



School accountability and teacher stress: international evidence from the OECD TALIS study

John Jerrim^{1,2} · Sam Sims²

Received: 25 April 2020 / Accepted: 18 March 2021
© The Author(s) 2021

Abstract

Accountability—the monitoring and use of student performance data to make judgements about school and teacher effectiveness—is increasing within school systems across the globe. In theory, by increasing accountability, the aims and incentives of governments, parents, school leaders and teachers become more closely aligned, potentially improving student achievement as a result. Yet, in practice, concerns are mounting about the stress that accountability is putting schools and teachers under. This paper presents new evidence on this issue, drawing upon data from more than 100,000 teachers across over 40 countries. We find evidence of a modest, positive correlation between school system accountability and how stressed teachers and headteachers are about this aspect of their job. When looking within schools, there is little evidence that the management practices of headteachers differ when they report feeling stressed about accountability, or that they transmit these feelings onto their staff. However, we do find strong evidence of ‘emotional contagion’ of stress amongst colleagues within schools, with teachers more likely to feel stressed by accountability if their colleagues do as well.

Keywords Accountability · Stress · Well-being · Mental health

1 Introduction

Within school systems across the globe, the issue of accountability is gaining in prominence and importance. Although ‘accountability’ in education can be conceptualised and operationalised in different ways (Gilbert, 2011), it essentially boils down to key agents within the school system (teachers, headteachers, schools) being held responsible for student achievement (Brill et al., 2018). In this paper, we

✉ John Jerrim
J.Jerrim@ucl.ac.uk

¹ Department of Social Science, UCL Institute of Education, University College London, 20 Bedford Way, London WC1H 0AL, UK

² UCL Institute of Education, London, UK

specifically focus upon such *contractual* accountability systems that serve a formative purpose, with other possible forms of accountability (most notably professional accountability mechanisms) beyond the scope of this work. Accompanied by the rise in a data-driven culture (Schildkamp, 2019), such contractual accountability systems in many countries have been synonymous with greater monitoring of student test scores (Hamilton & Koretz, 2002), which are increasingly being used to make judgements about the ‘effectiveness’ of individual teachers (Bitler et al., 2019) and their schools (Goldstein, 1997). One of the reasons why officials across the world have increased such scrutiny within the education system is due to a belief that such monitoring of schools and teachers is associated with higher levels of student performance (Hanushek & Raymond, 2005), a notion that has been supported by influential international organisations such as the OECD (OECD, 2011). With countries competing against one another in the global education arms race, having a strong system of school accountability—underpinned by the use of student assessment data—is now seen by many as a key ingredient to achieving educational success.

Yet this close monitoring of student, teacher and school performance—based largely upon student assessment data—may also be having unintended and undesirable consequences. Some countries with particularly intensive accountability regimes are now facing serious issues with the recruitment and retention of teachers (Craig, 2017), due to the increasing workloads and the negative impact that this may have upon well-being (Perryman & Calvert, 2019). England is a prime example. It has one of the most data-driven systems of school accountability anywhere in the world (Lough, 2019), yet also has one of the lowest levels of teacher job satisfaction and well-being (Jerrim & Sims, 2019) and is consequently struggling to recruit and retain enough staff within the profession (Foster, 2019).

Consequently, developing a better understanding of the unintended negative side effects of intensive data-driven methods of school and teacher accountability is key. We therefore explore this issue within this paper, producing new evidence on the correlates and consequences of accountability-induced stress amongst more than 100,000 teachers and 8000 school leaders from across the globe.

A number of previous studies have investigated the issue of accountability-driven stress amongst teachers, though often based upon relatively small samples drawn from within a single national setting (usually the USA). Using data drawn from three states within the USA, Ryan et al. (2017, p. 1) found that ‘accountability policies may affect teacher stress’, which in turn leads to greater levels of teacher turnover. Berryhill et al. (2009), also drawing upon data from the USA, investigated the link between teacher’s perceptions of school accountability and their job engagement. They suggested that certain types of accountability can lead to role conflict and reduced self-efficacy amongst staff. After reviewing a range of literature, Saeki et al. (2015, p. 95) conclude that ‘accumulating research suggests that test-based accountability practices have unintended, negative effects on teacher well-being, instructional practices, and student learning’. In a qualitative study of 22 science teachers from Indiana, Donnelly and Sadler (2009) found that some teachers felt accountability challenged their professionalism, led to teachers teaching to the test and had a negative impact upon the quality of instruction within their school. Jones and Egley (2004) found that teachers in Florida felt that accountability was having a negative effect upon the curriculum,

teaching and learning and teacher motivation. Valli and Buese (2007, p. 519) concluded that accountability had increased the expectations placed upon primary school teachers in the USA, with negative, unintended consequences for ‘teachers’ relationships with students, pedagogy, and sense of professional well-being’. In a survey of teachers mainly working in California, Richards (2012, p. 302) found that the ‘constant pressure of being accountable’ was one of the top-five sources of stress in their job. Similarly, qualitative research within Illinois (Byrd-Blake et al., 2010) suggested that the pressures of test-driven instruction and high-stakes testing were the parts of the job that teachers disliked the most.

Although insightful, many important questions about the link between accountability and teacher stress have yet to be addressed. For instance, do countries with more intensive, data-driven accountability systems have more stressed teachers and school leaders? Are teachers more likely to feel stressed about being held accountable for student achievement if their colleagues (and, particularly, senior colleagues) also feel under pressure? If school leaders feel stressed by the accountability system, how do their practices—and approaches to school management—change? And is senior management use of test score data in teacher appraisals increasing accountability-induced stress amongst their staff?

This paper will provide new insights into these issues, using data gathered as part of the Organisation for Economic Co-Operation and Development’s (OECD) 2018 Teaching and Learning International Study (TALIS). To trail our key findings, we find a modest, positive correlation between school system accountability and how stressed teachers and headteachers are about this aspect of their job. Yet this association is far from perfect, with several examples of high-accountability school systems where only a comparatively small proportion of staff report feeling stressed (e.g. the USA). When looking within schools, there is little evidence that management practices differ when headteachers report feeling stressed about accountability, or that they transmit these feelings of stress onto their staff. Similarly, in most countries, teachers are no more likely to feel stressed by accountability when student performance data is regularly used by senior management as part of their appraisal (even when a poor appraisal may lead to dismissal), though with some exceptions (e.g. England). We do however find strong evidence of ‘emotional contagion’ of accountability-driven stress amongst colleagues within schools across several countries, with teachers more likely to feel stressed by accountability if their colleagues do as well.

The paper now proceeds as follows. An overview of the background that underpins our research questions and empirical analysis is provided in Section 2. Section 3 then describes the TALIS 2018 data, with an overview of our methodological approach presented in Section 4. Results are then documented in Section 5, with discussion and conclusions following in Section 6.

2 Background and literature

The economic theory of the principal-agent problem (Grossman & Hart, 1983) provides one explanation for the increasing use of data-driven accountability in schools (Figlio & Loeb, 2011). The principal-agent problem occurs when one person or

group—the agents (e.g. schools, teachers)—take actions on behalf of (and/or which may have an impact upon) another group—the principals (e.g. parents, governments). It is thought that such a situation can lead to sub-optimal outcomes if the goals and incentives of the principals and of the agents are not well-aligned. Specifically, because ‘agents’ may act in their own self-interest, which may differ from the interests of the principal, then the goals of the principals may not be achieved. An example of this problem within education might be the allocation of instructional time to different subjects. For instance, agents (schools, teachers) may place greater value upon education in the Arts than the principal (e.g. the government). If left to their own devices, agents (schools, teachers) may thus devote a greater amount of instructional time to the Arts than the principals (the government) might wish.

A simplified illustration of the principal-agent relationship in education can be found in Fig. 1. This highlights how parents, government and school governors are the key ‘principals’ in the education system, while teachers are the key agents. In other words, teachers are the key group who ‘take action’ (i.e. educate children) on behalf of others (parents, government). Headteachers—and other members of the School Management Team (SMT)—fall in-between, with a role as both a principal and as an agent. Specifically, senior school leaders will be acting on behalf of parents and the government as part of their overarching responsibility to ensure children in their school are receiving a good education (making them the agent in this relationship). Yet they will be the ‘principal’ for more junior members of staff in their school, who are acting (i.e. educating children) upon behalf of them as the headteacher. Of course, the goals and incentives of a school SMT and teachers may also not be aligned, giving rise to the principal-agent problem between headteachers and their staff.

Data-driven methods of accountability are seen as a way of dealing with this possible misalignment of incentives in education, thus solving the principal-agent problem. As Figlio and Loeb (2011, p. 386) note:

The information content in school accountability systems can provide a powerful mechanism for overcoming the principal-agent problem. Assessing schools against the common metric of standardized student test scores provides policy makers and members of the general public with independent information regarding how well schools and school districts (and potentially teachers) are doing in comparison to their peers and outside performance standards. Measuring and reporting school performance and attaching positive and negative

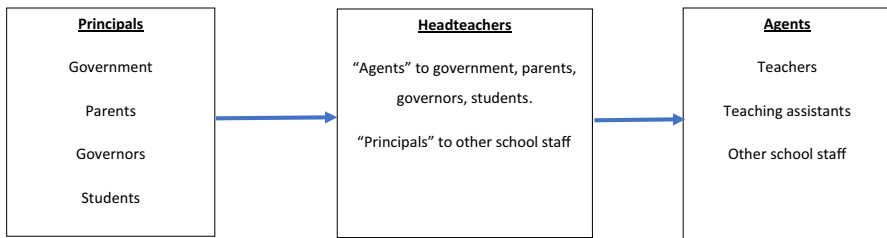


Fig. 1 A simplified illustration of principal-agent relationships within education

consequences to meeting or failing to meet performance objectives provides incentives that encourage educators to concentrate on the subjects and materials that are being measured and to potentially alter the methods through which they educate students. The measurement and reporting of a school's progress allows policy makers to assess how successful a school has been in meeting the state's achievement goals.

In other words, data-driven methods of accountability provide a means by which principals can monitor the performance of their agents, to make sure that their incentives are aligned, and that the agents are working to meet the principal's goals. However, one of the unintended negative consequences of data-driven accountability is that it may increase stress amongst teachers and school leaders. There are several channels through which this might occur.

First, data-driven accountability explicitly entails closer monitoring of the performance of teachers, using some kind of performance standard or metric. If the standards set by the principal are excessive, then this may 'produce stress when employees fail to meet performance requirements' (Smith & Amick, 1989, p. 280). Such monitoring may also increase fear amongst workers that they are not working up to the required standard, or may feel a pressure to work above the average of their peers. As Smith and Amick (1989, p. 280) note, although principals may see this as a desirable effect of accountability, in that it pressurises teachers and school leaders into raising performance standards, 'such work pressure can bring about adverse health challenges'. Such problems may be particularly acute in education—in comparison to other industries—due to the fact that there is a great deal of uncertainty surrounding accountability results; achievement scores are not simply 'produced' by schools and teachers, but also depend upon the input and effort of students and their parents. Previous research has found that uncertainty about (and lack of control over) outcomes is related to an increase in anxiety (Grupe & Nitschke, 2013). This may in turn make teachers and school leaders feel more stressed by the close monitoring that accountability entails. If poor performance according to accountability metrics also has material consequences (e.g. potential job loss, harmful to career advancement), then teachers may find this particularly stressful (Smith & Amick, 1989).

Second, data-driven accountability can lead to a loss of job autonomy. School leaders and teachers may no longer feel free to teach what they feel important, but instead focus upon what is being measured. Moreover, they may feel obliged to teach students in a certain way (i.e. using a method that they—or the principal—believes maximises student performance). A host of research has suggested that lacking autonomy in the workplace is associated with higher stress levels (Spector, 1986), and that being 'controlled by others can be a major contributor to high stress levels' (Relias Media, 1998). Weston (2011) argues that this may be due to our neurological threat-reward systems being activated by the lack of control teachers feel they have when their autonomy is reduced. A lack of control is also a central pillar of Karasek and Theorell's (1990) control-demand model of work-related strain, with stress potentially induced in teachers by the high demands of the job coupled with a low decision latitude (Michie, 2002). This is supported by recent qualitative research in

England (Perryman & Calvert, 2019, p. 18), with one teacher noting how: ‘I do not think it is the children/behaviour that drives teachers away from the profession – it is the lack of support and trust from management that ultimately is directed from the state – pressure of constant tests, assessments and targets. Teachers needed to be trusted more’ (emphasis our own).

Third, accountability may lead to teachers working longer hours, particularly upon auxiliary tasks such as testing, marking and administration. For instance, children may need to be tested more regularly, their work more regularly marked, more regular reviews and meetings around performance targets and extra lessons provided for those pupils struggling to meet their potential. Indeed, Perryman and Calvert (2019) highlight how, for many teachers, it is the nature of the extra workload generated by the extensive, data-driven accountability measures that are in place in England that is having an impact upon the well-being of teachers and forcing many to leave the profession. Within the Job-Demand-Control model (Karasek, 1979), this increase in workload and time-pressure will increase accountability-induced stressed via increasing the demands of the job. Yet it may also decrease teachers feeling of control and decision latitude, as they are increasingly required by principals (governments and senior leaders) to spend more time upon unfulfilling tasks (such as administration and marking). In this sense, the increase in total workload—as well as the nature of the work—is likely to result in teachers feeling more (accountability-induced) stress.

Fourth, accountability may change the atmosphere of a school as a workplace. For instance, teachers may become more stressed about accountability if their colleagues—particularly senior colleagues—are feeling stressed by accountability as well. This transfer of emotions within a group is known as ‘emotional contagion’, with the stress being felt by a teacher spreading like a virus to their colleagues (Hatfield et al., 2014). Previous research from Canada has suggested that such emotional contagion may exist within schools (Oberle & Schonert-Reichl, 2016), though this focuses upon the flow between teachers and children. Nevertheless, there are clear reasons to believe that teachers may transmit stress onto their colleagues, including by increasing the likelihood and severity of workplace conflict, transferring unwanted workload onto others, greater levels of stress-induced absenteeism leading to increased workloads and a general lowering of workplace moral. Previous research has also suggested that the emotions of some team members—most notably senior leaders—may be more contagious than others (Sy et al., 2005). Consequently, pressure and stress felt by school leaders may have a particularly big, negative effect upon sub-ordinate staff.

Finally, when principals (e.g. senior leaders) feel the pressure of the accountability system, they may change the way that they manage their school. This might include, for instance, greater monitoring of staff performance (e.g. via more frequent reviewing test score data and conducting teacher appraisals), becoming more autocratic in their management (further reducing teacher autonomy) and imposing harsher material sections upon teachers for below-par performance. This of course has the potential to reduce well-being and increase stress amongst teachers, which are thus potential mechanisms via which the emotional contagion of stress from senior to junior staff may occur.

2.1 Research questions

Based upon the background literature overviewed above, we have developed a set of four research questions to investigate in this paper.

We begin by focusing upon the big picture; do school systems with more extensive monitoring of schools and teachers through (mainly data-driven) accountability practices have more stressed teachers and school leaders? One would expect this to be the case as high-accountability systems will involve closer monitoring of teachers and schools, with staff under greater pressure to meet their targets. As noted above, such additional monitoring may lead to stress due to a fear of failure, school leaders and teachers feeling the need to produce ‘above-average’ results and due to consequences they may face if their results are deemed below-par. Our first research question is therefore:

Research question 1: Do countries that place more emphasis upon school performance accountability measures have more stressed teachers and headteachers?

In the second research question, we turn our attention to teachers and whether they feel more stressed by accountability when student test scores are used to judge their performance. The theoretical background and literature reviewed above noted how headteachers might feel obliged to use test score data in teacher appraisals as one solution to the principal-agent problem. Yet this is also likely to cause teachers stress due to the additional surveillance that it entails, the increase—and change in nature—of teacher workloads, the only partial control teachers have over student outcomes and due to a potential reduction in their autonomy. As implied by the work of Smith and Amick (1989), we also investigate whether teachers are particularly stressed when test score data are used in their appraisal and this may have material consequences for their career (e.g. they may face dismissal).

Research question 2: Are teachers more stressed by accountability when senior leaders regularly use achievement data to make judgements about their performance? (And when this may have consequences for their career)?

Next, we explore whether there is any evidence of ‘emotional contagion’ of stress amongst staff within schools? Specifically, do teachers feel more stressed by accountability when (a) their colleagues and (b) their headteacher also feel stressed by this aspect of their job?

Research question 3: Are teachers more stressed by accountability when their colleagues (including their headteachers) feel stressed by accountability as well?

Finally, when headteachers feel stressed by accountability, what is done differently within their school? Is there a less collaborative—and generally more toxic—atmosphere amongst staff? Are they more likely to use test score data in teacher appraisal (as a solution to the principal-agent problem)? Might headteachers be more likely to implement material sanctions against their staff for poor

performance, including dismissals and withholding pay rises? Or do they become more autocratic, with more junior staff less likely to be involved in decision-making processes? Answering such questions is important as, although many headteachers say that they feel stressed from the pressures of accountability (Jerrim & Sims, 2019), we currently know very little about how this changes their management (and the general environment) of their school. Our final research question is therefore:

Research question 4: When headteachers are stressed by accountability, how do their school management practices change, and does it worsen the environment in the school?

3 Data

The data we use are drawn from the 2018 round of the Teaching and Learning International Study (TALIS). This is an international survey of lower-secondary (ISCED level 2) teachers and headteachers conducted across more than 40 countries.¹ It was conducted in most Northern Hemisphere nations between March and May 2018, though took place slightly earlier (the end of 2017) in some Southern Hemisphere nations. Within each country, a nationally representative sample of 200 schools was drawn with probability proportional to size. Within each of these schools, the headteacher and 20 randomly selected teachers were asked to complete a questionnaire. The TALIS data therefore have a nested structure, with teachers (level 1), nested within schools (level 2) and within countries (level 3). The OECD has strict requirements around response rates, with 75% of schools and 75% of teachers required to complete the survey. (Details on the response rate for each country can be found in OECD (2019, Chapter 10)). Teacher and Balanced-Repeated-Replication (BRR) weights are provided as part of the TALIS database. These fully account for the complex sampling design in the estimation of population parameters and the associated standard errors (Micklewright et al., 2014). Unless stated otherwise, these weights are applied within the analysis.

As part of the TALIS questionnaire, teachers and headteachers were asked about their sources of stress at work. Our specific interest is in the stress teachers and headteachers report being caused by accountability, operationalised by how they responded to the statement: ‘Being held responsible for students’ achievement’. Responses to this question have been recoded into binary format, combining the two top (‘a lot’ and ‘quite a bit’) and the bottom two (‘not at all’, ‘to some extent’) categories. See Table 1 row 1 for further details.

Headteachers were also presented a series of questions about the management of their school. First, they were asked to provide information on the frequency with which each teacher in their school is appraised by different groups (see Table 1 row

¹ In some countries, teachers in primary and upper-secondary schools were also surveyed. We have chosen to focus upon the lower-secondary sample in which all participating TALIS countries were required to take part.

2). We draw upon this information when addressing research question 2, with a particular focus upon whether teachers feel more stressed by accountability when school management regularly evaluate their performance drawing upon national examination or other test score data. Second, they were asked what information is used to judge the performance of teachers as part of these appraisals (see Table 1 row 3). Here, our particular interest is in whether school management use of test score data is linked to teachers' accountability-induced stress. Finally, headteachers were also asked about the potential consequences of teacher appraisal (see Table 1 row 4). Our particular interest using these data is in whether teachers find accountability particularly stressful when test score metrics are used by senior management in their appraisal *and* when this has potentially serious consequences for their career.

3.1 PISA 2018

In addition to TALIS, we also draw upon information from the PISA 2018 headteacher survey in order to address research question 1 (whether school systems with more data-driven accountability have more stressed teachers and school leaders). As part of PISA 2018, headteachers were asked about their use of assessment data (see Table 1 row 5), whether external evaluation (e.g. inspections) is used as part of the quality assurance process for their school (see Table 1 row 6) and how achievement data about the school is tracked/disseminated (see Table 1 row 7). These indicators are combined into a single accountability scale,² which has been standardised to mean zero and standard deviation one across countries. The mean of this accountability scale is then calculated for each TALIS country, with greater values indicating greater use of (test score driven) school accountability. The final PISA senate weights have been applied throughout this process.

4 Methodology

4.1 Research question 1: Do countries that place more emphasis upon school performance accountability measures have more stressed teachers and headteachers?

Research question 1 focuses upon analysis at the country level, linking the stress suffered by teachers and school leaders to the intensity of the national system of school accountability. Our analysis will begin by simply presenting a scatterplot of our school accountability scale (based upon PISA data, as described in the previous section) against the proportion of teachers and headteachers who report feeling 'quite a bit' or 'a lot' of stress from being held responsible for pupil achievement.

This will be supplemented by estimation of a multi-level linear probability model, with teachers (level 1) nested within schools (level 2) nested within countries (level 3). A multi-level model has been used as an additional method to obtain the associations

² This has been done via estimation of a two-parameter Item-Response Theory model.

Table 1 Key measures used in the paper

Row	Question	Response options	Source
1	Thinking about your job at this school, to what extent are the following sources of stress in your work? <i>Being held responsible for students' achievement</i>	A lot/quite a bit; to some extent/not at all	TALIS 2018
2	On average, how often is each teacher formally appraised in this school by the following people? <i>(Headteachers, other members of the school management team, assigned mentors, teachers not part of the management team, external individuals or bodies (e.g. school inspectors))</i>	Never, less than biannually, biannually, annually, twice or more per year	TALIS 2018
3	Who uses the following types of information as part of the formal appraisal of teachers' work in this school? <i>Review of students' national test scores, school-based and classroom-based results (e.g. performance results, test scores)</i>	Yes/no	TALIS 2018
4	Please indicate the frequency that each of the following occurs in this school following a formal teacher appraisal. <i>Material sanctions such as reduced annual increases in pay are imposed, a change in the likelihood of a teacher's career development, dismissal or non-renewal of contract</i>	Never, sometimes, most of the time, always	TALIS 2018
5	In your school, are assessments of students for 15-year-olds used for any of the following purposes? <i>Compare their school to national performance, monitor school progress, judge teacher effectiveness, draw comparisons to other schools</i>	Yes/no	PISA 2018
6	Do the following arrangements aimed at quality assurance and improvements exist in your school and where do they come from? <i>External evaluation</i>	Yes (this is mandatory; based on school initiative)/no	PISA 2018
7	In your school, are achievement data used in any of the following accountability procedures? <i>Gets posted publicly, is tracked over time by an administrative authority, is provided directly to parents</i>	Yes/no	PISA 2018

as a more formal way of accounting for the clustered nature of these data. These models will control for teacher and headteacher demographic characteristics (gender, qualifications, whether work part-time, experience) and an OECD indicator at the country level. Specifically, the model estimated for the relationship between school accountability and teacher stress is specified as the following three-level model:

$$S_{ijk} = \alpha + \beta.T_{ijk} + \delta.H_{jk} + \theta.OECD_k + \gamma.Account_k + v_k + u_{jk} + \epsilon_{ijk}$$

S_{ijk}	whether the <i>teacher</i> feels ‘a lot’ or ‘quite a bit’ of stress due to accountability (1) or not (0).
T_{ijk}	demographic characteristics of teacher i in school j in country k
H_{jk}	headteacher of school j in country k
$OECD_k$	a dummy indicator of whether an OECD
$Account_k$	the measure of school system accountability described in Section 3
v_k	a country-level (level 3) random effect
u_{jk}	a school-level (level 2) random effect
ϵ_{ijk}	unexplained residual variance at the teacher level (level i)
i	teacher i
j	school j
k	country k

A similar two-level model is estimated for the link between school system accountability and headteacher stress:

$$S_{jk} = \alpha + \delta.H_{jk} + \theta.OECD_k + \gamma.Account_k + v_k + \epsilon_{jk}$$

where all variables are defined as above, except:

S_{jk}	whether the <i>headteacher</i> feels ‘a lot’ or ‘quite a bit’ of stress due to accountability (1) or not (0)
ϵ_{jk}	unexplained residual variance at the headteacher level (level i)

Estimates from these models will formalise the descriptive relationship illustrated by the scatterplots. At this point, it is worthwhile reminding readers that such a cross-national analysis can only reveal a correlational relationship, rather than detecting cause and effect.

4.2 Research question 2: Are teachers more stressed by accountability when senior leaders use achievement data to make judgements about their performance? (And when this may have consequences for their career)?

To begin with, we construct an indicator variable for whether senior management regularly monitors teachers using test score/achievement data. This is operationalised as (a) the headteacher/senior management team (SMT) conducting an appraisal with teachers at least once per year and (b) test score/achievement data being used

as part of this appraisal. If both of these conditions are met, the indicator is coded as one, and zero otherwise. We then estimate a logistic regression model, exploring whether this variable is associated with the stress teachers feel due to accountability. The model is specified as:

$$\text{logistic}(S_{ij}) = \alpha + \beta \cdot \text{App}_j + \gamma \cdot \text{Teacher_Controls}_i + \delta \cdot \text{School_Controls}_j$$

where

S_{ij}	whether the teacher feels ‘a lot’ or ‘quite a bit’ of stress due to accountability (1) or not (0)
App_j	a dummy variable for whether senior management use achievement data as part of an annual appraisal of teachers (1) or not (0)
$\text{Teacher_Controls}_i$	a set of controls for teacher characteristics.
School_Controls_j	a set of controls for background school characteristics, including the percent of SEN, disadvantaged, immigrant and foreign language pupils, headteacher experience (years), pupil-teacher ratio, teacher-teaching assistant ratio, teacher-administrator ratio, location and a scale capturing whether the headteacher believes instruction in the school is hindered by a lack of resources.
i	teacher i
j	school j

Two different specifications of this model are estimated, including a different set of teacher controls to test the robustness of the results. The base specification will include controls for teacher age, gender, experience and length of tenure in current school. In the second specification, we include additional controls for whether teachers say they feel stressed by other aspects of their job which are unlikely to be caused by accountability. This includes whether teachers feel stressed by (a) classroom discipline, (b) intimidation or verbal abuse from students, (c) having too many lessons to teach and (d) having to modify lessons for Special Educational Needs (SEN) pupils. This second model specification hence attempts to tease out the stress teachers feel from accountability (because they are being monitored by senior management) from other stressful aspects of their job.

Finally, we re-estimate this model separately for sub-samples of teachers depending upon whether they may face serious consequences if they receive a poor appraisal. Specifically, we divide the sample into two groups depending upon whether the headteacher indicated that a poor teacher appraisal can at least sometimes lead to dismissal, versus those who said that dismissal due to a poor appraisal never occurs. The second model specification is then estimated separately for these two groups, to investigate whether this factor moderates the results. The robustness of these findings is tested in Appendix A where we also include possible (a) material sanctions and (b) impact upon career advancement (as well as the possibility of dismissal) when dividing the sample into the two sub-groups.

Results from the pooled cross-country model (which also includes country dummy variables) will be presented in the main text, with the country-by-country results presented in Appendix B. Estimates will be presented as odd ratios in the main text, with marginal effects (probability differences) based upon a linear probability model presented in Appendix C.

4.3 Research question 3: Are teachers more stressed by accountability when their colleagues (including their headteachers) feel stressed by accountability as well?

To begin with, we investigate whether teachers within a school feel more stressed by accountability when their headteacher feels stressed by this aspect of their job. This is done via estimation of the logistic regression model:

$$\text{logistic}(S_{ij}) = \alpha + \beta.\text{Head_Stress}_j + \gamma.\text{Teacher_Controls}_i + \delta.\text{School_Controls}_j$$

where

Head_Stress_j a single, linear term indicating how stressed the headteacher said they were due to accountability (4-point scale)

All other variables are specified as per the logistic regression model presented under research question 2. Three specifications of this model will be estimated using different sets of controls. The first two specifications are the same as under research question 2 above (where teacher stress in other aspects of their job is added to the baseline model). Additionally, in the third model, we will also add controls for headteacher stress in other aspects of their job (school discipline, abuse from students, accommodating SEN students). The purpose of these additional controls is again to separate out the ‘impact’ of headteachers being stressed by accountability from them being stressed about other aspects of their job.

A similar procedure is followed when we investigate whether there appears to be ‘emotional contagion’ of accountability stress amongst teaching staff. First, for each teacher, we calculate the stress levels of their peers (i.e. other teachers who completed the survey within their school). This is taken as the school average (mean) of the 4-point question teachers were asked about stress due to accountability—having excluded each teacher’s own individual response. (Appendix H presents alternative estimates using the school modal value of accountability stress rather than the mean. Substantive results remain unchanged). A logistic regression model is then estimated, with this ‘peer stress’ variable entered as a linear term:

$$\text{logistic}(S_{ij}) = \alpha + \beta.\text{Peer_Stress}_j + \gamma.\text{Teacher_Controls}_i + \delta.\text{School_Controls}_j$$

where

Peer_Stress the accountability-induced stress reported by the colleagues of each teacher

The same three model specifications will be estimated as discussed above. This will be supplemented by a fourth specification, where we also control for their colleagues' reports of stress in other areas of their job (school discipline, abuse from students, having too many lessons to teach, having to modify lessons for SEN students) and their colleagues' overall levels of job satisfaction. Again, this will help illustrate whether it is their colleague's accountability-induced stress that is driving the association, or if this may be driven by other aspects of how their colleagues feel about their job.

4.4 Research question 4: When headteachers are stressed by accountability, how do their school management practices change, and does it worsen the environment in the school?

Finally, when headteachers are stressed by accountability, what changes within their school? We examine the following based upon teacher and headteacher responses to the TALIS background questionnaire:

- Whether the headteacher feels they have a need for Continual Professional Development (CPD) in 'using data for improving the quality of the school'
- Whether, over the last 12 months, the headteacher 'took action to ensure that teachers feel responsible for their students' learning outcomes'
- Whether, over the last 12 months, the headteacher 'provided parents or guardians with information on the school and student performance'
- Whether school management undertook more frequent teacher appraisals
- Whether senior management were more likely to use test scores in teachers' appraisals
- Whether headteachers were more likely to take material sanctions against teachers for a poor appraisal
- Whether teachers were more or less likely to involve staff in decision-making
- Whether the headteacher's management of the school is more autocratic
- Whether there is a collaborative school culture
- Whether teachers hold high expectations for student achievement
- Whether headteachers are more likely to feedback test score results to teachers

Regression modelling is again used, where we control for a set of basic background characteristics of the teacher and of the headteacher/school. Formally, this model is specified as:

$$P_{ijk} = \alpha + \beta \cdot \text{Stress}_{jk} + \gamma \cdot D_{ijk} + \delta \cdot P_{jk} + \theta \cdot E_{jk} + \tau \cdot R_{jk} + \pi \cdot L_{jk} + u_k$$

where

P_{ijk} teacher/headteacher report of whether the practice is followed in the school (0=no; 1=yes)

Stress_{jk} the accountability stress of headteacher j in country k

D_{ijk}	teacher demographic characteristics (age, experience, gender, length of tenure at current school)
P_{jk}	characteristics of pupils at school j in country k —the percent who (a) have special educational needs, (b) come from disadvantaged background, (c) come from immigrant backgrounds, (d) are non-native language speakers
E_{jk}	the number of years of experience of headteacher j in country k
R_{jk}	the resources of school j in country k , as captured by pupil:teacher ratio, teacher:teaching assistant ratio, teacher:administrator ratio and a continuous scale capturing whether headteachers believe that instruction in the school is hindered by a lack of resources
L_{jk}	the location (e.g. urban/rural) of school j in country k
u_k	country fixed effects
i	teacher i
j	School/headteacher j
k	country k

The intuition is that these models attempt to establish how school management practices and school environment differ when the headteacher feels stressed by accountability, within what are otherwise similar schools.

4.5 A note of caution

Regression models are used to address each of our research questions. These are, of course, limited by the data available within the TALIS dataset, meaning some models include a relatively standard set of school and teacher controls. Although a more detailed set of controls are included where possible (e.g. controlling for other sources of stress outside of that attributed to accountability), all estimates are still subject to an (untestable) selection-upon-observables assumption. We therefore feel it is prudent to remind readers that our results refer to conditional associations only, and do not necessarily capture cause and effect.

5 Results

5.1 Research question 1: Do countries that place more emphasis upon school performance accountability measures have more stressed teachers and headteachers?

To begin with, Fig. 2 illustrates the relationship between the scale of school system accountability and the percentage of headteachers (panel a) and teachers (panel b) who report being stressed by accountability at the country level. In both graphs, there is a moderate, positive correlation (Pearson $r=0.31$, $p=0.04$ for headteachers and $r=0.32$, $p=0.04$ for teachers). Consistent with our hypothesis, countries with more extensive, data-driven systems of school accountability also have staff who feel more stressed by this aspect of their job. Yet there are some clear exceptions

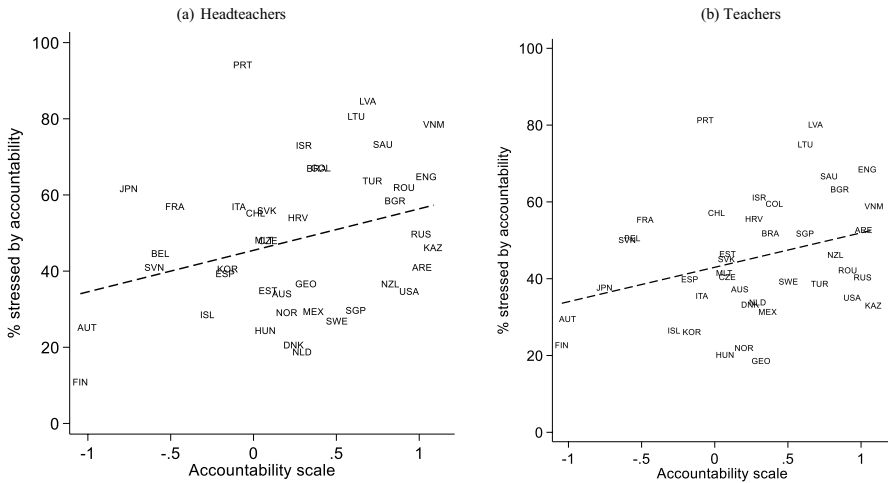


Fig. 2 The cross-national relationship between the extent of school accountability and the percent of staff stress by accountability. Accountability scale derived using PISA 2018 data, based upon how headteachers use student assessment data, how achievement data are disseminated to stakeholders and whether external evaluation is used in quality assurance. Higher values on this scale indicate greater levels of school accountability. OLS regression estimate illustrated by dashed line. Pearson correlation=0.31 in panel a and 0.32 in panel b

to this relationship as well. For instance, despite its extensive use of data-driven accountability, the USA sits just below the international average in terms of the proportion of teachers and headteachers reporting high levels of accountability-induced stress. On the other hand, in Portugal, many more teachers and headteachers report high levels of stress due to accountability than one would anticipate, given the level of accountability in its school system.

These results are formalised in Table 2 where we present results from a two-level (headteacher nested within countries) and a three-level (teachers nested within schools nested within countries) multi-level model. For both headteachers and teachers, there is a positive association. However, the results for headteachers (in particular) are imprecisely determined, with a large standard error. This reflects the limited sample size, at both the headteacher and country levels. For teachers, a one-standard deviation increase in the school system accountability scale is associated with a 4 percentage point increase in the percentage of teachers who say they feel stressed by being held accountable for student achievement. This is a moderate association, which is statistically significant at the 10% level.

5.2 Research question 2: Are teachers more stressed by accountability when senior leaders use achievement data to make judgements about their performance? (And when this may have consequences for their career)?

Table 3 presents the estimates which address our second research question, using TALIS data that has been pooled across all countries (see Appendix

Table 2 Estimates from a three-level multi-level model exploring the link between school system accountability and teacher/headteacher stress

	Headteacher		Teacher	
	Marginal effect	SE	Marginal effect	SE
Intensity of school system accountability				
Change in stress per standard deviation increase in school system accountability scale	4.2%	3.3%	4.2%	2.4%
<i>N</i>	7565		120,169	
Controls				
Teacher gender	-		Y	
Teacher experience	-		Y	
Teacher qualifications	-		Y	
Teacher part-time	-		Y	
Head gender	Y		Y	
Head experience	Y		Y	
Head part-time	Y		Y	
Head qualification	Y		Y	
OECD country	Y		Y	
Variance explained				
Level 1 (teacher level)	N/A		0.9%	
Level 2 (school level)	1.1%		0%	
Level 3 (country level)	6.6%		11%	

Estimates from multi-level linear probability models. Headteacher results are two-level models (headteachers nested within countries) while teacher results are a three-level model (teachers nested within schools within countries). Number of countries is 42. Variance explained estimated using the approach of Snijders and Bosker (1994)

G for parameter estimates and standard errors for the control variables). Overall, there is little evidence that senior school leaders regularly using student performance data when conducting appraisals leads to teachers feeling more stressed about accountability. The estimated odds ratio from both model specifications falls around one, suggesting that there is no overall, systematic difference in teacher stress associated with senior leaders regularly using student performance data when appraising their staff. A country-by-country breakdown of results is provided in Appendix B, with little evidence of a clear relationship between annual SMT use of test/exam score data in appraisals and accountability-driven stress in most. Potential exceptions include Columbia, Kazakhstan, Slovak Republic, South Africa, Sweden and England. In these nations, the estimated (log) odds ratios do appear more sizable across the two model specifications, and are usually statistically significant at least at the 10% level. For instance, in England, secondary teachers are around 12 percentage points more likely to say that they

Table 3 The association between whether regular use of test score data in appraisals by senior school leaders and accountability-induced stress amongst teachers

	Model 1		Model 2	
	OR	SE	OR	SE
SMT regularly use test scores in appraisal (Ref: No)				
Yes	1.064	0.038	1.064	0.079
N	133,757		132,587	
Controls				
Country dummies	Y		Y	
School location dummies	Y		Y	
School lack of resources scale	Y		Y	
Pupil:teacher ratio	Y		Y	
Teacher:TA ratio	Y		Y	
Teacher:admin ratio	Y		Y	
% of immigrant/disadvantaged pupils	Y		Y	
Headteacher experience	Y		Y	
Teacher experience + demographics	Y		Y	
Teacher feels stress by school discipline	-		Y	
Teacher feels stress from abuse from student	-		Y	
Teacher feels stress from too many lessons	-		Y	
Teacher stress modifying lessons for SEN pupils	-		Y	

SMT use of test scores in appraisal is defined as a SMT conducting an appraisal with teachers at least annually, which includes a review of the examination/test score data of pupils that they teach. Estimates using data pooled across all countries, with teacher (senate) weights and BRR weights applied. Estimates refer to odds ratios

feel stressed due to accountability if SMT use student performance data in teacher appraisals.³

Table 4 extends this analysis by dividing the pooled TALIS data into two sub-groups—those schools where teachers never face dismissal following an appraisal (column 1) and those where dismissal is a possibility (column 2). In other words, do we find teachers being more stressed by accountability when test score data is used in their appraisal *and* when this could have serious consequences for their career? We find little evidence that this is the case. The estimated odds ratios reported in Table 4 are again close to one and do not differ substantially between the two sub-groups (see Appendix G for parameter estimates and standard errors for the control variables). A similar finding emerges in our robustness tests presented in Appendix A, where we have split the sample into two groups using additional variables. Country-by-country results are again presented in Appendix B, flagging some potential exceptions to this broad, cross-national finding. In particular, in Columbia, Croatia, Italy, New Zealand, Slovak Republic, Slovenia,

³ This estimate has been produced using a linear probability model based upon the second model specification. It is also worth noting that around 90% of teachers in England are evaluated at least annually by a member of senior school management, where student performance data is reviewed.

Table 4 The association between whether regular use of test score data in appraisals by senior school leaders and accountability-induced stress amongst teachers. Sub-group estimates by whether the teacher potentially faces dismissal

	Won't face sack		Might face sack	
	OR	SE	OR	SE
SMT use test scores in appraisal (Ref: No)				
Yes	1.042	0.085	1.091	0.148
<i>N</i>	64,954		59,643	
Controls				
Country dummies	Y		Y	
School location dummies	Y		Y	
School lack of resources scale	Y		Y	
Pupil:teacher ratio	Y		Y	
Teacher:TA ratio	Y		Y	
Teacher:admin ratio	Y		Y	
% of immigrant/disadvantaged pupils	Y		Y	
Headteacher experience	Y		Y	
Teacher experience + demographics	Y		Y	

SMT use of test scores in appraisal is defined as a SMT conducting an appraisal with teachers at least annually, which includes a review of the examination/test score data of pupils that they teach. Estimates using data pooled across all countries, with teacher (senate) weights and BRR weights applied. Estimates refer to odds ratios. Estimates in column (1) refer to the sub-sample of teachers working in schools where the headteacher reports that staff are 'never' dismissed following an appraisal; estimates in column (2) are whether headteachers said dismissal 'sometimes', 'often' or 'always' occurs

South Africa, Sweden and the UAE, annual SMT review of student performance data in teacher appraisal is associated with higher levels of accountability-induced stress amongst teachers when dismissal is a possibility.

5.3 Research question 3: Are teachers more stressed by accountability when their colleagues (including their headteachers) feel stressed by accountability as well?

To begin with, we consider whether teachers report being more stressed by accountability when their headteachers also feel under more pressure from this aspect of their job. These results—for three different model specifications—can be found in Table 5 (see Appendix G for parameter estimates and standard errors for the control variables). Estimates are presented as odds ratios and refer to the increase in accountability-induced stress amongst teachers per each category increase in headteacher stress (e.g. the headteacher moving from selecting 'to some extent' to 'quite a bit' when reporting their stress due to accountability). Supplementary estimates entering each category as a separate dummy variable can be found in Appendix D.

There appears to be a modest, positive association between the accountability-induced stress reported by headteachers and by their staff. Across all model

Table 5 The association between headteacher and teacher stress about accountability

	Model M1		Model M2		Model M3	
	OR	SE	OR	SE	OR	SE
Headteacher stressed by accountability						
Change per one category increase in headteacher stress	1.077*	0.036	1.058	0.044	1.101*	0.051
<i>N</i>	133,471		132,311		132,049	
Controls						
Country dummies	Y		Y		Y	
School location dummies	Y		Y		Y	
School lack of resources scale	Y		Y		Y	
Pupil:teacher ratio	Y		Y		Y	
Teacher:TA ratio	Y		Y		Y	
Teacher:admin ratio	Y		Y		Y	
% of immigrant/disadvantaged pupils	Y		Y		Y	
Headteacher experience	Y		Y		Y	
Teacher experience + demographics	Y		Y		Y	
Other sources of stress for the teacher	-		Y		Y	
Other sources of stress for the headteacher	-		-		Y	

Outcome variable = whether the teacher reports feel 'quite a bit' or 'a lot' or stress (1) or not (0). Analysis based upon pooled TALIS 2018 lower secondary data pooled across countries. Final teacher (senate) and BRR weights applied. Estimates refer to odds ratios per one-category increase in headteacher stress about accountability (e.g. the headteacher typically saying that they suffer stress due to accountability 'quite a bit' rather than 'to some extent'). *Statistical significance at the 5% level

specifications, the odds ratio sits above one, with the coefficient statistically significant in M1 and M3.⁴ We should, however, emphasise that the magnitude of the estimated association is relatively modest; the results imply that the headteacher moving from the lowest stress category ('not at all') to the highest ('a lot') is associated with around a 6 percentage point increase in the percentage of teachers who report that accountability causes them stress.⁵ The country-by-country results presented in Appendix E also illustrate how emotional contagion of stress between headteachers and staff is only strong in certain countries. Specifically, Australia, Belgium, Colombia, Denmark, Estonia, Japan, Kazakhstan, South Korea, New Zealand, Portugal, Russia, Slovak Republic, Slovenia and Sweden are examples of countries where teachers are particularly likely to report feeling more stressed about accountability when their headteachers also feel stressed by this responsibility.

In Table 6, we turn to the analogous results for the emotional contagion of stress between teachers and their colleagues within a school (see Appendix G for parameter estimates and standard errors for the control variables). Here we do find

⁴ The estimated odds ratio is of similar magnitude across the three model specifications, though the standard error is slightly inflated in specification two. This is part of the explanation as to why results from the second model are not 'statistically significant'.

⁵ Estimates based upon a linear probability model using model specification 3.

Table 6 Emotional contagion. Are teachers more stressed about accountability when their colleagues are also stressed by it?

	Model M1		Model M2		Model M3		Model M4		
	OR	SE	OR	SE	OR	SE	OR	SE	
Accountability stress of colleagues									
Per one category increase	2.06*	0.62	1.82*	0.41	1.83*	0.42	2.14*	0.65	
N	134,036		132,863		132,151		131,521		
Controls									
Country dummies	Y		Y		Y		Y		
School location dummies	Y		Y		Y		Y		
School lack of resources scale	Y		Y		Y		Y		
Pupil:teacher ratio	Y		Y		Y		Y		
Teacher:TA ratio	Y		Y		Y		Y		
Teacher:admin ratio	Y		Y		Y		Y		
% of immigrant/disadvantaged pupils	Y		Y		Y		Y		
Headteacher experience	Y		Y		Y		Y		
Teacher experience + demographics	Y		Y		Y		Y		
Other sources of stress for the teacher	-		Y		Y		Y		
Other sources of stress for the headteacher	-		-		Y		Y		
Other sources of stress for colleagues	-		-		-		Y		

Outcome variable = whether the teacher reports feel 'quite a bit' or 'a lot' or stress (1) or not (0). Analysis based upon pooled TALIS 2018 lower secondary data pooled across countries. Final teacher (senate) and BRR weights applied. Estimates refer to odds ratios per one category increase in colleagues stress about accountability (e.g. the colleagues of a teacher typically saying that they suffer stress due to accountability 'quite a bit' rather than 'to some extent'). *Statistical significance at the 5% level

consistently strong evidence of emotional contagion. Across the four model specifications, the odds ratio using the pooled cross-national data is around two and is consistently statistically significant at the 5% level. For instance, a one-category increase in colleagues' accountability stress (e.g. a teachers' colleagues typically saying they suffer 'a lot' of stress due to accountability rather than 'quite a bit') is associated with a 14 percentage point increase in a teacher's own level of stress.⁶ This holds true even once we control for how stressed the teacher in question and their colleagues feel about other aspects of their job (e.g. school discipline, number of lessons) and their colleagues' overall level of job satisfaction. Moreover, Appendix H illustrates how this result is robust to using the modal value (rather than the mean) within each school to measure the accountability stress suffered by teachers' peers. It hence seems that, when a teacher's colleagues are more stressed by the pressures of accountability, they become more stressed about being held responsible for pupil's achievement themselves.

⁶ This estimate is based upon a linear probability model, using model specification 4.

Appendix F illustrates the cross-national variation in this result, finding strong evidence of such emotional contagion of stress in some countries, but not in others. Examples of countries where accountability stress amongst secondary teachers seems to be particularly contagious include England, Spain, Singapore, Denmark, Brazil and Hungary. On the other hand, in nations such as Finland, Italy, Austria, Norway and Sweden, there is little evidence that emotional contagion of accountability stress occurs.

Putting these results together, we find strong evidence that stress about accountability gets transmitted between teaching staff and their colleagues. Yet emotional contagion between headteachers and their subordinates seems, in most countries, to be relatively weak. This is counter to previous work (outside of education) by Sy et al. (2005), who suggested the transfer of emotions from senior staff to those more junior is particularly strong. One possible interpretation of this finding is that headteachers generally do a good job in trying to protect their staff when they themselves feel stressed about accountability. Yet this does then not seem to stop concern spreading amongst teachers, once an atmosphere of fear starts to take hold in a school.

5.4 Research question 4: When headteachers are stressed by accountability, how do their school management practices change, and does it worsen the environment in the school?

To conclude, Table 7 investigates what changes about a school when the headteacher feels stressed about accountability. The results presented are based upon pooled data across all countries, with unadjusted descriptive statistics provided on the left, and regression model estimates on the right (see Appendix G for parameter estimates and standard errors for the control variables). The latter reflect the change in the probability of the headteacher taking the action, for each category increase in headteacher stress.

Interestingly, most differences are small and fail to reach statistical significance at conventional thresholds. For instance, there is little evidence that headteachers become more autocratic in their management (see rows 8–10), become more likely to use test scores in teacher appraisals (row 4) and more frequently feedback test score data to staff (row 14) or that it leads to a less collaborative environment within the school (rows 11–12). This is broadly consistent with the results presented within the sub-section on emotional contagion above; if school leaders do not alter their approach to management when they are stressed—and it does not worsen the environment within the school—then it is perhaps not surprising that the link between headteacher and teacher stress surrounding accountability is relatively weak. Moreover, it again suggests that, even when headteachers themselves feel stressed about accountability, they try to not take negative actions (e.g. become more autocratic) which might put additional pressure on staff. One interpretation of this result—and more generally of those presented within this paper—is that teachers feel the pressure of accountability directly from the system, rather than it being driven by the actions of headteachers in response to the stress they themselves feel from accountability-driven pressures.

Table 7 What changes within a school when the headteacher feels stressed about accountability?

Head stressed by accountability →	Unconditional			Model		
	Not at all	To some extent	Quite a bit	A lot	Marginal effect	SE
1.Head feels in need of CPD in using data to improve school quality	57%	67%	74%	75%	3.8%	2.9%
2.Head takes action to ensure teachers feel responsible for pupil learning outcomes	21%	14%	19%	29%	2.0%	3.7%
3.Head provides parents with information on student performance	61%	57%	60%	68%	0.6%	2.5%
4.Head undertakes more frequent teacher appraisals	37%	34%	40%	44%	0.4%	1.1%
5.SMT more likely to use test scores in teacher appraisal	92%	92%	92%	91%	0.3%	0.5%
6.More likely to take material sanctions against teachers for poor appraisal	16%	19%	20%	20%	0.8%	1.8%
7.More likely to dismiss teacher following appraisal	53%	52%	47%	43%	1.7%	3.0%
8.Whether involve staff in decision-making (headteacher report)	43%	34%	36%	42%	-0.2%	1.4%
9.Whether involve staff in decision-making (teacher report)	17%	16%	16%	16%	-0.5%	0.4%
10.Whether management more autocratic	28%	28%	31%	30%	0.3%	1.1%
11.Whether there is a collaborative school culture (headteacher report)	39%	31%	28%	33%	-1.8%	1.1%
12.Whether there is a collaborative school culture (teacher report)	21%	19%	17%	18%	-0.6%	0.6%
13.Whether teachers hold high expectations for student achievement	41%	31%	30%	37%	-2.2%	4.0%
14. Whether headteachers feedback test score results to teachers	52%	51%	55%	60%	0.3%	1.4%

Analysis based upon pooled TALIS 2018 lower secondary data pooled across countries. Final teacher (senate) and BRR weights applied. Figures on the left refer to unconditional descriptive statistics. The model estimates illustrate the change in the probability for a one-category increase in headteacher stress due to accountability (e.g. the headteacher typically saying that they suffer stress due to accountability 'quite a bit' rather than 'to some extent'). Model controls for teacher demographics, school intake, school resources (e.g. pupil-teacher ratios) and country fixed effects

6 Conclusions

Accountability, and the close monitoring of student achievement data, is becoming increasingly common within school systems across the world. In theory, this additional scrutiny of schools and teachers should help in aligning the goals and incentives of governments and parents with those of school leaders and teachers, leading to gains in student learning (Figlio & Loeb, 2011). Yet many are concerned about the impact that accountability is having upon the workload, well-being and mental health of school staff (Saeki et al., 2015) and if this is turning people away from the teaching profession (Ryan et al., 2017). Thus, although increasing accountability may bring about short-run improvements in student performance, this could be counterproductive in the long term if it reduces teacher supply, with shortages of high-quality teachers failing to keep up with demand.

Despite the widespread interest in accountability in education, previous research on how it is related to teacher stress and well-being is limited, particularly outside the USA. This paper has therefore explored this issue, using recently released data from TALIS 2018 and PISA 2018. Specifically, we have conceptualised accountability occurring at different levels, including both when looking at the whole school system (i.e. do countries with more accountability in the school system have more stressed teachers?) and within schools (e.g. how do headteachers hold staff to account within their school, and does their approach differ when they themselves feel stressed by being held to account?). This has, in turn, provided important new evidence on the correlates and consequences of accountability-induced stress that is occurring within schools across the world.

Our results suggest that there is a cross-national relationship between school system accountability and how stressed school staff feel about this aspect of their job. Yet the strength of this relationship is modest (correlation ≈ 0.3 , $p \approx 0.04$), with some clear examples of countries with extensive, data-driven accountability in schools where comparatively few teachers and school leaders say that they feel stressed. We also find there to be only a weak relationship between how stressed headteachers feel about accountability and the stress felt by staff. One potential explanation for this finding is that the management practices of headteachers who feel under pressure from accountability do not seem to differ much from those that do not feel stressed by this part of their work. However, there is clear evidence of ‘emotional contagion’ of accountability-induced stress amongst staff within schools; an individual is much more likely to feel under pressure from this aspect of their job if their colleagues do as well.

These findings should be interpreted in light of the limitations of this research. First, all of the analyses have been conducted using cross-sectional data, and have demonstrated the presence (or absence) of a correlation, rather than establishing causation. There is a pressing worldwide need for more longitudinal data on teachers, allowing researchers to monitor how their levels of stress and well-being changes as they get promoted, when school management changes or they move to another job. Such longitudinal data would allow researchers to generate stronger evidence of there being a causal relationship with respect to many of the research questions we have posed.

Second, stress due to accountability has been captured using a single question across a large number of countries. This question could suffer with issues of cross-national comparability; does being ‘stressed’ in one country mean the same as being ‘stressed’ in another? Further waves of TALIS might seek to ask additional questions—forming an ‘accountability stress scale’ with at least metric measurement invariance across countries—in order for us to better understand the pressure that this increasingly prominent factor is affecting the well-being of teachers at work.

Third, relatedly, the questions included in the TALIS questionnaire provide only limited information about some of the constructs of interest. For instance, TALIS did not include any questions capturing the aims that accountability systems are expected to serve, with formative-based accountability systems being almost non-existent in some countries. Fourth, the focus of this paper has been contracted forms of school accountability that serve a formative purpose. Yet we recognise other forms of school system accountability are possible—including professional mechanisms of accountability—which may have a different relationship with the stress felt by teachers. Fifth, as noted by an anonymous referee, we have not considered differences between centralised and decentralised school systems. Given that strong mechanisms of school accountability may be particularly needed in decentralised systems (where schools and teachers are expected to reach specific aims, but are free to decide how), such systems may also have a different relationship between accountability and teacher stress. Finally, in parts of our analysis, we have been faced with a limited sample size. Indeed, it is important to remember that only around 200 schools (and hence headteachers) are surveyed in each country. This means our ability to detect cross-national differences in the relationship between headteachers’ actions and the stress felt by their staff has been limited. Larger data collections—or the ability to combine information across multiple survey waves—will in the future help researchers generate more precise estimates of the link between headteacher actions/behaviour and teachers’ levels of stress.

Despite these limitations, we believe the findings presented in this paper may hold some important implications for education policy and practice. For government officials, it is important that they recognise that increasing accountability within the school system is unlikely to be a one-way street to ‘school improvement’. Although it may, according to previous research (e.g. Hanushek & Raymond, 2005), lead to increases in student test scores in the short run, our evidence suggests it might also be associated with higher levels of teacher stress, which could ultimately drive individuals out of the profession. This could, in turn, have negative implications for student achievement over a longer time horizon. Benevolent education policymakers must weigh up the risks and rewards of these possibilities before deciding whether to increase (or decrease) school system accountability is the best route for their country to follow. For organisations looking to improve the mental health of teachers—and reduce stress induced by accountability—our finding of ‘emotional contagion’ is likely to be relevant. In particular, it suggests that there will be specific schools where there is an atmosphere of stress amongst staff about accountability, and where it will be important for such organisations to intervene. It may also indicate that whole school approaches to reducing accountability stress amongst staff may be particularly efficient and effective,

with a reduction in the stress levels of one staff member likely to bring benefits to others. Finally, SMTs are fine to continue the common practice of reviewing student performance data as part of annual teacher appraisals; we find little evidence that this increases stress levels amongst staff. However, it is important that school leaders continue to use student performance data appropriately, and do not make inappropriate inferences about it capturing the ‘quality’ or ‘performance’ of any individual member of staff.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11092-021-09360-0>.

Acknowledgements The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and qualitative methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics and the Ada Lovelace Institute. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org. We are grateful for their support. Helpful comments have been received on the draft from our project steering group, who we would like to thank for their input and support.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Berryhill, J., Linney, J., & Fromewick, J. (2009). The effects of educational accountability on teachers: Are policies too stress provoking for their own good? *International Journal of Education Policy and Leadership*, 4(5), 1–13.
- Bitler, M., Corcoran, S., Domina, T. and Penner, E. (2019). *Teacher effects on student achievement and height: A cautionary tale*. NBER Working Paper No. 26480.
- Brill, F., Grayson, H., Kuhn, L. and O’Donnell, S. (2018). *What impact does accountability have on curriculum, standards and engagement in education? A literature review*. NFER. Accessed 09/01/2020 from https://www.nfer.ac.uk/media/3032/nfer_accountability_literature_review_2018.pdf.
- Byrd-Blake, M., Afolayan, M., Hunt, J., Fabunmi, M., Pryor, B., & Leander, R. (2010). Morale of teachers in high poverty schools: A post-NCLB mixed methods analysis. *Education and Urban Society*, 42, 450–472.
- Craig, C. (2017). International teacher attrition: Multiperspective views. *Teachers and Teaching*, 23(8), 859–862. <https://doi.org/10.1080/13540602.2017.1360860>.
- Donnelly, L. A., & Sadler, T. D. (2009). High school science teachers’ views of standards and accountability. *Science Education*, 93, 1050–1075. <https://doi.org/10.1002/sce.20347>.
- Figlio, D. and Loeb, S. (2011). School Accountability. pp 383–421 in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann (eds) *Handbooks in Economics* Volume 3, North-Holland.
- Foster, D. (2019). *Teacher recruitment and retention in England*. House of Commons briefing paper 7222. Accessed 17/07/2019 from <https://researchbriefings.files.parliament.uk/documents/CBP-7222/CBP-7222.pdf>.

- Gilbert, C. (2011). *Towards a self-improving system: The role of school accountability*. Accessed 09/01/2020 from <https://dera.ioe.ac.uk/14919/1/towards-a-self-improving-system-school-accountability-thinkpiece%5B1%5D.pdf>.
- Goldstein, H. (1997). Methods in school effectiveness research. *School effectiveness and school improvement*, 8, 369–395.
- Grossman, S., & Hart, O. (1983). An analysis of the principal-agent problem. *Econometrica*, 51(1), 7–46.
- Grupe, D., & Nitschke, J. (2013). Uncertainty and anticipation in anxiety. An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, 14(7), 488–501.
- Hamilton, L. and Koretz, D. (2002). Tests and their use in test-based accountability systems in L. Hamilton, B. Stecher and S. Klein (eds) *Making sense of test-based accountability in education*. RAND.
- Hanushek, E., & Raymond, M. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24(2), 297–327.
- Hatfield, E., Bensman, L., Thornton, P., & Rapson, R. (2014). New perspectives on emotional contagion: A review of classic and recent research on facial mimicry and contagion. *Interpersona: An International Journal on Personal Relationships*, 8, 159–179.
- Jerrim, J., & Sims, S. (2019). Teachers in primary and secondary schools: TALIS 2018. Department for Education research report. Accessed 09/04/2021 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/919064/TALIS_2018_research.pdf.
- Jones, B. D. & Egle, R. J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education Policy Analysis Archives*, 12(39). Accessed 20/01/2020 from <https://epaa.asu.edu/ojs/article/view/194/320>.
- Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly*, 24, 285–307.
- Karasek, R. A., & Theorell, T. (1990). *Healthy work: Stress, productivity, and the reconstruction of working life*. Basic Books.
- Lough, C. (2019). *England's heads 'the most data-driven'*. TES 15th October 2019. Accessed 09/01/2020 from <https://www.tes.com/news/englands-heads-most-data-driven>.
- Michie, S. (2002). Causes and management of stress at work. *Occupational and Environmental Medicine*, 59, 67–72.
- Micklewright, J., Jerrim, J., Vignoles, A., Jenkins, A., Allen, R., Ilie, S., Bellarbre, E., Barrera, F., & Hein, C. (2014). Teachers in England's secondary schools: Evidence from TALIS 2013. Department for Education Research Report DFE-RR302. Accessed 09/04/2021 from <https://www.gov.uk/government/publications/teachers-in-secondary-schools-evidence-from-talis-2013>.
- Oberle, E. & Schonert-Reichl, K. (2016). Stress contagion in the classroom? The link between classroom teacher burnout and morning cortisol in elementary school students. *Social Science & Medicine*, 159. <https://doi.org/10.1016/j.socscimed.2016.04.031>.
- OECD. (2011). *PISA in focus 9. School autonomy and accountability: Are they related to student performance?*. Accessed 09/01/2020 from <https://www.oecd.org/pisa/pisaproducts/pisainfocus/48910490.pdf>.
- OECD. (2019). *TALIS 2018 Technical Report*. OECD. Accessed 20/01/2020 from https://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf.
- Perryman, J. and Calvert, G. (2019). What motivates people to teach, and why do they leave? Accountability, performativity and teacher retention. *British Journal of Educational Studies*. <https://doi.org/10.1080/00071005.2019.1589417>.
- Relias Media. (1998). *Lack of job autonomy can have deadly effect*. Accessed 09/01/2020 from <https://www.reliasmedia.com/articles/33797-lack-of-job-autonomy-can-have-deadly-effect>.
- Richards, J. (2012). Teacher stress and coping strategies: A national snapshot. *The Educational Forum*, 76(3), 299–316.
- Ryan, S., von der Embse, N., Pendergast, L., Saeki, E., Segool, N., & Schwing, S. (2017). Leaving the teaching profession: The role of teacher stress and educational accountability policies on turnover intent. *Teaching and Teacher Education*, 66, 1–11.
- Saeki, E., Pendergast, L., Segool, N., & von der Embse, N. (2015). Potential psychosocial and instructional consequences of the common core state standards: Implications for research and practice. *Contemporary School Psychology*, 19(2), 89–97.
- Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research*, 61(3), 257–273. <https://doi.org/10.1080/00131881.2019.1625716>.

- Smith, M. and Amick, B. (1989). Electronic monitoring at the workplace: Implications for employee control and job stress pp 275–289 in S. L. Sauter, J. J. Hurrell Jr and C. L. Cooper (eds) *Job Control and Worker Health*. Wiley.
- Snijders, T., & Bosker, R. (1994). Modeled variance in two-level models. *Sociological Methods and Research*, 22(3), 342–363.
- Spector, P. (1986). Perceived control by employees: A meta-analysis of studies concerning autonomy and participation at work. *Human Relations*, 39(11), 1005–1016.
- Sy, T., Cote, S., & Saavedra, R. (2005). The contagious leader: Impact of the leader's mood on the mood of group members, group affective tone, and group processes. *Journal of Applied Psychology*, 90(2), 295–305.
- Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, 44(3), 519–558.
- Weston, D. (2011). *Taking control: Why autonomy reduces stress*. Accessed 09/01/2020 from www.inform.ededucation.com/taking-control-why-autonomy-reduces-stress/.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.