The Incredible Years® Teacher Classroom Management programme and its impact on

teachers' professional self-efficacy, work related stress and general well-being: results from

the STARS randomised controlled trial

Rachel Hayes[1], Daniel Titheradge[1], Kate Allen [1], Matt Allwood [1], Sarah Byford[2], Vanessa

Edwards[1], Lorraine Hansford [1], Bryony Longdon [1], Shelley Norman [1], Brahm Norwich [3],

Abigail Emma Russell[1], Anna Price[1], Obioha C Ukoumunne[4], Tamsin Ford[1]

1. University of Exeter Medical School, South Cloisters, St Luke's Campus, Exeter,

   EX1 2LU

2. King's College London, King's Health Economics

3. Graduate School of Education, University of Exeter, North Cloisters, St Luke's

   Campus, Exeter, EX1 2LU

4. NIHR CLAHRC South West Peninsula (PenCLAHRC), University of Exeter, South

   Cloisters, St Luke's Campus, Exeter, EX1 2LU

* Corresponding author:

1. Dr Rachel Hayes, College of Medicine and Health, rm 2.05 College House, St Luke's

   Campus, Exeter, EX1 2LU

   Tel: +44 (0) 1392 722978

   E-mail:R.A.Hayes@exeter.ac.uk

Conflicts of interest: The authors have no conflicts to declare.

1

KEY WORDS

Teacher mental health; burnout; self-efficacy; classroom management

ABSTRACT

Background

Teaching is a stressful occupation with poor retention. The Incredible Years® Teacher Classroom Management (TCM) programme is a training program that past research has demonstrated may be an effective intervention for children's mental health, but little research has explored any impacts there may be on the teachers' own professional confidence and mental health.

Aims

In this paper we evaluate whether TCM may lead to changes in teachers' wellbeing, namely a reduction in burnout and an improvement in self-efficacy and mental health.

Sample

Eighty schools across the South West of England were recruited between September 2012 and September 2014. Headteachers were asked to nominate one class teacher to take part.

Methods

Eighty teachers were randomised to either attend a TCM course (intervention) or not (control). TCM was delivered to groups of up to twelve teachers in six whole-day workshops that were evenly spread between October and April. At baseline and nine months follow-up

we measured teachers' mental health using the Everyday Feelings Questionnaire (EFQ), burnout using the Maslach Burnout Inventory-General Survey (MBI-GS) and self-efficacy using the Teachers' Sense of Efficacy Scale-Short (TSES-Short).

Results

Using linear regression models there was little evidence of differences at follow-up between the intervention and control teachers on the outcomes (the smallest p-value was 0.09).

Conclusions

Our findings did not replicate previous research that TCM improved teachers' sense of efficacy. However, there were limitations with this study including low sample size.

Words in abstract: 235/250

INTRODUCTION

Teaching is commonly acknowledged to be an extremely stressful occupation, with poor and possibly worsening retention in the UK (Industrial Injuries Advisory Council, 2017; Worth, Bamford, & Durbin, 2015). The prolonged response to chronic emotional and interpersonal stressors in the workplace may lead to "*burnout*": a syndrome of exhaustion, cynicism and reduced professional self-efficacy, which may underlie or exacerbate poor mental health and exit from the profession (Maslach, Schaufeli, & Leiter, 2001).

Mental illness is one of the leading causes of disability (OECD, 2014) and a major cause of absence from work due to sickness in the UK (Office of National Statistics, 2017). The economic impact of poor mental health is not just through lost working days and absenteeism, but also through "presenteeism" (reduced productivity of employees still present at work). Several studies report high rates of psychological distress among teachers at both primary and secondary schools in comparison with the general population (Hinz et al., 2014; Kidger et al., 2016; Titheradge et al., 2019). The Well-being In Secondary Education (WISE) study, which involved 555 UK secondary school teachers, detected high levels of moderate to severe depression (19.4%); reporting depression was associated with being female, dissatisfaction with work, presenteeism, sickness absence, interpersonal difficulties, and low pupil attendance (Kidger et al., 2016). The educational and care sectors have the highest rates of presenteeism of any employment sector, which in turn, predicts high levels of absenteeism (Aronsson, Gustafsson, & Dallner, 2000).

Teachers in the UK are required to respond rapidly to shifts in policy, of which there have been many over the last decade, which may impact on resilience, morale and self-efficacy (Day & Gu, 2007). Alongside lost productivity due to disability, sickness, and presenteeism, the staffing of schools is also challenged by the premature loss of experienced teachers from the workforce, which may become an increasing issue as pupil numbers are expected to rise in the UK (Chiong, Menzies, & Parameshwaran, 2017; Worth et al., 2015). Poor mental health is a leading cause of exit from the teaching profession; for example, it was cited by 46% and 37% of Irish and Scottish teachers, respectively, as the reason for leaving teaching (Brown, Gilmour, & Macdonald, 2006; Maguire & O'Connell, 2007). A survey of more than 900 teachers who had taught for a decade or more in the UK suggested that both school level and policy level influences impacted on job satisfaction and retention, but that the two most important factors were professional mastery and altruistic factors (Chiong et al., 2017).

Pupil misbehaviour significantly interferes with a teacher's ability to teach their class. It was reported that an average of 13% of lesson time was required to maintain classroom control among teachers from 23 different countries, while up to 25% of teachers reported regularly losing at least 30% of their lesson time to managing disruptions or administrative tasks (OECD, 2009). Analysis of data from the Longitudinal Study of Young People in England (LSYPE) found that, after controlling for a wide range of pupil and school factors, self-reported classroom misbehaviour reduced a student's GCSE scores by the equivalent of one grade in five subjects (Department of Education, 2012a).Teachers report that disruptive behaviour and the task of managing the classroom can lead to high levels of stress and burnout (Beltman, Mansfield, & Price, 2011; Brouwers & Tomic, 2000; Collie, Shapka, &

Perry, 2012; Kokkinos, 2007) and they also comment on a lack of training in this area (Merrett & Wheldall, 1993).

Jennings and Greenberg (2009) proposed a 'burnout cascade', whereby a teacher's difficulties with behavioural management in the classroom leads to a decreased sense of self-efficacy that results in negative consequences for their relationship with the children, a more reactive and more negative classroom environment that amplifies disruptive behaviour and in turn increases distress and burnout for the teacher. Stressed teachers have more negative interactions with their pupils, but when supported with behaviour management training they report reduced personal emotional difficulties and decreased disruptive behaviour among their pupils (McGilloway et al., 2011). A meta-analysis of 16 studies suggested that self-efficacy in relation to classroom management was related to burnout (Aloe, Amo, & Shanahan, 2014), which in turn predicted reported pupil misbehaviour (Aloe, Shisler, Norris, Nickerson, & Rinker, 2014). Teachers with less developed classroom management skills have been found to have classrooms with higher reported overall levels of child aggression, peer rejection and exclusion among their pupils (Kellam et al., 2008; Kellam, Ling, Merisca, Brown, & Ialongo, 1998).

In contrast, teachers with highly developed classroom management skills may obtain better results both academically and socially, and may reverse the "burnout cascade" to the benefit of themselves and their pupils (Jennings & Greenberg, 2009; Webster-Stratton, Reid, & Stoolmiller, 2008). Teachers' mental health is an important issue; and may also adversely impact on the mental health, attendance and academic attainment of their pupils (Kidger et al., 2016; McLean & Connor, 2015; Oberle & Schonert-Reichl, 2016). Further to this, in a

study of the School-Wide Positive Behaviour Interventions and Supports program, teacher-reported burnout was lower in schools where the intervention was delivered with the highest fidelity (Ross, Romer, & Horner, 2012).

Therefore, an intervention that supports teachers to manage disruptive behaviour and promote socio-emotional competence could potentially benefit not only their pupils, but also the teachers themselves. The Incredible Years® (IY) foundation has developed three interlinked programmes for parents, teachers and children that aim to promote children's social, emotional and academic competence (Herman, Borden, Reinke, & Webster-Stratton, 2011). The Teacher Classroom Management (TCM) course has been identified as the most promising school-based intervention with the best evidence for improving children's mental health (Nye, 2017; Pidano & Allen, 2015; Whear et al., 2012). A recent meta-analysis (Nye, 2017) demonstrated a small, statistically significant effect of the TCM intervention on reducing child conduct problems (effect size (Hedges g) = -0.20, 95% CI: -0.38 to -0.01). The Supporting Teachers And childRen in Schools (STARS) trial (Ford et al., 2018) compared the mental health of over 2,000 children across 80 UK schools and found that TCM training reduced the Strengths and Difficulties Questionnaire Total Difficulties score (SDQ-TD) (mean difference = -1.0; 95% CI: -1.9 to -0.1). This reduction was greater for children whose baseline SDQ-TD indicated that they scored above the 80th centile for the British school age population (>11; classified as *struggling;* mean difference = -2·6; 95% CI: -4·6 to -0·6) (Goodman, 2015). The STARS study included a parallel process evaluation (Ford et al., 2018) where teachers were invited to attend focus groups and interviews to share their experiences of the training, what facilitated or hindered change in the classroom and asking about any personal impacts the teachers had experienced. Murray, Rabiner, Kuhn, Pan, and

Sabet (2018) compared the classrooms of 91 elementary teachers in North Carolina and found that TCM made a positive impact on independently observed school climate but did not produce any universal improvements in child outcomes. However, a priori subgroup analysis also suggested that children scoring above the SDQ-TD cut-point of 12 did benefit from exposure to TCM-trained teachers.

The focus of the TCM training is on collaborative learning, reflections about teachers' own experiences and group work to find solutions to problems encountered in the classroom, with time between each workshop for teachers to practice the new strategies they have learned (Webster-Stratton, Reinke, Herman, & Newcomer, 2011). The key concepts which are covered in each of the six TCM workshops are outlined in Table 1. TCM draws on cognitive social learning theory, particularly Patterson's theories (Patterson, 1982) about how coercive cycles of interaction between adults and children reinforce unwanted behaviour patterns, Bandura's ideas (Bandura, 1977) about the importance of modelling and self-efficacy, and Piaget's developmental interactive learning methods (Piaget & Inhelder, 1962). In addition, it also incorporates strategies for challenging angry, negative, and depressive internal dialogue in adults whilst interacting with children, drawn from cognitive behavioural approaches (Beck, 1976).

Our feasibility study which involved 40 UK primary school teachers (Marlow et al., 2015), reported that the most important aspects of the training were sharing experiences, the support of colleagues in the group and time out to reflect on, practice and rehearse new techniques. This study did not include a control group, but teachers' scores on the Teachers Sense of Efficacy scale (Tschannen-Moran & Wollfolk, 2001) suggested perceived improvements in

classroom management (mean change pre- to post-: 2.1, 95% CI: 0.4 to 3.8) and in professional efficacy on the Maslach Burn Out Inventory (mean change pre- to post-: 2.3, 95% CI: -0.3 to 4.9). No other randomised trials of TCM have measured the potential impact on teachers' mental health or professional confidence (Hickey et al., 2017; Hutchings, Martin-Forbes, Daley, & Williams, 2013; Martin, 2009; Murray et al., 2018; Reinke, Herman, & Dong, Under revision).

Rational for study

Given that teachers experience high levels of work-based stress and have been shown to often have poor mental health alongside an increased subsequent potential for burn-out, reduced self-efficacy, absenteeism and ultimately exit from the profession, it is important to test interventions that could support teachers in these domains. The STARS (Supporting Teachers and childRen in Schools) study was a five year, two-arm, pragmatic, parallel group, superiority, cluster randomised controlled trial designed to evaluate whether the TCM course delivered to teachers improved the mental health of individual children as well as investigating any personal effects on teachers (Ford et al., 2018). This paper reports teacher outcomes from the STARS study to test the following hypothesis: exposure to TCM training would improve teachers' mental health and sense of self-efficacy and reduce their levels of burnout.

METHODS

The trial is reported in accordance with CONSORT and TIDieR guidelines (Campbell, Piaggio, Elbourne, Altman, & Grp, 2012; Hoffmann et al., 2014; Schulz, Moher, & Altman,

2010). The study design and procedures are presented in full in the published trial protocol (Ford et al., 2012) which was approved by the Trial Steering Committee (TSC) and Data Monitoring Committee (DMC). Ethical approval for the conduct of the trial was obtained from the Peninsula College of Medicine and Dentistry Research Ethics Committee (12/03/141).

Study design and participants

Primary school teachers were recruited in three separate cohorts for baseline data collection in September 2012 (Cohort 1), September 2013 (Cohort 2) and September 2014 (Cohort 3) from schools across the South West of England (see Figure 1). Schools were approached through unsolicited contact with headteachers and publicity at local education conferences. To be eligible for inclusion, schools needed a single-year class with 15 or more children in Reception up to Year 4 (children aged between 4 and 8 years), with a teacher who held classroom responsibility for at least four days per week. Schools were excluded if they primarily taught pupils with special educational needs, had an acting headteacher or none in post, or were judged as inadequate and requiring "special measures" (additional support) in their last Office for Standards in Education, Children's Services and Skills (Ofsted) inspection. Headteachers were asked to nominate one class teacher that met the inclusion criteria and this class teacher was then invited to take part in the STARS trial and be randomised to either attend a TCM course or not. Headteachers nominated class teachers independently of the research team and for a variety of different reasons, including: being newly qualified, allocation of a class with known behavioural challenges or because the class teacher had a particular interest in behaviour management. Written consent was obtained

from the headteacher for the school's participation and from the class teacher for their involvement after nomination by their headteacher.

Randomisation and masking

Randomisation at the school level using computer generated random numbers was completed after baseline data collection to avoid recruitment and response bias (Eldridge, Kerry, & Torgerson, 2009). It was conducted by an independent researcher based at the University of Exeter who was masked to the identity of the schools to ensure allocation concealment. Allocation was balanced on the following school factors: urban versus rural/semi-rural area; Key Stage 1 (KS1: Reception to Year 2, aged 4-7) versus Key Stage 2 (KS2: Year 3 or 4, aged 8-9); and deprivation (whether % of children eligible for free school meals was greater than 19%, the UK national average in 2012 (Department of Education, 2012b)).

We were unable to mask allocation to the schools or teachers, since the school needed to release the class teacher to attend the training. The main research team were not masked as feasibility work indicated that visual cues in the classroom and enthusiastic comments from teachers would undermine attempts to do so. However, teachers completed their measures independently of researchers directly onto an on-line web based database, so the opportunity for researchers to influence teachers' responses was minimal.

Procedures and Intervention

TCM was evaluated in isolation from other IY® Programs and was delivered to groups of up to twelve teachers in six whole-day workshops that were evenly spread between October and

April of the same academic year. The workshops took place during the school day at a venue external to the teachers' schools. The facilitating group leaders, who delivered the training in pairs, were behaviour support practitioners with a teaching background, had completed the mandatory three-day basic TCM delivery training, and had led at least two previous courses prior to the start of the trial. No restrictions were placed on schools regarding access to other training and support services during the course of the study.

As part of our parallel process evaluation (Ford et al., In press) all teachers who attended the TCM training were also invited to take part in a focus group immediately after their training had finished. The focus groups were explained to teachers as an opportunity to explore their experiences and views of the TCM training such as how the course was run, positive and negative aspects of the course, and whether or not they had used or discussed the TCM strategies with colleagues.

Group leaders were supervised by the TCM programme developer, Carolyn Webster Stratton (CWS), to ensure fidelity to the TCM model. All TCM workshops were video-recorded for use as part of the supervision process. After each workshop group leaders completed IY® standardised 'agendas and checklists' to report which 'activities' they had delivered during the workshop. They also selected a 10-30 minute video section of the day (either one long clip or a variety of shorter clips) to be reviewed by CWS. The videos could be examples of a section they felt went well, or found challenging. These materials were sent to CWS who returned a detailed review of the video clips with comments about how to improve the delivery of the course that were subsequently discussed in a one hour video conference call

with all six group leaders. The supervision process was very successful and CWS was confident that the group leaders were delivering the course with fidelity to the model.

Outcomes

Teachers completed the Everyday Feelings Questionnaire (EFQ) (Uher & Goodman, 2010), Teachers' Sense of Efficacy Scale-Short (TSES-Short) (Tschannen-Moran & Wollfolk, 2001) and the Maslach Burnout Inventory-General Survey (MBI-GS) (Maslach, Jackson, & Leiter, 1996) on two occasions: once at the beginning of the academic year in September/October (baseline) before training began and once in the final half-term of the same academic year in June two months after training ended (follow-up). Whilst additional timepoints during the training year would have been beneficial to track potential change, including any mediators of this change, we were mindful of the need to reduce the response burden on teachers and therefore chose to limit data collection to just pre- and post-training.

The EFQ is a 10-item validated measure (Uher & Goodman, 2010) that records how the respondent has felt over the previous four weeks. Half of the items focus on well-being and half on distress. Items are scored from 0 to 4 for items with distress content and from 4 to 0 for items with wellbeing content. The EFQ has a maximum score of 40; a higher score indicates higher levels of distress with scores above 19 being indicative of at least moderate levels of clinical depression (Titheradge et al., 2019). The mean (SD) score for a population-representative sample of 2,109 professionals was 11.4 (SD 5.9) (Green, McGinnity, Meltzer, Ford, & Goodman, 2005). Uher and Goodman (2010) demonstrated in a large population-based sample of adults that the EFQ was internally consistent, with Cronbach's alpha 0.89

and average inter-item covariance 0.42. The corrected item-total correlations ranged between 0.52 and 0.71. When the EFQ was compared to the twelve-item version of the General Health Questionnaire, a brief validated measure of the symptoms of common mental disorder (Goldberg et al., 1997), concurrent validity was assessed as good, with a correlation between the measures of 0.74 (95% CI: 0.73 to 0.75).

The TSES is a validated measure that can be used either as the long (24 items) or the shortened 12-item version (Tschannen-Moran & Wollfolk, 2001). In order to reduce burden on the teachers, we chose to use the shortened version to assess the teacher's perception of their sense of effectiveness. The TSES-Short contains three subscales, each with 4 items, Student Engagement, Instructional Practice and Classroom Management. Responses are rated on a nine-point scale for each item with anchors at 1 (*nothing*), 3 (*very little*), 5 (*some influence*), 7 (*quite a bit*) and 9 (*a great deal*). Mean scores are calculated for each subscale with a higher score indicating a greater sense of efficacy. A sample (N=410) of teachers reported the following average sub-scale scores: 7.2 for Student Engagement, 7.3 for Instructional Practice and 6.7 for Classroom Management. Internal consistency is very good with a Cronbach's alpha of 0.90 overall and ranging from 0.81 to 0.86 for the subscales (Tschannen-Moran & Wollfolk, 2001). Construct validity was demonstrated by comparing the TSES-Short with two other measures of teacher efficacy by Kerlinger (1986) and Hoy and Woolfolk (1993) with significant correlations between the scales (r = 0.18 and 0.52, p<0.01) as well as with both the personal teaching efficacy factor (r = 0.61; p<0.01) and the general teacher efficacy (GTE) factor (r = 0.16; p<0.01).

The MBI-GS is a leading measure of burn-out, consisting of 16-items across three separate subscales: Exhaustion, Cynicism and Professional Efficacy. Respondents choose from seven options: 0 (*never*), 1 (*sporadic*), 2 (*now and then*), 3 (*regular*), 4 (*often*), 5 (*very often*) and 6 (*daily*). Mean scores are calculated for each subscale with a resultant range of 0 to 6. A high degree of burnout is reflected in high scores on Exhaustion and Cynicism and low scores on Professional Efficacy. A large population sample (N=9055) found that the average sub-scale scores were as follows: 1.48 for Exhaustion, 4.66 for Professional Efficacy and 1.48 for Cynicism (Schutte, Toppinen, Kalimo, & Schaufeli, 2010). The MBI-GS has high internal consistency for all three sub-scales with Cronbach's alpha 0.86 for Exhaustion, 0.83 for Professional Efficacy and 0.75 for Cynicism (Schutte et al., 2010). Maslach et al. (1996) demonstrated good external validity for each of the sub-scales.

Table 1 here

Statistical analysis

Although randomisation occurred at the school level, since only one teacher from each school was recruited and we are analysing teacher specific outcomes, there is no need to allow for correlation between responses from the same school.

We compared the intervention and control arms using the intention-to-treat principle, analysing the teachers according to the trial arm they were randomised to. Complete case analyses were carried out (i.e., missing data were not imputed). Linear regression models were fitted to analyse the follow-up EFQ, TSES-Short and MBI-GS measures, adjusting for

the randomisation balancing factors (urban versus rural status, KS1 versus KS2 status, % of children on free school meals), study cohort (whether recruited in 2012, 2013 or 2014) and the baseline outcome score. In addition, we investigated whether the following variables were predictive of the outcome: school level Index of Multiple Deprivation (IMD), % of children identified as having special education needs at the class level, teacher's gender, and whether the teacher had more than five years of teaching experience. Where they were predictive we also adjusted for them. All analyses were performed using Stata software v14.2 (StataCorp, 2015).

RESULTS

We recruited a total of 80 trial teachers (one from each of 80 schools); 40 were randomised to attend a TCM training programme and 40 to the control arm (Figure 1). Compared with the national average (Department of Education, 2012b), participating schools had similar class sizes (means 27.4 vs 26.8) and eligibility for free school meals (18.3% vs 19.3%), but included fewer voluntary controlled schools (5% vs 14.4%), and more community (61.3% vs 55.3%) and academy schools (10% vs 6%). TCM training was well attended; 36 (90%) of the 40 teachers in the intervention arm attended four or more TCM workshops; 23 (58%) attended all six workshops.

Baseline data were collected for all 80 teachers and characteristics were generally balanced between the two arms (Table 2), although control teachers tended to be more experienced (68% vs 50% had taught for 5 or more years) and have lower EFQ scores (13.9 vs 17.2) than

intervention teachers. Teachers in the intervention arm also reported lower self-efficacy and higher levels of burnout at baseline. Our analyses adjusted for these baseline differences.

Both intervention and control teachers scored quite highly on the exhaustion sub-scale of the MBI-GS (mean (SD) 2.9 (1.4) and 2.5 (1.4) respectively). In a population study of over 9000 teachers the mean score on this sub-scale was 1.48 (Schutte et al., 2010). Similarly, the mean EFQ total score was quite high amongst the teachers (intervention mean=17.2, control mean=13.9) compared to an otherwise similar population of professionals (mean=11.4) (Green et al., 2005).

Follow-up data on the EFQ, TSES-Short and MBI-GS were collected for 74 teachers (93%). Three teachers were lost to follow-up as they no longer worked at the school (2 control and 1 intervention), one control teacher was absent due to long-term sickness (1) and two intervention teachers were on maternity leave (2).

Figure 1 here

Table 3 summarises the comparison between the trial arms at follow-up for the three sub-scales of the Teacher Self-Efficacy Scale, the 3 sub-scales of the Maslach Burn-Out Inventory and for the Everyday Feelings Questionnaire. After adjusting for the baseline score and other factors, there were no statistically significant differences between the intervention and control teachers' on these measures.

Table 3 here

DISCUSSION

Our results do not suggest that TCM training reduces teacher-reported 'burnout', nor that it improves self-efficacy or wellbeing. Despite this, teachers from the STARS process evaluation reported that the TCM training had helped them to create a more positive cycle of behaviour through modelling and led to a range of important impacts on their practice (Ford et al., in press). These included helping them see things from the child's point of view, which they believed improved their relationship with the children; being able to take time to think more before responding, as well as enabling them to feel more positive, confident and in control.

These quantitative findings are therefore surprising, particularly in the context of the previous feasibility work (Marlow et al., 2015) that suggested that TCM may lead to improved self-efficacy. The feasibility study used the original longer 24-item version of the TSES scale (Gibson & Dembo, 1984; Marlow et al., 2015) whilst the main STARS trial used the short form of the TSES scale. The decision to use the short form in the main STARS study was taken to minimise teacher burden, but this shorter scale may have been insufficiently sensitive. We would recommend that future studies use the longer version.

Participating teachers also commented that whilst they were able to identify positive changes as a result of the training, these were restricted directly to their teaching and relationships with the children as opposed to the other non-teaching activities present in their role (Ford et al., in press). Teachers also reported that changes in the wider context of their professional lives were having a greater and more negative impact on their wellbeing (Ford et al., in

press). Our measure of psychological distress and burnout reported on teachers experience in general, although arguably all three subscales of the self-efficacy measure (instructional practice, classroom management and student engagement) focus on the classroom.

There have been a large number of changes to the English education system since 2011/12 when Marlow et al. (2015)'s original study was completed (Roberts, 2017), and the inevitable disruption of their implementation may have reduced the potential impact of TCM on teachers well-being. Many teachers who participated in the process evaluation commented on the increasing pressures they faced at work, which may have contributed to the shockingly high and sustained levels of psychological distress that we detected. Titheradge et al. (2019) demonstrated that a large percentage of teachers who participated in STARS had an EFQ score in the range indicative of moderate depression (EFQ 20 or above) at baseline (29%) and at follow-up (20%).

STARS was powered to detect a difference in the mental health of the children, while teacher outcomes were important secondary measures. Therefore, we may be missing important changes that occur in teachers due to a lack of statistical power. Given the width of the confidence intervals in our analyses we cannot rule out the possibility of effects that would be meaningful if they could be tested in a larger sample.

In our process evaluation teachers reported varying experiences of how well TCM fitted with the wider school behaviour policy (Ford et al., in press). Some teachers reported difficulties deploying it effectively within their classrooms if teaching assistants did not agree or follow the same strategies (Ford et al., 2018). A consistent whole-school approach to TCM, that

involves teaching and playground assistants as well as teachers, may be more effective, particularly as a recent trial of two evidence-based approaches to children's behavioural management suggested that organisational level factors were the strongest predictors of teachers' stress and job satisfaction (Ouellette et al., 2017). It is intuitively plausible that if behavioural management strategies perceived to be helpful and effective were conflicting with the wider school approach, access to TCM training might undermine rather than improve teacher well-being.

Considering Jennings and Greenberg (2009)'s proposed 'burnout cascade', whereby difficulties managing the behaviour of pupils led to a reduced sense of self-efficacy, which in turn results in poorer quality relationships with pupils and the subsequent amplification of disruptive behaviour, it is feasible that personal impacts for teachers will only begin to emerge after positive classroom behaviours have become firmly embedded. This means that these changes may not be apparent until sometime after TCM training has been completed. In this study we only have teacher measures two months after training ended, it is possible this is not long enough for the 'burnout cascade' to be reduced.

Study Strengths, Limitations and Further Research

This study benefits from high retention of teachers over the follow-up period, the delivery of TCM with fidelity by experienced practitioners with a background in teaching who were trained and supervised by the programme developer, concealed randomisation, and the use of strongly validated and widely used outcome measures. Trial arms were balanced in terms of school characteristics. There were, however, imbalances in terms of teachers (higher levels of

psychological distress, lower self-efficacy, fewer qualified for more than five years and older in the intervention arm), which may have made it harder to detect an effect on teacher outcomes.

High levels of attendance suggest that teachers valued TCM, while the participating schools were generalisable to the UK population in terms of class size and eligibility for free school meals.

Like all studies, there are some methodological limitations. It was not possible to mask teachers to their allocation, risking response bias. The exclusion of schools that lacked a substantive headteacher, or were judged as failing in their last OFSTED inspection was adopted to protect the internal validity of the study. The burden of research was seen as inappropriate for struggling schools by our advisory group, who also reported that many schools in this situation change senior management team, with a high likelihood that the incoming leads might withdraw the school from the study. In addition, such schools would already have had the involvement of a number of external organisations who would be requiring changes and we wanted to evaluate TCM without these additional influences. This choice reduces the generalisability of STARS as it is highly likely that such schools experience greater difficulty in managing classroom behaviour.

In our feasibility work headteachers were very clear that they needed to make the decision about which teacher took part in the trial, they felt that any attempt to control the selection of teachers would be a major disincentive to their school's participation in the study. This opens up the possibility of two potential biases. If teachers were selected because they struggled

with behaviour management, we might overestimate the impact of the intervention, while if selected because of a particular interest in behavioural management, we might underestimate the impact if interest correlates with skills, or overestimate the impact if interest correlates with receptiveness. However, as selection of teachers preceded randomisation, it should not have compromised the internal validity of the study and reasonable balance was obtained on teacher characteristics (see Table 2). Our process evaluation (Ford et al., In press) involved interviews with headteachers and suggested a number of reasons for their choice of teacher to nominate, which included newly qualified teacher status, allocation of a class known to be particularly challenging or known interest in behaviour management. Only training one teacher in each school is also a limitation of this study, particularly if organisational level support was not put in place to support the teacher's implementation of their recent training.

Our process evaluation suggested that the major focus of change for teachers occurred in the classroom, which in their opinion was primarily driven by large improvements with their relationships with pupils (Ford et al., in press). These teachers also reported that TCM did not alleviate the sadly plentiful challenges outside the classroom that they experienced as major sources of stress and distress. We lacked an explicit measure of teacher-pupil relationships and given that our parallel process evaluation suggests that improved teacher pupil-relationships were a key component of change in children's outcomes (Ford et al., in press), we would strongly recommend that future studies include one. Poor teacher-pupil relationships predict subsequent poor child mental health, particularly the onset of behaviour problems, and are also related to poor family function (primary school age only) and future exclusion (secondary school age children only) (Lang et al., 2013) as well as poor academic attainment (Cadima et al., 2015).

Conclusions

Our results do not support previous work that suggested TCM improves teachers' well-being in relation to mental health, self-efficacy or burn out. While we could not recommend TCM as an intervention to support teachers, it does appear to be an effective intervention for child mental health, particularly among those whose behaviour causes significant difficulties (Ford et al., 2018; Murray et al., 2018; Nye, 2017). However, the parallel process evaluation suggests that teachers did find it beneficial for both their classroom management skills and that it reduced stress emanating from disruptive behaviour within the classroom, primarily by improving their relationships with the children. Future studies should include more teachers so as to be adequately powered for teacher outcomes, as well as assess teacher-pupil relationships as a potential mechanism of action.

REFERENCES

Aloe, A. M., Amo, L., C., & Shanahan, M., E. (2014). Classroom Management Self-Efficacy and Burnout: A Multivariate Meta-analysis. *Educational Psychology Review, 26*(1), 101-126. doi:10.1007/s10648-013-9244-0

Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B., & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review, 12*, 30-44. doi:http://dx.doi.org/10.1016/j.edurev.2014.05.003

Aronsson, G., Gustafsson, K., & Dallner, M. (2000). Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Community Health, 54*(7), 502-509.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev, 84*(2), 191-215.

Beck, A. T. (1976). *Cognitive therapy and emotional disorders*. New York: Meridian.

Beltman, S., Mansfield, C., & Price, A. (2011). Thriving not just surviving: A review of research on teacher resilience. *Educational Research Review, 6*(3), 185-207.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education, 16*(2), 239-253. doi:Doi 10.1016/S0742-051x(99)00057-8

Brown, J., Gilmour, W. H., & Macdonald, E. B. (2006). Ill health retirement in Scottish teachers: process, outcomes and re-employment. *Int Arch Occup Environ Health, 79*(5), 433-440. doi:10.1007/s00420-005-0060-9

Cadima, J., Doumen, S., Verschueren, K., & Buyse, E. (2015). Child engagement in the transition to school: Contributions of self-regulation, teacher–child relationships and classroom climate. *Early Childhood Research Quarterly, 32*, 1-12. doi:https://doi.org/10.1016/j.ecresq.2015.01.008

Campbell, M. K., Piaggio, G., Elbourne, D. R., Altman, D. G., & Grp, C. (2012). Consort 2010
statement: extension to cluster randomised trials. *British Medical Journal, 345*. doi:DOI
10.1136/bmj.e5661

Chiong, C., Menzies, L., & Parameshwaran, M. (2017). Why do long-serving teachers stay in the
teaching profession? Analysing the motivations of teachers with 10 or more years' experience
in England. *British Educational Research Journal, 43*(6), 1083-1110. doi:10.1002/berj.3302

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social–emotional learning:
Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational
Psychology, 104*(4), 1189.

Day, C., & Gu, Q. (2007). Variations in the conditions for teachers' professional learning and
development: sustaining commitment and effectiveness over a career. *Oxford Review of
Education, 33*(4), 423-443. doi:10.1080/03054980701450746

Department of Education. (2012a). *Pupil behaviour in schools in England*. (DFE-RR218). London.

Department of Education. (2012b). *Statistical First Release: Schools, Pupils, and their
Characteristics, January 2012*. (SFR 10 2012). London, UK.

Eldridge, S., Kerry, S., & Torgerson, D. J. (2009). Bias in identifying and recruiting participants in
cluster randomised trials: what can be done? *British Medical Journal, 339*. doi:DOI
10.1136/bmj.b4006

Ford, T., Edwards, V., Sharkey, S., Ukoumunne, O. C., Byford, S., Norwich, B., & Logan, S. (2012).
Supporting teachers and children in schools: the effectiveness and cost-effectiveness of the
Incredible Years teacher classroom management programme in primary school children: a
cluster randomised controlled trial, with parallel economic and process evaluations. *BMC
Public Health, 12*, 719. doi:10.1186/1471-2458-12-719

Ford, T., Hayes, R., Byford, S., Edwards, V., Fletcher, M., Logan, S., . . . Ukoumunne, O. C. (2018).
The effectiveness and cost-effectiveness of the Incredible Years® Teacher Classroom

Management programme in primary school children: results of the STARS cluster

randomised controlled trial. *Psychol Med*. doi:10.1017/S0033291718001484

Ford, T., Hayes, R. A., Byford, S., Edwards, V., Fletcher, M., Logan, S., . . . Ukoumunne, O. (In

press). The effectiveness and cost-effectiveness of the Incredible Years® Teacher Classroom

Management programme in primary school children: results of the STARS cluster

randomised controlled trial with parallel economic and process evaluations. *National Institute*

for Health Research.

Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational*

Psychology, 76(4), 569-582. doi:10.1037/0022-0663.76.4.569

Goldberg, D. P., Gatrer, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997).

The validity of two versions of the GHQ in the WHO study of mental illness in general health

care. *Psychological Medicine, 27*, 191-197.

Goodman, R. (2015). Scoring the Strengths & Difficulties Questionnaire for age 4-17. Retrieved from

http://www.sdqinfo.org/py/sdqinfo/c0.py

Green, H., McGinnity, A., Meltzer, H., Ford, T., & Goodman, R. (2005). Mental health of children

and young people in Great Britain, 2004. In. London: Palgrave Macmillan.

Herman, K. C., Borden, L. A., Reinke, W. M., & Webster-Stratton, C. (2011). The impact of the

Incredible Years parent, child, and teacher training programs on children's co-occurring

internalizing symptoms. *School Psychology Quarterly, 26*(3), 189.

Hickey, G., McGilloway, S., Hyland, L., Leckey, Y., Kelly, P., Bywater, T., . . . O'Neill, D. (2017).

Exploring the effects of a universal classroom management training programme on teacher

and child behaviour: A group randomised controlled trial and cost analysis. *Journal of Early*

Childhood Research, 15(2), 174-194. doi:DOI: 10.1177/1476718X15579747

Hinz, A., Zenger, M., Brahler, E., Spitzer, S., Scheuch, K., & Seibt, R. (2014). Effort-Reward

Imbalance and Mental Health Problems in 1074 German Teachers, Compared with Those in

the General Population. *Stress Health*. doi:10.1002/smi.2596

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., . . . Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ : British Medical Journal, 348*. doi:10.1136/bmj.g1687

Hoy, W., & Woolfolk, A. (1993). *Teachers' Sense of Efficacy and the Organizational Health of Schools* (Vol. 93).

Hutchings, J., Martin-Forbes, P., Daley, D., & Williams, M. E. (2013). A randomized controlled trial of the impact of a teacher classroom management program on the classroom behavior of children with and without behavior problems. *Journal of School Psychology, 51*(5), 571-585. doi:http://dx.doi.org/10.1016/j.jsp.2013.08.001

Industrial Injuries Advisory Council. (2017). *Anxiety and depression in teachers and healthcare workers: IIAC position paper 37*. Retrieved from London: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/642478/anxiety-and-depression-in-teachers-and-healthcare-workers-iiac-position-paper-37.pdf

Jennings, P. A., & Greenberg, M. T. (2009). The Prosocial Classroom: Teacher Social and Emotional Competence in Relation to Student and Classroom Outcomes. *Review of Educational Research, 79*(1), 491-525. doi:10.3102/0034654308325693

Kellam, S. G., Brown, C. H., Poduska, J., Ialongo, N., Wang, W., Toyinbo, P., . . . Wilcox, H. C. (2008). Effects of a Universal Classroom Behavior Management Program in First and Second Grades on Young Adult Behavioral, Psychiatric, and Social Outcomes. *Drug Alcohol Dependence, 95*(S5-S28). doi:10.1016/j.drugalcdep.2008.01.004

Kellam, S. G., Ling, X. G., Merisca, R., Brown, C. H., & Ialongo, N. (1998). The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology, 10*(2), 165-185. doi:Doi 10.1017/S0954579498001564

Kerlinger, F. N. (1986). *Foundations of behavioral research*. New York: Holt, Rinehart & Winston.

Kidger, J., Brockman, R., Tilling, K., Campbell, R., Ford, T., Araya, R., . . . Gunnell, D. (2016).

    Teachers' wellbeing and depressive symptoms, and associated risk factors: A large cross

    sectional study in English secondary schools. *J Affect Disord, 192*, 76-82.

    doi:10.1016/j.jad.2015.11.054

Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers. *British*

    Journal of Educational Psychology, 77(1), 229-243. doi:10.1348/000709905X90344

Lang, I. A., Marlow, R., Goodman, R., Meltzer, H., & Ford, T. (2013). Influence of problematic

    child-teacher relationships on future psychiatric disorder: population survey with 3-year

    follow-up. *Br J Psychiatry, 202*(5), 336-341. doi:10.1192/bjp.bp.112.120741

Maguire, M., & O'Connell, T. (2007). Ill-health retirement of schoolteachers in the Republic of

    Ireland. *Occup Med (Lond), 57*(3), 191-193. doi:10.1093/occmed/kqm001

Marlow, R., Hansford, L., Edwards, V., Ukoumunne, O., Norman, S., Ingarfield, S., . . . Ford, T.

    (2015). Teaching classroom management – a potential public health intervention? *Health*

    Education, 115(3/4), 230-248. doi:doi:10.1108/HE-03-2014-0030

Martin, P. A. (2009). *From small acorns: The positive impact of adopting simple teacher classroom*

    management strategies on global classroom behaviour and teacher-pupil relationships.

    (PhD), University of Wales, Bangor,

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *Maslach Burnout Inventory Manual* (3rd ed.).

    Palo Alto, CA: Consulting Psychologists Press.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology, 52*,

    397-422. doi:10.1146/annurev.psych.52.1.397

McGilloway, S., Hyland, L., Mháille, G., Lodge, A., O'Neill, D., Kelly, P., . . . Donnelly, M. (2011).

    Positive classrooms, positive children: A randomised controlled trial to investigate the

    effectiveness of the incredible years teacher classroom management programme in an Irish

    context (short-term outcomes). Retrieved from

McLean, L., & Connor, C. M. (2015). Depressive Symptoms in Third-Grade Teachers: Relations to Classroom Quality and Student Achievement. *Child development, 86*(3), 945-954. doi:10.1111/cdev.12344

Merrett, F., & Wheldall, K. (1993). How Do Teachers Learn to Manage Classroom Behaviour? A study of teachers' opinions about their initial training with special reference to classroom behaviour management. *Educational Studies, 19*(1), 91-106. doi:10.1080/0305569930190106

Murray, D. W., Rabiner, D. L., Kuhn, L., Pan, Y., & Sabet, R. F. (2018). Investigating teacher and student effects of the Incredible Years Classroom Management Program in early elementary school. *Journal of School Psychology*. doi:https://doi.org/10.1016/j.jsp.2017.10.004

Nye, E. (2017). *Classroom behaviour management to support children's social, emotional, and behavioural development.* (PhD), University of Oxford,

Oberle, E., & Schonert-Reichl, K. A. (2016). Stress contagion in the classroom? The link between classroom teacher burnout and morning cortisol in elementary school students. *Social Science and Medicine, 159*, 30-37. doi:doi.org/10.1016/j.socscimed.2016.04.031

OECD. (2009). *Creating Effective Teaching and Learning Environments: First Results from TALIS*. Paris: OECD Publishing Retrieved from http://www.oecd.org/dataoecd/17/51/43023606.pdf.

OECD. (2014). *Mental Health and Work: United Kingdom*.

Office of National Statistics. (2017). *Sickness absence in the labour market: 2016*. London: Office of National Statistics.

Ouellette, R. R., Frazier, S. L., Shernoff, E. S., Cappella, E., Mehta, T. G., Maríñez-Lora, A., . . . Atkins, M. S. (2017). Teacher Job Stress and Satisfaction in Urban Schools: Disentangling Individual-, Classroom-, and Organizational-Level Influences. *Behavior Therapy*. doi:https://doi.org/10.1016/j.beth.2017.11.011

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Piaget, J., & Inhelder, B. (1962). *The Psychology of the Child*. New York, NY: Basic Books.

Pidano, A. E., & Allen, A. R. (2015). The Incredible Years series: A review of the independent research base. *Journal of Child and Family Studies, 24*(7), 1898-1916.

Reinke, W. M., Herman, K. C., & , & Dong, N. (Under revision). The Incredible Years Teacher Classroom Management Program: Outcomes from a Group Randomized Trial. *Prevention science (revise and resumbit).*

Roberts, N. (2017). *The school curriculum in England.*

Ross, S., Romer, N., & Horner, R. (2012). Teacher Well-Being and the Implementation of School-Wide Positive Behavior Interventions and Supports *Journal of Positive Behavior Interventions, 14*(2), 118-128.

Schulz, K. F., Moher, D., & Altman, D. G. (2010). Consort 2010 Comments. *Lancet, 376*(9748), 1222-1223. doi:Doi 10.1016/S0140-6736(10)61879-X

Schutte, N., Toppinen, S., Kalimo, R., & Schaufeli, W. (2010). *The factorial validity of the Maslach Burnout Inventory—General Survey (MBI-GS) across nations and occupations* (Vol. 73).

StataCorp. (2015). Stata Statistical Software: Release 14. In. College Station: TX: StataCorp LP.

Titheradge, D., Hayes, R., Longdon, B., Allen, K., Price, A., Hansford, L., . . . Ford, T. (2019). Psychological distress amongst primary school teachers; a comparison with clinical and population samples. *Public Health, 166*, 53-56. doi:https://doi.org/10.1016/j.puhe.2018.09.022

Tschannen-Moran, M., & Wollfolk, H. (2001). A Teacher efficacy; capturing an elusive construct *Teaching and Teacher Education, 17*, 783-805.

Uher, R., & Goodman, R. (2010). The Everyday Feeling Questionnaire: the structure and validation of a measure of general psychological well-being and distress. *Soc Psychiatry Psychiatr Epidemiol, 45*(3), 413-423. doi:10.1007/s00127-009-0074-9

Webster-Stratton, C., Reid, M. J., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: evaluation of the Incredible Years Teacher and Child Training

Programs in high-risk schools. *Journal of Child Psychology and Psychiatry, 49*(5), 471-488.

doi:DOI 10.1111/j.1469-7610.2007.01861.x

Webster-Stratton, C., Reinke, W. M., Herman, K. C., & Newcomer, L. L. (2011). The incredible years

teacher classroom management training: the methods and principles that support fidelity and

training delivery. *School Psychology Review, 40*(4), 509-529.

Whear, R., Thompson-Coon, J., Boddy, K., Ford, T., Racey, D., & Stein, K. (2012). The effect of

teacher-led interventions on social and emotional behaviour in primary school children: a

systematic review. *British Educational Research Journal*, 1-38.

doi:10.1080/01411926.2011.650680

Worth, J., Bamford, S., & Durbin, B. (2015). *Should I Stay or Should I Go? NFER Analysis of

Teachers Joining and Leaving the Profession*. Retrieved from Slough:

Table 1*: Table of the key concepts covered in each of the six TCM workshops

Workshop	Workshop title	Key concepts
1	Building positive relationships with students and the proactive teacher	Building relationships: Value of showing attention and appreciation to increase positive child behaviours Importance of getting to know parents to develop relationship with child Proactive strategies: Fostering caring through the notion of classroom as community and as family Teacher as model – caring for and respecting all children Emphasising the importance of predictable routines and schedules for difficult students Clear, respectful, positive commands/requests
2	Teacher attention, coaching, encouragement and praise	Value of praise and encouragement being used by teachers to increase children's positive self-talk, to help them learn to self-evaluate and to promote prosocial behaviours

		Help teachers understand the perspective of children, and the importance of using academic, persistence, social and emotion coaching with children
		Model ways to promote positive self-praise
3	Motivating students through incentives	Dispel notion that praise and tangibles are bad for children
		Explain pitfalls of negative messages and negative notes to parents
		Importance of positive messages going home to parents
		Discuss different incentive systems and how to set them up
		Discuss teachers reinforcing themselves and other teachers
4	Decreasing inappropriate behaviour – ignoring and redirecting	Discipline hierarchies
		How to give effective instructions, and use distractions and redirections
		Understanding the importance of starting with the least intrusive approach
		Teaching both teachers and children to understand how to ignore inappropriate behaviour effectively
5	Decreasing inappropriate behaviour – follow	Helping children learn to self-regulate using calm down areas in the classroom

	through with consequences	The importance of the ignoring technique as a strength
		How to use logical and/or natural consequences (not loss of privileges or work chores)
6	Emotional regulation, social skills and problem solving training	Children need lots of practice to learn social skills
		The importance of encouraging children's responsibility and cooperative behaviour in classroom
		Social, emotion and persistence coaching to help children learn self-regulation and maintain focus
		Recognition of how powerful a child's reputation is on other people's interactions with them

Table 2: Summary of baseline characteristics

Variable	Intervention (TCM)	Control (TAU)
School characteristics	$N_S = 40$	$N_S = 40$
Rural versus urban school, n (%)		
rural, n (%)	18 (45·0)	19 (47·5)
urban, n (%)	22 (55·0)	21 (52·5)
Education Key Stage[a]		
Key stage 1, n (%)	20 (50·0)	21 (52·5)
Key stage 2, n (%)	20 (50·0)	19 (47·5)
% eligible for free school meals, median (IQR)	12 (8 to 24)	14 (10 to 23)
Index of multiple deprivation score, median (IQR)	0·17 (0·08 to 0·24)	0·16 (0·10 to 0·27)
Teacher characteristics	$N_T = 40$	$N_T = 40$
More than 5 years of teaching, n (%)	20 (50)	27 (68)
Age in years, mean (SD)	34·5 (9·0)	31·4 (8·7)
Female, n (%)	32 (80)	33 (83)
Permanent appointment, n (%)	32 (80)	34 (85)
Has a leadership position, n (%)	4 (10)	2 (5)
Key Stage[a] taught		

Key stage 1, n (%)	20 (50)	21 (53)
Key stage 2, n (%)	20 (50)	19 (48)
Teaching qualification		
PGCE, n (%)	17 (43%)	22 (55%)
BA, BSC, BEd including QTS, n (%)	18 (45%)	16 (40%)
Other, n (%)	1 (3%)	0 (0%)
Missing, n (%)	4 (10%)	2 (5%)
Whole classroom support		
Full Time, n (%)	23 (58%)	22 (55%)
Part Time, n (%)	17 (43%)	15 (38%)
1-2-1 Teaching support, n (%)	18 (45%)	22 (55%)
Percentage of SEN children in class, mean (SD)	18 (45%)	22 (55%)
Teacher Self-efficacy Questionnaire		
Student Engagement subscale, mean (SD)	6·8 (1·0)	7·1 (1·0)
Instructional Practice subscale, mean (SD)	6·9 (1·0)	7·2 (0·9)
Classroom Management subscale, mean (SD)	7·3 (0·9)	7·5 (0·9)
Maslach Burn-Out Inventory		
Exhaustion, mean (SD)	2·9 (1·4)	2·5 (1·4)
Cynicism, mean (SD)	1·2 (1·0)	1·1 (1·0)
Professional Efficacy, mean (SD)	4·2 (1·0)	4·6 (0·8)
Everyday Feelings Questionnaire (teacher well-being), mean (SD)	17·2 (6·9)	13·9 (6·6)

Pupil characteristics

	$N_P = 1037$	$N_P = 1038$
Female, n (%)	483 (46·6)	491 (47·3)
Age in years at last birthday, mean (SD; range)	6.2 (1.4; 4 to 9)	6.4 (1.3; 4 to 8)
Year group		
Reception	182 (17·6)	88 (8·5)
Year 1	176 (17·0)	192 (18·5)
Year 2	135 (13·0)	275 (26·5)
Year 3	389 (37·5)	220 (21·2)
Year 4	155 (14·9)	263 (25·3)
	$N_P = 721$	$N_P = 701$
Ethnicity		
White, n (%)	689 (95·6)	663 (94·6)
Black, n (%)	4 (0·6)	4 (0·6)
Asian, n (%)	5 (0·7)	11 (1·6)
Mixed, n (%)	20 (2·8)	18 (2·6)
Other, n (%)	3 (0·4)	5 (0·7)
	$N_P = 595$	$N_P = 502$
Eligible for free school meals, n (%)	70 (11·8)	64 (12·7)
	$N_P = 860$	$N_P = 844$
Index of multiple deprivation score, median (IQR)	0·16 (0·08 to 0·64)	0·15 (0·09 to 0·25)

[a] Education Key Stage 1 covers Reception to Year 2 for children aged 4-7; Key Stage 2 covers

Years 3-6 for children aged 7-8

N_S – denominator for schools; N_T – denominator for teachers; N_P – denominator for pupils

Table 3: Main comparisons of teacher outcomes at 9 month follow-up

Outcome	Intervention arm (I)	Control arm (C)	Unadjusted	Adjusted mean difference		
	N=37	N=37	mean difference			
	mean (SD)	mean (SD)	estimate	estimate	95% CI	p-value
Teacher Self-Efficacy Questionnaire						
Student Engagement	7.5 (0.9)	7.4 (1.0)	0.2	0.3	-0.05 to 0.7	0.09
Instructional Practice	7.6 (0.7)	7.6 (0.8)	0.03	0.1	-0.2 to 0.4	0.53
Classroom Management	7.9 (0.7)	7.9 (0.8)	0.007	0.1	-0.2 to 0.4	0.43
Maslach Burn-Out Inventory						
Exhaustion	2.5 (1.4)	2.3 (1.5)	0.2	-0.1	-0.6 to 0.4	0.72
Cynicism	1.2 (1.1)	1.2 (1.0)	0.005	-0.1	-0.5 to 0.3	0.74

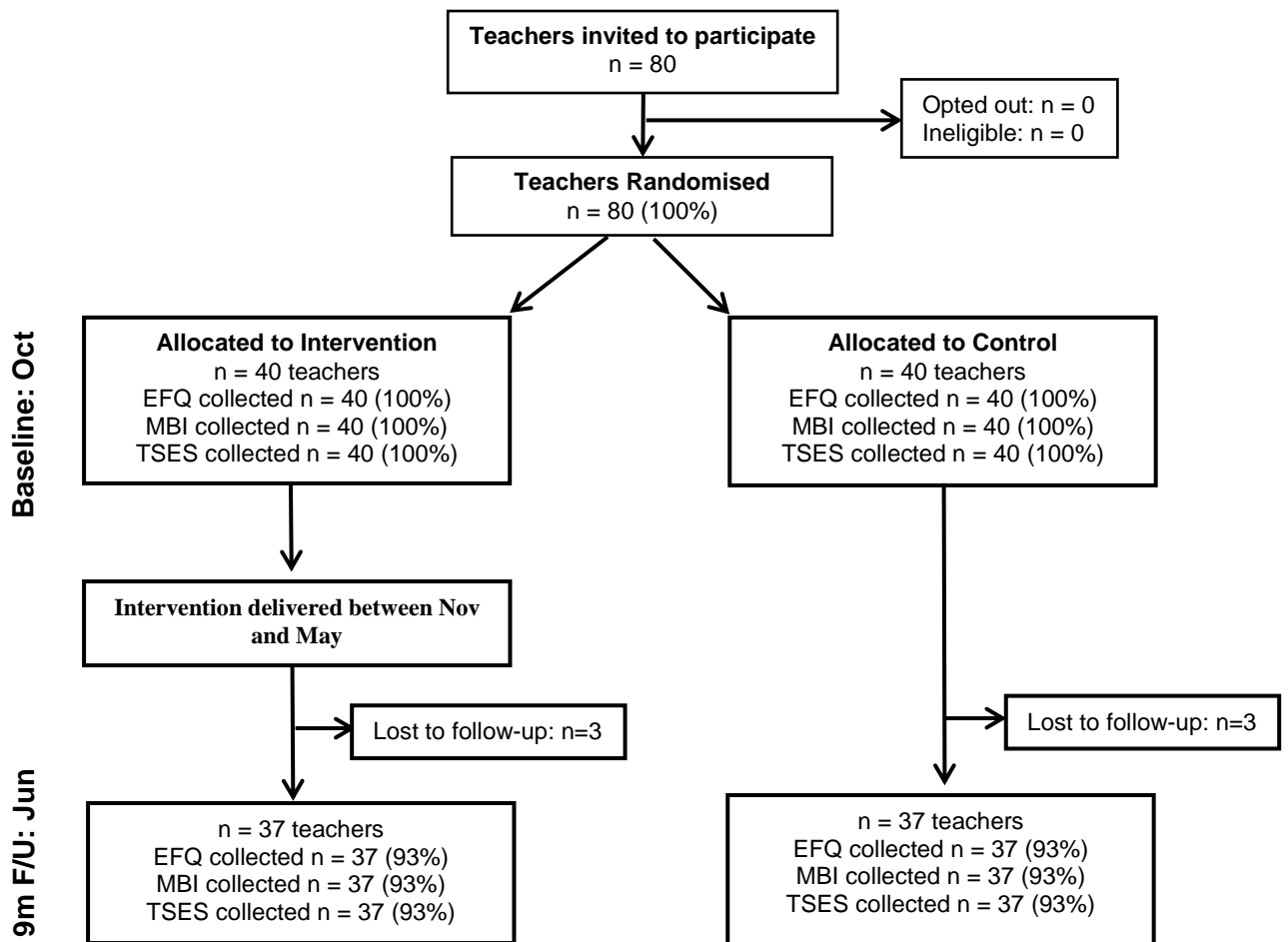|						
Professional Efficacy	4.2 (1.0)	4.5 (0.8)	-0.3	-0.06	-0.4 to 0.3	0.73
Everyday Feelings Questionnaire	15.6 (7.6)	13.6 (6.0)	2.0	-0.2	-2.7 to 2.2	0.85

Figure 1: CONSORT diagram depicting participant flow through the trial