0.36$). Principal components analysis identified a single factor that explained 68.1% of the variance and had factor loadings of +0.71 for 5A–C and -0.71 for absence residuals. This continuous variable was termed ‘value-added education’, and reflects both attainment and attendance rates.

Student commitment to school in terms of learning and the pastoral community was measured respectively using the 4-item ‘commitment to academic values’ and the 8-item ‘sense of belonging’ subscales of the Beyond Blue School Climate Questionnaire (BBSCQ) (Sawyer et al., 2010). This asked respondents to rate their level of agreement with statements about their school using the following responses: “YES! Totally agree”, “yes, I agree a bit”, “no, I don’t really agree” or “NO! Totally disagree!!”. Higher scores on these scales indicate a lower sense of belonging or commitment. The scale was developed in Australia (Sawyer et al., 2010) using items from the Gatehouse, (Bond, Thomas, Coffey, & Glover, 2004) Quality of School Life (Epstein & McPartland, 1976), Patterns of Adaptive Learning (Roeser, Midgley, & Urdan, 1996), Manitoba School Improvement Survey (Earl & Lee, 1998) and Psychological Sense of School Membership (Goodenow, 1993) questionnaires. Cronbach’s alphas for the belonging and academic commitment sub-scales of 0.85 and 0.82 were reported for a sample of similar age (personal communication Lyndal Bond 21 July 2011). Smoking was measured using a single self-report question asking students whether they have ever having tried smoking, previously used in the Ripple study (Stephenson et al., 2008).

The correlation between our new measure of unhealthy boundaries and ‘value-added education’ was assessed at the school level. School-level average reports of boundaries were created by collapsing responses within schools. Correlations with student-reported school commitment and smoking were assessed using multi-level models, adjusting for student sex, ethnicity (categorised based on student reports as White British or Irish, Asian or Asian British, Black or Black British, Chinese or Chinese British, Mixed ethnicity or other), religion

(categorised based on student reports as none, Christian, Jewish, Muslim, Hindu, Sikh, don't know/not sure or other), family structure (based upon student report of the adults they live with dichotomised into single parent/two parent households), parental work (based upon student report of whether any adults in the household were in paid work or not), housing tenure (based on student responses to whether their house or flat was rented from the Council/housing association, rented from a landlord, owned by their family, other, or whether they didn't know), and student responses to the family affluence scale (Currie et al., 2008). The scores on the boundaries scale were standardised to facilitate interpretation. Finally, analysis examined whether adjustment for measures of school commitment reduced the association between the new scale and student smoking to establish whether the latter association might be confounded (or more likely mediated) by student commitment.

Results

In total, 101 staff in 38 schools responded (85% response rate). Of these, 35 were senior leaders, 51 were middle leaders, 14 were classroom teachers and one was a pastoral support worker. Response rates to all items were very good. Of eligible students, 6667 (93.6%) completed questionnaires.

Reliability

The staff/student boundaries scale, the academic/broader student development scale and the school/community boundaries scale had acceptable but not good alpha values (table 2). The alpha values for the student/student boundaries scale were unacceptably low. Confirmatory factor analysis suggested that the model which specified all five subscales allowing for covariance between subscales ($\chi^2(619)=771.42$, RMSEA=0.05(0.04 – 0.06), CFI=0.84), resulted in better fit than the model where all items loaded onto a single factor ($\chi^2(629)=1113.47$, RMSEA=0.09(0.08 – 0.10), CFI=0.52). Further evidence on model fit was obtained from the factor loadings and standard errors (table 3). The factor loadings were very low for several of the items (C4, D3, D4, E1, E2, E4 and F7), with particularly low loadings for several of the items on the student/student boundaries scale. The individual subscales were then tested (table 4). The academic/broader student development scale showed close to acceptable fit and the school/community boundaries scale showed good fit. None of the other subscales had acceptable values on the fit indices.

[Table 2 near here.]

[Table 3 near here.]

[Table 4 near here.]

Exploratory factor analysis was undertaken on the polychoric correlation matrix (web appendix 1). The average inter-item correlation was 0.17. For the subscales, KMO values ranged from 0.53 for the student/student boundaries up to 0.75 for the school/community boundaries subscale, with a value of 0.69 for all items. Bartlett's test of sphericity was statistically significant within each subscale and across all items. However, given the combination of small sample size with the extent of low correlations in the correlation matrix (even within subscales), the high uniqueness values, and the reasonably low KMO values (especially for the student/student boundaries subscale), these data may not be appropriate for exploratory factor analysis (MacCallum, Widaman, Zhang, & Hong, 1999).

The boundaries scale was modified in the following way. Low loading items were removed from subscales to improve alpha values and increase inter-item correlations. Removing item C4 from the staff/student boundaries subscale, and item F7 from the academic/broader student development boundaries scale improved alpha values to 0.73 and 0.70 respectively. Item G4 had low discriminatory power and was removed from the school/community boundaries scale, increasing alpha to 0.74. The average inter-item correlation in the student/student boundaries subscale was just 0.11. The scree plot for the student/student boundaries subscale indicated that each additional factor extracted explained a similar amount of variance. One-, two- and three-factor solutions did not produce consistent or meaningful results within this subscale. Hence, this subscale along with items C4, F7 and G4 were removed.

The variance in the subscales and full scales was partitioned into the within- and between-school variance (web appendix 2). Across all subscales there was a similar amount of variability within and between schools. Teachers within schools varied in their responses as much as teachers from different schools.

Validity

Between schools, VAE ranged from -2.83 to 2.25 (mean=0, SD=1). The correlations between school value-added education and our unhealthy school boundaries scale using all original items ($r=-0.17$) and the modified total boundaries scale ($r=-0.15$) were low, as were all subscales (staff/staff $r=0.07$; staff/student $r=-0.01$; student/student $r=-0.12$; academic/broader student development $r=-0.25$; school/community boundaries $r=-0.14$). However, the unhealthy school boundaries scale correlated strongly with reduced student-reported school commitment and increased student-reported smoking. Overall 5.35% of students had ever tried smoking. Between students, belonging ranged from 0 to 24 (mean=8.07, SD=4.40), and academic commitment ranged from 0 to 12 (mean=1.45, SD=1.65) with higher scores equating to a lower belonging and academic commitment. A standard deviation increase in the unhealthy school boundaries scale that included all original items was associated with increased odds of students ever having smoked (OR=1.21, SE=0.11), a lower student commitment to academic values ($\beta=0.08$, SE=0.03) and a lower student sense of belonging ($\beta=0.49$, SE=0.13). The modified total boundaries scale showed similar associations with smoking (OR=1.18, SE=0.11), commitment to academic values ($\beta=0.08$, SE=0.03) and sense of belonging ($\beta=0.45$, SE=0.13). Adjusting for both measures of student commitment reduced the association of the original all-item and modified scale with smoking only very marginally (OR=1.18; SE 0.10 and OR=1.15; SE 0.10 respectively).

Discussion

Summary of key findings

Overall, our new measure of unhealthy school boundaries had sub-optimal inter-item reliability. The staff/student boundaries scale, the academic/broader student development scale and the school/community boundaries scale had acceptable but not good internal reliability while that for the student/student boundaries scale was unacceptably low. The scale performed substantially better when the subscales were included in the specification, rather than when all items were presumed to measure a single construct. There was little agreement in responses between different teachers in the same schools which further suggests the measure had limited reliability, though these analyses drew on data from a maximum of only three individuals per school so that normal variation might explain this finding.

In terms of criterion validity, there was no correlation between value-added education and our unhealthy school boundaries scale but significant adjusted associations between the new scale and reduced student-reported commitment to academic values and sense of belonging, and increased student-reported smoking. The lack of effect on the association with smoking suggest that our new scale measures something distinctive about schools not captured by the existing measures of student commitment, and that any association between our new scale and smoking is not mediated by measures of students' commitment to school.

Limitations

In terms of methodological limitations, we did not undertake new qualitative research to inform wording of our items, nor cognitive tests to assess teachers' understanding of the questions. Our research also relied on convenience samples of only three staff-members per school. Our assessment of criterion validity relied on cross-sectional data and so could not examine temporality. Our research occurred in a sample of 40 schools in south-east England, which although not manifestly atypical of other schools in England, may not be representative of other settings.

More fundamentally, our measure of unhealthy school boundaries is static and so cannot hope to represent the socially constructed and processual nature of school organisation and young people's social practices. Furthermore, the study employs measures of student commitment to school and smoking which focus on behaviours rather than on social practices; we recognise these are inattentive to meaning, purpose or context. But despite these inevitable limitations, our belief in undertaking this study is that it should be a useful contribution to developing a quantitative 'summary snapshot' of the extent to which different schools appear to be eroding disparities of power and resource allocation, and to examine whether such a measure is associated with measures of student commitment to school and engagement in smoking. The study cannot explain why young people themselves feel committed to school or not, or decide to smoke or not, or how this varies with context. Nonetheless, the study's finding that there are associations between our measures of eroded school boundaries (despite the limited reliability of this measure) and reduced student smoking is an interesting one which might encourage further quantitative research to see if this applies to other settings, as well as further qualitative research to better understand questions of process, meaning and purpose.

Implications for policy and research

Our measure is the first that has attempted to assess the extent of unhealthy boundaries in schools between and among staff and students, between academic and broader learning and between schools and their local communities. Despite the limited internal reliability of our measure, our findings do provide some evidence that school-level boundaries may influence rates of smoking even after adjusting for student characteristics. It thus provides the first direct evidence that schools might be able to promote student health by becoming more participative and eroding boundaries within and around the school. However, further work is required to optimise our measure's reliability. We recommend further qualitative research to inform the wording of new items before repeating tests for internal reliability and criterion validity and assessing the measure in terms of test-retest reliability.

What then are the implications of our analysis for health interventions in schools? There has been much critique of health promotion in schools. Such interventions have been criticised both in terms of their lack of impact and their coercive potential (Gard & Wright, 2014). Some have described such interventions as a form of fascism (Fitzpatrick & Tinning, 2014) and some as a component of governmentality, whereby schools contribute to the regulation of bodies and behaviours (Galitz & Robert, 2014). The theory of human functioning might be viewed in different ways. It aims to move thinking about schools and health away from health education lessons and even away from ensuring that aspects of the school environments such as canteens and smoking policies make the healthy choice the easy (or even only) choice. It aims to refocus attention on the health implications of how the school is run and teaches its students: encouraging students' active participation and commitment. Some might see in this an ever more pervasive system of health control within schools so that this now encompass all aspects of their work (Gard & Wright, 2014; Leahy, 2014). But others might see in these developments a more benign recognition of the practical and ethical limits to trying to direct students to healthy behaviours either via health education or environmental nudges. These developments might imply a recognition that schools can best promote young people's health by helping them develop real autonomy and trying to ensure students are not so alienated, marginalised or unsafe within schools that they choose to engage in health-harming actions such as smoking as forms of self-medication, cognitive escape, rebellion or protection because alternative options are so scarce. However, it is also important to recognise the evidence that school effects on health outcomes appear to

be small and not create unrealistic expectations that schools alone might remedy the much broader societal determinants of mental and physical ill health.

Clearly, further quantitative and qualitative research is needed to explore the theory of human functioning and school organisation. Quantitative research should examine the extent to which the theory seems to predict rates of health behaviours in different schools in different settings. The recognition that smoking and drug use etc. are best regarded as socially located and meaningful social practices, with meanings specific to certain times and place, and not merely behaviours (Blue et al., 2016), means they may be implicated in student responses to school very differently across time and place. Thus, in addition to quantitative research examining how well the theory appears to predict patterns of such practices in different settings, qualitative research is also needed to provide insights into questions of process and meanings.

Theoretical refinement is also required. The theory of human functioning and school organisation currently treats all boundaries within schools as unhealthy. This is useful in shifting the focus of research on young people's health from individual deficits in knowledge or attitudes, etc. towards deficits in school organisational priorities (and by implication neo-liberal government priorities). However, it may be that certain boundaries within schools actually help promote student commitment. For example, different students choosing different subjects to study in the later years of their secondary education to reflect their interests and aspirations or a school providing students with a critical and in-depth understanding of particular academic subjects might respectively represent healthy inter-student and inter-subject boundaries. The key to distinguishing this probably lies in the extent to which they promote or suppress student autonomy, but more reflection is required.

Acknowledgements

We would like to thank the other co-investigators working on the INCLUSIVE trial as well as the students and staff of the participating schools.

References

- ALSPAC. (2002). Questionnaire for Head teacher <http://www.bristol.ac.uk/alspac/researchers/resources-available/data-details/questionnaires/documents/ques-s07-questionnaire-for-the-head-teacher.pdf>.
- Aveyard, P., Markham, W. A., Lancashire, E., Bullock, A., MacArthur, C., & Cheng, K. K. (2004). The influence of school culture on smoking among pupils. *Social Science & Medicine*, 58(9), 1767-1870.
- Bacharach, S. B., Bamberger, P., & McKinney, V. (2000). Boundary management tactics and logics in action: the case of peer-support providers. *Administrative Science Quarterly*, 45(4), 704-736.
- Bartlett, M. S. (1937). The statistical conception of mental factors. *British Journal of Psychology*, 28, 97-104.
- Bernstein, B. (1975). *Class, codes and control, vol. 3: towards a theory of educational transmission*. London: Routledge.
- Bisset, S., Markham, W. A., & Aveyard, P. (2007). School culture as an influencing factor on youth substance use. *Journal of Epidemiology and Community Health*, 61(6), 485-490.
- Blue, S., Shove, E., Carmona, C., & Kelly, M. P. (2016). Theories of practice and public health: understanding (un)healthy practices. *Critical Public Health*, 26(1), 36-50.
- Bond, L., Thomas, L., Coffey, C., & Glover, S. (2004). Long-term impact of the gatehouse project on cannabis use of 16-year-olds in Australia. *Journal of School Health*, 74(1), 23-30.
- Bonell, C., Fletcher, A., Jamal, F., Wells, H., Harden, A., Murphy, S., & Thomas, J. (2013). Theories of how the school environment impacts on student health: systematic review. *Health & Place*, 24, 242-249.
- Bonell, C., Fletcher, A., Sorhaindo, A., Wells, H., & McKee, M. (2012). How market-oriented education policies might influence young people's health: development of a logic model from qualitative case studies in English secondary schools. *Journal of Epidemiology and Community Health*, 66(7), e24.
- Bonell, C., Jamal, F., Harden, A., Wells, H., Parry, W., Fletcher, A., . . . Moore, L. (2013). Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis. *Public Health Research*, 1(1).
- Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behav Res*, 1(2), 245-276. doi: 10.1207/s15327906mbr0102_10
- Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16(3), 297-334.
- Currie, C., Molcho, M., Boyce, W., Holstein, B., Torsheim, T., & Richter, M. (2008). Researching health inequalities in adolescents: the development of the Health Behaviour in School-Aged Children (HBSC) Family Affluence Scale. *Social Science & Medicine* (), 66, 1429-1436.
- Day, C., Sammons, P., Hopkins, D., Harris, A., Leithwood, K., Gu, Q., . . . Kington, A. (2007). *The Impact of School Leadership on Pupil Outcomes Interim Report*. London: Department for Education.
- DECIPHER. (2014). Schools Health Research Network <http://man301110a.decipher.uk.net/en/content/cms/research/research-projects/shrn/>.
- Department for Education. (2015). IDACI score and rank of IDACI from <http://www.education.gov.uk/cgi-bin/inyourarea/idaci.pl>
- Diez-Roux, A. V. (2007). Environnement résidentiel et santé: état de la question et perspectives pour le futur. *Rev Epidemiol Sante Publique*, 55(1), 13-21.
- Earl, L. M., & Lee, L. E. (1998). *Evaluation of the Manitoba School Improvement Program*. Toronto: University of Toronto/Proactive Information Services Inc.
- Epstein, J. L., & McPartland, J. M. (1976). The concept and measurement of the quality of school life. *American Educational Research Journal*, 13(1), 15-30.
- Fitzpatrick, K., & Tinning, R. (2014). Health education's fascist tendencies: a cautionary exposition. *Critical Public Health*, 24(2), 132-142.
- Fletcher, A., & Bonell, C. (2008). "Escaping from prison": Female students' substance use and the school environment. *Journal of Epidemiology and Community Health*, 62 (Suppl. 1), A10.
- Fletcher, A., Bonell, C., & Rhodes, T. (2009). New counter-school cultures: female students' drug use at a high-achieving secondary school. *British Journal of Sociology of Education*, 30(5), 549-562.
- Fletcher, A., Bonell, C., Sorhaindo, A., & Rhodes, T. (2009). Cannabis use and 'safe' identities in an inner-city school risk environment. *International Journal of Drug Policy*, 20(3), 244-250.
- Freire, P. (1989). *Pedagogy of the oppressed*. Trowbridge: Sheed and Ward.
- Gademmann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, 17(3), 1-13.

- Galitz, T., & Robert, D. (2014). Governing bullying through the new public health model: a Foucaultian analysis of a school anti-bullying programme. *Critical Public Health*, 24(2), 182-195.
- Gard, M., & Wright, J. (2014). Schools and critical public health: towards dialogue, collaboration and action. *Critical Public Health*, 24(2), 109-114.
- Goodenow, C. (1993). Classroom belonging among early adolescent students relationships to motivation and achievement. *The Journal of Early Adolescence*, 13(1), 21-43.
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.
- Jamal, F., Fletcher, A., Harden, A., Wells, H., Thomas, J., & Bonell, C. (2013). The school environment and student health: a systematic review and meta-ethnography of qualitative research. *BMC Public Health*, 13, 798.
- Kaiser, H. F. (1974). An index of factor simplicity. *Psychometrika*, 39, 31-36.
- Leahy, D. (2014). Assembling a health(y) subject: risky and shameful pedagogies in health education. *Critical Public Health*, 24(2), 171-181.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychol Methods*, 4(1), 84.
- Markham, W. A. (2015). School culture and teenage substance use: a conceptual and operational framework. *Educational Review*, 67(3), 282-299.
- Markham, W. A., & Aveyard, P. (2003). A new theory of health promoting schools based on human functioning, school organisation and pedagogic practice. *Social Science & Medicine*, 56(6), 1209-1220.
- Markham, W. A., Aveyard, P., Bisset, S. L., Lancashire, E. R., Bridle, C., & Deakin, S. (2008). Value-added education and smoking uptake in schools: a cohort study. *Addiction*, 103(1), 155-161.
- Markham, W. A., Young, R., Sweeting, H., West, P., & Aveyard, P. (2012). Does school ethos explain the relationship between value-added education and teenage substance use? A cohort study. *Social Science and Medicine*, 75(1), 69-76.
- Marmot, M. G. (2010). *Fair Society, Healthy Lives. The Marmot Review. Strategic Review of Health Inequalities in England post-2010*. London: The Marmot Review.
- Moore, G., Littlecott, H., Fletcher, A., Hewitt, G., & Murphy, S. (2016). Variations in schools' organisational commitment to student health and implementation of health improvement activities: a cross-sectional survey of secondary schools in Wales. *BMC Public Health*, 16(138).
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.
- Nussbaum, M. C. (1990). Aristotelian social democracy. In R. B. Douglas, G. M. Mara & H. Richardson (Eds.), *Liberalism and the good* (pp. 203-251). London: Routledge.
- Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology*, 88(3), 408.
- Sawyer, M. G., Pfeiffer, S., Spence, S. H., Bond, L., Graetz, G., Kay, D., . . . Sheffield, J. (2010). School-based prevention of depression: a randomised controlled study of the beyondblue schools research initiative. *Journal of Child Psychology and Psychiatry*, 51(2), 199-209.
- StataCorp. (2013). *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
- Stephenson, J., Strange, V., Allen, E., Copas, A., Johnson, A., Bonell, C., . . . RIPPLE Study Team. (2008). The long-term effects of a peer-led sex education programme (RIPPLE): a cluster randomised trial in schools in England. *PLoS Med*, 5(11), e224.
- Tobler, A. L., Komro, K. A., Dabroski, A., Aveyard, P., & Markham, W. A. (2011). Preventing the link between SES and high-risk behaviors: "Value-added" education, drug use and delinquency in high-risk, urban schools. *Prevention Science*, 12(2), 211-221.
- Williams, P. (2002). The competent boundary spanner. *Public Administration*, 80(1), 103-124.

Text box 1

SETA: The first set of questions is concerned with staff authority and responsibility in your school (Hierarchical or distributed staff authority).

- A1. The head teacher takes most of the decisions with little staff consultation³
- A2. Teachers participate on a regular basis in the development of school policies³
- A3. The senior leadership team consult with staff when making decisions⁶
- A4. Teachers in this school have a sense of collective responsibility for student learning¹
- A5. Teachers in this school have a sense of collective responsibility for student wellbeing²
- A6. Teachers and other staff in the classroom work collaboratively¹

SETB: The next set of questions is concerned with teacher-student collaboration at your school (Teacher authority or teacher-student collaboration).

- B1. In my school students participate in decision making²
- B2. Teachers in this school always show respect towards students⁴
- B3. Students' views are listened to and taken seriously by staff in this school³
- B4. Teaching strategies at this school enable students to build their own knowledge²
- B5. There are opportunities for students to take responsibilities for their own learning in school²
- B6. In this school the senior leadership team makes decisions without consulting students⁶

SETC: The next set of questions is concerned with teachers support for students in your school (Teacher support for students across school or restriction to classroom)

- C1. Teachers at this school are often involved in extracurricular activities⁴
- C2. In my school teachers mix with students at break times⁶
- C3. In my school teachers mix with students at lunch time⁶
- C4. In my school, a lot of student pastoral care is delegated to non teachers⁶
- C5. In my school, teachers avoid intervening in students disputes outside the classroom⁶

SETD: The next set of questions tries to get a sense of how student learning is organised (Dividing up or bringing together students (learning)).

- D1. Teachers at this school are more interested in the students with potential to do well in tests and examinations⁴
- D2. The school has a system for rewarding students who work hard and/or make good progress even if they do not reach high standards³
- D3. Students of similar academic ability are grouped together for teaching in most subject areas¹
- D4. This school targets resources on the students on the borderline of achieving 5 good GCSEs⁶

SETE: The next set of questions tries to get a sense of the nature of discipline and pastoral care practices in your school (Dividing up or bringing together students (discipline and pastoral))

- E1. My school mixes together students who are of different ages e.g. through tutor groups or extracurricular activities⁶
- E2. Certain students in my school are repeatedly isolated from other students in response to misbehaviour⁶
- E3. My school has a strong system of peer mentoring or peer buddying⁶
- E4. My school runs conflict resolution programmes for students⁶

SETF: The next set of questions is concerned with activities, practices or policies around student development (Focus on academic or broader development).

- F1. The school has a system for rewarding students who achieve in non academic areas e.g. sport, arts⁴
- F2. Our school provides a broad range of extracurricular activities for students (e.g. plays, athletics, music, dance)¹
- F3. The school development/improvement plan has targets related to student health and wellbeing
- F4. School INSET/training days often focus on student health⁵
- F5. The school has a comprehensive written policy to address student smoking, drugs or alcohol use⁵
- F6. The school teaches a social and emotional learning curriculum⁵
- F7. My school offers a range of non traditional subjects for students in years 10 and 11⁶

SETG: The final set of questions is concerned with the extent to which your school is linked with the local community (Links to or separation from the local community).

G1. Parents often visit the school¹

G2. This school engages parents in school improvement efforts²

G3. This school aims to build community support for the school's improvement efforts²

G4. Parents are regularly informed about the progress and achievements at school of their child

G5. Parents give a lot of support to the work of the school³

1 Agree strongly

2 Agree moderately

3 Agree slightly

4 Disagree slightly

5 Disagree moderately

6 Disagree strongly

Sources

1. Taken from The Impact of School Leadership on Pupil Outcomes Key Staff Questionnaire – Secondary(Day et al., 2007)
2. Adapted from The Impact of School Leadership on Pupil Outcomes Key Staff Questionnaire - Secondary(Day et al., 2007)
3. Taken from ALSPAC heads questionnaire(ALSPAC, 2002)
4. Adapted from ALSPAC heads questionnaire(ALSPAC, 2002)
5. Adapted from SHRN school questionnaire(DECIPHer, 2014)
6. New question

Table 1: Characteristics of schools

<i>Characteristic</i>	<i>Schools</i>	
<u><i>Inspection rating</i></u>	<i>Not yet inspected</i>	1
	<i>Requires improvement</i>	4
	<i>Good</i>	24
	<i>Outstanding</i>	11
<u><i>Type</i></u>	<i>Voluntary aided</i>	4
	<i>Foundation</i>	6
	<i>Academy</i>	25
	<i>Community</i>	5
<u><i>Sex</i></u>	<i>Boys</i>	3
	<i>Girls</i>	7
	<i>Mixed</i>	30
<u><i>Free school meals</i></u>	<i>0-20</i>	12
	<i>21-40</i>	11
	<i>41-60</i>	12
	<i>61-80</i>	5
<u><i>Attainment (best 8 GCSEs value-added)*</i></u>	<i>> 1000</i>	24
	<i><1000</i>	16

* This is a measure of the progress students make from entry to performance in GCSE exams with 1000 being the national median.

Table 2. Measures of scale response and reliability

<i>Boundaries</i>	<i>Response rates</i>		<i>Internal consistency</i>	
	<i>Completed all items (n)</i>	<i>Completed half of items (n)</i>	<i>Alpha (standardised)</i>	<i>Ordinal alpha</i>
Staff/staff boundaries	100	101	0.63	0.73
Staff/student boundaries	100	101	0.72	0.79
Student/student boundaries	98	101	0.51	0.58
Academic/broader student development boundaries	99	101	0.69	0.78
School/community boundaries	97	101	0.73	0.82
total boundaries (sum of subscales)	93	101	0.82	0.82

Table 3: Factor loading estimates (standardised) from confirmatory factor analysis testing the theoretical measurement models

Items	Five factors specified ¹			One factor specified		
Staff/Staff boundaries						
	<i>Estimate</i>	<i>S.E</i>	<i>uniqueness</i>	<i>Estimate</i>	<i>S.E</i>	<i>uniqueness</i>
A1	0.39	0.09	0.85	0.32	0.10	0.90
A2	0.59	0.10	0.65	0.53	0.12	0.72
A3	0.64	0.08	0.59	0.61	0.09	0.63
A4	0.63	0.07	0.61	0.49	0.07	0.76
A5	0.56	0.08	0.68	0.41	0.10	0.84
A6	0.50	0.09	0.75	0.42	0.12	0.82
Staff/student boundaries						
B1	0.62	0.08	0.62	0.49	0.10	0.76
B2	0.55	0.08	0.70	0.46	0.10	0.79
B3	0.69	0.07	0.53	0.57	0.09	0.68
B4	0.65	0.08	0.58	0.52	0.08	0.73
B5	0.55	0.10	0.70	0.43	0.09	0.82
B6	0.63	0.07	0.60	0.52	0.09	0.73
C1	0.43	0.09	0.82	0.31	0.09	0.90
C2	0.72	0.06	0.49	0.41	0.10	0.84
C3	0.78	0.05	0.39	0.48	0.08	0.77
C4	0.10	0.10	0.99	0.09	0.12	0.99
C5	0.46	0.09	0.79	0.40	0.09	0.84
Student/student boundaries						
D1	0.54	0.08	0.71	0.39	0.11	0.85
D2	0.57	0.06	0.68	0.47	0.08	0.78
D3	0.16	0.12	0.97	0.10	0.14	0.99
D4	0.07	0.08	1.00	0.03	0.08	1.00
E1	0.22	0.12	0.95	0.18	0.13	0.97
E2	0.27	0.10	0.93	0.21	0.12	0.96
E3	0.66	0.07	0.57	0.53	0.08	0.72
E4	0.37	0.10	0.86	0.31	0.10	0.90
Academic/broader student development						
F1	0.80	0.08	0.36	0.57	0.07	0.67
F2	0.64	0.09	0.60	0.41	0.10	0.83
F3	0.65	0.07	0.58	0.46	0.08	0.79
F4	0.66	0.09	0.57	0.51	0.09	0.75
F5	0.43	0.11	0.81	0.27	0.10	0.93
F6	0.57	0.11	0.67	0.38	0.12	0.86
F7	0.24	0.09	0.94	0.15	0.09	0.98
School/community boundaries						
G1	0.79	0.06	0.38	0.61	0.08	0.63
G2	0.70	0.06	0.51	0.62	0.08	0.62
G3	0.71	0.08	0.50	0.58	0.08	0.67
G4	0.69	0.08	0.52	0.42	0.08	0.83
G5	0.59	0.08	0.65	0.50	0.09	0.75
Model fit						
	$(\chi^2(619)=771.42, RMSEA=0.05(0.04 -$			$(\chi^2(629)=1113.47, RMSEA=0.09(0.08 -$		

¹ Allowing for covariance between all five factors.

Table 4. Factor loadings (standardised) and model fit for the subscales tested individually

<i>Item</i>	<i>Staff/staff boundaries</i>	<i>Staff/student boundaries</i>	<i>Student/student boundaries</i>	<i>Student development boundaries</i>	<i>School/community boundaries</i>	<i>Total boundaries (summed subscales)</i>
A1	0.52					
A2	0.59					
A3	0.60					
A4	0.60					
A5	0.61					
A6	0.45					
Subscale test: $\chi^2=51.880$ (9df), RMSEA=0.22 (0.16 - 0.28), CFI=0.64						0.75
B1		0.61				
B2		0.41				
B3		0.57				
B4		0.36				
B5		0.38				
B6		0.65				
C1		0.35				
C2		0.85				
C3		0.85				
C4		0.07				
C5		0.45				
Subscale test: $\chi^2=140.366$ (44df), RMSEA=0.15 (0.12 - 0.18), CFI=0.80						0.79
D1			0.62			
D2			0.61			
D3			0.29			
D4			0.15			
E1			0.25			
E2			0.32			
E3			0.50			
E4			0.30			
Subscale test: $\chi^2=66.302$(20df), RMSEA=0.15 (0.11-0.19), CFI=0.42						0.56
F1				0.68		
F2				0.59		
F3				0.62		
F4				0.70		
F5				0.58		
F6				0.59		
F7				0.34		
Subscale test: $\chi^2=25.746$ (14df), RMSEA=0.09 (0.03 - 0.15), CFI=0.922						0.60
G1					0.85	
G2					0.68	
G3					0.72	
G4					0.64	
G5					0.58	
Subscale test: $\chi^2=8.617$(5df), RMSEA=0.09 (0.00-0.18), CFI=0.984						0.72

Web appendix 1: Correlation coefficients for the items

	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2	C3	C4	C5	D1	D2	D3	D4	E1	E2	E3	E4	F1	F2	F3	F4	F5	F6	F7	G1	G2	G3	G4	G5	
A1	100																																					
A2	0.53	100																																				
A3	0.41	0.57	100																																			
A4	-0.04	0.23	0.30	100																																		
A5	0.17	0.16	0.15	0.60	100																																	
A6	0.07	0.13	0.26	0.41	0.41	100																																
B1	0.39	0.29	0.29	0.18	0.18	0.22	100																															
B2	0.24	0.28	0.39	0.59	0.59	0.30	0.20	100																														
B3	0.23	0.27	0.36	0.55	0.56	0.30	0.57	0.57	100																													
B4	0.21	0.43	0.48	0.76	0.29	0.43	0.16	0.38	0.45	100																												
B5	0.24	0.37	0.19	0.48	0.09	0.25	0.28	0.08	0.26	0.57	100																											
B6	0.42	0.47	0.53	0.04	0.11	0.24	0.62	0.20	0.46	0.20	0.19	100																										
C1	-0.02	0.21	0.31	0.38	0.23	0.01	0.09	0.16	0.24	0.35	0.25	-0.02	100																									
C2	0.06	0.18	0.26	0.24	0.15	0.20	0.26	0.17	0.17	-0.01	0.10	0.38	0.17	100																								
C3	0.07	0.27	0.36	0.30	0.27	0.25	0.16	0.22	0.18	0.13	0.19	0.21	0.32	0.80	100																							
C4	0.14	0.18	0.08	0.08	0.03	-0.22	0.15	0.08	0.13	0.00	0.20	0.22	0.06	-0.17	-0.19	100																						
C5	-0.03	0.14	0.13	0.37	0.36	0.33	0.23	0.34	0.26	0.22	0.34	0.24	0.35	0.26	0.29	0.28	100																					
D1	0.28	0.14	0.30	0.25	0.21	0.25	0.17	0.40	0.41	0.28	0.41	0.16	0.04	-0.02	0.03	0.08	0.27	100																				
D2	0.29	0.24	0.35	0.31	0.07	0.23	0.26	0.12	0.20	0.41	0.32	0.23	0.19	0.18	0.10	0.08	0.27	0.39	100																			
D3	-0.02	0.01	0.24	0.13	-0.03	0.14	0.17	0.09	0.21	-0.04	0.20	0.10	0.00	-0.03	-0.04	0.04	-0.14	0.33	0.01	100																		
D4	0.10	0.01	-0.07	0.08	0.29	0.16	0.19	0.13	0.04	-0.03	0.28	-0.01	-0.09	-0.05	0.01	0.21	0.00	0.29	-0.32	0.31	100																	
E1	0.01	0.06	0.07	0.12	-0.09	0.12	0.37	-0.02	0.19	0.01	0.18	0.26	0.07	0.22	0.16	-0.02	0.07	0.00	0.12	-0.04	0.08	100																
E2	0.20	0.08	0.18	0.02	0.08	0.10	0.39	0.24	0.35	-0.05	0.11	0.37	-0.02	0.06	-0.01	0.19	-0.04	0.27	0.15	0.19	0.17	0.21	100															
E3	0.13	0.39	0.42	0.35	0.31	0.42	0.41	0.35	0.32	0.31	0.42	0.28	0.24	0.25	0.26	0.12	0.35	0.26	0.26	0.12	0.10	0.32	0.23	100														
E4	0.08	0.10	0.22	0.21	0.24	0.32	0.07	0.18	0.18	0.24	0.07	0.32	0.04	0.22	0.15	-0.03	0.11	0.09	0.47	-0.02	-0.19	0.02	-0.17	0.16	100													
F1	0.32	0.36	0.47	0.15	0.28	0.39	0.32	0.19	0.26	0.32	0.34	0.37	0.36	0.24	0.39	0.12	0.39	0.46	0.58	0.09	-0.04	-0.01	-0.01	0.34	0.27	100												
F2	0.09	0.47	0.39	0.34	0.30	0.12	0.07	0.26	0.16	0.38	0.37	0.14	0.56	0.18	0.26	-0.05	0.32	0.51	0.32	0.03	0.03	0.24	0.08	0.37	0.10	0.46	100											
F3	0.11	0.35	0.32	0.01	0.17	0.31	0.35	0.23	0.39	0.18	0.26	0.48	-0.01	0.20	0.23	0.08	0.16	0.56	0.30	0.11	-0.02	0.00	0.38	0.27	0.23	0.44	0.37	100										
F4	0.24	0.43	0.40	0.24	0.43	0.29	0.22	0.38	0.52	0.39	0.27	0.39	0.31	0.05	0.27	0.12	0.12	0.31	0.20	0.21	0.14	0.11	0.14	0.34	0.15	0.48	0.37	0.47	100									
F5	-0.08	0.26	0.28	0.12	0.13	0.26	0.13	0.00	0.18	0.17	0.26	0.24	0.13	-0.03	0.00	0.11	0.20	0.18	0.20	0.14	0.00	-0.09	-0.03	0.07	0.22	0.37	0.16	0.43	0.39	100								
F6	-0.12	0.12	0.22	0.30	0.42	0.26	0.26	0.02	0.34	0.18	0.22	0.21	0.16	0.16	0.21	-0.04	0.07	0.08	0.07	0.20	0.12	0.04	0.10	0.40	0.24	0.36	0.38	0.39	0.27	0.47	100							
F7	0.11	0.28	0.16	0.06	-0.03	0.11	0.04	-0.01	-0.03	0.07	0.01	-0.05	0.03	-0.04	0.01	0.02	-0.02	0.13	0.20	0.05	0.02	0.30	-0.03	0.30	0.04	0.24	0.42	-0.07	0.34	0.07	0.12	0.10	100					
G1	0.20	0.47	0.40	0.55	0.33	0.31	0.37	0.40	0.24	0.50	0.35	0.40	0.27	0.41	0.56	-0.02	0.40	0.23	0.45	0.03	-0.08	0.08	-0.01	0.28	0.25	0.51	0.38	0.25	0.16	0.25	0.16	0.07	100					
G2	0.21	0.46	0.46	0.49	0.31	0.34	0.31	0.34	0.34	0.42	0.37	0.23	0.27	0.16	0.32	0.42	-0.08	0.11	0.16	0.29	-0.04	-0.07	0.09	0.18	0.48	0.09	0.40	0.11	0.31	0.26	0.35	0.49	0.15	0.50	100			
G3	0.23	0.33	0.43	0.37	0.40	0.24	0.35	0.21	0.46	0.23	0.26	0.39	0.24	0.51	0.51	-0.19	0.19	0.03	0.29	-0.04	-0.32	0.19	0.00	0.29	0.26	0.45	0.20	0.38	0.23	0.12	0.43	0.00	0.62	0.56	100			
G4	0.33	0.55	0.53	0.07	0.28	0.23	0.53	-0.08	0.36	0.32	0.20	0.51	0.20	0.35	0.27	0.22	0.16	0.14	0.51	0.00	-0.10	-0.08	0.21	0.27	0.40	0.55	0.42	0.41	0.22	0.33	0.45	0.04	0.58	0.39	0.50	100		
G5	0.13	0.16	0.22	0.43	0.38	0.25	0.42	0.33	0.36	0.53	0.43	0.18	0.22	0.15	0.31	0.26	0.35	0.22	0.17	-0.05	0.23	-0.01	0.29	0.45	0.01	0.28	0.09	0.21	0.18	0.14	0.26	0.02	0.53	0.47	0.27	0.38	100	

Web appendix 2: Partitioning variance within and between schools

	Mean	Std. Dev.	Min	Max	Observations
<i>Staff/staff boundaries</i>					
overall	2.22	0.58	1.35	3.79	N = 100
between		0.34	1.53	2.90	n = 38
within		0.47	0.91	3.47	T-bar = 2.63
<i>Staff/student boundaries</i>					
overall	3.02	0.70	1.67	4.89	N = 100
between		0.47	2.03	4.16	n = 38
within		0.52	1.57	4.12	T-bar = 2.63
<i>Student development boundaries</i>					
overall	2.33	0.67	1.26	4.66	N = 100
between		0.54	1.40	4.66	n = 38
within		0.50	1.24	3.57	T-bar = 2.63
<i>School/community boundaries</i>					
overall	2.14	0.70	1.11	3.89	N = 97
between		0.54	1.26	3.56	n = 38
within		0.48	1.16	3.73	T-bar = 2.55
<i>Total boundaries</i>					
overall	0.00	0.91	-1.68	2.48	N = 94
between		0.60	-1.40	1.20	n = 38
within		0.67	-1.74	1.88	T-bar = 2.47

The variance in the boundaries is partitioned into between and within school variance using the *xtsum* command in Stata version 12. The overall row shows the global mean and standard deviation with the total range in scores across all teachers in all schools. For the staff/staff boundaries the mean was 2.22, the standard deviation was 0.58, and the range was 1.35-3.79. The between row refers to the variability of school level means in scale scores. For the staff/staff boundaries subscale school means ranged from 1.53 to 2.90. The within row refers to the teachers' deviations from each school's mean. To interpret the within variance, the global mean must be added back in, for example some teachers did not deviate by 3.74 points from the school level mean on the Staff/staff boundaries subscale, they deviated by 1.25 points (3.47(max)-2.22(global mean)). There is a very similar amount of variance within and between schools, as shown by the similar standard deviations within schools and between schools. The observations column indicates the number of observations overall (N), the number of schools (n) and the average number of observations (teachers) per school (T-bar).