

Training teachers in classroom management to improve mental health in primary school children: the STARS cluster RCT

Tamsin Ford, Rachel Hayes, Sarah Byford, Vanessa Edwards, Malcolm Fletcher, Stuart Logan, Brahm Norwich, Will Pritchard, Kate Allen, Matthew Allwood, Poushali Ganguli, Katie Grimes, Lorraine Hansford, Bryony Longdon, Shelley Norman, Anna Price, Abigail Emma Russell and Obioha C Ukoumunne



**National Institute for
Health Research**

Training teachers in classroom management to improve mental health in primary school children: the STARS cluster RCT

Tamsin Ford,^{1*} Rachel Hayes,¹ Sarah Byford,² Vanessa Edwards,¹ Malcolm Fletcher,¹ Stuart Logan,¹ Brahm Norwich,³ Will Pritchard,⁴ Kate Allen,¹ Matthew Allwood,¹ Poushali Ganguli,² Katie Grimes,⁵ Lorraine Hansford,¹ Bryony Longdon,¹ Shelley Norman,⁶ Anna Price,¹ Abigail Emma Russell¹ and Obioha C Ukoumunne⁷

¹University of Exeter Medical School, Exeter, UK

²King's Health Economics, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

³Graduate School of Education, University of Exeter, Exeter, UK

⁴Education and Early Years, Cornwall County Council, Truro, UK

⁵Educational and Counselling Psychology and Special Education, University of British Columbia, Vancouver, BC, Canada

⁶School of Psychology, University of Exeter, Exeter, UK

⁷National Institute for Health Research Collaborations for Leadership in Applied Health Research and Care South West Peninsula (PenCLAHRC), University of Exeter, Exeter, UK

*Corresponding author

Declared competing interests of authors: Stuart Logan was a member of the Health Technology Assessment (HTA) Commissioning Board 2003–10, the HTA Medicines for Children Themed Call 2005–6 and the Rapid Trials and Add on Studies Board 2012.

Published March 2019

DOI: 10.3310/phr07060

This report should be referenced as follows:

Ford T, Hayes R, Byford S, Edwards V, Fletcher M, Logan S, *et al.* Training teachers in classroom management to improve mental health in primary school children: the STARS cluster RCT. *Public Health Res* 2019;**7**(6).

Public Health Research

ISSN 2050-4381 (Print)

ISSN 2050-439X (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full PHR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/phr. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Public Health Research* journal

Reports are published in *Public Health Research* (PHR) if (1) they have resulted from work for the PHR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Public Health Research* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

PHR programme

The Public Health Research (PHR) programme, part of the National Institute for Health Research (NIHR), evaluates public health interventions, providing new knowledge on the benefits, costs, acceptability and wider impacts of non-NHS interventions intended to improve the health of the public and reduce inequalities in health. The scope of the programme is multi-disciplinary and broad, covering a range of interventions that improve public health. The Public Health Research programme also complements the NIHR Health Technology Assessment programme which has a growing portfolio evaluating NHS public health interventions.

For more information about the PHR programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/phr>

This report

The research reported in this issue of the journal was funded by the PHR programme as project number 10/3006/07. The contractual start date was in June 2012. The final report began editorial review in March 2018 and was accepted for publication in September 2018. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The PHR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the PHR programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the PHR programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Ford *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

NIHR Journals Library Editor-in-Chief

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals)

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

Training teachers in classroom management to improve mental health in primary school children: the STARS cluster RCT

Tamsin Ford,^{1*} Rachel Hayes,¹ Sarah Byford,² Vanessa Edwards,¹ Malcolm Fletcher,¹ Stuart Logan,¹ Brahm Norwich,³ Will Pritchard,⁴ Kate Allen,¹ Matthew Allwood,¹ Poushali Ganguli,² Katie Grimes,⁵ Lorraine Hansford,¹ Bryony Longdon,¹ Shelley Norman,⁶ Anna Price,¹ Abigail Emma Russell¹ and Obioha C Ukoumunne⁷

¹University of Exeter Medical School, Exeter, UK

²King's Health Economics, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

³Graduate School of Education, University of Exeter, Exeter, UK

⁴Education and Early Years, Cornwall County Council, Truro, UK

⁵Educational and Counselling Psychology and Special Education, University of British Columbia, Vancouver, BC, Canada

⁶School of Psychology, University of Exeter, Exeter, UK

⁷National Institute for Health Research Collaborations for Leadership in Applied Health Research and Care South West Peninsula (PenCLAHRC), University of Exeter, Exeter, UK

*Corresponding author T.J.Ford@exeter.ac.uk

Background: Poor mental health in childhood is common, persistent and associated with a range of adverse outcomes that include persistent psychopathology, as well as risk-taking behaviour, criminality and educational failure, all of which may also compromise health. There is a growing policy focus on children's mental health and the role of schools in particular in addressing this.

Objectives: To evaluate whether or not the Incredible Years® (IY) Teacher Classroom Management (TCM) training improved children's mental health, behaviour, educational attainment and enjoyment of school, improved teachers' mental health and relationship with work, and was cost-effective in relation to potential improvements.

Design: A two-arm, pragmatic, parallel-group, superiority, cluster randomised controlled trial.

Setting: A total of 80 UK schools (clusters) were recruited in three distinct cohorts between 2012 and 2014 and randomised to TCM (intervention) or teaching as usual [(TAU) control] with follow-ups at 9, 18 and 30 months. Schools and teachers were not masked to allocation.

Participants: Eighty schools ($n = 2075$ children) were randomised: 40 ($n = 1037$ children) to TCM and 40 ($n = 1038$ children) to TAU.

Interventions: TCM was delivered to teachers in six whole-day sessions, spread over 6 months. The explicit goals of TCM are to enhance classroom management skills and improve teacher–student relationships.

Main outcome measures: The primary planned outcome was the teacher-reported Strengths and Difficulties Questionnaire Total Difficulties (SDQ-TD) score. Random-effects linear regression and marginal logistic regression models using generalized estimating equations were used to analyse outcomes.

Results: The intervention reduced the SDQ-TD score at 9 months [adjusted mean difference (AMD) -1.0 , 95% confidence interval (CI) -1.9 to -0.1 ; $p = 0.03$] but there was little evidence of effects at 18 months (AMD -0.1 , 95% CI -1.5 to 1.2 ; $p = 0.85$) and 30 months (AMD -0.7 , 95% CI -1.9 to 0.4 ; $p = 0.23$). Planned subgroup analyses suggested that TCM is more effective than TAU for children with poor mental health. Cost-effectiveness analysis using the SDQ-TD suggested that the probability of TCM being cost-effective compared with TAU was associated with some uncertainty (range of 40% to 80% depending on the willingness to pay for a unit improvement in SDQ-TD score). In terms of quality-adjusted life-years (QALYs), there was evidence to suggest that TCM was cost-effective compared with TAU at the National Institute for Health and Care Excellence thresholds of £20,000–30,000 per QALY at 9- and 18-month follow-up, but not at 30-month follow-up. There was evidence of reduced disruptive behaviour ($p = 0.04$) and reductions in inattention and overactivity ($p = 0.02$) at the 30-month follow-up. Despite no main effect on educational attainment, subgroup analysis indicated that the intervention's effect differed between those who did and those who did not have poor mental health for both literacy (interaction $p = 0.04$) and numeracy (interaction $p = 0.03$). Independent blind observations and qualitative feedback from teachers suggested that teachers' behaviour in the classroom changed as a result of attending TCM training.

Limitations: Teachers were not masked to allocation and attrition was marked for parent-reported data.

Conclusions: Our findings provide tentative evidence that TCM may be an effective universal child mental health intervention in the short term, particularly for primary school children who are identified as struggling, and it may be a cost-effective intervention in the short term.

Future work: Further research should explore TCM as a whole-school approach by training all school staff and should evaluate the impact of TCM on academic progress in a more thorough and systematic manner.

Trial registration: Current Controlled Trials ISRCTN84130388.

Funding: This project was funded by the National Institute for Health Research (NIHR) Public Health Research programme and will be published in full in *Public Health Research*; Vol. 7, No. 6. See the NIHR Journals Library website for further project information. Funding was also provided by the NIHR Collaboration for Leadership in Applied Health Research and Care South West Peninsula (NIHR CLAHRC South West Peninsula).

Contents

List of tables	xi
List of figures	xiii
List of supplementary material	xv
List of abbreviations	xvii
Plain English summary	xix
Scientific summary	xxi
Chapter 1 Introduction	1
Disruption in school	1
Universal interventions in the school setting: the Incredible Years® Teacher Classroom Management programme	1
Description of Teacher Classroom Management training	2
Existing research on the Teacher Classroom Management programme	5
<i>Pre–post studies of Teacher Classroom Management</i>	5
<i>Randomised controlled trials of the Teacher Classroom Management programme in primary school-aged children</i>	5
<i>Need for the current study</i>	5
Chapter 2 Child outcome methods	9
Study design	9
Setting, participants and recruitment	9
Inclusion and exclusion criteria	9
<i>Recruitment of schools</i>	9
Randomisation and concealment	10
<i>Intervention</i>	10
Data collection	11
Outcome measures	11
<i>Primary outcome</i>	11
<i>Secondary outcomes</i>	12
Sample size	14
Statistical analysis	15
Data management	16
<i>Missing outcome measures</i>	16
Ethics approval and research governance	16
<i>Confidentiality</i>	17
<i>Informed consent</i>	17
<i>Assessment of harms and adverse effects</i>	17
<i>Adverse events</i>	18
Patient and public involvement	18
Chapter 3 Child outcome results	19
Participants	19
Baseline comparability	19

Adherence to intervention	19
Data completeness	19
Primary outcome	25
Subgroup analysis	25
Secondary outcomes	25
National Pupil Database analysis	25
Sensitivity analysis of pupil progress	30
Chapter 4 Economic evaluation	31
Aim	31
Methods	31
<i>Perspective</i>	31
<i>Method of economic evaluation</i>	31
<i>Costs</i>	31
Data analysis	34
<i>Resource use</i>	35
<i>Difference in costs</i>	35
<i>Short-term cost-effectiveness analysis</i>	35
<i>Long-term cost-effectiveness analysis</i>	35
Results	36
<i>Data completeness</i>	36
<i>Resource use over follow-up</i>	36
<i>Costs over the 30-month follow-up period</i>	37
<i>Outcomes over the 30-month follow-up period</i>	37
<i>Short-term cost-effectiveness analysis using Strengths and Difficulties Questionnaire</i>	
<i>Total Difficulties score</i>	38
<i>Short-term cost-effectiveness analysis using quality-adjusted life-years</i>	41
<i>Long-term cost-effectiveness analysis</i>	45
Discussion	45
Chapter 5 Fidelity to the model	47
Overview	47
Supervision with programme developer	47
Group leader-completed agendas and checklists	47
Development of new fidelity tool	49
Teacher–pupil Observation Tool	51
<i>Introduction</i>	51
<i>Core principles of Teacher Classroom Management that the Teacher–Pupil</i>	
<i>Observation Tool can evaluate include</i>	52
<i>Categorisation of Teacher–Pupil Observation Tool variables</i>	52
<i>Teacher–Pupil Observation Tool observation methods</i>	52
<i>Analysis</i>	53
<i>Teacher–Pupil Observation Tool observation results</i>	54
<i>Teacher–Pupil Observation Tool observation conclusion</i>	56
Discussion	56
Chapter 6 Process evaluation	59
Background	59
Methods	59
<i>Teacher focus groups/interviews</i>	59
<i>Headteacher interviews</i>	59
<i>Teacher Classroom Management group leader focus group/interviews</i>	60
<i>Procedure for focus groups/interviews</i>	60

<i>Teacher-completed measures</i>	60
<i>Data analysis</i>	60
Results of teacher, headteacher and group leader focus groups/interviews	60
<i>Participant details</i>	60
<i>Teachers</i>	61
<i>Headteachers</i>	61
<i>Group leaders</i>	61
Results and discussion	61
Influences on teachers' learning: what were the key elements of the Teacher Classroom Management training from the teachers' perspective?	63
<i>Structure</i>	63
<i>Training as a group</i>	63
<i>Group leadership</i>	65
<i>Delivery methods: modelling, rehearsing, using in practice and reviewing</i>	66
<i>Materials</i>	66
<i>Time to reflect</i>	67
Barriers to and facilitators of the use of Teacher Classroom Management strategies in the classroom	68
<i>Facilitators</i>	68
<i>Barriers</i>	68
<i>Course content</i>	69
<i>Context</i>	70
What do teachers feel has changed in their approach? What impact has it had on others?	72
<i>Impact on teachers' practice</i>	72
<i>Change of 'mindset'</i>	73
<i>A child's eye view</i>	73
<i>Building relationships</i>	74
<i>Thinking before responding</i>	75
<i>Positivity</i>	75
<i>Feeling calmer, more confident and in control</i>	76
<i>Creating a positive cycle through role modelling</i>	78
Impact on children	78
<i>Changes in behaviour: children becoming ready to learn</i>	78
<i>Promoting independence and taking responsibility</i>	79
<i>Impact on children with behavioural needs</i>	79
<i>Raising self-esteem</i>	80
<i>No impact</i>	80
Impact on parent-teacher relationships	80
<i>No impact/little impact/unknown impact</i>	81
Use of Teacher Classroom Management strategies	82
<i>Specific and proximal praise</i>	82
<i>Ignoring</i>	83
<i>Rewards and consequences</i>	83
<i>Time-out to calm down</i>	84
<i>Emotional regulation</i>	84
<i>Behaviour plans</i>	85
<i>Coaching and commentary</i>	86
<i>Comparison of feedback across Teacher Classroom Management training groups</i>	87
Written feedback completed by the teachers	87
<i>Participant details</i>	87
<i>Results and discussion</i>	87
Overall discussion	90

Chapter 7 The impact of Teacher Classroom Management training on teachers' well-being and their relationships with work	91
Introduction	91
Aims	92
Methods	92
<i>Outcomes</i>	92
<i>Statistical analysis</i>	92
Results	93
Discussion	93
Chapter 8 Comparison of outcomes between children who are taught by teachers in the academic year following Teacher Classroom Management training and those who are taught by teachers while they receive Teacher Classroom Management training	97
Introduction	97
Aims	97
Methods	97
<i>Outcomes</i>	98
<i>Statistical analysis</i>	98
Results	99
Discussion	103
Chapter 9 Discussion	105
Effectiveness and cost-effectiveness as a universal child mental health intervention	105
Fidelity of delivery	108
Academic attainment	109
Teacher mental health, burnout and self-efficacy	109
The feasibility and acceptability of the Teacher Classroom Management course	110
Methodological issues	111
Future research	113
Conclusions	114
Acknowledgements	115
References	119
Appendix 1 Topic guides used in the process evaluation	129
Appendix 2 Number of teachers attending focus groups/interviews	137
Appendix 3 Participant characteristics	139
Appendix 4 Additional examples of each theme/subtheme	141

List of tables

TABLE 1	Table of the key concepts covered in each of the six TCM workshops	3
TABLE 2	Overview of RCTs of the TCM programme in primary school-aged children	6
TABLE 3	Summary of changes to original STARS protocol approved by the Peninsula College of Medicine and Dentistry Research Ethics Committee	17
TABLE 4	Baseline characteristics by trial arm status	21
TABLE 5	Distribution of number of TCM sessions attended by teachers	23
TABLE 6	Numbers and percentages of participants with available data	24
TABLE 7	Main comparison on teacher-reported SDQ-TD score (primary outcome) using different approaches for handling missing data	26
TABLE 8	Comparison of teacher-reported secondary outcomes	27
TABLE 9	Comparison of parent- and child-reported secondary outcomes	28
TABLE 10	Summary of number of absences in the first and second academic year of the trial for intervention and TAU groups	29
TABLE 11	Subgroup analysis of SDQ-TD score (teacher reported) at each follow-up	29
TABLE 12	Unit costs and sources used in economic analyses	33
TABLE 13	Availability of service use data at follow-up	36
TABLE 14	Service use over the 30-month follow-up period	36
TABLE 15	Mean costs (£) per participant over the 30-month follow-up period (CC1)	37
TABLE 16	Mean outcomes per participant over follow-up (CC1)	38
TABLE 17	Mean costs (£) and SDQ-TD scores per participant over the 30-month follow-up period for sensitivity analyses (CC2 and MI)	40
TABLE 18	Summary of costs, QALYs and cost-effectiveness over time	44
TABLE 19	Total numbers of DVDs shown in each TCM course group and workshop	48
TABLE 20	Percentage of activities completed by group leaders in each TCM group and workshop from the standardised checklists	48
TABLE 21	T-POT variables and behaviour coded, grouping of composite variables	53

TABLE 22 Summary of the frequency of teacher behaviours during the 30-minute observation measured by the T-POT at baseline and follow-up by trial arm status	54
TABLE 23 Comparison of frequency of teacher behaviours at follow-up between the trial arms	55
TABLE 24 Intraclass correlation coefficients quantifying inter-rater agreement for all T-POT variables	57
TABLE 25 Summary of teacher characteristics at baseline	94
TABLE 26 Main comparisons of teacher outcomes at the 9-month follow-up (T1)	95
TABLE 27 Summary of children's characteristics at baseline	100
TABLE 28 The comparison between the intervention and post-TCM arms at the 9-month follow-up	102

List of figures

FIGURE 1 Teacher Classroom Management training logic model highlighting the proposed mechanisms of change	4
FIGURE 2 Schematic detailing the timing of outcome measures	12
FIGURE 3 Consolidated Standards of Reporting Trials diagram	20
FIGURE 4 Bootstrapped mean differences in costs and effects (SDQ-TD score) of intervention compared with TAU at 30-month follow-up (CC1)	39
FIGURE 5 Cost-effectiveness acceptability curve showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using the SDQ-TD score at 30-month follow-up (CC1)	39
FIGURE 6 Cost-effectiveness acceptability curves showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using SDQ-TD score (all analyses)	41
FIGURE 7 Bootstrapped mean differences in costs and effects (QALYs) of intervention compared with TAU at 30-month follow-up (CC1)	41
FIGURE 8 Cost-effectiveness acceptability curve showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using QALYs at 30-month follow-up (CC1)	42
FIGURE 9 Bootstrapped mean differences in costs and effects (QALYs) of intervention compared with TAU at 9-month follow-up (CC1)	42
FIGURE 10 Cost-effectiveness acceptability curve showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using QALYs at 9-month follow-up (CC1)	43
FIGURE 11 Bootstrapped mean differences in costs and effects (QALYs) of intervention compared with TAU at 18-month follow-up (CC1)	43
FIGURE 12 Cost-effectiveness acceptability curve showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using QALYs at 18-month follow-up (CC1)	44
FIGURE 13 Summary of costs and outcome over follow-up	45
FIGURE 14 The final tally sheet used to assess the fidelity of the delivery of TCM time-out	50
FIGURE 15 Graph of the frequency with which the group leaders highlighted principles of time-out in the five TCM courses (from explanation of time-out section of the tally sheet)	51

FIGURE 16 Graph of the frequency with which group leaders showed empathy in the five TCM courses (from supportive atmosphere section of the tally sheet)	51
FIGURE 17 Number of teachers who attended the TCM training in each year	62
FIGURE 18 Teaching Pyramid®	69
FIGURE 19 Percentage of teachers who found TCM 'helpful' or 'very helpful'	88
FIGURE 20 Percentage of teachers who rated course content, video clips, teaching and discussion as 'helpful'/'very helpful'	89
FIGURE 21 Percentage of teachers who rated the course groups as 'helpful' or 'very helpful'	89
FIGURE 22 Teachers' substudy CONSORT diagram	93
FIGURE 23 Schematic detailing the timing of outcome measures for the Post-TCM substudy	98
FIGURE 24 Post-TCM substudy flow of participants	99

List of supplementary material

Report Supplementary Material 1 Case report form for data about children

Report Supplementary Material 2 Final approved version of protocol

Report Supplementary Material 3 Teacher Satisfaction Questionnaire

Report Supplementary Material 4 Case report form for teacher-related outcomes

Supplementary material can be found on the NIHR Journals Library report project page (www.journalslibrary.nihr.ac.uk/programmes/phr/10300607/#/documentation).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

List of abbreviations

ADHD	attention deficit hyperactivity disorder	IY	Incredible Years®
AMD	adjusted mean difference	KS	key stage
APP	assessment of pupil progress	MBI	Maslach Burnout Inventory
CACE	complier average causal effect	MI	multiple imputation
CAMHS	Child and Adolescent Mental Health Services	NICE	National Institute for Health and Care Excellence
CA-SUS	Child and Adolescent Service Use Schedule	NIHR	National Institute for Health Research
CC1	complete-case analysis	NPD	National Pupil Database
CC2	complete-case sensitivity analysis	NQT	newly qualified teacher
CEAC	cost-effectiveness acceptability curve	Ofsted	Office for Standards in Education, Children's Services and Skills
CHU-9D	Child Health Utility Index 9 Dimensions	ONS	Office for National Statistics
CI	confidence interval	OR	odds ratio
CTU	Clinical Trials Unit	PBQ	Pupil Behaviour Questionnaire
DMC	Data Monitoring Committee	QALY	quality-adjusted life-year
DVD	digital versatile disc	RCT	randomised controlled trial
EFQ	Everyday Feelings Questionnaire	RR	rate ratio
EQ-5D	EuroQol-5 Dimensions	SD	standard deviation
EQ-5D-Y	EuroQol-5 Dimensions Youth version	SDQ	Strength and Difficulties Questionnaire
FG	focus group	SDQ-TD	Strengths and Difficulties Questionnaire Total Difficulties
GCSE	General Certificate of Secondary Education	SEN	special educational needs
GP	general practitioner	STARS	Supporting Teachers And childRen in Schools
HIFAMS	How I Feel About My School	TA	teaching assistant
HTA	Health Technology Assessment	TAU	teaching as usual
ICC	intracluster (intraschool) correlation coefficient	TCM	Teacher Classroom Management
ICER	incremental cost-effectiveness ratio	T-POT	Teacher–Pupil Observational Tool
IDACI	Income Deprivation Affecting Children Index	TSC	Trial Steering Committee
IMD	Index of Multiple Deprivation	TSES	Teachers' Sense of Efficacy Scale
IQR	interquartile range	TSQ	Teacher Satisfaction Questionnaire
		UAG	User Advisory Group
		VIF	variance inflation factor

Plain English summary

Between 5% and 6% of primary school-aged children, an average of one or two in every class, can be described as having extremely challenging behaviour, which can be difficult for teachers to manage and makes it hard for all children in the class to learn effectively. The Supporting Teachers And childRen in Schools (STARS) trial wanted to know if attending Teacher Classroom Management (TCM) training helped teachers to improve their behaviour-management strategies and if this in turn would lead to better mental health for the children they teach. The trial involved 80 primary school teachers, and their classes, across the south-west of England. Half of the teachers attended TCM training (intervention) and half did not (control). The mental health and behaviour of the children they taught was compared before training started and after it had finished, 9 months later. The children continued to be compared for the next two academic years, at 18 and 30 months from the start of the study. A total of 2075 children aged between 4 and 9 years took part, and teachers reported that they enjoyed the training and found it helpful. Children whose teachers attended the training had slightly better mental health at 9 months than children whose teachers had not attended the training. These differences were not found 18 and 30 months later when the children were taught by different teachers; both sets of children scored similarly. Although the difference found was not large, the children who initially had the worst mental health made bigger improvements in their mental health, and these improvements were still present 30 months later. The children in the intervention arm had better classroom-based behaviour at the 9-, 18- and 30-month comparisons. It was also found that the intervention may be better value for money than teaching as usual, particularly at the 9- and 18-month follow-up points. TCM training should be tested further for primary school teachers but the trial indicated that it would work best if it were to involve all teachers and teaching assistants in a school.

Scientific summary

Background

Extremely challenging behaviour that interferes with a child's ability to engage in normal social and educational settings is relatively common. Between 5% and 6% of primary school-aged children, or one or two children in the average school classroom, will display behaviours that reduce their ability to function, and conduct disorder is one of the strongest predictors of later adult mental health difficulties. Poor behaviour not only affects the individual child's ability to engage with education but also disrupts the learning of other children in the class. Teachers report that managing challenging behaviour can be a cause of additional stress in their role and that their initial training did not provide enough practical support for them to feel confident in their ability to control and engage their class. The Supporting Teachers And childRen in Schools (STARS) trial aimed to find out if the Incredible Years® (IY) Teacher Classroom Management (TCM) training programme could help teachers to improve the behaviour of the children in their classes. There is a lot of previous research that shows that the corresponding parent and child IY training programmes are very effective, but TCM training had not yet been tested in isolation in UK schools in a large randomised trial.

Teacher Classroom Management training is delivered to groups of 12 teachers over 6 full days over a 6-month period, a format that enables teachers to practice the new strategies that they have learnt between sessions. TCM has four explicit goals:

1. to enhance teacher management skills and improve teacher–pupil relationships
2. to assist teachers to develop effective individual and group behaviour plans in order to enable proactive (as opposed to reactive) classroom management
3. to encourage teachers to adopt and promote social and emotional regulation skills
4. to encourage teachers to strengthen positive teacher–parent relationships.

Teacher Classroom Management uses a range of methods to deliver the training of these four principles and is based on cognitive social learning theory as well as on several theoretical perspectives that are widely used in developmental psychology.

Objectives

To determine if TCM is an effective universal intervention for improving children's mental health in the context of the UK primary school system.

Secondary objectives included examining whether or not access to TCM training:

- improved teachers' mental health and professional self-efficacy and reduced their burnout
- improved children's academic attainment and classroom-based behaviour
- improved children's happiness in school and behaviour at home
- was cost-effective in relation to potential improvements in children's mental health
- affected teachers' practice in the classroom (investigating what factors supported or hindered any potential changes).

Methods

The STARS trial was a two-arm, pragmatic, cluster randomised controlled trial designed to evaluate whether or not the TCM course (delivered at class level) improves the mental health of individual children. We recruited 80 primary schools across the south-west of England between May 2012 and July 2014; one class (teacher and all pupils) from each school was selected by the headteacher. Schools were randomised with 40 of the class teachers receiving TCM training (intervention) and 40 being asked to continue teaching as usual [(TAU) control].

Schools were eligible for inclusion if they had a single-year class with ≥ 15 children aged between 4 and 9 years who were taught by a teacher who held classroom responsibility for at least 4 days per week. Schools were excluded if they primarily taught pupils with special educational needs, lacked a substantive headteacher or had been judged as failing at their last Office for Standards in Education, Children's Services and Skills inspection. All children in the selected classes were eligible for inclusion provided that the class teacher judged that they and their parents had sufficient English-language comprehension to understand recruitment information and complete outcome measures.

Written consent was obtained from the headteacher for the school's participation and from the class teacher for their involvement. Parent information leaflets were sent home with children and parents were given 2 weeks to 'opt out' their child from the research. Verbal assent was obtained from children each time they were asked to complete a questionnaire.

Outcome measures were collected on four occasions: baseline measures were collected in October, and follow-up measures were collected 9 (June), 18 (February) and 30 (February) months later. Baseline and 9-month assessments took place during the first academic year of participation, before and after the intervention, respectively, so were completed by the same teacher. The 18-month and 30-month assessments were completed by different teachers.

At each assessment point we asked teachers and parents to complete the Strengths and Difficulties Questionnaire (SDQ), which measures children's mental health. The SDQ measure provides a Total Difficulties (SDQ-TD) score comprising the Behaviour, Emotions, Inattention/Overactivity and Peer Relationships subscale scores, as well as Pro-social and Impact scores. Higher scores indicate poorer mental health, except on the Pro-social subscale. Our primary outcome was the teacher-completed SDQ-TD score. In addition, teachers were asked to complete the Pupil Behaviour Questionnaire (PBQ), which measures children's classroom-based disruptive behaviours, and to rate each child's academic progress in literacy and numeracy. Parents were also asked to complete a brief questionnaire about their child's use of key NHS services. Children were asked to complete the How I Feel About My School (HIFAMS) measure, which assesses children's attitudes towards school, with higher scores indicating greater happiness. The study also linked to the National Pupil Database (NPD) so that accurate records of attendance for all included children could be obtained.

Researchers who did not know which teachers had attended TCM training observed lessons in just over one-quarter of the schools. They were recording behaviours that are specifically targeted in TCM training. Researchers met with teachers in focus groups and asked them to say in their own words if they felt that anything had changed in their approach to teaching and what impact, if any, this had had on the children they teach and on other staff and parents.

The trial outcomes at follow-up were compared using the intention-to-treat principle. Random-effects regression models were used to compare the intervention and TAU children, allowing for the similarity of data (clustering) within one class of children compared with a different class of children.

The study tested to see if the effect of TCM training might be stronger or weaker for children falling into subgroups based on school- or child-level deprivation status (in bottom two deciles vs. otherwise), whether

or not the child scored in the struggling range on the teacher-reported SDQ-TD score at baseline, the length of the study teacher's experience (> 5 years vs. ≤ 5 years), key stage (KS) status (KS 1 vs. KS 2), the child's gender and the cohort status (cohort 1, 2 or 3). Given the relationship between emotional health and educational progress, the study also explored if any effect of the intervention on educational progress was modified by whether or not the child had scored in the struggling range on the teacher-reported SDQ-TD score at baseline.

This study was granted ethics approval by the Peninsula College of Medicine and Dentistry Research Ethics Committee, now under the auspices of the University of Exeter Medical School Committee, on 8 March 2012 (reference number 12/03/141). The University of Exeter acted as the sponsor for the study. The trial was registered with the International Standard Randomised Controlled Trial Register with the reference number ISRCTN84130388 and was funded by the National Institute for Health Research Public Health Research programme and the National Institute for Health Research Collaborations for Leadership in Applied Health Research and Care South West Peninsula. All of the information collected was kept strictly confidential and held in accordance with the principles of the Data Protection Act 1998 (Great Britain. *Data Protection Act 1998*. London: The Stationery Office; 1998).

Results

A total of 2075 children were recruited to the trial (1037 in the TCM arm and 1038 in the TAU arm). A further 113 children were either opted out by their parents ($n = 107$) or ineligible ($n = 6$). The study lost contact with 271 (13%) children over the 30-month follow-up period, and two parents withdrew permission for parent-reported outcomes but permitted the collection of teacher- and child-reported outcomes. During the trial some schools did not provide teacher-completed data on child outcomes at the 9-month ($n = 1$), 18-month ($n = 2$) and 30-month ($n = 1$) assessments. In addition, one intervention school withdrew from the trial after completing the 18-month assessment. Primary outcome data were collected at 9-, 18- and 30-month follow-up for 96%, 89% and 85% of participants, respectively. Thirty-six (90%) of the 40 teachers in the intervention arm attended four or more TCM sessions; 23 teachers (58%) attended all six.

The study found that TCM improved child mental health, according to the teacher-reported SDQ-TD score, by 1.0 point [95% confidence interval (CI) 0.1 to 1.9; $p = 0.03$] at the 9-month follow-up. There was little evidence, however, of an effect at the 18-month ($p = 0.85$) and 30-month follow-ups ($p = 0.23$).

Tests of interaction indicated that TCM led to greater reductions in the teacher-reported SDQ-TD score at 9 months (interaction $p < 0.001$) for children who were classified by their teacher as struggling with their mental health at baseline (mean difference -2.6 , 95% CI -4.6 to -0.6) than for children who were not (mean difference -0.4 , 95% CI -1.2 to 0.4). A subgroup effect was also found at 30 months ($p < 0.001$) but not at 18 months ($p = 0.10$).

There was evidence, based on the PBQ score, of reduced disruptive behaviour across all 30 months of follow-up ($p = 0.04$). Likewise, there was evidence that TCM reduces the percentage of children who are classified as struggling according to the SDQ-TD score ($p = 0.05$) and reduces the Inattention/Overactivity score ($p = 0.02$) across the full 30-month follow-up. At 9 months only, there was also evidence of a reduction in peer relationship problems ($p = 0.02$) and an improvement in pro-social behaviour ($p = 0.02$). Finally, there was little evidence of effects on teacher-reported emotions and impact, assessment of pupil progress (APP), parents' assessment of their child's mental health or the child-reported outcome HIFAMS.

There was little evidence that the intervention had any effect on either the rate of overall absence during the first [adjusted rate ratio (RR) 1.08, 95% CI 0.95 to 1.24; $p = 0.24$] or the second (adjusted RR 1.10, 95% CI 0.72 to 1.70; $p = 0.65$) year of the trial or the number of unauthorised absences during the first (adjusted RR 1.03, 95% CI 0.90 to 1.18; $p = 0.62$) or second (adjusted RR 0.96, 95% CI 0.75 to 1.22; $p = 0.74$)

year of the trial. School exclusions were reported on 22 separate occasions, two in the intervention arm and 20 in the TAU arm, which resulted in a total loss of 64 school sessions: 3 sessions in the intervention arm and 61 sessions in the TAU arm. These exclusions were issued to a total of six children, two from the intervention arm and four from the TAU arm of the trial.

Although there was no overall effect of the intervention on academic progress in either literacy or numeracy, subgroup analysis did indicate that the intervention effect differed between those who were and were not classified by their teacher as struggling with their mental health at baseline for both literacy (interaction $p = 0.04$) and numeracy (interaction $p = 0.03$). The intervention arm had lower odds than the TAU arm of below-expectation assessments in literacy [odds ratio (OR) 0.77, 95% CI 0.53 to 1.12] and numeracy (OR 0.82, 95% CI 0.59 to 1.14) among those children not classified as struggling, whereas it had greater odds of below-expectation assessments for literacy (OR 1.17, 95% CI 0.70 to 1.94) and numeracy (OR 1.35, 95% CI 0.88 to 2.06) among those children who were classified as struggling. This suggests that children who were classified as struggling performed worse in the intervention arm than in the TAU arm, whereas children who were not classified as struggling performed better in the intervention than in the TAU arm. However, all four of these CIs include unity, so it is difficult to interpret these findings, other than to comment that there seems to be a differential effect according to baseline mental health.

The cost of the TCM course was calculated to be £11.52 per child and was applied to the intervention arm. Observed mean total costs of services used over the 30-month follow-up period were very slightly lower for the intervention arm (£524.16) than for the TAU arm (£528.14). However, this difference was not statistically significant (adjusted mean difference £30.24, 95% CI -£140.98 to £201.47; $p = 0.7$). For the primary cost-effectiveness analysis using the SDQ-TD score, the lower costs and better outcomes in the intervention group generate an incremental cost-effectiveness ratio of -£19.90 per unit improvement in SDQ-TD and suggested that the probability of TCM being cost-effective compared with TAU was associated with some uncertainty (range of 40% to 80% depending on the willingness to pay for a unit improvement in SDQ-TD score). In terms of quality-adjusted life-years (QALYs), there was evidence to suggest that TCM was cost-effective compared with TAU at the National Institute for Health and Care Excellence threshold of £20,000–30,000 per QALY at the 9-month and 18-month follow-up, but not at the 30-month follow-up.

There was little evidence that TCM had any impact on teachers' self-efficacy, burnout and well-being scores. However, teachers did say that they felt that the TCM training had helped them in other ways, including helping them to see things from the child's point of view, improving their relationships with the children, and helping themselves to feel more positive, confident and in control. In addition, we demonstrated that in those schools in which researchers observed lessons, teachers' behaviour had changed following training: teachers used more praise and were more positive in their behaviour towards the children.

Conclusions

A small but statistically significant improvement in teacher-reported children's mental health was detected at 9 months (SDQ-TD). The findings provide tentative evidence that TCM may provide teachers with strategies that can have an effect on children's mental health, and, although small, the effect detected suggests that this intervention could usefully be tested further. The planned subgroup analyses suggest that children with poorer mental health at baseline derived the most benefit according to teacher report.

Economic evaluation, carried out at the 30-month follow-up using the SDQ-TD, suggests that TCM may be cost-effective compared with TAU. However, it was not possible to draw a firm conclusion without knowing society's willingness to pay for improvements in SDQ-TD score. In terms of QALYs, there was evidence to suggest that TCM was cost-effective compared with TAU in the short term, but not at the 30-month follow-up.

The effect of TCM on the primary outcome was not maintained at 18 and 30 months, which could mean that TCM has no longer-term impact or could be a result of the children's reaction to the teaching style of their subsequent teachers who had not accessed TCM training. Most effective universal programmes employ a whole-school approach, and the findings from this study would suggest that training all school staff to use the same strategies might amplify and sustain any initial impact on children's mental health that training a single teacher might have.

The small but sustained effects on disruptive behaviour and inattention/hyperactivity as measured by the teacher SDQ across all three follow-ups are interesting and warrant replication. The linkage to the NPD demonstrated some interesting differences in relation to exclusion between the trial arms, which suggests that these data may be a useful vehicle for conducting longer-term follow-up of the STARS participants.

No intervention effect on parent-reported measures was detected, but this was not unexpected. Children respond differently in different situations and TCM targets classroom behaviour rather than behaviour at home.

In the process evaluation, teachers reported that children were calmer, more motivated and more ready to learn; however, these reported experiences did not translate into empirical evidence of an improvement in academic attainment at a universal level. The interaction with baseline mental health, which is difficult to interpret in terms of direction, is interesting and suggests that the impact on attainment is worth studying in more depth. It is intuitively plausible that a more settled classroom would allow accelerated progress for children who were otherwise thriving. Given the crudeness of the measure available, it was not surprising that there was no main effect on attainment detected.

Implications and directions for future research

1. The findings of the STARS trial provide strong evidence that TCM is feasible and acceptable in the UK context.
2. Our findings provide early evidence that TCM may be an effective universal child mental health intervention in the short term, particularly for children who are struggling.
3. TCM should be explored as a whole-school approach.
4. The impact of TCM on teachers' relationships with children and on children's academic progress should be evaluated.

Trial registration

This trial is registered as ISRCTN84130388.

Funding

This project was funded by the Public Health Research programme of the National Institute for Health Research (NIHR) and the NIHR Collaboration for Leadership in Applied Health Research and Care South West Peninsula (NIHR CLAHRC South West Peninsula).

Chapter 1 Introduction

This chapter explains the problem that the study was investigating, what previous research has been carried out on this, the intervention that it was thought might help and the reasons why this was thought to be the right intervention to study.

Disruption in school

Disruption in the classroom can undermine the quality of teaching and interrupt the learning of all children in a class.¹ Conduct disorder (or behaviour that consistently violates social norms to the extent that a child's ability to function is impaired) is the most common childhood mental health problem, with a prevalence of 5–6% in the UK and high levels of comorbidity.^{2,3} Furthermore, levels of disruptive behaviours are normally distributed throughout the population.⁴ This means that in each classroom in the UK there is likely to be at least one child with a diagnosable conduct disorder, as well as several others who are disruptive to their classmates and teachers at lower levels.⁵ Parent training is the evidence-based treatment for childhood behaviour problems⁶ but this has limited impact on school-based problems.⁷ A recent systematic review of teacher-led interventions for internalising and externalising symptoms reported weak evidence for the effectiveness of interventions on internalising problems (Cohen's $d = 0.13$) and no evidence for programmes on externalising problems.⁸

Teachers report that disruptive behaviour and the task of managing the classroom can lead to high levels of stress and burnout.^{9–12} Teachers themselves note the lack of training that they receive in the area of classroom management.¹³ Improving the skills of one teacher to manage their classroom effectively would have an impact not only on the children they currently teach but also on children they will teach in the future: assuming one primary school teacher has a new class of 30 children every 2–3 years, within 10 years they will teach > 100 individuals.

A universal intervention that trains teachers has the potential to reach many thousands of children. Interventions that promote socioemotional competencies will affect children at high risk of later mental health problems, as well as their peers who may not have early indicators of risk but who still make up a substantial proportion of those with later mental health difficulties.¹⁴ Promoting resilience at an early stage in life could lead to a population-level reduction in the burden and cost associated with mental ill health. In addition, poor mental health is highly associated with lower levels of academic attainment,^{15–17} and recent work has demonstrated that improving a child's mental health will lead to subsequent improvements in their levels of attainment.¹⁸ Therefore, an effective intervention that can be applied and delivered in an ordinary school setting, with no cost other than to train the teachers delivering it, would incur minimal costs for a wide-reaching population health benefit.

Universal interventions in the school setting: the Incredible Years® Teacher Classroom Management programme

There are a variety of universal interventions in existence that are delivered in school settings in order to promote child socioemotional competencies and minimise disruptive behaviour. Those that are led by teachers, and, therefore, may benefit many children rather than targeting individual children who are 'high risk', are more limited. A systematic review of teacher-led interventions that target children's social and emotional behaviour found that only two existing programmes had been evaluated in a controlled design (randomised and non-randomised or pre–post quasi-experimental designs): the Incredible Years® (IY) Teacher Classroom Management (TCM) programme,¹⁹ and the Good Behaviour Game.^{20,21} Of these two programmes, the systematic review found that the evidence for TCM was the most robust.²¹ Importantly, this programme aims not only to improve the socioemotional skills and behaviour of children with externalising disorders but also to promote socioemotional skills in all children, with the aim of preventing poor outcomes in the future.¹⁹

Description of Teacher Classroom Management training

Like the IY parent and child programmes, TCM training was initially developed for individual children with conduct disorders but has been extended to children with risk factors that increase their likelihood of going on to develop a psychiatric disorder or to have other poor social outcomes, such as delinquency or substance use.²² It is believed that, by helping teachers to support these children, TCM will reduce disruption to the rest of the class, leading to a general improvement in all of their pupils' socioemotional skills. The focus of the TCM training is on collaborative learning, discussions of teachers' own experiences and group work to find solutions to problems encountered in the classroom.

The TCM training has four explicit goals:

1. to enhance teacher behaviour-management skills and improve teacher–pupil relationships
2. to help teachers to develop effective individual and group behaviour plans in order to enable proactive (as opposed to reactive) classroom management
3. to encourage teachers to adopt and promote social and emotional regulation skills
4. to encourage teachers to strengthen positive teacher–parent relationships (the promotion of positive relationships between parents, children and teachers is a central tenet of the IY series).²³

Research to date suggests that, although teachers already use a lot of these techniques, TCM training can increase their skills and confidence in using them sufficiently to produce significant improvements in the behaviour of their pupils. TCM uses a range of methods to deliver manualised training designed to improve classroom management. TCM draws on cognitive social learning theory, particularly Patterson's theories²⁴ about how coercive cycles of interaction between adults and children reinforce unwanted behaviour patterns, Bandura's ideas²⁵ about the importance of modelling and self-efficacy, and Piaget and Inhelder's developmental interactive learning methods.²⁶ In addition, it incorporates strategies for challenging angry, negative and depressive internal dialogue in adults while interacting with children that are drawn from cognitive behavioural approaches.²⁷

The delivery of the TCM teaching objectives follows these theoretical perspectives and includes problem-solving, role-play, modelling, goal setting, reflective learning, group discussion and support. Cognitive and emotional self-regulation training is also included in the course.²³

The manualised version of TCM is intended to be delivered in a collaborative style to groups of teachers by a trained group leader over 6 full days, with time between each session for teachers to practise the new strategies they have learnt.²³ *Table 1* outlines the key concepts that are covered in each of the six TCM workshops.

We have proposed a number of mechanisms of change in the logic model shown in *Figure 1*. We hypothesise that TCM will produce changes in the teachers' behaviour in the classroom that will lead to positive changes in the children within the classroom. We anticipate that any changes that the teacher makes will have an impact both on individual children and on their class as a whole. It is likely that a reinforcement loop will be in operation, exemplifying positive changes between all three groups. It is important to consider the wider context and what impact this may have on TCM's mechanisms of change. Certainly, the identified external factors may support or damage this proposed theory of change, and these factors are ones that are explored as part of the process evaluation work.

Published studies using TCM vary in terms of whether they adhere strictly to the advised training and implementation or adapt the TCM training for individual contexts, as well as in terms of whether TCM is applied as a stand-alone intervention or in conjunction with other IY components or adaptations.

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	<p>Building relationships:</p> <ul style="list-style-type: none"> • Value of showing attention and appreciation to increase positive child behaviours • Importance of getting to know parents to develop relationship with child <p>Proactive strategies:</p> <ul style="list-style-type: none"> • 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	<p>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</p> <p>Help teachers understand the perspective of children, and the importance of using academic, persistence, social and emotion coaching with children</p> <p>Model ways to promote positive self-praise</p>
3	Motivating students through incentives	<p>Dispel notion that praise and tangibles are bad for children</p> <p>Explain pitfalls of negative messages and negative notes to parents</p> <p>Importance of positive messages going home to parents</p> <p>Discuss different incentive systems and how to set them up</p> <p>Discuss teachers reinforcing themselves and other teachers</p>
4	Decreasing inappropriate behaviour – ignoring and redirecting	<p>Discipline hierarchies</p> <p>How to give effective instructions and use distractions and redirections</p> <p>Understanding the importance of starting with the least intrusive approach</p> <p>Teaching both teachers and children to understand how to ignore inappropriate behaviour effectively</p>
5	Decreasing inappropriate behaviour – follow through with consequences	<p>Helping children learn to self-regulate using calm-down areas in the classroom</p> <p>The importance of the ignoring technique as a strength</p> <p>How to use logical and/or natural consequences (not loss of privileges or work chores)</p>
6	Emotional regulation, social skills and problem-solving training	<p>Children need lots of practice to learn social skills</p> <p>The importance of encouraging children's responsibility and co-operative behaviour in classroom</p> <p>Social, emotion and persistence coaching to help children learn self-regulation and maintain focus</p> <p>Recognition of how powerful a child's reputation is on other people's interactions with them</p>

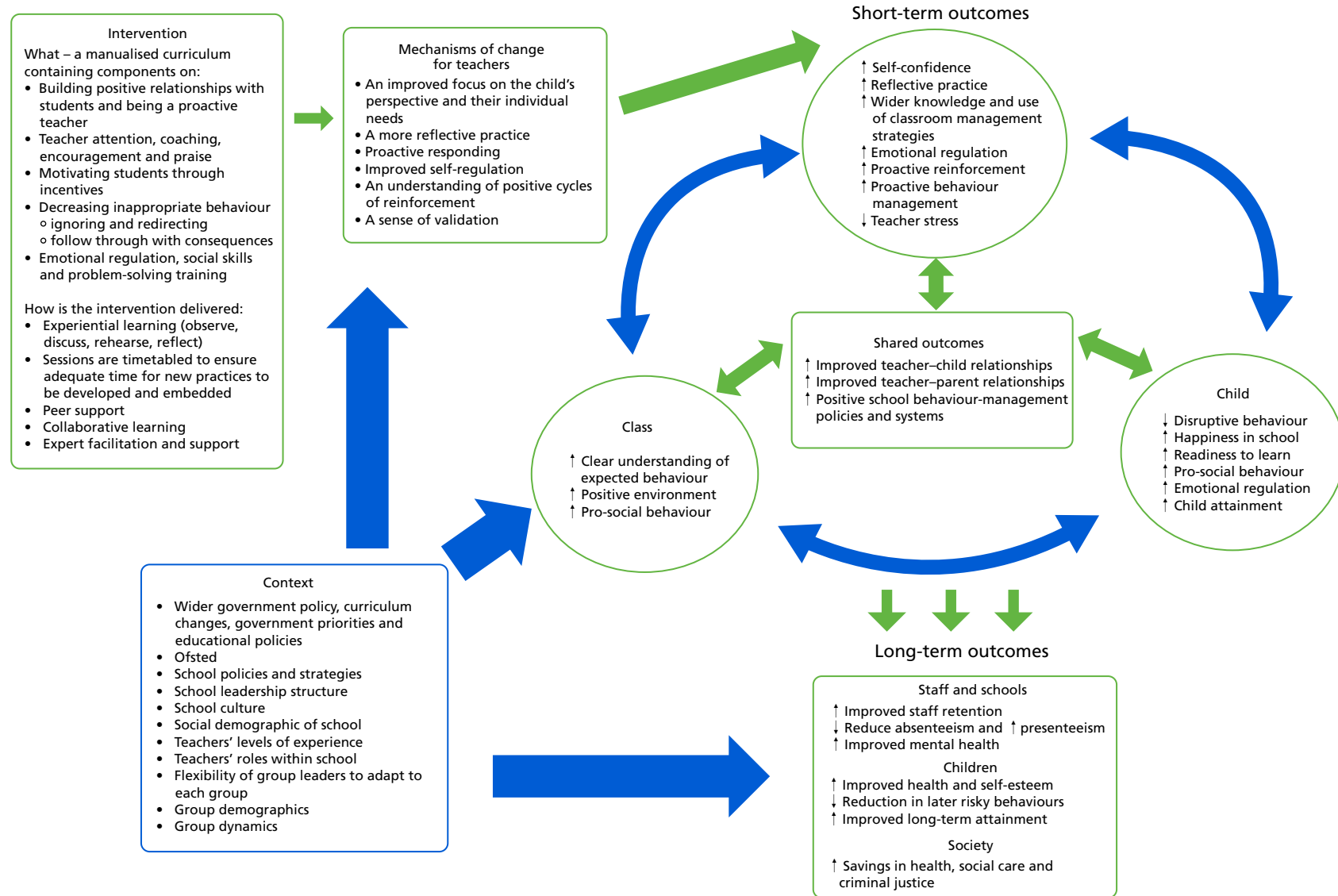


FIGURE 1 Teacher Classroom Management training logic model highlighting the proposed mechanisms of change. ↑, increased; ↓, reduced.

Existing research on the Teacher Classroom Management programme

At the time of writing, 21 independent data sets had been identified that report on the use of the TCM training, and two reviews had been identified that each aimed to collate the evidence around TCM.^{28,29}

Three initial randomised controlled trials (RCTs) conducted by the IY developers tested the effects of the TCM training in conjunction with other IY components in high-risk samples. The studies reported that children in the intervention arms had significantly fewer teacher-reported symptoms of hyperactivity and antisocial behaviour,²² more social competencies^{22,30} and better emotion regulation³⁰ and were more co-operative with their teachers and less aggressive with peers.³¹ Teachers who received the TCM training were observed to use more positive and fewer harsh and critical techniques,²² and more praise, as well as to be more confident, consistent and nurturing.³¹ There was tentative evidence that classrooms with higher levels of conduct problems benefited most from the intervention.³⁰ As these initial studies evaluated TCM with the parallel parent programmes or mentoring support, it is impossible to know which IY components were responsible for changes in child and teacher behaviour.

Pre-post studies of Teacher Classroom Management

Four non-RCTs utilising pre-post study designs,^{32–34} in one case with a control group,³⁵ have reported on TCM delivered as intended by the developer.^{32–35} In one study, the mental health, ability to concentrate, and peer relationships of target children who were selected by teachers to test their new classroom-management strategies significantly improved.^{33,36,37} In contrast, Kirkhaug *et al.*³⁵ did not detect significant improvements in their sample of 83 children with severe externalising problems following TCM. Children in the intervention group had improved academic performance and less student-teacher conflict than children in the control group.³⁵

Fergusson *et al.*³⁴ made use of data collected on 237 primary school teachers who received TCM training and reported a significant increase in the frequency and usefulness of positive strategies to manage behaviour. Similarly, a study of 24 pre-school teachers reported that they used more positive strategies and praise following TCM training.³² In contrast, a small ($n = 15$) non-peer-reviewed study reported no significant improvements in teacher strategies.³³

Randomised controlled trials of the Teacher Classroom Management programme in primary school-aged children

We identified four RCTs of TCM as a stand-alone intervention that have explored the impact of TCM on primary school-aged pupils and their teachers.^{38–41} A summary of these trials is given in *Table 2*.

These studies suggest that TCM training may improve teachers' abilities to manage their classroom positively by increasing the use of praise and proactive strategies, and by applying fewer harsh, negative and critical strategies. Children in TCM classrooms may derive benefits in the areas of social skills, in particular peer relationship and emotional regulation skills. Disruptive behaviour and negative responses towards peers and teachers may decrease following TCM. Specific findings vary by study design, population and context, with some studies reporting that TCM benefits those considered high risk most, and some finding that TCM prevents deterioration in the classroom environment across the school year. Many existing studies focus solely on the benefits of TCM for high-risk children, pre-school-age groups and education systems outside the UK.

Need for the current study

There are substantial gaps in the evidence base for TCM as a stand-alone intervention delivered as intended by the programme developers. The most recent systematic review of TCM²⁹ includes two of the studies discussed in this chapter^{38,41} in quantitative meta-analyses and is supported by the wider research base on TCM. There is evidence that TCM reduces negative and increases positive classroom-management strategies, reduces child conduct problems and child behaviour difficulties more broadly, and is effective in both high-risk and community samples of children. Both of these studies were published after the current study began.

TABLE 2 Overview of RCTs of the TCM programme in primary school-aged children

RCT (country, author and year)	Participants/method	Measures	Findings
Wales, Hutchings <i>et al.</i> ³⁸ (2013)	12 teachers; 107 selected target children aged 3–7 years Schools paired – one of each pair randomly assigned to TCM or wait-list control	T-POT ⁴² Teacher-completed SDQ ³⁶ IY TSQ and workshop evaluations	Teachers Fewer commands following TCM Teacher behaviours explicitly targeted by TCM (e.g. use of specific commands) changed in expected direction No impact on teacher behaviour towards whole class. May be because teachers were aware of the target children and may have focused on the study children to the detriment of the wider class during observations Children Compliance to commands increased in intervention group; no change in control group Significantly less off-task behaviour in intervention group following TCM Negative responses to teachers significantly decreased among target children
USA, Murray <i>et al.</i> ⁴¹ (2018)	45 intervention teachers and 46 control teachers; 1276 students aged 5–7 years In addition to training, each TCM teacher received two consultation sessions within their classroom during the delivery of the programme	CLASS ⁴³ Teacher coder impressions measure ⁴³	Teachers High levels of satisfaction with the training Significant effect on positive climate from the CLASS for the TCM group was reported post intervention, but this did not sustain into the following year Children Those who had high baseline social or behavioural difficulties improved relative to those in the control group
Ireland, Hickey <i>et al.</i> ³⁹ (2017)	22 teachers (one intervention and one control from each of the 11 participating schools); 12 children from each class	T-POT ⁴² Teacher-completed SDQ ³⁶ TSQ	Teachers Reported significantly more frequent use, and perceived usefulness, of positive management strategies Significantly less use of harsh and critical strategies among intervention teachers relative to control teachers Few changes were observed using the T-POT: TCM teachers used significantly fewer negative strategies at follow-up; however, they also used more negative strategies than the control group at baseline

TABLE 2 Overview of RCTs of the TCM programme in primary school-aged children (*continued*)

RCT (country, author and year)	Participants/method	Measures	Findings
			<p>Children</p> <p>Children in the control group had significantly higher SDQ scores at follow-up on the internalising problems subscale than intervention children</p> <p>High-risk children in TCM classrooms scored significantly lower than the control high-risk children on the SDQ-TD and impairment scores at follow-up</p> <p>Cost analysis</p> <p>TCM programme estimated to cost an average of £1682.31 per teacher</p>
USA, Reinke <i>et al.</i> (2018) ⁴⁰	105 teachers across nine schools; 1817 children aged 5–8 years	<p>The TOCA-C⁴⁴</p> <p>T-COMP⁴⁵</p> <p>Children completed standardised maths and reading academic assessments (there were no blinded observations)</p>	<p>Children</p> <p>Children in the TCM group significantly improved on emotional self-regulation, pro-social behaviour and social competence relative to control children</p> <p>No significant effects on teacher-reported conduct or disruptive behaviours</p> <p>Children who initially scored poorly on teacher-reported measures of social and academic competence improved more in the TCM condition than their peers in the control condition; however, this did not hold when using standardised academic measures of competence</p>

CLASS, Classroom Assessment Scoring System; SDQ, Strengths and Difficulties Questionnaire; SDQ-TD, Strengths and Difficulties Questionnaire Total Difficulties; T-COMP, Revised Social Competence Scale-Teacher version; TOCA-C, Teacher Observation of Classroom Adaptation-Checklist; T-POT, Teacher-Pupil Observation Tool; TSQ, Teacher Strategies Questionnaire.

The current study was needed to demonstrate whether or not TCM may be an effective universal intervention for child and teacher mental health in the context of the UK primary school system. Existing studies of TCM conducted in Wales and Ireland^{38,39} are limited by small sample sizes (12 and 22 teachers in total), and, to date, there are only four RCTs of TCM applied in the primary school context.^{38–41} Academic outcomes of children whose teachers have received the TCM training have been studied on only one other occasion³⁵ and no study has investigated cost-effectiveness.

The Supporting Teachers And childRen in Schools (STARS) trial aimed to address these evidence gaps by including children up to 9 years of age, including child- and parent-reported measures of child mental health, examining the impact of TCM on child academic attainment, examining the effect of TCM on teachers and all children within their classes, and exploring whether or not TCM has more impact on children who are considered as being at high risk for later poor outcomes.

Chapter 2 Child outcome methods

This chapter details what happened to the children in the trial and demonstrates what questions were asked and how often we asked them.

Study design

The STARS trial was a two-arm, pragmatic, parallel-group, superiority, cluster RCT designed to evaluate whether TCM training (delivered at class level) improves the mental health of individual children. Schools were randomly allocated to the intervention arm (TCM training) or the control arm [teaching as usual (TAU)]. The trial included a parallel economic evaluation to examine the cost-effectiveness of the TCM training and a mixed-method process evaluation that used qualitative methods to assess the acceptability of the TCM training; these are reported on in *Chapters 4* and *6*, respectively.

Setting, participants and recruitment

We recruited 80 primary schools across the south-west of England in Devon, Torbay and Plymouth in three separate cohorts: September 2012 (cohort 1), September 2013 (cohort 2) and September 2014 (cohort 3). One class (teacher and all pupils) was selected by the headteacher from each school independently of the research team.

Inclusion and exclusion criteria

Schools were considered for inclusion if they had a single-year class with ≥ 15 children aged between 4 and 9 years who were taught by a teacher who held classroom responsibility for at least 4 days per week. Schools were excluded if they primarily taught pupils with special educational needs (SEN), lacked a substantive headteacher or had been judged as failing at their last Office for Standards in Education, Children's Services and Skills (Ofsted) inspection.

All children in the selected classes were eligible for inclusion provided that the class teacher judged that they and their parents had sufficient English-language comprehension to understand recruitment information and complete outcome measures.

Recruitment of schools

Schools were approached through unsolicited contact with headteachers, publicity at local education conferences and contacting learning communities. Schools were recruited in three separate cohorts between April 2012 and June 2014, with a waiting list of schools for each cohort in the event of last-minute drop-out before the first data collection period.

Written consent was obtained from the headteacher for the school's participation and from the class teacher for their involvement after nomination by the headteacher. Parent information leaflets were sent home with children and parents were given 2 weeks to 'opt out' their child from the research. In order to opt out their child, parents could return a ready-prepared letter or contact the research team directly. Verbal assent was obtained from children each time they were asked to complete a questionnaire.

Randomisation and concealment

Randomisation of schools, using computer-generated random numbers, was carried out by an independent researcher based at the University of Exeter who was masked to the identity of the schools to ensure allocation concealment. The allocation was passed on to the trial manager, who then informed the schools. Randomisation was completed after baseline data collection to avoid recruitment and response bias.⁴⁶ Allocation was completed separately for each cohort, with all schools in a given cohort allocated en bloc. An equal number of schools were allocated to each arm overall, but unequal allocation ratios were used for cohorts 1 and 3 to ensure that there was an adequate number of intervention teachers to fill each TCM training group, with the ratio of intervention to TAU schools being 10 : 5 in cohort 1, 15 : 15 in cohort 2 and 15 : 20 in cohort 3.

Allocation was balanced on the following school factors: urban versus rural/semi-rural area, key stage (KS) 1 (Reception to Year 2) versus KS 2 (Years 3 or 4) and deprivation (whether the percentage of children eligible for free school meals was > 19%, the UK national average in 2012⁴⁷). Because there were relatively few clusters within each cohort, there was a high chance that random allocation using standard stratification or minimisation methods would not have yielded trial arms that were similar on the balancing factors. To overcome this, we used the approach of randomly selecting an allocation sequence for each cohort from the 5% with the least imbalance out of one million randomly generated potential allocations (permutations). Imbalance was quantified by the sum of the mean differences on the three balancing factors between trial arms weighted by the inverse of the variance of those factors.⁴⁸

We were unable to mask the schools and teachers because the school needed to release the class teacher to attend the training. Children and parents were not informed about whether or not the teacher attended training. Baseline measures were completed before randomisation. Even following randomisation, parents and children were unlikely to be aware of whether or not their child's teacher had completed TCM training. The follow-up measures were questionnaires that were completed independently of the researchers (with the exception of the service-related interviews, which were completed by a subsample of parents, and the child measures, as younger children might require support) and thus difficult for the core team of researchers to influence. In addition, the teacher-completed follow-up measures in the second and third years of each school's participation in the study (18 and 30 months post baseline) were completed by a teacher who did not access the intervention, although they were likely to have known whether their colleague who taught the class in the first academic year of the study did or did not.

Intervention

In the STARS trial, TCM was delivered to groups of up to 12 teachers in 6 whole-day sessions between October and April of each academic year. The sessions took place during the school day but at an external venue. The facilitating group leaders, who delivered the training in pairs, were behaviour support practitioners, had completed the mandatory TCM basic training and had led at least two previous courses prior to the start of the trial. They received monthly supervision from the programme developers, which included video reviews of each session, to ensure fidelity.

As recommended by the education community, and to incentivise recruitment and retention, TAU schools were offered TCM training during their second year of involvement in STARS as long as the attending teacher did not teach the study children during the 30-month follow-up period. All training costs were met for both intervention and TAU schools, including the provision of a £160 payment for each day the teacher attended TCM training to support the provision of a replacement teacher for their class. Our process evaluation involved interviews with headteachers and suggested a number of factors to consider when making their choice of teacher to nominate, including newly qualified teacher (NQT) status, allocation of a class known to be particularly challenging or a known teacher interest in behaviour management. No restrictions were placed on schools about access to other training and support services.

A total of eight TCM groups were delivered; in order that each group contained at least 10 teachers, the number of intervention and TAU teachers varied by cohort and year of the trial. In cohort 1, 15 schools were randomised: 10 to the intervention group and 5 to the TAU group. During this first year of the trial only one TCM group was delivered to these 10 intervention teachers. During cohort 2, 30 schools were randomised: 15 to the intervention group and 15 to the TAU group. During this second year of the trial, two TCM groups were run, which contained a mix of the 15 cohort 2 intervention teachers and the five cohort 1 TAU teachers. During cohort 3, 35 schools were randomised: 15 to the intervention group and 20 to the TAU group. During this third year of the trial, three TCM groups were run, which contained a mix of the 15 cohort 3 intervention teachers and the 15 cohort 2 TAU teachers. During the fourth year of the trial, two TCM groups were delivered to the cohort 3 TAU teachers.

Data collection

All baseline (T0) measures were collected before the end of the first academic half term of the schools' involvement, usually in October. Follow-up measures were then completed 9 [June, time point 1 (T1)], 18 [February, time point 2 (T2)] and 30 [February, time point 3 (T3)] months later. The 9-month time point was chosen to capture any initial impact the training may have had. Longer-term follow-up was important to see if potential impacts were sustained, reduced or potentially accelerated and the 18- and 30-month time points were chosen as they were mid-way through the following two academic years. Although additional time points during the first 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, parents and children, and, therefore, we chose fewer time points that were optimally placed to capture change. Baseline and 9-month assessments took place during the first academic year of participation, before and after the intervention, respectively, so were completed by the same teacher at both these time points. The 18-month and 30-month assessments occurred during the children's second and third academic year of the trial and were therefore completed by different teachers (*Figure 2*). It was not possible to ask the original teacher to complete the 18- and 30-month assessments because, in order to complete the measures, teachers must spend a large amount of time with the child to accurately assess their current development. Teacher measures were completed on an online database built for the trial (see *Report Supplementary Material 1*). To enable a supply teacher to supervise the teacher's class, schools received £80 for each time point at which teachers completed the outcomes (£320 in total) and £160 for each training day attended (£960 in total). Teachers also personally received a £10 gift voucher after outcome completion at each wave. Parent questionnaires were sent home with participating children. Parents received reminders via the school office and, where possible, second questionnaires were posted directly to the home. Parents received a £5 gift voucher for every completed questionnaire (£20 in total). Child-reported outcome data were collected during school time by researchers as a classroom activity for children aged ≥ 7 years, or individually for younger children. School staff were present but instructed not to assist the children.

Outcome measures

Primary outcome

Strengths and Difficulties Questionnaire Total Difficulties score (teacher completed)

The Strengths and Difficulties Questionnaire (SDQ) is a widely used measure of mental health in childhood containing 25 Likert items (each scored 0 to 2) comprising five scales (each with five items).³⁷ Our primary outcome was the Strengths and Difficulties Questionnaire Total Difficulties (SDQ-TD) score completed by the children's class teacher, which sums four of these five scales and has a possible score ranging from 0 to 40.

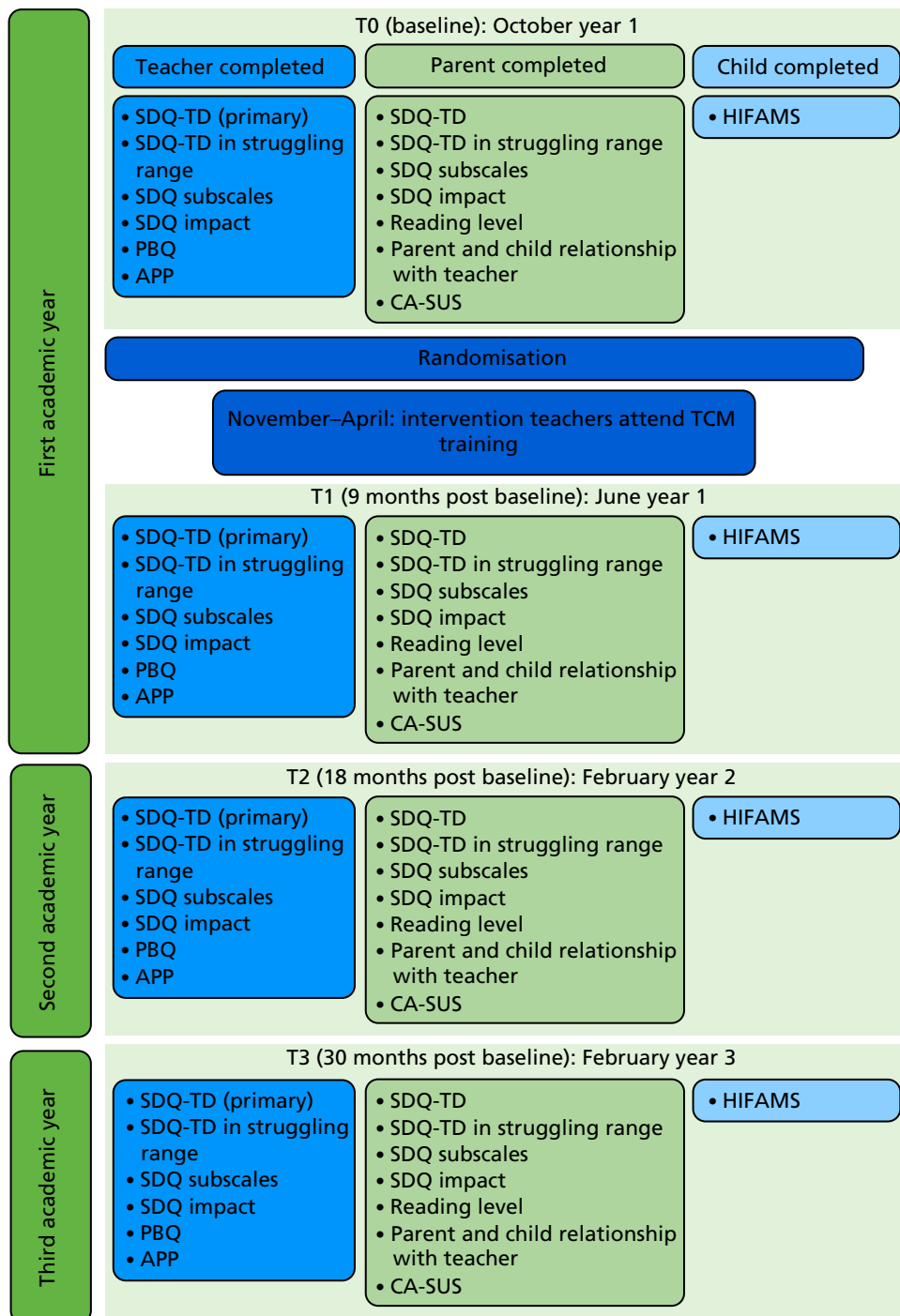


FIGURE 2 Schematic detailing the timing of outcome measures. APP, assessment of pupil progress; CA-SUS, Child and Adolescent Service Use Schedule; HIFAMS, How I Feel About My School; PBQ, Pupil Behaviour Questionnaire.

Secondary outcomes

Strengths and Difficulties Questionnaire Total Difficulties score (parent completed)

Parents and teachers were asked to complete the SDQ about children at each time point, and the parents' SDQ-TD score was a secondary outcome in the analysis.

Strengths and Difficulties Questionnaire Total Difficulties score above the clinical cut-off point (parent and teacher completed)

The SDQ-TD score was dichotomised to indicate those with clinically high scores (scored ≥ 12 on the teacher report and ≥ 14 on the parent report) as struggling compared with those with scores in the normal range (scored ≤ 11 on the teacher report and ≤ 13 on the parent report) as normal.⁴⁹

Strengths and Difficulties Questionnaire subscale scores (parent and teacher completed)

The following subscale scores of the SDQ (each containing five items with a possible score ranging from 0 to 10) were compared between the TCM and TAU trial arms: Behaviour, Emotions, Hyperactivity, Peer Relationships and Pro-social.³⁷

Teacher-reported Strengths and Difficulties Questionnaire impact score (teacher completed)

The teacher-reported SDQ impact score (three items with a possible total score from 0 to 9) quantifies the extent to which difficulties in the areas of emotions, concentration, behaviour or being able to get on with other people have an impact on a child's everyday life in terms of peer relations and classroom learning. The measure was dichotomised into those whose life was affected by difficulties (scoring 1 to 9) and those whose life was not (scoring 0).

Parent-reported Strengths and Difficulties Questionnaire impact score (parent completed)

The parent-reported SDQ impact scale (five items with a total possible score from 0 to 15) quantifies the extent to which difficulties in the areas of emotions, concentration, behaviour or being able to get on with other people affect a child's everyday life in terms of home life, friendships, classroom learning and leisure activities. The measure was dichotomised into those whose life was affected by difficulties (scoring 1 to 15) and those whose life was not (scoring 0).

Adapted Pupil Behaviour Questionnaire (teacher completed)

The Pupil Behaviour Questionnaire (PBQ) was developed for and used extensively in school effectiveness studies and is based on the findings of the Elton Report.⁵⁰ It measures the types of classroom-based disruptive behaviours of particular concern to school staff and has been validated as part of the trial.⁵¹ The adapted version contains six items with the following scoring categories: 0 = never, 1 = occasionally, 2 = frequently. Items are summed, with a higher total score (possible range: 0–12) indicating more disruptive behaviour.

How I Feel About My School (child completed)

Children completed the How I Feel About My School (HIFAMS),⁵² which measures children's attitudes towards school, with higher scores indicating greater happiness. The HIFAMS is a 7-item measure of a child's attitude towards school, with scores ranging from a possible 0 to 14 (summed across seven items each scored from 0 to 2).

Teacher assessment of pupil progress (teacher completed)

Teachers rated the children on sublevels according to their academic progress. These sublevels were mapped on to two categories (below expectation or at or above expectation) and analysed as a binary outcome.

Parent assessment of reading level (parent completed)

Parents rated the children on sublevels according to their reading ability. These sublevels were mapped on to six levels (categories) ranging from 'cannot read' to 'reads very well' and then grouped further into two categories (developing reader vs. fluent reader) and analysed as a binary outcome.

Assessment of pupils' relationships with teacher (parent completed)

Parents rated their child's relationship with their teacher on one of three ordinal categories. This was analysed as a binary outcome [poor or satisfactory (categories one and two) vs. good (category three)].

Assessment of parents' relationships with teachers (parent completed)

Parents rated their own relationship with their child's teacher on one of three ordinal categories. This was analysed as a binary outcome [poor or satisfactory (categories one and two) vs. good (category three)].

Child and Adolescent Service Use Schedule (parent completed)

Parents completed a brief, self-report version of the Child and Adolescent Service Use Schedule (CA-SUS)⁵³⁻⁵⁵ to collect data on children's use of key services (high cost, high frequency of use) of relevance to this population; full details are presented in *Chapter 4*.

Child- and school-level demographics

Parents were asked to provide the following demographic details: child's eligibility for free school meals, postcode [to link to the Income Deprivation Affecting Children Index (IDACI)], the number of children living in the household, housing tenure (rented or not) and the highest level of qualification of the parent(s) or carer(s). Data on a number of socioeconomic indices that might be related to outcomes were collected for all trial schools. We gathered school-level data on the percentage of children eligible for free school meals at recruitment and the IDACI at lower super output area as a proxy for the school catchment area according to the school's postcode.⁵⁶

National Pupil Database

We asked parents for opt-in consent to access their child's details from the National Pupil Database (NPD). This is a nationally held database to which all schools in the UK are required to submit data about the children in their school on a termly basis. The NPD includes information about characteristics such as gender, ethnicity, first language, eligibility for free school meals and SEN requirements, and detailed information about any absences and exclusions. Individual-level data on school attendance, exclusions and SEN status were obtained from the NPD for all children whose parents provided appropriate consent.

Sample size

As previously reported in our protocol,⁵⁷ 40 schools (clusters) were randomised to each of the intervention and TAU arms, using one class (teacher and pupils) from each school. Assuming that each class contains 30 pupils and that the recruitment rate is 70% (achieved among parents in the Helping Children Achieve trial⁵⁸), we anticipated that 21 (i.e. 30×0.7) children from each class of 30 and a total of 840 (i.e. 21×40) children in each trial arm would participate in the study. Assuming 10% attrition for the children, we expected 19 of them to be followed up at T3 in each class and a total of 760 (i.e. 19×40) children to be followed up at T3 in each trial arm. As clusters were randomised, the sample size was calculated taking account of the correlation between pupils' responses within clusters (or equivalently the variation between clusters). The variance inflation factor (VIF) (design effect) is given by $VIF = 1 + [(n-1) \times ICC]$, presented in Donner and Klar,⁵⁹ where n is the number of pupils providing outcome data at follow-up in each school and ICC is the intraclass (intraschool) correlation coefficient for the primary outcome. Nineteen children in each school were expected to provide follow-up data at T3 and the ICC for the SDQ-TD score (the primary outcome) was assumed to be 0.15 based on analysis of data from Sayal *et al.*⁶⁰ The assumed ICC value takes account of both the inherent variability across schools and the additional variability resulting from the fact that only one classroom was included in the study from each school. The VIF was 3.7 {i.e. $1 + [(19 - 1) \times 0.15]$ } and the target sample size therefore has the same effective sample size as an individually randomised trial with 205 ($= 760/3.7$) participating pupils at follow-up in each arm, providing 85% power at the 5% level of significance to detect a difference between trial arms of 0.3 standard deviation (SD) units (Cohen's $d = 0.3$), or a difference of 1.8 points on the raw SDQ scale (the SD of the teacher-reported SDQ has been estimated to be 5.9 for 4801 UK children aged 5–10 years).⁶¹ This effect

would reduce the percentage of children classified in the borderline/abnormal range from 19.7% to 13.7%,⁶¹ where borderline/abnormal is defined as those scoring 12 or more out of 40 on the teacher-reported SDQ-TD score. Data from Goodman and Goodman⁴⁹ suggest that the odds of psychiatric disorder decrease by 33% for each two-point decrease on the teacher SDQ and by 40% for each two-point decrease on the parent SDQ.

Statistical analysis

All analyses, performed using Stata[®] version 14.2 (StataCorp LP, College Station, TX, USA) software, were pre-specified in a statistical analysis plan that was reviewed by the independent Data Monitoring Committee (DMC) and the Trial Steering Committee (TSC).

Baseline characteristics of the schools, teachers and children were summarised for each trial arm, reporting means and SDs [or medians and interquartile ranges (IQRs)] for quantitative variables, and numbers and percentages for categorical variables. The characteristics of participating schools were compared with those reported in the 2012 school census for England.⁴⁷

The trial outcomes at follow-up were compared using the intention-to-treat principle; children were analysed strictly according to the trial arm to which their school was randomised. The main findings presented are based on analyses of complete cases. In addition, we carried out sensitivity analyses based on 50 multiply imputed data sets using the chained equations approach.⁶²

Quantitative outcomes were compared between trial arms using random-effects linear regression models, and binary outcomes were compared using marginal logistic regression models using generalised estimating equations with information sandwich ('robust') estimates of standard error assuming an exchangeable correlation structure. These methods allow for the correlation of children's outcome scores within schools. The primary analyses were those in which potential confounders were adjusted for, specified a priori in the analysis plan as the following: the three school-/class-level factors used to balance the randomisation, cohort, child gender, baseline score of the outcome, Index of Multiple Deprivation (IMD) score based on the child's address, number of children living in their household and whether or not the child's household was rented. The last three of these nine prognostic factors had a large number of missing data because they were parent reported. Adjusting for these would have resulted in the loss of one-quarter of the sample in the complete-case analyses (CC1). On this basis, and after discussion with our DMC (5 June 2017), we agreed the primary analysis as the CC1 adjusted for only the six a priori prognostic factors that were not parent reported (CC1). For completeness, we also report the findings from the complete-case sensitivity analysis (CC2) and the analysis of imputed data with all nine prognostic factors included [multiple imputation (MI)].

For all outcomes, tests of interaction were used to assess whether or not there was evidence that the effect of the TCM intervention differed across the three follow-up time points. When the interaction effect is statistically significant at the 5% level, we report the effect at each wave; otherwise, we report an estimate of the average intervention effect across the three follow-up waves.

In an ancillary analysis we used the two-stage least squares instrumental variable method⁶³ to calculate the complier average causal effect (CACE) estimate of the intervention effect on the primary outcome teacher-reported SDQ-TD score that would have occurred if all the teachers in the intervention arm had attended all six TCM training sessions.

Tests of interaction were used in pre-specified exploratory analyses to assess whether or not the effect of TCM on the primary outcome differs across subgroups defined by the following potential moderator variables: school- or child-level deprivation status (in bottom two deciles vs. otherwise), whether or not the child scored in the struggling range on the teacher-reported SDQ-TD score at baseline, the length of the

study teacher's experience (> 5 years vs. ≤ 5 years), KS status (KS 1 vs. KS 2), the child's gender and cohort status. In a further sensitivity analysis, a test of interaction was used to assess whether or not the effect of TCM on the primary outcome differed between the subgroup defined as having a primary or secondary SEN category of social, emotional or mental health, as classified by the NPD, versus those who did not.

Random-effects Poisson regression was used to compare the pupil-level rates of absence obtained from the NPD between the intervention and TAU arms in years 1 and 2 of the trial. We report crude rate ratios (RRs) and RRs and 95% confidence intervals (CIs) that are adjusted for school-/class-level factors [urban vs. rural/semi-rural area, KS 1 vs. KS 2, deprivation (% of children eligible for free school meals in 2012)], cohort and child gender.

Given the relationship between emotional health and educational attainment,¹⁵⁻¹⁸ tests of interaction were also carried out to assess whether or not the effect of the intervention on the assessment of pupil progress (APP) outcomes is modified by whether or not the child scored in the struggling range on the teacher-reported SDQ-TD score at baseline.

Data management

Data entry and cleaning was overseen by the trial manager. All data were stored on a custom-built password-protected database maintained by the Peninsula Clinical Trials Unit (CTU), a UK Clinical Research Collaboration-accredited CTU.

Missing outcome measures

The majority of teachers used a web-based electronic data-capture system to complete all questionnaire measures on the children, which did not allow for any items to be missed. Where this was not possible, teachers were asked to complete paper measures, which were double-entered on to the web-based database. Parents and children completed paper measures, which, again, were double-entered on to the database.

Paper questionnaires were checked for missing data on receipt and efforts were made to obtain these data from participants. In cases in which ambiguous data were not clarified with the participant, items were marked as 'spoiled' and recorded as missing.

Missing items within outcome measures were marked in accordance with the established conventions for that measure and overall totals and subtotals were imputed as instructed.

Where possible, teacher-reported outcomes were collected for children who had left their study school. This was achieved by asking the child's former teacher to complete the measures if the period between the child leaving the study school and data collection was less than one academic term. If the period between the child leaving the study school and data collection was greater than one academic term, we attempted to trace their current teacher in their new school and ask them to complete the measures.

Ethics approval and research governance

This study was granted ethics approval by the Peninsula College of Medicine and Dentistry Research Ethics Committee, now under the auspices of The University of Exeter Medical School Committee on 8 March 2012 (reference number: 12/03/141). The University of Exeter acted as the sponsor for the study. The trial was registered with the International Standard Randomised Controlled Trial Register with the reference number ISRCTN84130388. STARS was hosted in the Child Health Research Group at the University of Exeter, which has experience in the successful delivery of community-based paediatric trials. The trial was overseen by the independent TSC [Paul Stallard (chairperson), Gail Seymour, Shirley Larkin, Tobit Emmens

and David Glenny] and the DMC [Paul Ewings (chairperson), Siobhan Creanor and Andrew Richards]. A summary of the changes made to the original protocol⁵⁷ is given in *Table 3*, and the complete final trial protocol can be found in *Report Supplementary Material 2*.

Confidentiality

All of the information collected was kept strictly confidential and held in accordance with the principles of the Data Protection Act 1998.⁶⁴ Each participant was assigned a research number and all data were stored without subject identification. Data were held on a secure database on a password-protected computer at the University of Exeter. Access to data was, and continues to be, restricted to the research team.

Informed consent

Obtaining consent for this trial was a four-stage process:

1. headteachers – after receiving the information leaflet and having the opportunity to discuss the implications of the study, headteachers provided written consent for the school to participate in the trial.
2. teachers – after receiving the information leaflet and having the opportunity to discuss the implications of the study, the nominated teacher provided written consent for their participation in the trial.
3. parents – an information leaflet about the trial was sent via the schools to the parents of all children in the nominated teacher's class. This explained that if parents wished to opt their child out of the trial, they needed to return a form by a specified date (2 weeks later), otherwise consent would be inferred. Parents were able to opt themselves and their child out of the measurements but were not able to opt the teacher or school out of the study.
4. children – if parents had not opted their child out of the trial, the child's verbal assent was obtained before they completed the questionnaire measure on each occasion. If a child became distressed or appeared reluctant during data collection, this was assumed to indicate their wish not to complete the measure.

Assessment of harms and adverse effects

Child-completed questionnaires were screened for signs of severe distress and where children reported feeling sad in response to all seven questions of the HIFAMS measure, a conversation was held with the class teacher, headteacher or nominated deputy to ensure that they were aware of any difficulties that the child was facing and that they could put in place plans to support the child.

TABLE 3 Summary of changes to original STARS protocol approved by the Peninsula College of Medicine and Dentistry Research Ethics Committee

Changes to protocol	Date
Removal of child quality-of-life measure, extra detail of randomisation procedures and adaptation of two questionnaires	September 2012
TAU schools can choose to send a different teacher to the TCM training, as long as this teacher does not teach the children in the trial	May 2013
Collection of class-level attendance was added as an outcome	July 2013
Alternative method of gaining consent for telephone interviews with parents or teachers using either consent via e-mail (parent CA-SUS interviews) or verbally (teacher or SENCo interviews) was added	May 2014
Qualitative interviews to be completed with teaching assistants and an additional process evaluation FG with the TCM group leaders were added. Additional consenting of parents to allow the research team access to the NPD was added	January 2015
Post-TCM extension study was added	August 2016

FG, focus group; SENCo, Special Educational Needs Co-ordinator.

All researchers in contact with children and schools had the necessary Criminal Records Bureau checks and received training in child safeguarding. The trial had a specific safeguarding policy, but this was never enacted.

Adverse events

We followed Good Clinical Practice guidelines for identifying, acting on and reporting adverse events: we adopted the guideline definitions of adverse events and serious adverse reactions, reporting a total of one adverse event, which the TSC and DMC classified as unrelated to the trial.

Patient and public involvement

Public involvement has been key at all stages of the design, planning and implementation of the STARS programme of work. A User Advisory Group (UAG) was established comprising parents, teachers, headteachers and behavioural support network teachers who were delivering the TCM intervention. The UAG provided key advice on the acceptability of the study to parents and teachers, recruitment and data collection procedures. Specifically, they provided advice on:

- trial design – the UAG proposed that the TAU schools should be offered the opportunity to receive TCM training as an incentive for recruitment
- costing – the UAG was insistent that the funding for supply teacher cover should be provided as part of the trial as it was considered that schools would be unable to cover these costs themselves and that this would be a barrier to recruitment
- consent procedure – the UAG was fully consulted about the proposed consent process; they commented from their own (teacher/parental) perspective, and the process was adapted as a result of their advice
- outcome measures/questionnaires – the UAG commented on the outcome measures and instructions for completion
- trial literature – the UAG was involved in the development of the trial literature, including information literature and consent forms.

In addition to the UAG, we worked with the school councils of three local primary schools to develop the HIFAMS measure to ask for children's perceptions of attending school and we asked the children for their views on how it could be adapted to make it more 'user-friendly' for children.

Chapter 3 Child outcome results

This chapter describes the results that were found when considering the impacts of TCM training on children.

Participants

During the 26 months of recruitment, between April 2012 and June 2014, we recruited a total of 80 schools; 40 were allocated to each trial arm across all three cohorts (10 and 5 schools in the TCM and TAU arms, respectively, for cohort 1; 15 and 15 schools in the TCM and TAU arms, respectively, for cohort 2; and 15 and 20 schools in the TCM and TAU arms, respectively, for cohort 3). During the trial some schools did not provide teacher-completed data on child outcomes at the 9-month ($n = 1$), 18-month ($n = 2$) and 30-month ($n = 1$) assessments. In addition, one intervention school withdrew from the trial after completing the 18-month assessment (*Figure 3*) as a result of a change in headteacher who did not wish to uphold the agreement made by their predecessor. A total of 2075 children were recruited to the trial (1037 in the TCM arm and 1038 in the TAU arm). A further 113 children were either opted out by their parents ($n = 107$) or ineligible ($n = 6$). We lost contact with 271 (13%) children over the 30-month follow-up period and two parents withdrew permission for parent-reported outcomes but permitted the collection of teacher- and child-reported outcomes.

Baseline comparability

Compared with the national average,⁴⁷ participating schools had similar class sizes (trial mean 27.4 vs. national mean 26.8 children) and eligibility for free school meals (trial mean 18.3% vs. national mean 19.3%), but the sample included fewer voluntary controlled schools (5% trial schools vs. 14.4% national schools) and more community (61.3% trial schools vs. 55.3% national schools) and academy schools (10% vs. 6%). Baseline characteristics were generally balanced between the two arms (*Table 4*). Primary outcome data were collected at 9-, 18- and 30-month follow-up for 96%, 89% and 85% of participants, respectively. The proportion of children scoring in the struggling range on the teacher-reported SDQ-TD questionnaire in both arms approached the expected 20% (cut-off point at the 80th centile)⁴⁹ but was lower according to parent-reported SDQ-TD (TCM, 16.5%; TAU, 15.5%), which suggests that we lacked parental data on some vulnerable children. No serious adverse events were reported in either trial arm.

Adherence to intervention

Thirty-six (90%) of the 40 teachers in the intervention arm attended four or more TCM sessions; 23 teachers (58%) attended all six sessions (*Table 5*).

Data completeness

The numbers and percentages of participants with completed data are reported for each time point in *Table 6*.

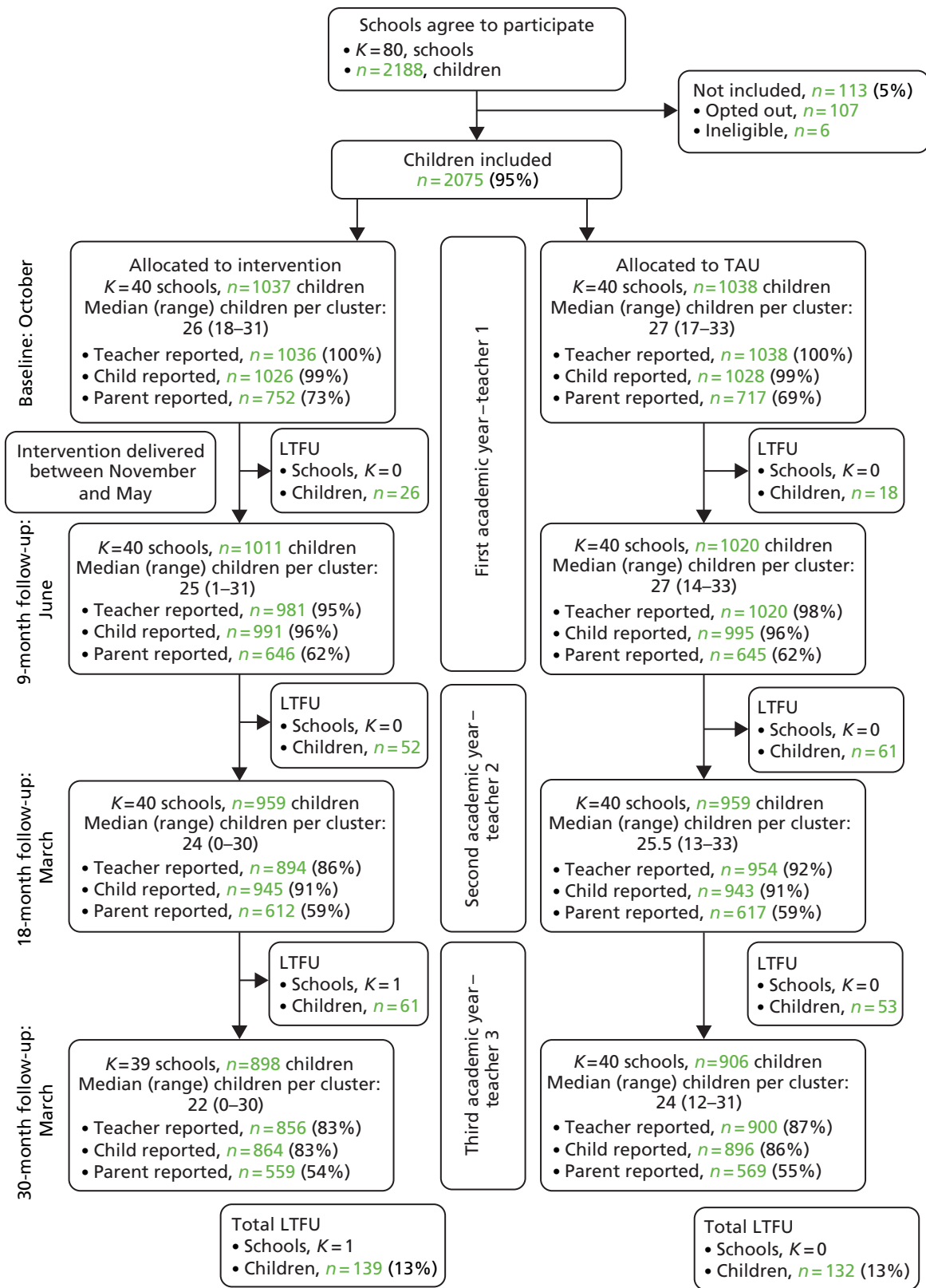


FIGURE 3 Consolidated Standards of Reporting Trials diagram. K, number of schools (clusters); LTFU, lost to follow-up; n, number of children.

TABLE 4 Baseline characteristics by trial arm status

Variable	Intervention (TCM)	Control (TAU)
School (cluster) characteristics	N_s = 40	N_s = 40
Rural/semi-rural vs. urban school, <i>n</i> (%)		
Urban	22 (55)	21 (53)
KS, <i>n</i> (%)		
1	20 (50)	21 (53)
2	20 (50)	19 (48)
Percentage of children eligible for free school meals, median (IQR)	12 (8–24)	14 (10–23)
IDACI, median (IQR)	0.17 (0.08–0.24)	0.16 (0.10–0.27)
Teacher (cluster) characteristics	N_t = 40	N_t = 40
> 5 years of teaching, <i>n</i> (%)	20 (50)	27 (68)
Age (years), mean (SD)	34.5 (9)	31.4 (9)
Female, <i>n</i> (%)	32 (80)	33 (83)
NQT, <i>n</i> (%)	2 (5)	0 (0)
Management position, <i>n</i> (%)	4 (10)	2 (5)
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)
MBI		
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)
EFQ (teacher well-being), mean (SD)	17.2 (6.9)	13.9 (6.6)
Pupil characteristics	N_p = 1037	N_p = 1038
Female, <i>n</i> (%)	483 (47)	491 (47)
Age in years at last birthday, mean (SD; range)	6.2 (1.4; 4–9)	6.4 (1.3; 4–8)
Ethnicity, <i>n</i> (%)	N_p = 721	N_p = 701
White	689 (96)	663 (95)
Black	4 (1)	4 (1)
Asian	5 (1)	11 (2)
Mixed	20 (3)	18 (3)
Other	3 (0)	5 (1)

continued

TABLE 4 Baseline characteristics by trial arm status (continued)

Variable	Intervention (TCM)	Control (TAU)
	<i>N_p</i> = 595	<i>N_p</i> = 502
Eligible for free school meals, <i>n</i> (%)	70 (12)	64 (13)
	<i>N_p</i> = 860	<i>N_p</i> = 844
IDACI, median (IQR)	0.16 (0.08–0.64)	0.15 (0.09–0.25)
Number of children in household, <i>n</i> (%)	<i>N_p</i> = 770	<i>N_p</i> = 747
1	125 (16)	122 (16)
2	403 (52)	389 (52)
3	175 (23)	158 (21)
4	45 (6)	49 (7)
≥ 5	22 (3)	29 (4)
Housing	<i>N_p</i> = 766	<i>N_p</i> = 744
Lives in rented housing, <i>n</i> (%)	475 (62)	423 (57)
Qualifications	<i>N_p</i> = 758	<i>N_p</i> = 734
Parent's highest qualification, <i>n</i> (%)		
None	29 (4)	46 (6)
GCSE or equivalent/A level or equivalent	377 (50)	377 (51)
University degree or equivalent and above	352 (46)	311 (42)
SDQ score	<i>N_p</i> = 1036	<i>N_p</i> = 1038
SDQ-TD score (teacher report), mean (SD) (clinical cut-off point ≥ 12)	6.8 (5.6)	6.6 (6.1)
SDQ-TD in struggling range ^a (teacher report), <i>n</i> (%)	206 (20)	200 (19)
SDQ behaviour score (teacher report), mean (SD) (clinical cut-off point ≥ 3)	0.8 (1.5)	0.9 (1.6)
SDQ Emotions score (teacher report), mean (SD) (clinical cut-off point ≥ 3)	1.5 (2.0)	1.4 (2.1)
SDQ Overactivity score (teacher report), mean (SD) (clinical cut-off point ≥ 6)	3.3 (3.0)	3.1 (3.2)
SDQ Peer Relationships score (teacher report), mean (SD) (clinical cut-off point ≥ 3)	1.2 (1.5)	1.1 (1.7)
SDQ Pro-social score (teacher report), mean (SD) (clinical cut-off point < 6)	7.3 (2.5)	7.6 (2.4)
SDQ Impact score > 0 (teacher report), <i>n</i> (%) (clinical cut-off point ≥ 1)	395 (38.1)	373 (35.9)
SDQ score	<i>N_p</i> = 733 to 752	<i>N_p</i> = 706 to 715
SDQ-TD score (parent report), mean (SD) (clinical cut-off point ≥ 14)	7.8 (6.0)	7.8 (5.9)
SDQ-TD in struggling range ^b (parent report), <i>n</i> (%)	124 (16.5)	111 (15.5)
SDQ Behaviour score (parent report), mean (SD) (clinical cut-off point ≥ 3)	1.5 (1.6)	1.4 (1.6)
SDQ Emotions score (parent report), mean (SD) (clinical cut-off point ≥ 4)	1.8 (2.0)	1.8 (2.0)
SDQ Overactivity score (parent report), mean (SD) (clinical cut-off point ≥ 6)	3.3 (2.6)	3.3 (2.6)
SDQ Peer Relationships score (parent report), mean (SD) (clinical cut-off point ≥ 3)	1.2 (1.6)	1.3 (1.6)
SDQ Pro-social score (parent report), mean (SD) (clinical cut-off point < 6)	8.4 (1.7)	8.5 (1.7)
SDQ Impact score > 0 (parent report), <i>n</i> (%) (clinical cut-off point ≥ 1)	244 (33.3)	209 (29.6)

TABLE 4 Baseline characteristics by trial arm status (*continued*)

Variable	Intervention (TCM)	Control (TAU)
	$N_p = 1036$	$N_p = 1038$
PBQ, mean (SD)	2.0 (2.4)	1.9 (2.4)
	$N_p = 1025$	$N_p = 1028$
HIFAMS, mean (SD)	10.9 (2.5)	11.1 (2.3)
Assessment of pupil reading level (parent report)	$N_p = 746$	$N_p = 713$
Fluent reader, <i>n</i> (%)	320 (43)	349 (49)
Assessment of pupil relationship with teacher (parent report)	$N_p = 753$	$N_p = 713$
Good relationship, <i>n</i> (%)	632 (84)	622 (87)
Assessment of parent relationship with teacher (parent report)	$N_p = 731$	$N_p = 703$
Good relationship, <i>n</i> (%)	465 (64)	485 (69)
Literacy and numeracy	$N_p = 1036$	$N_p = 1037$
Below average on literacy, <i>n</i> (%)	440 (43)	451 (44)
Below average on numeracy, <i>n</i> (%)	343 (33)	353 (34)
<p>A level; Advanced level; EFQ, Everyday Feelings Questionnaire; GCSE, General Certificate of Secondary Education; MBI, Maslach Burnout Inventory; N_p, pupils denominator; N_s, schools denominator; N_t, teachers denominator. a Struggling defined as scoring ≥ 12 out of 40. b Struggling defined as scoring ≥ 14 out of 40.</p>		

TABLE 5 Distribution of number of TCM sessions attended by teachers

Number of TCM sessions attended	Number of teachers (%)
0	4 (5)
1	2 (2)
2	1 (1)
3	2 (2)
4	4 (5)
5	16 (20)
6	52 (64)
Mean: 6	

TABLE 6 Numbers and percentages of participants with available data

Outcome	Time point, n (%)							
	Baseline		9 months		18 months		30 months	
	TCM	TAU	TCM	TAU	TCM	TAU	TCM	TAU
Primary outcome: teacher-reported SDQ-TD score	1036 (100)	1038 (100)	981 (95)	1020 (98)	894 (86)	954 (92)	856 (83)	900 (87)
<i>Secondary teacher-reported outcomes</i>								
SDQ Behaviour score	1036 (100)	1038 (100)	981 (95)	1020 (98)	894 (86)	954 (92)	856 (83)	900 (87)
SDQ Emotions score	1036 (100)	1038 (100)	981 (95)	1020 (98)	894 (86)	954 (92)	856 (83)	900 (87)
SDQ Overactivity score	1036 (100)	1038 (100)	981 (95)	1020 (98)	894 (86)	954 (92)	856 (83)	900 (87)
SDQ Peer Relationships score	1036 (100)	1038 (100)	981 (95)	1020 (98)	894 (86)	954 (92)	856 (83)	900 (87)
SDQ Pro-social score	1036 (100)	1038 (100)	981 (95)	1020 (98)	894 (86)	954 (92)	856 (83)	900 (87)
SDQ Impact score > 0	1036 (100)	1038 (100)	981 (95)	1020 (98)	894 (86)	954 (92)	856 (83)	900 (87)
PBQ score	1036 (100)	1038 (100)	981 (95)	1020 (98)	894 (86)	954 (92)	856 (83)	900 (87)
APP – Literacy	1036 (100)	1037 (100)	953 (92)	1019 (98)	862 (83)	897 (86)	794 (77)	843 (81)
APP – Numeracy	1036 (100)	1037 (100)	953 (92)	1019 (98)	862 (83)	897 (86)	794 (77)	843 (81)
<i>Secondary parent-reported outcomes</i>								
SDQ-TD score	751 (72)	715 (69)	644 (62)	641 (62)	610 (59)	615 (59)	558 (54)	567 (55)
SDQ Behaviour score	752 (73)	715 (69)	645 (62)	642 (62)	611 (59)	617 (59)	558 (54)	569 (55)
SDQ Emotions score	752 (73)	715 (69)	645 (62)	641 (62)	610 (59)	617 (59)	558 (54)	568 (55)
SDQ Overactivity score	751 (72)	715 (69)	645 (62)	642 (62)	611 (59)	616 (59)	558 (54)	569 (55)
SDQ Peer Relationships score	751 (72)	715 (69)	644 (62)	642 (62)	611 (59)	616 (59)	558 (54)	568 (55)
SDQ Pro-social score	752 (73)	715 (69)	645 (62)	642 (62)	611 (59)	617 (59)	558 (54)	569 (55)
SDQ Impact score > 0	733 (71)	706 (68)	624 (60)	637 (61)	600 (58)	606 (58)	550 (53)	557 (54)
Assessment of pupil reading level	746 (72)	713 (69)	639 (62)	638 (61)	605 (58)	610 (59)	557 (54)	567 (55)
Assessment of pupil relationship with teacher	753 (73)	713 (69)	643 (62)	642 (62)	608 (59)	617 (59)	556 (54)	565 (54)
Assessment of parent relationship with teacher	731 (70)	703 (68)	646 (62)	641 (62)	609 (59)	612 (59)	555 (54)	565 (54)
<i>Secondary child-reported outcomes</i>								
HIFAMS score	1026 (99)	1028 (99)	991 (96)	995 (96)	945 (91)	943 (91)	864 (83)	896 (86)

Primary outcome

Table 7 summarises the comparison between the trial arms at follow-up for the primary outcome measure. The primary complete-case analysis is labelled CC1 in the table. TCM improved child mental health according to the teacher-reported SDQ-TD score by 1.0 point (95% CI 0.1 to 1.9; $p = 0.03$) at the 9-month follow-up. There was little evidence, however, of an effect at the 18-month ($p = 0.85$) and 30-month follow-ups ($p = 0.23$). The findings from the fully adjusted CC2 analysis were similar, except for the fact that there was only weak evidence of an effect at 9 months on the teacher-reported SDQ-TD (adjusted mean reduction 0.8, 95% CI 0.1 to 1.6; $p = 0.09$). Post hoc analysis showed that this is because the large number of children lost from the fully adjusted analysis, lost as a result of missing data on the three parent-reported potential confounders, were also those in whom the TCM effect was greatest. The intervention effect on teacher-reported SDQ-TD was -1.6 (95% CI -2.8 to -0.4) for the 534 children with missing data on the three parent-reported potential confounders and -0.8 (95% CI -1.7 to 0.1) for the remaining 1467 children with complete data. Finally, the fully adjusted analysis of imputed data (MI analysis) provided very similar results to our primary partially adjusted analysis (adjusted mean reduction 1.0, 95% CI 0.2 to 1.9; $p = 0.02$). All of the remaining findings are based on the approach used in the partially adjusted analysis (CC1 analysis). Findings from the CACE analysis were almost identical to those from the primary intention-to-treat analysis, which suggests that the estimated effects would have been no different had all the teachers in the TCM arm attended all six sessions.

Subgroup analysis

Tests of interaction indicated that TCM led to greater reductions in the teacher-reported SDQ-TD score at 9 months (interaction $p < 0.001$) for children who were classified by their teacher as struggling with their mental health at baseline (mean difference -2.6 , 95% CI -4.6 to -0.6) than for children who were not (mean difference -0.4 , 95% CI -1.2 to 0.4). A subgroup effect was also found at 30 months ($p < 0.001$) but not at 18 months ($p = 0.10$). TCM may also have greater benefits at 30 months (interaction test p -value of 0.02) for children taught by teachers with > 5 years' experience (mean difference on teacher-reported SDQ-TD score -2.1 , 95% CI -3.8 to -0.4) than for children taught by teachers with ≤ 5 years' experience (mean difference 0.3, 95% CI -1.3 to 1.9). TCM appeared more effective for cohort 2 schools than for cohorts 1 and 3 schools (interaction p -value of 0.02), but there was little evidence of subgroup effects for the other potential moderator variables.

Secondary outcomes

Table 8 summarises the findings from the teacher-reported secondary outcomes. There was evidence, based on the PBQ score, of reduced disruptive behaviour across all three follow-ups ($p = 0.04$). Likewise, there was evidence that TCM reduces the percentage of children who are classified as struggling according to the SDQ-TD score ($p = 0.05$) and reduces the Inattention/Overactivity score ($p = 0.02$) across all waves. At 9 months only, there was also evidence of a reduction in peer relationship problems ($p = 0.02$) and an improvement in pro-social behaviour ($p = 0.02$). There was little evidence of effects on teacher-reported Emotions and Impact scores, APP, parents' assessment of their child's mental health or the child-reported outcome HIFAMS (Table 9).

National Pupil Database analysis

We received parental consent to access the NPD for a total of 1178 children; this represents 57% of all participating children, with only 71 (3%) parents refusing consent; the remaining 826 parents (40%) did not respond to the invitation.

TABLE 7 Main comparison on teacher-reported SDQ-TD score (primary outcome) using different approaches for handling missing data

Follow-up	Trial arm		Analysis, AMD (I – C)												
	Intervention	TAU	CC1: primary analysis						CC2: sensitivity analysis				MI: sensitivity analysis		
	Mean (SD)	Mean (SD)	<i>n</i>	Estimate	95% CI	<i>p</i>	ICC ^a	<i>n</i>	Estimate	95% CI	<i>p</i>	<i>n</i>	Estimate	95% CI	<i>p</i>
9 months	5.5 (5.4)	6.2 (6.2)	2001	-1.0	-1.9 to -0.1	0.03	0.18	1467	-0.8	-1.6 to 0.1	0.09	2075	-1.0	-1.9 to -0.2	0.02
18 months	6.7 (6.9)	6.5 (6.3)	1848	-0.1	-1.5 to 1.2	0.85	0.18	1371	-0.2	-1.5 to 1.1	0.75	2075	-0.1	-1.4 to 1.1	0.82
30 months	6.1 (6.0)	6.5 (6.6)	1756	-0.7	-1.9 to 0.4	0.23	0.12	1318	-0.6	-1.8 to 0.5	0.30	2075	-0.8	-1.9 to 0.3	0.14

AMD, adjusted mean difference; C, control; I, intervention.
^a ICC from crude (unadjusted) analysis.

TABLE 8 Comparison of teacher-reported secondary outcomes

Outcome	Intervention, mean (SD) or %	TAU, mean (SD) or %	AMD/OR ^a			
			Estimate	95% CI	p-value	ICC
SDQ-TD score in struggling range^b						
9–30 months ^c	16.7	19.2	0.70	0.48 to 0.99	0.05	0.06
SDQ Behaviour score						
9 months	0.7 (1.5)	0.9 (1.6)	–0.1	–0.3 to 0.1	0.27	0.09
18 months	1.0 (1.8)	0.9 (1.7)	–0.03	–0.3 to 0.3	0.86	0.12
30 months	0.9 (1.5)	1.0 (1.8)	–0.2	–0.5 to 0.1	0.18	0.10
SDQ Emotions score						
9 months	1.3 (1.9)	1.5 (2.2)	–0.3	–0.6 to 0.1	0.14	0.20
18 months	1.7 (2.2)	1.6 (2.1)	0.1	–0.3 to 0.6	0.63	0.18
30 months	1.6 (2.1)	1.6 (2.1)	–0.005	–0.4 to 0.3	0.98	0.09
SDQ Overactivity score						
9–30 months	2.7 (2.9)	2.8 (3.0)	–0.4	–0.7 to –0.1	0.02	0.08
SDQ Peer Relationships score						
9 months	0.8 (1.4)	1.0 (1.7)	–0.2	–0.4 to –0.03	0.02	0.12
18 months	1.1 (1.7)	1.0 (1.6)	0.1	–0.2 to 0.4	0.62	0.13
30 months	1.1 (1.6)	1.1 (1.7)	–0.07	–0.4 to 0.2	0.60	0.10
SDQ Pro-social score						
9 months	8.2 (2.3)	8.0 (2.3)	0.4	0.1 to 0.8	0.02	0.25
18 months	7.8 (2.4)	8.0 (2.3)	–0.1	–0.6 to 0.4	0.67	0.20
30 months	8.1 (2.2)	7.6 (2.3)	0.5	–0.03 to 1.0	0.06	0.16
SDQ Impact score > 0						
9–30 months	34.5	37.3	0.80	0.61 to 1.05	0.11	0.05
PBQ score						
9–30 months	1.8 (2.4)	1.9 (2.6)	–0.3	–0.5 to –0.01	0.04	0.04
Below average on literacy						
9–30 months (T1 to T3)	32	36	0.91	0.64 to 1.31	0.62	0.07
Below average on numeracy						
9–30 months (T1 to T3)	30	33	0.91	0.64 to 1.31	0.62	0.07

AMD, adjusted mean difference; OR, odds ratio.

^a Mean difference reported for quantitative outcomes and odds ratios reported for binary outcomes.

^b Struggling indicated by scoring ≥ 12 out of 40.

^c The sample size for 9-month assessments is 981 in the intervention arm and 1020 in the TAU arm. The sample size for 18-month assessment is 894 in the intervention arm and 954 in the TAU arm. The sample size for 30-month assessment is 856 in the intervention arm and 900 in the TAU arm.

TABLE 9 Comparison of parent- and child-reported secondary outcomes

Outcome	Intervention arm, mean (SD) or %	TAU arm, mean (SD) or %	AMD/OR ^a			ICC
			Estimate	95% CI	p-value	
Parent-reported outcomes						
<i>SDQ-TD score</i>						
9–30 months	7.7 (6.5)	7.6 (6.4)	0.1	–0.3 to 0.5	0.64	0.01
<i>SDQ-TD score in struggling range^b</i>						
9–30 months	18	15	1.24	0.92 to 1.67	0.16	0.03
<i>SDQ Behaviour score</i>						
9–30 months	1.4 (1.7)	1.3 (1.6)	0.03	–0.1 to 0.1	0.63	0.03
<i>SDQ Emotions score</i>						
9–30 months	2.1 (2.3)	2.0 (2.3)	0.02	–0.2 to 0.2	0.86	0.02
<i>SDQ Overactivity score</i>						
9–30 months	3.0 (2.7)	3.0 (2.6)	0.1	–0.1 to 0.3	0.30	0.004
<i>SDQ Peer Relationships score</i>						
9–30 months	1.3 (1.7)	1.3 (1.8)	–0.03	–0.2 to 0.1	0.66	0.03
<i>SDQ Pro-social score</i>						
9–30 months	8.6 (1.7)	8.7 (1.7)	–0.04	–0.2 to 0.1	0.54	0.002
<i>SDQ Impact score > 0</i>						
9–30 months	33	31	0.94	0.71 to 1.25	0.69	0.04
<i>Assessment of pupil reading level</i>						
Fluent reader						
9–30 months	64	68	0.96	0.75 to 1.23	0.75	0.08
<i>Assessment of pupil relationship with teacher, good relationship</i>						
9–30 months	86	84	1.20	0.93 to 1.54	0.16	0.03
<i>Assessment of parent relationship with teacher, good relationship</i>						
9–30 months	73	73	1.12	0.87 to 1.45	0.39	0.04
Child-reported outcomes						
<i>HIFAMS score</i>						
9 months	10.8 (2.5)	10.9 (2.4)	0.02	–0.3 to 0.3	0.89	0.08
18 months	10.5 (2.5)	10.4 (2.8)	0.3	–0.1 to 0.7	0.17	0.11
30 months	10.4 (2.8)	10.3 (2.8)	0.2	–0.2 to 0.7	0.35	0.11

AMD, adjusted mean difference; OR, odds ratio.

a Mean difference reported for quantitative outcomes and odds ratios reported for binary outcomes.

b Struggling indicated by scoring ≥ 14 out of 40.

Notes

The sample size for 9-month parent-reported assessments ranges from 624 to 646 in the intervention arm and 637 to 642 in the TAU arm.

The sample size for 18-month parent-reported assessments ranges from 600 to 611 in the intervention arm and 606 to 617 in the TAU arm.

The sample size for 30-month parent-reported assessments ranges from 550 to 558 in the intervention arm and 557 to 569 in the TAU arm.

The sample size for 9-month child-reported assessment is 991 in the intervention arm and 995 in the TAU arm.

The sample size for 18-month child-reported assessment is 943 in the intervention arm and 943 in the TAU arm.

The sample size for 30-month child-reported assessment is 864 in the intervention arm and 896 in the TAU arm.

There was little evidence that the intervention had any effect on the rate of overall absence during either the first (adjusted RR 1.08, 95% CI 0.95 to 1.24; $p = 0.24$) or second (adjusted RR 1.1, 95% CI 0.72 to 1.7; $p = 0.65$) year of the trial or on the number of unauthorised absences during the first (adjusted RR 1.03, 95% CI 0.90 to 1.18; $p = 0.62$) or second (adjusted RR 0.96, 95% CI 0.75 to 1.22; $p = 0.74$) year of the trial (Table 10). School exclusions were reported on 22 separate occasions (2 in the intervention arm and 20 in the TAU arm), which resulted in a total loss of 64 separate school sessions, morning or afternoon (3 sessions in the intervention arm and 61 sessions in the TAU arm). These exclusions were issued to a total of six children, two from the intervention and four from the TAU arm of the trial. Tests of interaction did not indicate any subgroup effects for children whose primary or secondary category of SEN was social, emotional or mental health, but this analysis lacked power owing to the small number of children with social, emotional and mental health SEN (Table 11).

TABLE 10 Summary of number of absences in the first and second academic year of the trial for intervention and TAU groups

Year	Median	IQR	Maximum	Median	IQR	Maximum	Crude RR (I/C)	Adjusted		
								RR (I/C)	95% CI	p-value
Year 1	Intervention arm (N = 490)			TAU arm (N = 522)						
Absences	10	4–16	83	9	4–16	93	1.04	1.08	0.95 to 1.24	0.24
Unauthorised absences	0	0–4	55	0	0–2	68	1.44	1.10	0.72 to 1.70	0.65
Year 2	Intervention arm (N = 591)			TAU arm (N = 586)						
Absences	10	4–18	87	9	4–17	93	1.04	1.03	0.90 to 1.18	0.62
Unauthorised absences	4	0–12	65	2	0–10	65	1.25	0.96	0.75 to 1.22	0.74

TABLE 11 Subgroup analysis of SDQ-TD score (teacher reported) at each follow-up

Outcome	Presence	Intervention		TAU		AMD (95% CI)	p-value for interaction
		n	Mean (SD)	n	Mean (SD)		
9-month follow-up							
Social, emotional and mental health SEN	Present	17	13.1 (8.1)	14	18.2 (5.4)	–0.1 (–5.4 to 5.1)	0.89
	Not present	554	4.6 (4.5)	572	4.9 (5.0)	–0.9 (–1.7 to 0.0)	
18-month follow-up							
Social, emotional and mental health SEN	Present	15	14.8 (5.0)	14	15.9 (4.5)	0.5 (–3.4 to 4.4)	0.33
	Not present	530	5.7 (6.2)	564	5.3 (5.4)	–0.3 (–1.7 to 1.0)	
30-month follow-up							
Social, emotional and mental health SEN	Present	16	14.8 (7.2)	12	19.8 (5.4)	–1.4 (–5.4 to 2.7)	0.91
	Not present	526	5.2 (5.4)	548	5.4 (6.1)	–0.7 (–1.8 to 0.4)	

AMD, adjusted mean difference.

Sensitivity analysis of pupil progress

Although there was no main effect of the intervention on APP in either literacy or numeracy, subgroup analysis did indicate that the intervention's effect on APP differs between those who were and were not classified by their teacher as struggling with their mental health at baseline, for both literacy (interaction $p = 0.04$) and numeracy (interaction $p = 0.03$). The intervention arm had lower odds than the TAU arm of below-expectation assessments in literacy [odds ratio (OR) 0.77, 95% CI 0.53 to 1.12] and numeracy (OR 0.82, 95% CI 0.59 to 1.14) among those not classified as struggling, but it had greater odds of below-expectation assessments for literacy (OR 1.17, 95% CI 0.70 to 1.94) and numeracy (OR 1.35, 95% CI 0.88 to 2.06) among those who are classified as struggling. All four of these CIs, however, include unity, so it is difficult to interpret these findings, other than to comment that there seems to be a differential effect according to baseline mental health.

Chapter 4 Economic evaluation

Aim

The aim of the economic evaluation was to assess the cost-effectiveness of the TCM course compared with that of TAU over the short- and long-term periods.

Methods

Perspective

The prespecified economic evaluation took a broad public-sector perspective and included the use of all health, education and social care services, plus criminal justice sector resources and criminal activity. In addition, productivity losses of parents relating to the needs of their child were identified as relevant by the research team and subsequently included in the broad perspective.

Method of economic evaluation

Within-trial cost-effectiveness analysis

The primary economic evaluation was a trial-based cost-effectiveness analysis of TCM (intervention) compared with TAU (control) at 30-month follow-up, explored in terms of the primary outcome measure, the SDQ-TD score. A secondary cost-utility analysis using quality-adjusted life-years (QALYs) based on the EuroQol-5 Dimensions Youth version (EQ-5D-Y)⁶⁵ was proposed, on the assumption that feasibility testing of the measure prior to the start of the trial proved successful.

Long-term cost-effectiveness analysis

A secondary economic evaluation aimed to examine the expected cost-utility of TCM compared with TAU over the longer term, by extrapolating data from the RCT supplemented with data from the literature using decision-modelling techniques,⁶⁶ should the cost-effectiveness analysis of TCM prove promising in the short term.

Costs

Measurement of resources

Service use was recorded for all trial participants using a modified version of the CA-SUS, which has been successfully applied in various youth populations.^{53,54,67,68} The service use schedule was designed to be taken home by the children for self-completion by their parents. It therefore needed to be brief and focused on a set of key resource items, namely those that are high cost and/or high volume of use. Information collected at baseline in the brief CA-SUS covered service use in the 6 months prior to the start of the trial. At subsequent follow-ups, service use since the previous time point was recorded to ensure that the entire duration of the trial had been captured if participants missed intermediate assessments. The measure was tested prior to implementation on a sample of families to ensure face validity and adequate coverage.

Data included in the brief CA-SUS were collected from parents at baseline and at 9-, 18- and 30-month follow-up and included the following health, social care and education-based services:

- use of NHS hospital services – inpatient stays (total number of nights for any reason), outpatient or day-patient appointments (total number of appointments for any reason), and accident and emergency attendances (total number of attendances for any reason and number of times by ambulance)

- use of community health and social care services – primary health care (total number of appointments with any health professional in a general practice surgery), and community mental health and well-being services [total number of contacts with, e.g. Child and Adolescent Mental Health Services (CAMHS), Family Support Workers]
- use of prescribed, psychotropic medication including methylphenidate, dexamphetamine and atomoxetine
- use of local authority accommodation such as foster care, respite care and residential care (total number of days in any of these accommodation types)
- parental productivity loss – time taken off work by parents as a result of child’s behavioural difficulties or other worries.

A standard interview-based version of the CA-SUS, covering a broader range of health, education and social care services, plus criminal justice sector resources, was also developed for completion in interview with a random sample of 50 parents at either the 18-month or 30-month follow-up, to assess the validity of the brief self-report data. Data were collected on the resources outlined below, some of which overlapped with the brief self-report questionnaire but commonly with more detail requested, plus additional resources to ensure coverage of our specified perspective.

Services overlapping with the brief self-report questionnaire

- Use of local authority accommodation such as foster care, respite care and residential care (total number of days in any of these accommodation types).
- Use of NHS hospital services: hospital admissions (name of hospital, medical specialty and number of nights for each admission), outpatient or day-patient appointments, including CAMHS appointments if based in a hospital setting (name of hospital, medical specialty and number of appointments for each reason), and accident and emergency attendances (name of hospital, reason for attendance, number of contacts for each reason and use of ambulance for each reason).
- Use of community health and social care services: general practitioner (GP) at home, surgery or by telephone (number of contacts and average duration per contact for each type of appointment); practice nurse (total number of contacts and average duration per contact); district nurse, health visitor, midwife or school/college nurse (total number of contacts and average duration per contact); community paediatrician or school doctor (total number of contacts and average duration per contact); care co-ordinator, case manager or key worker (total number of contacts and average duration per contact); psychiatrist (total number of contacts and average duration per contact); clinical psychologist (total number of contacts and average duration per contact); CAMHS worker (total number of contacts and average duration per contact); counsellor (total number of contacts and average duration per contact); family therapist (total number of contacts and average duration per contact); social worker (total number of contacts and average duration per contact); family support worker (total number of contacts and average duration per contact); social services youth worker (total number of contacts and average duration per contact); educational psychologist (total number of contacts and average duration per contact); education welfare officer (total number of contacts and average duration per contact); helpline (total number of contacts and average duration per contact); minor injury unit (total number of contacts and average duration per contact); and walk-in centre (total number of contacts and average duration per contact).
- Use of prescribed, psychotropic medication for mental health problems such as depression, anxiety and attention deficit hyperactivity disorder (ADHD) (for each medication listed – name of medication, date medication use started, dose of medication, frequency of medication use, date medication use stopped).

Additional services not covered in the brief self-report questionnaire

- Use of *education facilities*: mainstream school, specialist school (e.g. for emotional and behavioural difficulties), residential school, pupil referral unit or other exclusion service and home tuition (hours attended per day, days attended per week and the number of weeks attended for each type of facility).
- Use of *criminal justice services*: contacts with criminal justice professionals or services such as youth offending team worker, police and solicitor or other legal representative (number of contacts and average duration per contact for each type of professional or service); whether or not the child had been a victim of a crime (details of crime and total number of offences for each type of crime); and whether or not the child had committed any crimes (details of crime and total number of offences for each type of crime).
- Data on the training and supervision requirements for teachers in the intervention group and on the delivery of the TCM training were collected directly from STARS trial records.

Valuation of resources

For each participant, a nationally applicable unit cost was applied to each item of service use reported in the self-report CA-SUS, to calculate the total cost for the duration of the trial, and in the interview-based CA-SUS, to calculate the total cost for the period of interest (as data collected only at either the 18- or the 30-month time point). Costs for NHS hospital contacts were sourced from *NHS Reference Costs 2014–15*.⁶⁹ Costs contained in the annual *Unit Costs of Health and Social Care* publication were applied to community-based health, social care and local authority accommodation services.⁷⁰ The costs of medications were based on averages listed in the *British National Formulary*⁷¹ for the generic drug and were calculated using national averages of dosage for young children. Productivity losses were valued using workers' median gross earnings as listed in the Office for National Statistics (ONS) *Annual Survey of Hours and Earnings: 2015*.⁷² Unit costs applied were for the 2014/15 financial year and were discounted by 3.5% if incurred after 12 months, as recommended by the National Institute for Health and Care Excellence (NICE).⁷³ All unit costs applied are listed in *Table 12* and are reported in Great British pounds.

TABLE 12 Unit costs and sources used in economic analyses

Service	Unit cost or range (£)	Source	Notes
Inpatient hospital (per night)	488.00–674.00	<i>NHS Reference Costs 2014–15</i> ⁶⁹	Weighted average of short and long stay
Outpatient hospital (per appointment)	180.00	<i>NHS Reference Costs 2014–15</i> ⁶⁹	
A&E (per attendance)	152.00	<i>NHS Reference Costs 2014–15</i> ⁶⁹	
Community mental health (per appointment)	43.00	Curtis 2015 ⁷⁰	Assuming 30-minute appointment with CAMHS/Family Support Worker
General practice (per appointment)	18.91	Curtis 2015 ⁷⁰	Average of GP and nurse; assuming 10-minute consultation
Medication (per mg)	0.02–0.19	BNF 70 ⁷¹	Assuming recommended national guidelines on dosage
Accommodation (per day)	317.79	Curtis 2015 ⁷⁰	Average of local authority foster care, private foster care and private residential care
Productivity loss (per day)	74.00	ONS <i>Annual Survey of Hours and Earnings: 2015</i> ⁷²	Based on part-time workers' median gross earnings

A&E, accident and emergency; BNF, British National Formulary.

For education facilities, all young people were reported to be in mainstream school for the duration of the trial; no specialist or residential facilities or home-based tuition were reported, so education facilities were excluded from the economic analyses. For criminal justice sector resources, no criminal justice contacts or criminal activity were reported, so criminal justice resources were also excluded.

The cost of TCM training delivery was calculated using a standard microcosting (bottom-up) approach⁷⁴ and was based on the initial costs for training the IY[®] TCM group leaders, as well as costs for the training and supervision of teachers receiving the TCM training. These included costs of travel, subsistence, course books and other materials, as well as administrative support provided to group leaders and teachers. Costs of teachers' salaries and appropriate capital, administrative and managerial overheads were not included in the calculation because the intervention was delivered in routine classes and as part of teachers' usual responsibilities. However, the costs to schools to provide supply cover in order for the teachers to attend TCM training was included. TCM is a universal intervention, and teachers commonly teach multiple classes and frequently change employment. Intervention costs were allocated across an estimated number of individual pupils who were likely to benefit from the teacher training, not just those included in the trial. The estimate assumed that teachers remained in the workforce for 5 years, as per published reports of trends in teaching,⁷⁵ and an average class size of 30 students.

Data analysis

The primary economic evaluation was a within-trial analysis exploring cost-effectiveness at the 30-month follow-up point using the SDQ-TD scores. A secondary analysis using EQ-5D-Y measure of health-related quality of life was proposed, provided that feasibility work to test the EQ-5D-Y in this younger population prior to the start of the trial proved successful. The youth version of the EuroQol 5-Dimensions (EQ-5D) is validated for 8- to 15-year-olds,⁶⁵ and the study team were testing a simplified version for 4- to 9-year-olds. Our feasibility work suggested that our very young sample struggled to understand the concepts involved, which, combined with a lack of evidence to support the reliability and validity of the EQ-5D-Y among children under the age of 8 years, led us to drop the measure from the protocol.⁷⁶

Instead, a published algorithm was used to map SDQ scores to the Child Health Utility Index 9 Dimensions (CHU-9D), a generic preference-based health-utility measure that combines children's length of life and quality of life into a single summary measure⁷⁷ and thus allows comparison of interventions in a decision-making context. The CHU-9D was not considered for direct measurement in the study because, at the time of designing and implementing the study, no valuation system was available and the measure was not validated for use with 5- to 6-year olds.

The mapping algorithm, developed in a sample of 200 caregivers of young people in Australia attending CAMHS,⁷⁸ transforms the five SDQ subscale scores (emotion, conduct, hyperactivity, peer and prosocial) to give utility weights and is presented in *Equation 1*:

$$\begin{aligned} \text{Utility} = & 0.88 + (-0.019 \times \text{emotion}) + (-0.009 \times \text{conduct}) + (-0.001 \times \text{hyperactivity}) \\ & + (-0.008 \times \text{peer}) + (0.005 \times \text{prosocial}). \end{aligned} \quad (1)$$

The resulting utility weights are used to calculate QALYs using *Equation 2*:

$$\text{QALY} = \text{utility weight} \times \text{years of life}. \quad (2)$$

So, for example, a child who lives for 2 years with quality of life weighted at 0.5 (2×0.5) will have one QALY. QALYs were calculated using the area-under-the-curve approach, in which changes in utility scores were assumed to follow a linear path.⁷⁹

Within-trial cost-effectiveness analyses using QALY outcomes were assessed at 9, 18 and 30 months to examine the trajectory of cost-effectiveness in the first instance and to support longer-term decision analysis.

Three scenarios were considered, as per the analysis of clinical data. The base-case analysis (CC1), conducted for within-trial analyses using both SDQ-TD score and QALY outcomes, was based on complete cases and partially adjusted for prespecified potential confounders (cohort, child gender and the three school-/class-level factors used to balance randomisation) and for baseline costs and outcomes, as appropriate.⁷⁹ In addition, fully adjusted sensitivity analyses (same confounders as CC1 plus IMD score based on child's address, number of children living in the household and whether or not the child's household was rented) were conducted for complete cases (CC2) and 50 multiply imputed data sets (MI) using the chained equations approach for within-trial analysis using SDQ-TD scores.

Resource use

Differences in the use of services between trial arms over the 30-month follow-up period were compared descriptively. These are reported for each service as the mean, SD and range by group and as a percentage of the group who had at least one contact. No statistical comparisons were made to avoid problems of multiple testing and to keep the focus of the evaluation on costs and cost-effectiveness.

Difference in costs

Costs for the use of individual services were summed to give total costs per participant. Differences in mean costs between trial arms at 30-month follow-up were compared using standard parametric *t*-tests, and bias-corrected accelerated non-parametric bootstrapping was used to assess the robustness of these results.^{80,81} Despite the skewed nature of cost data, this is the recommended approach as it allows inferences to be made about the arithmetic mean and is more meaningful from a budgetary perspective.⁸²

Short-term cost-effectiveness analysis

Incremental cost-effectiveness ratios (ICERs) (i.e. the additional cost of one intervention compared with another divided by the additional effect) were calculated based on parameter estimates from random-effects linear regression models that model the costs and SDQ-TD score (primary analysis) or QALY (secondary analysis) outcomes, taking account of the hierarchical structure of the data in cluster randomised trials. Although this method allows costs and outcomes to be considered together in a decision-making context, the ICER is calculated from four sample mean values and is therefore subject to statistical uncertainty.

The uncertainty of these estimates was explored first by bootstrapping 1000 resamples to generate a new distribution of estimates and plotting these on to a cost-effectiveness plane for interpretation and, second, by constructing cost-effectiveness acceptability curves (CEACs). The CEAC is a plot of the probability of the intervention being cost-effective (*y*-axis) for a range of willingness-to-pay thresholds per unit improvement in outcome (*x*-axis).⁸³

Long-term cost-effectiveness analysis

To explore the longer-term implications of TCM, results from the trial were to be extrapolated and supplemented with data from the literature using decision analytic modelling techniques. As this was an exploratory analysis, details of the model were not specified in advance because the need for and design of the model would be dependent on the outcome of the within-trial analyses using QALY outcomes, so these are not specified here.

Results

Data completeness

The availability of service use data at each follow-up is summarised in *Table 13*. Data for the entire follow-up were available for 507 (49%) participants in the intervention arm and 500 (48%) participants in the TAU arm. One participant was identified as an influential outlier⁸⁴ (i.e. with total costs in the 99th percentile) and was removed from the analysis.

Resource use over follow-up

Use of services over the 30-month follow-up period, as reported in the brief version of the CA-SUS, are summarised in *Table 14*. Services most commonly accessed by the young people in this trial included general practice professionals (accessed by 67% of the population) and outpatient services (accessed by approximately 37% of the population). Accident and emergency contacts were reported by almost 30% of the sample, whereas inpatient hospital admissions (reported by around 5% of the sample) and prescribed psychotropic medications (< 1% of the sample) were rare. The use of health services was broadly similar

TABLE 13 Availability of service use data at follow-up

Assessment period	Intervention, n (%)	TAU, n (%)
Baseline	754 (73)	720 (69)
9 months	646 (62)	645 (62)
18 months	612 (59)	617 (59)
30 months	559 (54)	569 (55)
All periods	508 ^a (49)	500 (48)

a Includes one participant who was identified as an influential outlier and removed from the analysis.

TABLE 14 Service use over the 30-month follow-up period

Service	Unit	Intervention (n = 507)			TAU (n = 500)		
		Mean (SD)	Range	Per cent using	Mean (SD)	Range	Per cent using
Inpatient hospital	Number of nights	0.1 (0.6)	0 to 7	5	0.1 (1.2)	0 to 23	4
Outpatient hospital	Number of appointments	1.3 (2.5)	0 to 20	39	1.1 (2.3)	0 to 17	35
A&E	Number of visits	0.5 (1.1)	0 to 13	29	0.5 (1.1)	0 to 10	30
Community mental health	Number of appointments	0.4 (2.1)	0 to 31	8	0.7 (5.0)	0 to 104	10
General practice ^a	Number of appointments	2.2 (2.9)	0 to 30	67	2.4 (3.1)	0 to 33	67
Medication ^b	Per cent using	N/A	N/A	0.4	N/A	.	1
Accommodation ^c	Number of days	0.0 (0.0)	0 to 0	0	0.2 (2.6)	0 to 56	0.4
Productivity loss	Number of days taken off work by parent as a result of child's behaviour or other worries	1.0 (16.5)	0 to 370	6	0.4 (2.8)	0 to 49	5

A&E, accident and emergency; N/A, not applicable.

a Includes appointments with any practitioner in general practice (e.g. GP, practice nurse, health visitor).

b Includes use of methylphenidate, dexamfetamine and atomoxetine.

c Includes time spent in local authority foster care, respite care and residential care.

across the two arms, as was parental time off work. Local authority accommodation was used only by those in the TAU arm, but this was used by < 1% of the sample.

A subsample of 48 participants completed the standard interview-based version of the CA-SUS. Service use reported in interviews was broadly similar to that reported in the brief self-report questionnaire for resources that overlapped in both. Additional services included in only the standard interview-based version of the CA-SUS were education facilities and criminal justice sector resources. For education facilities, all young people were reported to be in mainstream school for the duration of the trial; no specialist or residential facilities or home-based tuition were reported, so education facilities were excluded from the economic analyses. For criminal justice sector resources, no criminal justice contacts or criminal activity were reported, so criminal justice resources were also excluded.

Costs over the 30-month follow-up period

Table 15 summarises total costs over the 30-month follow-up period of the trial for those children with full economic data (service use and outcome data), including a breakdown of costs by type of service. Cost of the TCM training was calculated to be £11.52 per participant and was applied to the intervention arm. Observed mean total costs of services used over the 30-month follow-up period were very slightly lower for the intervention arm (£524.16) than for the TAU arm (£528.14). However, this difference was not statistically significant [adjusted mean difference (AMD) £30.24, 95% CI –£140.98 to £201.47; *p*-value = 0.7].

Outcomes over the 30-month follow-up period

Table 16 summarises outcomes over follow-up for those children with full economic data. Observed mean SDQ-TD scores were slightly better in the intervention arm (5.2) than in the TAU (5.4) arm at 30-month follow-up. However, this difference was not statistically significant (AMD –0.5, 95% CI –1.7 to 0.6; *p* = 0.4).

There was little difference in mapped utility scores between trial arms and across follow-up, and little change over time, which resulted in very small differences in QALYs across trial arms (see Table 16). Despite slightly better utilities in the intervention arm at 30 months, mean QALYs were slightly lower in the intervention arm (2.161) than the TAU arm at final follow-up (2.192). This is partly due to the very small differences between the two groups and the fact that QALYs are a cumulative measure calculated as the

TABLE 15 Mean costs (£) per participant over the 30-month follow-up period (CC1)^a

Costs	Intervention mean cost, £ (SE) (n = 507)	TAU mean cost, £ (SE) (n = 500)	Unadjusted mean difference	AMD		
				Estimate	95% CI	<i>p</i> -value
Baseline costs	119.82 (15.20)	115.99 (10.87)	3.83	–11.67	–49.90 to 26.55	0.6
Follow-up costs						
Intervention	11.52 (0.00)	0.00 (0.00)	11.52	N/A	N/A	.
Hospital	373.49 (0.70)	354.59 (36.80)	18.90	21.60	–70.13 to 113.33	0.6
Community	60.24 (0.70)	74.91 (8.26)	–14.67	–12.42	–24.12 to 9.28	0.3
Medication	2.93 (0.70)	19.08 (7.06)	–16.16	–9.11	–25.93 to 7.71	0.3
Accommodation	0.00 (0.00)	49.36 (27.66)	–49.63	–49.70	–127.21 to 27.80	0.2
Productivity loss	76.60 (57.18)	27.67 (40.58)	48.93	–70.41	–42.51 to 183.33	0.2
Total over follow-up	524.16 (90.60)	528.14 (64.31)	–3.98	30.24	–140.98 to 201.47	0.7

N/A, not applicable.

a CC1: partially adjusted complete-case analysis (base-case).

TABLE 16 Mean outcomes per participant over follow-up (CC1)^a

Outcome	Intervention, mean (SD) (n = 507)	TAU, mean (SD) (n = 500)	Unadjusted mean difference	AMD		
				Estimate	95% CI	p-value
SDQ-TD scores^b						
Baseline	5.9 (5.2)	5.6 (5.6)	0.3	0.3	-0.6 to 1.2	0.6
9 months	4.6 (4.5)	5.0 (5.3)	-0.4	-0.7	-1.5 to 0.1	0.1
18 months	5.6 (6.0)	5.3 (5.4)	0.3	-0.1	-1.3 to 1.1	0.9
30 months	5.2 (5.5)	5.4 (6.1)	-0.2	-0.5	-1.7 to 0.6	0.4
Utility scores						
Baseline	0.873 (0.052)	0.876 (0.057)				
9 months	0.886 (0.044)	0.881 (0.054)				
18 months	0.873 (0.061)	0.879 (0.055)				
30 months	0.878 (0.056)	0.876 (0.060)				
QALYs						
9 months	0.659 (0.032)	0.659 (0.038)	0.000	0.003	-0.001 to 0.006	0.1
18 months	1.318 (0.061)	1.319 (0.070)	-0.001	0.005	-0.003 to 0.013	0.3
30 months	2.161 (0.196)	2.192 (0.127)	-0.032	-0.048	-0.106 to 0.010	0.1

a CC1: partially adjusted complete-case analysis (base case).
b Higher SDQ scores reflect poorer outcomes.

area under the curve between baseline and 30-month follow-up. The difference was not statistically significant after adjusting for baseline variables (AMD -0.048, 95% CI -0.106 to 0.010; $p = 0.1$). AMD at 9 months and 18 months was 0.003 and 0.005, respectively, and these differences were also not statistically significant (p -value of 0.1 and 0.3, respectively).

Short-term cost-effectiveness analysis using Strengths and Difficulties Questionnaire Total Difficulties score

For the primary cost-effectiveness analysis using the SDQ-TD score, the lower costs and better outcomes in the intervention group generate an ICER of -£19.90 per unit improvement in SDQ-TD, which suggests that the intervention dominates TAU and is cost-effective at the 30-month follow-up point.

Figure 4 shows the scatterplot of bootstrapped mean differences in costs and outcome. The majority of the scatter points indicate that TCM is more effective than TAU (to the right of the y-axis) and lie in the north-east quadrant (50%), where TCM is more effective but more costly, and the south-east quadrant (29%), where TCM is more effective and less costly. The remaining scatter points show poorer outcomes for TCM than for TAU and fall in the north-west (14%; less effective, more costly) and south-west (7%; less effective, less costly) quadrants.

The CEAC for the partially adjusted primary analysis (SDQ-TD score; CC1) suggests that the probability of TCM being cost-effective compared with TAU ranges from just under 40% at a zero willingness to pay for a unit improvement in SDQ-TD score to nearly 80% at a £1500 willingness-to-pay threshold (Figure 5). The point at which TCM has a higher probability of being cost-effective compared with TAU (i.e. probability > 50%) is at a willingness-to-pay level of £70 per unit improvement in SDQ-TD score.

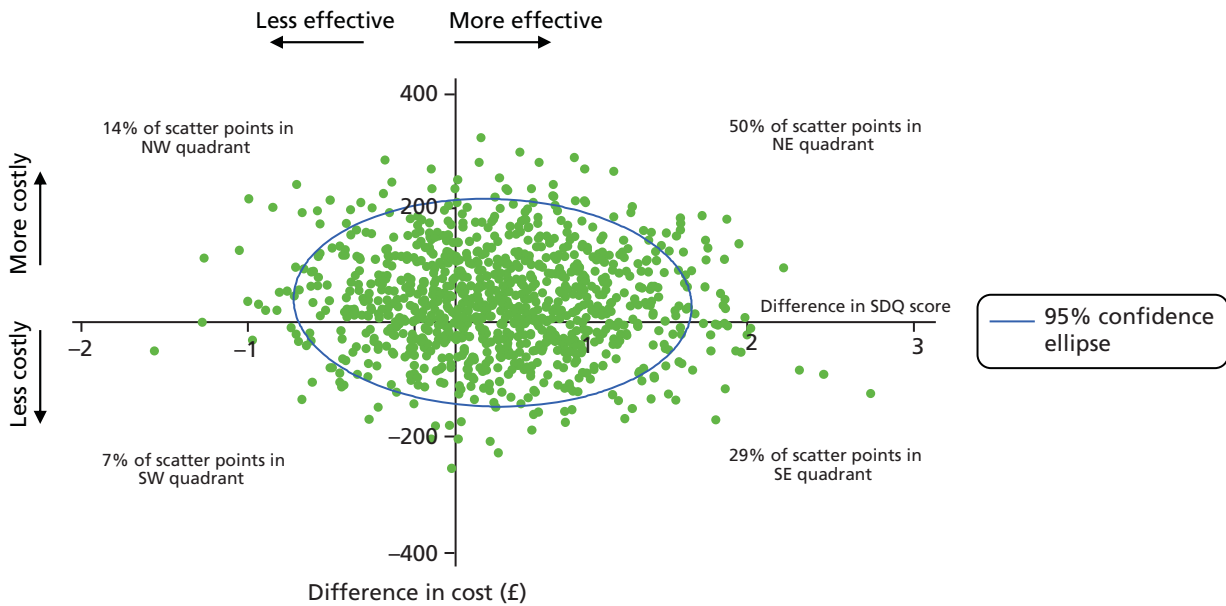


FIGURE 4 Bootstrapped mean differences in costs and effects (SDQ-TD score) of intervention compared with TAU at 30-month follow-up (CC1). NE, north-east; NW, north-west; SE, south-east; SW, south-west.

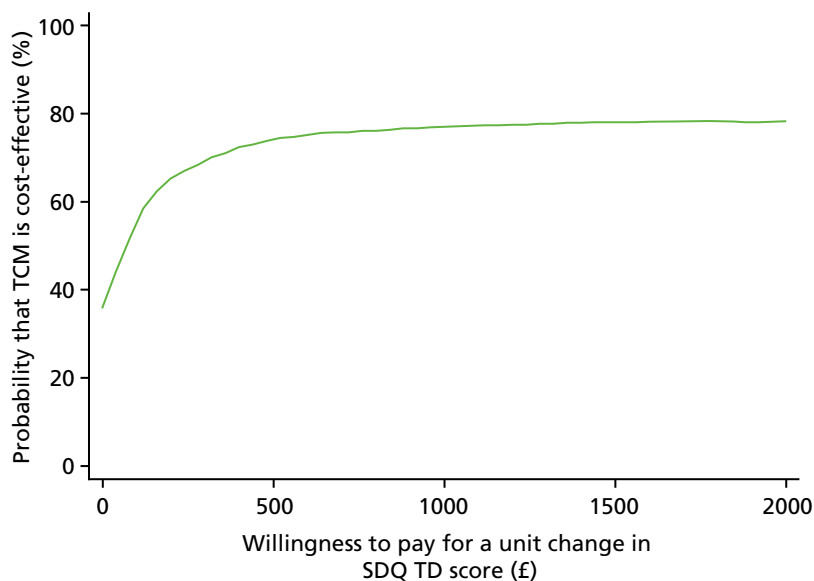


FIGURE 5 Cost-effectiveness acceptability curve showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using the SDQ-TD score at 30-month follow-up (CC1).

The cost and SDQ-TD results of the sensitivity analyses using fully adjusted (CC2) and imputed (MI) data sets are presented in *Table 17*. There were no statistically significant between-group differences in costs or outcomes for either of the sensitivity analyses. For CC2, AMD of total costs (£16.44, 95% CI -£155.32 to £188.21; $p = 0.9$) and outcomes (-0.5, 95% CI -1.7 to 0.6; $p = 0.4$) were similar to that of the primary analysis (CC1), supporting the high probability of TCM being cost-effective for a range of willingness-to-pay thresholds per unit improvement in SDQ-TD score. Imputation of missing data (MI) showed a larger difference in both costs (AMD £252.00, 95% CI -£221.00 to £725.00; $p = 0.3$) and outcomes (AMD -0.6, 95% CI -1.6 to 0.5; $p = 0.3$) and had a negative impact on the probability of TCM being cost-effective, but only at the lower levels of willingness-to-pay thresholds (*Figure 6*).

TABLE 17 Mean costs (£) and SDQ-TD scores per participant over the 30-month follow-up period for sensitivity analyses (CC2 and MI)

Outcome	Intervention, mean cost, £ (SE)	TAU, mean cost, £ (SE)	CC1 AMD (<i>n</i> = 1007)			CC2 AMD (<i>n</i> = 1007)			MI sensitivity analysis AMD (<i>n</i> = 2075)		
			Estimated	95% CI	<i>p</i> -value	Estimated	95% CI	<i>p</i> -value	Estimate	95% CI	<i>p</i> -value
Total costs	524.30 (91.75)	528.67 (67.34)	30.24	-140.98 to 201.47	0.7	16.44	-155.32 to 188.21	0.9	252.00	-221.00 to 725.00	0.3
SDQ-TD score	5.2 (0.6)	5.4 (0.5)	-0.5	-1.7 to 0.6	0.4	-0.5	-1.7 to 0.6	0.4	-0.6	-1.6 to 0.5	0.3

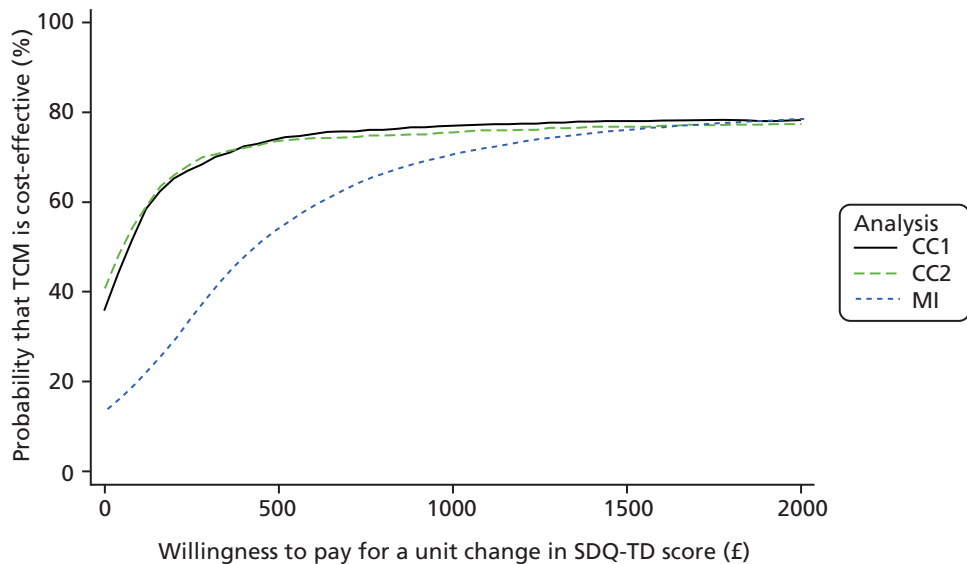


FIGURE 6 Cost-effectiveness acceptability curves showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using SDQ-TD score (all analyses).

Short-term cost-effectiveness analysis using quality-adjusted life-years

For the secondary cost-effectiveness analysis using QALYs, the poorer outcomes and higher costs suggest that the intervention is dominated by TAU at the 30-month follow-up point and is not cost-effective (Figure 7). The majority of scatter points lie in the north-west quadrant (61%) and indicate that TCM is less effective and more costly; 34% of scatter points fall in the south-west quadrant, showing that TCM is less effective and less costly, and the remaining scatter points fall in the north-east (3%; more effective, more costly) and south-east (2%; more effective, less costly) quadrants.

The CEAC for the short-term cost-effectiveness analysis using QALYs suggests that the probability of TCM being cost-effective compared with TAU at 30 months is < 10% at the NICE willingness-to-pay threshold of £20,000 per QALY (Figure 8).

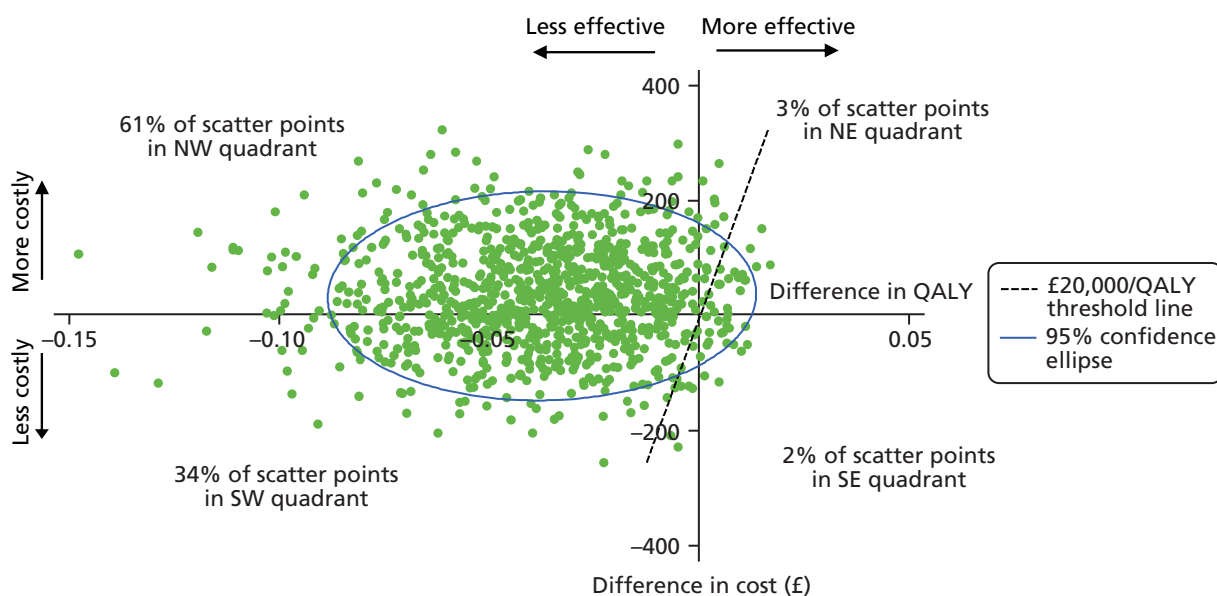


FIGURE 7 Bootstrapped mean differences in costs and effects (QALYs) of intervention compared with TAU at 30-month follow-up (CC1). NE, north-east; NW, north-west; SE, south-east; SW, south-west.

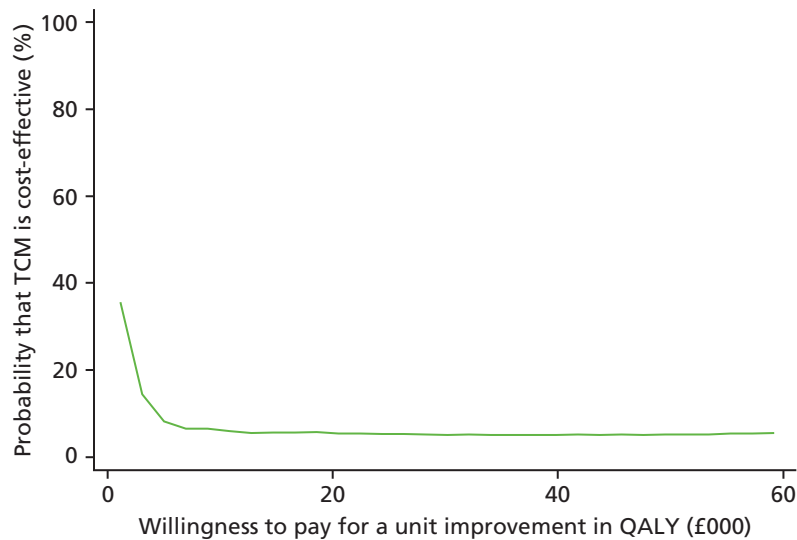


FIGURE 8 Cost-effectiveness acceptability curve showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using QALYs at 30-month follow-up (CC1).

To explore changes over time, cost-effectiveness using QALYs was also explored at 9 and 18 months in post hoc analyses.

The scatterplot of bootstrapped mean differences in costs and QALYs at 9-month follow-up indicates that TCM is more effective than TAU (Figure 9). The majority of the scatter points lie in the north-east quadrant (87%), where TCM is more effective but more costly. The very few remaining scatter points show poorer outcomes for TCM than for TAU and fall in the north-west (8%; less effective, more costly) and south-west (1%; less effective, less costly) quadrants.

The CEAC for cost–utility at 9-month follow-up suggests that the probability of TCM being cost-effective compared with TAU is 60% at a willingness-to-pay threshold of £20,000 (Figure 10).

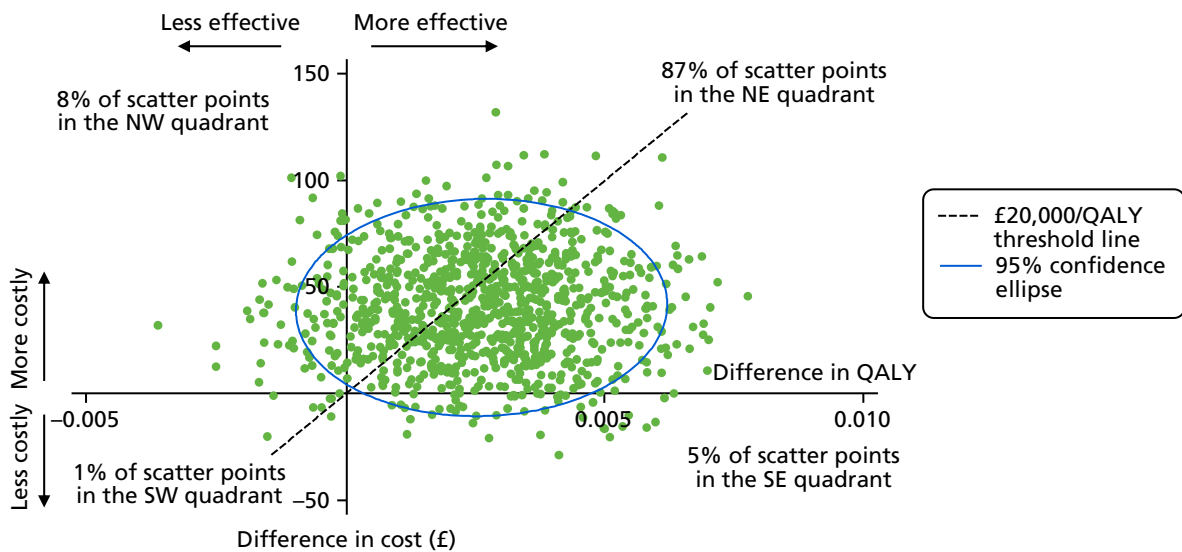


FIGURE 9 Bootstrapped mean differences in costs and effects (QALYs) of intervention compared with TAU at 9-month follow-up (CC1). NE, north-east; NW, north-west; SE, south-east; SW, south-west.

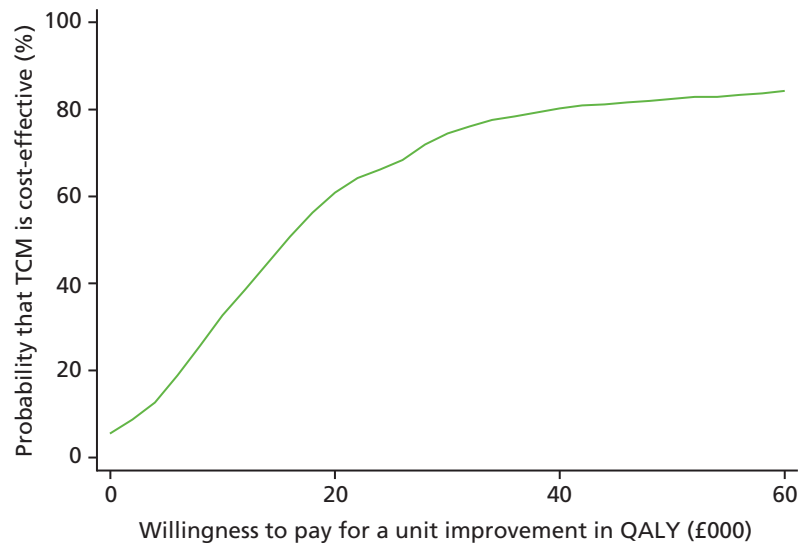


FIGURE 10 Cost-effectiveness acceptability curve showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using QALYs at 9-month follow-up (CC1).

Similarly, the scatterplot of bootstrapped mean differences in costs and QALYs at 18-month follow-up indicates that TCM is more effective than TAU (*Figure 11*). The majority of the scatter points lie in the north-east quadrant (40%), where TCM is more effective but more costly, and the south-east quadrant (41%), where TCM is more effective and less costly. The remaining scatter points show poorer outcomes for TCM than for TAU and fall in the north-west (10%; less effective, more costly) and south-west (9%; less effective, less costly) quadrants.

The CEAC for cost-utility at 18-month follow-up suggests that the probability of TCM being cost-effective compared with TAU is nearly 80% at a willingness-to-pay threshold of £20,000 (*Figure 12*), which is considerably higher than the results at 30-month follow-up (< 10%).

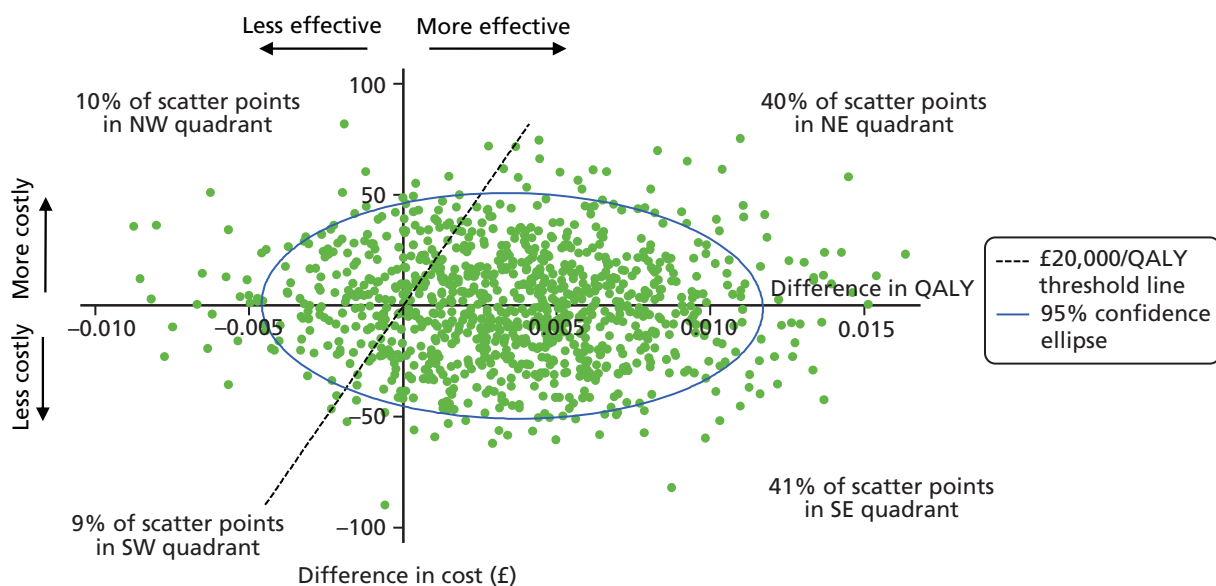


FIGURE 11 Bootstrapped mean differences in costs and effects (QALYs) of intervention compared with TAU at 18-month follow-up (CC1). NE, north-east; NW, north-west; SE, south-east; SW, south-west.

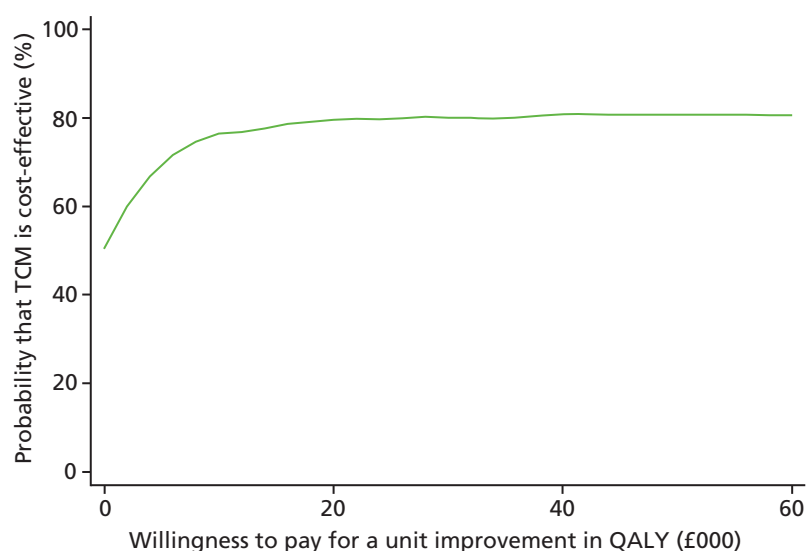


FIGURE 12 Cost-effectiveness acceptability curve showing the probability that TCM is cost-effective compared with TAU for different values of willingness-to-pay thresholds using QALYs at 18-month follow-up (CC1).

Table 18 shows a summary of costs, QALYs and cost-effectiveness over the follow-up period. At a willingness-to-pay threshold of £20,000 per QALY, the probability of TCM being cost-effective is 61% at 9 months and 80% at 18 months but drops to 5% at 30 months. This summary trend is illustrated in Figure 13 and shows that the AMD in costs (TCM – TAU) increases over time, whereas the AMD in QALYs (TCM – TAU) decreases, which has an impact on the probability of TCM being cost-effective after 18 months. The figure also presents trend lines for differences in costs and effects that assume a linear relationship with time and show one possibility of how changes may have occurred over the trial period. The cost-effectiveness analysis does not depend on this assumption, so the linear relationship is not tested.

TABLE 18 Summary of costs, QALYs and cost-effectiveness over time

Summary	Time point			
	Baseline	9 months	18 months	30 months
Costs (£)				
AMD (TCM – TAU)	-11.67	39.07	-1.08	30.24
95% CI	-49.90 to 26.55	-6.08 to 84.21	-51.31 to 49.15	-140.98 to 201.47
QALYs				
AMD (TCM – TAU)	-0.001	0.003	0.004	-0.05
95% CI	-0.006 to 0.004	-0.001 to 0.01	-0.003 to 0.01	-0.11 to 0.01
Cost-effectiveness				
Probability that TCM is cost-effective at WTP of £20,000 (%)	35	61	80	5
Cost/QALY when TCM > TAU (£)	N/A	16,000	10	N/A

N/A, not applicable; WTP, willingness to pay.

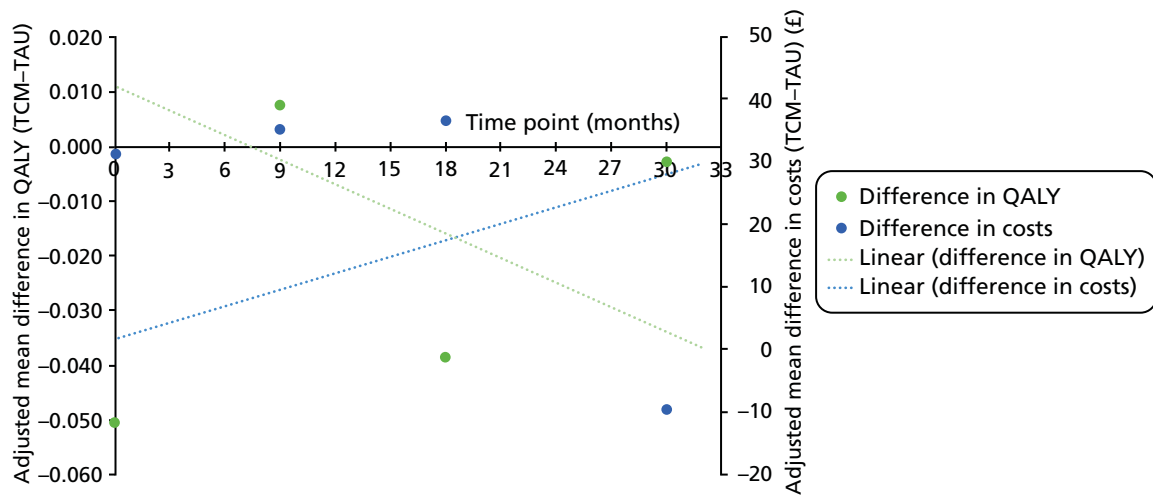


FIGURE 13 Summary of costs and outcome over follow-up.

Long-term cost-effectiveness analysis

Given the clear negative cost-effectiveness results using QALYs at the 30-month follow-up and very small differences in SDQ-TD scores between the two groups at 30 months, any extrapolation into the future using modelling techniques and additional data from longitudinal studies on costs and outcomes over the longer term would serve only to replicate the within-trial results presented. This analysis was therefore not undertaken.

Discussion

Although the results of the clinical analysis alone do not support the hypothesis that TCM is more effective than TAU, the primary cost-effectiveness analysis, comparing outcomes in terms of total SDQ-TD scores, suggested that TCM may be cost-effective compared with TAU at 30 months for a relatively wide range of values of willingness to pay for improvements in SDQ-TD score. This is due to the economic analysis being concerned with the joint difference in both costs and effects; costs were slightly lower and effects (SDQ-TD scores) were slightly better in the TCM group than in the TAU group. It is also important to note that although the clinical analysis focused on statistical significance, the recommended decision-making approach for economic evaluation is to explore cost-effectiveness irrespective of statistical significance in order to make the best use of the available evidence and to determine which intervention has the higher probability of being cost-effective.^{85,86}

The economic results using the SDQ-TD were not altered in the sensitivity analysis controlling for additional randomisation variables (CC2) but were affected by missing data (> 50% missing in CC1). The sensitivity analysis with MI data showed that the probability of TCM being cost-effective was reduced at lower willingness-to-pay thresholds (< £1500) for a unit improvement in total SDQ-TD scores. Although the SDQ is a commonly used measure of mental health in children and adolescents,³⁶ it does not have a set of utility weights and is, therefore, of limited use in allocating scarce health-care resources. With very small differences in costs and effects, as in this study, it is not possible to draw firm conclusions about the cost-effectiveness of TCM without knowing society's willingness to pay for improvements in SDQ-TD scores.

The development of mapping algorithms enables researchers to translate information gained from a non-preference-based measure (such as the SDQ) into health state utility weights for use in economic evaluations. In settings or populations in which it may not be possible to ascertain direct evidence on health state utility (e.g. in primary school children), mapping is considered the only feasible way to conduct cost-utility analyses. One such mapping algorithm, estimated in Australia using the CHU-9D, was used to estimate utility scores for a secondary within-trial analysis using QALYs as the measure of effect, following poor performance of the EQ-5D-Y in feasibility testing.

Mapped utility scores changed very little between trial arms and over time, and the mean score at 30-month follow-up was close to 0.88 in both groups, which approximates the weighted average mean utility score for a general population sample aged < 12 years from developed countries in health-care and school settings.⁸⁷ The mapped cost–utility analysis results suggest that TCM has a higher probability of being cost-effective than TAU at 9 and 18 months but not at 30 months, which raises questions of the appropriate time horizon for analyses of this type. Although health economic evaluations tend to focus on the longest time horizon available within a trial in order to provide the maximum opportunity for cost and outcome impacts to be recorded, in trial-based evaluations it is conceivable that the majority of evaluations demonstrating cost-effectiveness of interventions at early follow-up points would eventually reach a point at which cost-effectiveness was no longer evident, as the time horizon was extended, primarily because of the effects of interventions ‘wash-out’ over time.

Although NICE recommends the use of mapped health state utility estimates when directly collected data are not available,⁷³ the validity of these mapping methods has not been fully addressed. Mapping can result in a loss of information and increased uncertainty and thus requires a closer assessment of the agreement between target and source instruments. It has been argued that current methods for mapping are not known to be conceptually robust⁸⁸ and carry a ‘significant risk that may be harmful to population health’.⁸⁹ It is therefore important to consider the results of the mapped cost–utility analysis with greater uncertainty. It is possible that vital information is lost when mapping the five subscores of the SDQ (emotional symptoms, conduct problems, inattention/hyperactivity, peer problems, and a pro-social scale) to the nine dimensions of the CHU-9D (sad, worried, pain, annoyed, tired, homework or schoolwork, daily routine, activities and sleep) and could therefore lead to erroneous conclusions about cost-effectiveness. Authors of the mapping algorithm used in the present study noted that although the algorithm performed well in predicting mean group observed utility values, it did not perform well in predictions at individual participant level.⁷⁸ Despite published guidelines and a checklist for the reporting of mapping algorithms,^{90,91} only a few studies have been found to fulfil or partly fulfil all criteria on the checklist.⁹²

Small differences in SDQ-TD and mapped utility scores over the trial follow-up made extrapolation beyond the trial period unnecessary, as the baseline beyond which data would be extrapolated from the end of the trial period was almost identical in the two groups. Longer-term analysis would therefore serve only to replicate the within-trial results. Although small clinical differences can result in large changes in health outcomes over longer-term models, there was limited evidence to suggest this in the trial data, particularly given the sharp drop in the probability of TCM being cost-effective using QALYs from 79.5% at 18 months to 4.6% at 30 months. It is, therefore, fair to assume that any projections of costs and effects beyond 30 months would be of questionable validity. Appropriate temporal extrapolation of costs and utilities is reliant not only on clinical outcomes but also on many other long-term processes, such as adaptation to particular health states, the data on which are not currently available for children and adolescents.

Finally, the study was limited by failing to consider spillover effects in the quality of life of teachers. TCM has been designed to increase the well-being of teachers, as well as of students, and future research should attempt to capture these effects by directly administering a preference-based quality-of-life instrument such as the EQ-5D, which is commonly used in adults.⁹³

Chapter 5 Fidelity to the model

This chapter presents the various ways in which we ensured that the intervention was being delivered to teachers in the correct manner.

Overview

The STARS trial used a variety of IY-generated pre-existing materials to ensure that the course was delivered with fidelity to the TCM model, which included the use of manualised checklists of the topics covered by each session and monthly supervision with the programme developer. All group leaders attended the required 3-day basic training in TCM and delivered two courses each as part of the feasibility work to ensure their familiarity with the manual.⁹⁴ There is no existing tool to assess independently how faithfully group leaders deliver training, so, as part of this trial, we began the process of developing such a tool. In addition, we used a pre-existing observational measure, the Teacher–Pupil Observational Tool (T-POT),⁴² to assess if the TCM training had altered the behaviour of teachers in their classrooms. These different methods of assessing fidelity, and their findings, are described in turn below. The data reported in this chapter include only the six TCM groups that included intervention teachers. The final two TCM groups delivered in Year 4 of the study, which included only TAU teachers, are not represented in this chapter.

Supervision with programme developer

Throughout their time delivering TCM workshops for the STARS trial, the six group leaders were supervised by the TCM programme developer, Carolyn Webster Stratton, 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 from each TCM group selected a 10- to 30-minute video section of the day (either one long clip or a variety of shorter clips) to be reviewed by Carolyn Webster Stratton. The videos could be examples of a section that teachers felt went well, or a section that they felt did not go as well and on which they wanted advice, which they explained in a video review form. These materials, plus the teachers' evaluations of that workshop, were sent to Carolyn Webster Stratton, who returned a detailed review of the video clips with comments about how to improve the delivery of the course; these reviews were subsequently discussed in a 1-hour video conference call with all six group leaders. This method of supervision began during the 2 years of feasibility work, which allowed time to refine the process.

The supervision process was very successful and Carolyn Webster Stratton was confident that the group leaders were delivering the course with fidelity to the model. The process was also well received by all of the group leaders who, despite competing time pressures, always completed the necessary preparation.

Group leader-completed agendas and checklists

As part of the supervision process, group leaders completed standardised 'agendas and checklists' of the expected curriculum that they covered in each workshop. These were developed by the TCM programme developers to report on those 'activities' that they had delivered during the session. The checklist included information on:

- training vignettes that were shown in the workshop
- group activities that were used in the workshop, such as specific role-plays or buzz ideas, explaining the rationale for different TCM techniques
- basic housekeeping expected when running the TCM training, such as writing up the day's agenda and ensuring that the teachers completed their self-monitoring checklist, which was completed after each session.

The checklists and agendas were an important tool for ensuring fidelity to the model, as the video review supervision discussed in the section above was only a small segment of a workshop that lasted the whole day. It was, therefore, important to have a method to assess the extent to which important elements of the day's workshop were completed.

Analysis of the 'agendas and checklists' from the six courses found that the mean number of digital versatile discs (DVDs) shown per course group ranged from 14.8 to 19.5. The group leaders tended to show more DVDs in the first two workshops (workshop 1: building positive relationships with students, and the proactive teacher, mean videos shown = 24.9; workshop 2: teacher attention, coaching, encouragement and praise, mean videos shown = 28.9), whereas considerably fewer were shown in workshop 5 (decreasing inappropriate behaviour – follow through with consequences, mean videos shown = 7.8; *Table 19*). This difference is consistent with the number of DVDs that were available for group leaders to show in these workshops (70 for workshop 1, 57 for workshop 2 and 21 for workshop 5).

The number of DVDs shown in each workshop was fairly consistent across groups, with the exception of Group 5, which was shown slightly fewer DVDs in workshop 2 ($n = 15$), and Group 1, which was shown fewer in workshop 3 ($n = 8$) (see *Table 19*). This may be because the group leaders had decided that they needed more time for role-play and discussion.

Group leaders were required to cover certain activities during each workshop and mark whether or not they had done so using a simple checklist. Group leaders reported doing the majority of these activities (mean 92% of activities; range 73–100%) (*Table 20*).

TABLE 19 Total numbers of DVDs shown in each TCM course group and workshop

Workshop	Total number of DVDs shown (group)						Mean (total possible)
	1	2	3	4	5	6	
1	23	27	24	24	22	28	24.9 (70)
2	21	25	23	26	15	20	28.9 (57)
3	8	18	16	11	17	17	15.1 (35)
4	20	16	16	13	13	13	15.5 (33)
5	9	9	8	8	5	7	7.8 (21)
6	16	11	15	11	18	15	15.1 (32)
Mean	16.2	17.7	17	15.5	15	16.7	

TABLE 20 Percentage of activities completed by group leaders in each TCM group and workshop from the standardised checklists

Workshop	Percentage of activities completed (group)						Mean
	1	2	3	4	5	6	
1	85	85	92	92	92	100	91
2	88	94	100	100	88	100	95
3	100	83	100	83	73	100	90
4	100	93	100	93	100	100	98
5	74	84	89	100	82	79	85
6	93	86	87	93	100	87	91
Mean	90	88	95	94	89	94	

Development of new fidelity tool

The supervision process with the IY developer was very successful and demonstrated faithful running of the course; however, we lacked a tool to make independent assessments of fidelity. Using TCM recorded sessions, we developed a quantifiable measure that could be used in the future to assess fidelity formally.

Given that all of the TCM workshops in the STARS trial were video-recorded, this was the most suitable, and rich, source of data to use to construct the measure. The observational measure was designed to systematically analyse specific segments of these video clips, identifying what group leaders did using a tally sheet to code predefined behaviours/actions. Owing to time constraints, only one TCM topic could be studied, and the selected topic needed to be challenging, operationalised and to have a large proportion of the topic delivered by the group leaders.

The element of the course that best met these principles was the TCM delivery of time-out. We had observed that both teachers and the group leaders may initially struggle with the concept of time-out. Time-out in the TCM training is not a punishment but rather a method of giving a child time to calm down in order to regain emotional control to avoid negative behaviours; time-out is the most operationalised strategy in the TCM training. It is also pivotal because if it is not delivered well, and in line with the intervention, it has the potential to undermine the whole TCM training and the principles underpinning it.

An initial tally sheet of behaviours and actions that would be expected from group leaders was developed and trialled using TCM groups delivered in our feasibility work.⁹⁴ It is important to note that, as with all interventions, group leaders may tailor elements of the programme to meet the needs of each of the specific groups, provided that core components are delivered.²³

Based on the experience of applying the tally sheet to feasibility groups, we reduced the number of items to tally (removing duplicates or streamlining items), changed the layout of the sheet so that it flowed better and was easier to use, and edited the descriptions of the items so that they were more concise. The tally sheet was finalised in September 2016 (*Figure 14*) and contained 24 items based on the principle of time-out within TCM. The included items were all chosen for theoretical reasons. There were three sections:

1. explanation of time-out – including whether or not group leaders used questions and discussions and created a supportive atmosphere
2. general – this included free text on what methods were used to explain time-out to teachers and children (e.g. role-play, videos), what activities were used and if the group leaders displayed certain behaviours such as asking for specific feedback
3. principles of time-out – this was the more theoretical section of the tally sheet, which aimed to identify if the group leaders explained the principle with fidelity to the IY model and also acted as an aid for those who may not be familiar with the principle of time-out.

Once the tally sheet had been finalised, two researchers independently coded the session from Group 1 of the STARS trial and, once coding rules were agreed, good agreement was reached between the coders. In total, videos from the time-out section of five TCM courses in the STARS trial were assessed using the fidelity tool because of a failed recording of one course (Group 4). The median length of the videos on time-out was 1 hour 57 minutes (IQR 103–133 minutes). As shown in *Figures 15* and *16*, the tool was sufficiently sensitive to detect differences between the TCM groups in terms of different aspects covered by this section of the TCM training that triangulate with other data availability on the process of delivery. For example, group leaders from Group 5 sought advice in supervision about troublesome group dynamics and teacher feedback was less positive than in other groups. The raters were blind to these issues when completing the tally. The observed differences in frequencies may be because of differences in the personalities of the group leaders, the dynamics of each particular group and how quickly the group understood the principles not only of time-out but also of the whole TCM training.

Explanation of TO	Tally
Link to IY Pyramid	
Highlighted principles of TO (see column on right)	
Reminding teachers of the importance of their own emotional regulation	
Questions and discussions	
Open-ended questions asked	
Asked what could be done differently	
Buzz exercise/free thinking	
Provided a solution without asking other people's opinions/ideas first	
Wrote down things that none of the teachers contributed to	
Disagreed with an opinion/idea	
Supportive atmosphere	
Used labelled praise	
Used unlabelled praise	
Discussing application to own settings	
Used personal experience	
Showed empathy	
Laughter	
Agreed with teachers	

General	
How was TO explained to teachers (e.g. video, role-play, buzz)	
How was explaining TO to children explained (e.g. video, role-play, buzz)	
Demonstrated TO using two/three people to the whole group	
Asked for specific positive feedback about the teacher's involvement in role-play	
Did all teachers role-play?	
Teachers clearly understood what they should be doing in role-play	
Asked what went well in the role-play witnessed	

Comments

Principles of IY® TO

- Time to calm down
- A structured ignore
- Bad behaviour should be ignored by both the teacher(s) and the other children
- TO needs to initiate immediately after undesirable behaviour
- TO is for unsafe or violent behaviours only, including language
- TO should be used sparingly
- TO should be used as a last resort
- Other behaviour plans should be in place and used before TO is implemented (positive opposites, ignoring, etc.)
- Law of Least Disruptive Intervention - the least disruptive way of changing a child's behaviour should be used
- IY® TO is often misunderstood by teachers in the UK, it is different to how TO is normally viewed
- It is not a punishment (i.e. a 'naughty step')
- It is a space for a child to self-regulate their emotions
- TO is a consequence of a behaviour: children must have been previously taught that this type of behaviour will result in TO
- TO ends only when the child is calm in the chair
- If a child refuses to go to the chair and remains in the main group, a strategy suggested is to remove the remaining group of children and start a new activity with them elsewhere
- Teachers should provide verbal reminders to children about how to calm down, but these should be indirect comments, so the child in TO is not directly engaged
- TO should be rehearsed with children before it is used officially
- Children value being with the teacher, TO takes away this privilege
- Children should be welcomed back to the group once TO is completed
- TO needs to be intelligently applied, depending on the child in question
- TO should be ideally a classroom based activity
- The importance of the puppets in explaining TO to children

FIGURE 14 The final tally sheet used to assess the fidelity of the delivery of TCM time-out. TO, time-out.

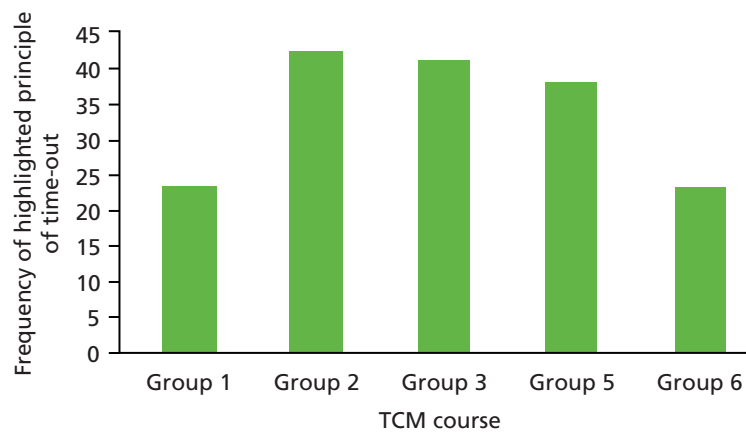


FIGURE 15 Graph of the frequency with which the group leaders highlighted principles of time-out in the five TCM courses (from explanation of time-out section of the tally sheet).

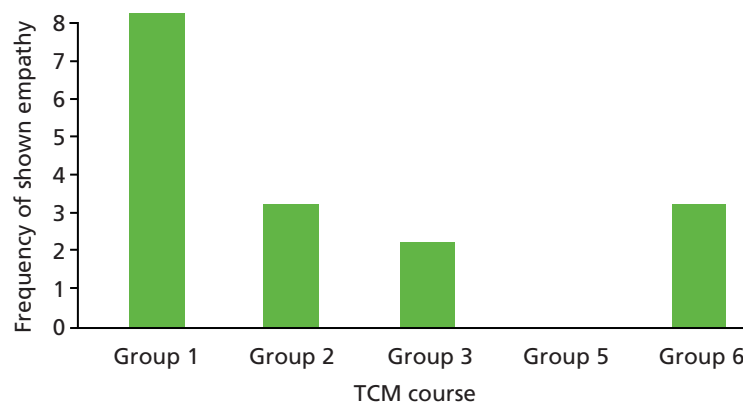


FIGURE 16 Graph of the frequency with which group leaders showed empathy in the five TCM courses (from supportive atmosphere section of the tally sheet).

Using the video data from the STARS trial, we have a prototype measure that could be applied in the future by different researchers to assess the fidelity of a group leader's delivery of the most challenging session within the TCM training. We acknowledge that the measure requires additional testing and benchmarking against other TCM training groups. There are also limitations of the measure that need to be addressed before the tool can be used more widely. The next stage in the development of the model is to develop more precise rules for using the measure and to test the reliability and replicability of the tool. Once complete, we hope that the measure will be used by other researchers when analysing the TCM training.

Teacher-pupil Observation Tool

Introduction

The T-POT⁹⁵ is an observational coding system that was designed to capture potential changes in teachers' classroom behaviour. The T-POT was developed specifically to evaluate components targeted in the TCM training that participating teachers are encouraged to implement, for example using directed, specific praise for appropriate pupil behaviours (e.g. 'You have put those books away in height order, well done') as opposed to non-specific comments (e.g. 'well done') that fail to explain what the desired behaviour is. The items included on the T-POT stem from two other observational systems: the Dyadic Parent-child Interaction Coding System⁹⁶ and the Multiple Option Observation System for Experimental Studies.⁹⁷ The measure has been shown to demonstrate good inter-rater reliability, concurrent and discriminant validity, and internal consistency⁴² and has been used in previous trials to demonstrate changes in teachers' classroom behaviour following attendance on a TCM course.³⁹

Core principles of Teacher Classroom Management that the Teacher–Pupil Observation Tool can evaluate include

1. Training should result in fewer commands being given, but these should be clear, specific and positive.³⁸
2. Teachers should reduce their negative behaviour, and children should increase their positive behaviour.⁹⁸
3. Teachers should use fewer harsh and critical strategies.³⁰
4. Teachers should ignore minor inappropriate behaviour.³⁰
5. Teachers should use specific labelled praise for appropriate behaviour.³⁰
6. Children should increase how compliant they are in response to teacher commands⁹⁹ and decrease off-task and disruptive behaviour.

Categorisation of Teacher–Pupil Observation Tool variables

We chose to use eight collapsed categories of the T-POT variables that focus on teacher behaviour: (1) positive teacher behaviours, (2) negative teacher behaviours, (3) neutral teacher behaviours, (4) labelled praise, (5) unlabelled praise, (6) indirect commands, (7) direct commands and (8) no opportunity given for pupil compliance after a request. Teacher positive behaviour encompassed general positive statements that did not fall under the categories of praise. We used four categories of child behaviour: (1) compliance to teacher requests, (2) non-compliance to teacher requests, (3) negative child behaviours and (4) positive child behaviours.

Using the T-POT, we aimed to test if the following behaviour changes occurred after attendance on a TCM course:

1. teachers increase their use of specific labelled praise
2. teachers use fewer instructions (commands) and those that are used are positive and specific in nature
3. teachers decrease their negative behaviour
4. teachers give more opportunity for children to respond to questions or commands
5. children are more compliant to teacher requests
6. children decrease their off-task and disruptive behaviour.

Teacher–Pupil Observation Tool observation methods

Observations were carried out with each class twice: once in October at the beginning of the school year, prior to TCM training (baseline), and 8 months later in June at the end of the same school year (follow-up). Twenty-three schools participated in the baseline phase; one teacher left their school during the year and so follow-up was completed for 22 schools. Data from the school that the teacher left were therefore excluded from the analysis. There were 13 schools in the intervention group, and 9 in the TAU group. Observers were blinded to the allocation of the schools, and the analyses were carried out by a researcher also blind to group allocation.

Teachers were observed teaching for 30 minutes by either one or two observers at baseline and at follow-up. Teachers were asked to deliver whole-class teaching where possible. The frequency of teacher behaviours was tallied independently by the observers. Child responses to teacher behaviours were also coded, but general child behaviour across the classroom was not, as the focus of the observation was the teacher.

The T-POT measured 36 variables, condensed into 19 main items (11 pertaining to teacher-initiated behaviour and 8 pertaining to child-initiated behaviour), with additional subresponses to code how the child responds to the teacher- or peer-initiated interaction (e.g. whether the child is compliant, non-compliant or has no opportunity to respond). *Table 21* summarises the items that were coded, with examples of behaviour that would be coded under these categories, and outlines the categories that are collapsed into the 12 composite variables used in this process evaluation.

TABLE 21 T-POT variables and behaviour coded, grouping of composite variables

Composite category	Other variables included	Behaviour coded under label
Teacher behaviours		
Teacher positive		General positive affect, encouragement, when/then statements
Teacher negative		Criticism, negative command, warning, 'sssh', negative physical behaviour
Teacher neutral	Acknowledgement	Brief acknowledgement, reflective statements or descriptive comments
	Problem-solving	Statement or question that attempts to encourage child to solve a problem. Double coded with question/command/statement
	Question	Questions
Labelled praise		Specific positive verbalisation
Unlabelled praise		Non-specific praise
Indirect command		Non-specific or implied request for a specific behavioural response (e.g. 'will you please do as I ask?')
Direct command		Specific direct command (e.g. 'come here')
No opportunity for compliance		Teacher gives no opportunity for child to comply after issuing a command or request
Child behaviours		
Positive		Positive verbal behaviour, affect or physical warmth displayed by child
Negative	Aggression to peer (verbal)	Child verbally aggressive to peer
	Aggression to peer (physical)	Child physically aggressive to peer
	Aggressive to teacher	Verbal or physical aggression aimed at teacher
	Destructive	Behaviour that causes damage to an object or intends to damage an object
	Disruptive	Inappropriate non-directed behaviour (i.e. would be disruptive to a classroom but is appropriate during play time)
	Off-task	≥ 30 seconds of off-task behaviour
Child compliance		Compliance or positive response to teacher questions or commands
Child non-compliance		Non-compliance or negative response to teacher questions or commands

Analysis

Composite variables were created from the T-POT items to generate the 12 variables used in the current analysis (see *Table 21*). As the data are count variables, Poisson regression was used to assess whether or not there was an increase in the rate of the measured behaviours at follow-up in the intervention group compared with the TAU group. We report crude RRs, and RRs and 95% CIs that are adjusted for the corresponding baseline behaviour count.

Intraclass correlation coefficients were calculated for the two observational raters on each variable using the *lone* command in Stata in order to quantify inter-rater agreement. The main reported analyses use the primary rater's scores. In addition, after the follow-up observation, each rater was asked to guess which group the teacher belonged to (intervention or TAU).

Teacher–Pupil Observation Tool observation results

Median frequencies and IQRs of each T-POT variable at baseline and follow-up are reported in *Table 22*. Some behaviours were rarely observed, for example child non-compliance and child negative behaviour. The results from the Poisson regression analyses comparing the intervention and TAU groups at follow-up are reported in *Table 23*.

Praise

The rate of labelled praise increased in the intervention group from baseline to follow-up and decreased very slightly in the TAU group. Poisson regression results indicated that after adjusting for the baseline rate, the intervention group used 72% more labelled praise at follow-up than the TAU group (adjusted RR 1.72, 95% CI 1.31 to 2.27; $p < 0.001$).

Commands

Teachers in both groups used fewer direct commands at follow-up than at baseline: there was little evidence that the rate of use at follow-up differed between the groups (adjusted RR 1.00 95% CI 0.86 to 1.17; $p = 0.96$). Teachers in the intervention group increased their use of indirect and total commands, whereas indirect and total command rates decreased in the TAU group. The rate at follow-up was greater in the intervention group than in the TAU group (adjusted RR 1.21 95% CI 1.07 to 1.34; $p < 0.001$).

TABLE 22 Summary of the frequency of teacher behaviours during the 30-minute observation measured by the T-POT at baseline and follow-up by trial arm status

Variable	TAU group ($n = 9$)				Intervention group ($n = 13$)			
	Baseline		Follow-up		Baseline		Follow-up	
	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Teacher								
Teacher positive	29	16–38	35	25–49	37	31–48	51	37–64
Teacher negative	36	30–38	19	12–27	26	18–38	15	7–21
Teacher neutral	180	141–236	180	173–216	196	185–221	181	155–230
Labelled praise	7	3–11	11	2–12	7	3–11	14	10–15
Unlabelled praise	41	36–45	32	31–43	32	28–62	42	35–48
Total praise	52	6–67	43	12–67	41	25–92	56	31–88
Indirect command	90	65–117	86	54–94	63	59–90	73	60–109
Direct command	47	31–52	39	27–46	41	33–28	28	15–38
Total commands	134	71–191	106	59–145	110	61–173	119	41–203
No opportunity for compliance	90	76–97	66	51–77	71	48–100	55	50–65
Child response								
Positives	80	72–87	90	69–98	97	84–106	89	76–106
Negatives	2	0–3	1	0–1	3	0–6	0	0–1
Compliance	141	117–159	158	99–169	120	116–165	162	109–179
Non-compliance	0	0–1	0	0–1	0	0–1	1	0–1

TABLE 23 Comparison of frequency of teacher behaviours at follow-up between the trial arms

Behaviour	Intervention		TAU		Crude RR (I/C)	Adjusted		
	Mean rate per minute (baseline)	Mean rate per minute (follow-up)	Mean rate per minute (baseline)	Mean rate per minute (follow-up)		RR (I/C)	95% CI	p-value
Total praise	1.65	1.96	1.53	1.45	1.35	1.31	1.16 to 1.48	<0.001
Labelled praise	0.24	0.45	0.29	0.27	1.66	1.72	1.31 to 2.27	<0.001
Unlabelled praise	1.42	1.51	1.25	1.18	1.28	1.20	1.05 to 1.38	0.009
Total commands	3.71	3.77	4.47	3.65	1.03	1.23	1.13 to 1.34	<0.001
Direct commands	1.29	1.02	1.50	1.15	0.89	1.00	0.86 to 1.17	0.96
Indirect commands	2.42	2.75	2.97	2.50	1.10	1.21	1.09 to 1.34	<0.001
Teacher								
Teacher positive	1.38	1.67	0.96	1.23	1.36	1.24	1.07 to 1.43	0.004
Teacher negative	0.93	0.59	1.09	0.67	0.88	1.03	0.85 to 1.26	0.75
Teacher neutral	6.69	6.29	6.16	6.50	0.97	0.94	0.88 to 1.00	0.05
No opportunity for compliance	2.31	1.95	3.13	2.16	0.90	1.13	1.00 to 1.27	0.04
Child response								
Positives	3.18	3.02	2.59	2.77	1.09	0.98	0.88 to 1.09	0.72
Negatives	0.14	0.04	0.09	0.04	0.95	0.98	0.44 to 2.15	0.95
Compliance	4.63	5.06	4.47	4.67	1.08	1.03	0.96 to 1.11	0.35
Non-compliance	0.04	0.03	0.06	0.02	1.15	1.24	0.44 to 3.50	0.69

C, control; I, intervention.

Teacher behaviour

Observation of negative teacher behaviours decreased from baseline to follow-up in both groups, and there was little evidence of a difference between groups (adjusted RR 1.03, 95% CI 0.85 to 1.26; $p = 0.75$). Teachers in the intervention group were, however, observed to display positive behaviour more frequently following the intervention (adjusted RR 1.24, 95% CI 1.07 to 1.43; $p = 0.004$).

Opportunity to respond

The frequency with which children were given no opportunity to comply with a teacher request decreased from baseline to follow-up in both groups. The intervention group had a higher rate of giving children no opportunity to comply than the TAU group at follow-up (adjusted RR 1.13, 95% CI 1.00 to 1.27; $p = 0.04$).

Compliance and positive responses

Child compliance increased from baseline to follow-up in both groups, but there was little evidence that there were differences between the two groups (adjusted RR 1.03, 95% CI 0.96 to 1.11; $p = 0.35$). Similarly, child positive responses were not different between the two groups (adjusted RR 0.98, 95% CI 0.88 to 1.09; $p = 0.72$).

Off-task and disruptive behaviour

There was no difference between the groups on child negative and non-compliant behaviour following intervention: both variables decreased from baseline to follow-up in both the intervention and TAU groups (child negative adjusted RR 0.98, 95% CI 0.44 to 2.15; $p = 0.95$; child non-compliance adjusted RR 1.24, 95% CI 0.44 to 3.50; $p = 0.69$).

Inter-rater agreement

Intraclass correlation coefficients for all T-POT variables prior to composite variable creation are reported in *Table 24*. Twenty-one variables had ICCs of > 0.7 , and 21 variables had ICCs of between 0.3 and 0.7. Only two variables had ICCs of < 0.3 . Seventeen of the T-POT variables had zero agreement between raters, but this may relate to the rarity of these behaviours, for example, child destructive behaviour and negative child responses to teacher requests. None of the variables with an ICC of < 0.3 related to teacher behaviour; all related to child responses to teacher or child behaviours that were not the observers' focus. Most variables had an ICC of > 0.5 , with 0.96 being the highest ICC for teacher indirect commands at baseline (see *Table 24*).

Observer guesses

Following their follow-up observations, the observers were asked to guess which trial group they thought the teacher was in (intervention or TAU). Of the intervention schools, 50% of the guesses made by the observers were correct (8/16 recorded guesses). Of the TAU schools, 67% of guesses were correct (8/12 recorded guesses).

Teacher–Pupil Observation Tool observation conclusion

For both groups, many scores moved in the direction hypothesised. The main findings indicated that the use of labelled praise and indirect commands increased more in the intervention group than in the TAU group. Teachers in the intervention group also used more positive and fewer neutral behaviours than teachers in the TAU group at follow-up. In contrast to expected findings, teachers in the intervention group gave children no opportunity to comply to their instructions more frequently than teachers in the TAU group, although in both groups the frequency of this decreased from baseline to follow-up. We did not find evidence of differences in the child behaviour variables between the intervention and TAU groups.

There is the potential that a Hawthorne effect was present. In order to recruit schools to the trial, headteachers and teachers were fully informed of the purpose of the study and told about the TCM programme. Thus, teachers in TAU classes may have applied these principles ad hoc to their own teaching methods if they were selected by headteachers for their interest in attending the course. Future studies using the T-POT should include observations of teachers not involved in the study in order to ascertain how participating in the trial may have affected control group behaviour.

The observers who rated the teacher behaviours were asked to guess which arm each teacher was in after observing and they were correct half of the time for intervention teachers and two-thirds of the time for TAU teachers. This indicates some success in maintaining blinded status for the observers.

Discussion

All intervention trials must measure fidelity to the model to ensure that the trial is truly evaluating the intervention as intended. Fidelity encompasses adherence to, exposure to and quality of the intervention;¹⁰⁰ all aspects should be assessed. In the STARS trial, a number of pre-existing fidelity tools were used: the 'agendas and checklists', high-quality training and regular supervision with the TCM developer to measure adherence and exposure to the intervention. Supervision with the programme developer was very successful and it was telling that her feedback was positive in terms of the fidelity and quality of delivery of the course. We used video data from the STARS trial to develop a tool that could be applied in the future by different researchers to assess the fidelity of a group leader's delivery of the most challenging session within the TCM training. The measure could be used to explore variability within delivery in future studies.

The impact of the TCM training on teachers' behaviour is also an important factor in the fidelity of the intervention. To assess this, we used the pre-existing T-POT measure, an observational tool, with independent observers rating teachers' behaviour in the classroom. Results from the T-POT show increased use of labelled praise as expected, but, counter to our hypothesis, increased indirect commands by

TABLE 24 Intraclass correlation coefficients quantifying inter-rater agreement for all T-POT variables

Description of variable	ICC	
	Baseline	Follow-up
Teacher positive (general positive comments, not praise or commands)	0.39	0.53
Teacher negative (general negative comments)	0.71	0.91
Teacher acknowledgement	0.50	0.74
Labelled praise	0.93	0.64
Unlabelled praise	0.87	0.78
Indirect command	0.96	0.85
Direct command	0.86	0.59
Teacher negative: positive child response	0.67	0.91
Teacher negative: negative child response	0.95	0.00
Teacher positive: positive child response	0.39	0.59
Teacher positive: negative child response	–	–
Problem solving	0.53	0.35
Question	0.78	0.53
Question: child complies	0.78	0.62
Question: child does not comply	0.86	0.70
Indirect command: child has no opportunity to comply	0.90	0.82
Indirect command: child complies	0.27	0.88
Indirect command: child does not comply	0.86	0.39
Direct command: child has no opportunity to comply	0.06	0.59
Direct command: child complies	0.36	0.32
Direct command: child does not comply	0.89	0.00
Child initiates contact with peer	0.72	0.59
Child initiates contact with peer: peer responds positively	0.78	0.58
Child initiates contact with peer: peer responds negatively	0.00	–
Child positive general behaviours	0.79	0.56
Child aggressive behaviour	0.00	–
Child destructive behaviour	–	–
Child disruptive behaviour	0.39	–
Child off task	0.92	0.63
Teacher gives time-out warning	0.00	–
Teacher gives time-out warning: child has no opportunity to comply	0.00	–
Teacher gives time-out warning: child complies	–	–

ICC = 0 where behaviours were very infrequently coded.
The dashes indicate that behaviours were not coded.

teachers in the intervention group compared with teachers in the TAU group. Teachers in the intervention group also used more positive and less neutral behaviours than the TAU group teachers at follow-up, which was predicted. These results suggest changes in the teachers' behaviour after attending the TCM training, but ideally observations would have been carried out in all schools and at all data points.

The endorsement of the programme developer that the programme was delivered with fidelity and to high quality, combined with the group leaders' completion of the TCM programme's checklists and agenda, means that we can be confident that teachers had access to the training as intended. Observations via the T-POT suggested that access to TCM predicted some, but not all, of the predicted changes in teachers' behaviour. This analysis lacked power, as not all classrooms were observed and future studies should observe all classrooms. Our measure of fidelity revealed some variance between groups on the most contentious and challenging aspect of the TCM course, but, with an invalidated measure and only six groups, we were not able to pursue it further. Additional research is needed to ensure that the fidelity tool we have developed is reliable and replicable, at which point studies could be pooled to see if variance in delivery mediates variability in outcome.

Chapter 6 Process evaluation

This chapter reports in detail what teachers thought about the TCM training and how they felt it might have benefited the children in their class.

Background

The process evaluation aimed to explore teachers', headteachers' and TCM group leaders' experiences of TCM, the use and uptake of TCM strategies in the classroom by teachers and other staff and influences on this, and also aimed to improve our understanding of the main outcomes more generally. Our main research questions included:

1. What are the key elements of the TCM training from a teacher's perspective?
2. What are the barriers to and facilitators of the use of TCM strategies in the classroom?
3. What do teachers feel has changed in their approach? What impact has it had on others?
4. Which TCM strategies have been used/not used?

To address these questions, both qualitative and quantitative methods were used (see Hansford *et al.*¹⁰¹) for the process evaluation protocol. These methods included:

1. teacher focus groups (FGs) interviews
2. headteacher interviews
3. group leader FGs/interviews
4. Teacher Satisfaction Questionnaire (TSQ).

This chapter describes the above methods in more detail, presents the results and discusses these in the context of the process evaluation. Although considered part of the process evaluation, information about fidelity is discussed in *Chapter 5*.

Methods

Teacher focus groups/interviews

All teachers who attended the TCM training, irrespective of if they were TAU or intervention teachers, were invited, by e-mail, to take part in a FG immediately after their training had finished. The FGs were explained to teachers as an opportunity to explore their experiences and views of the TCM training, in terms of, for example, how the course was run, positive and negative aspects of the course, and whether or not teachers had used or discussed the TCM strategies with colleagues.

Teachers who were unable to attend the FG were invited to take part in a short telephone interview at a time that suited them. Teachers received a £10 gift voucher to say thank you for attending a FG or interview.

Headteacher interviews

At the end of the first year of the trial, a subsample of headteachers was invited, via e-mail, to take part in a short, semistructured telephone interview. These interviews were designed to ascertain whether or not TCM had had any wider impact on the school, as well as headteachers' views on the use and uptake of the TCM strategies within the classroom more generally.

Teacher Classroom Management group leader focus group/interviews

Teacher Classroom Management group leaders were invited to take part in a FG after the final TCM course had been run (May 2015) and were asked for feedback on the intervention and its relevance to schools and teachers. If group leaders were unable to attend the FG, they were invited to take part in a telephone interview as an alternative.

Procedure for focus groups/interviews

Topic guides, designed by the research team, were used to run the FGs and interviews for teachers, headteachers and group leaders (see *Appendix 1*).

All focus groups lasted approximately 1.5 hours and started with participants setting up ground rules covering areas such as confidentiality. Each FG was run by a trained researcher-facilitator and, where possible, researchers did not facilitate the FGs attended by teachers from their named schools. A researcher, who acted as an observer, recorded the FG using a voice recorder for later transcription and analysis.

All interviews were conducted by a trained researcher and lasted no longer than 45 minutes.

Teacher-completed measures

Teachers were asked to complete a TSQ (see *Report Supplementary Material 3*) after their final TCM session, at 9 months and 21 months post course. The questionnaire asked teachers to rate how helpful the TCM training was in developing their classroom management skills on a three-point scale from 'helpful' to 'unhelpful'. Teachers were also asked to rate how useful they found, and how frequently they used, a list of 24 teaching strategies that were either supported/not supported by TCM on a five-point scale from 'not at all useful' to 'very useful' and 'none of the time' to 'all of the time', respectively. These 24 teaching strategies were combined to create two average scores: one for how useful teachers found the strategies and one for how frequently teachers used the strategies. Negative items (i.e. those that did not support TCM) were reverse coded and one item ('use green light/yellow light/red light as a warning system') was not included as it was deemed unclear whether or not it supported/did not support TCM.

Teachers were also given a shorter workshop evaluation form to complete after each TCM session. This form asked teachers to rate how helpful they found the content, videos, teaching and group discussion on a four-point scale from 'not helpful' to 'very helpful'.

Data analysis

All FGs and interviews were recorded and transcribed for later analysis. Framework Analysis¹⁰² was used to analyse these data. The analysis was conducted by a team of three researchers working on the STARS trial (LH, AP and KA) using NVivo 11 (QSR International, Warrington, UK). Data were analysed immediately after collection to enable data collection to continue only until saturation was reached (as judged by the researchers conducting the analysis).

Results of teacher, headteacher and group leader focus groups/interviews

Participant details

A total of 80 schools were given the opportunity for one teacher to attend TCM training. Intervention teachers ($n = 40$) received TCM training in their first year of involvement in the study and TAU teachers ($n = 40$) were offered a place on the TCM course 1 year later, when they were no longer teaching the trial children.

Of the 80 schools given the opportunity, four TAU schools chose not to send a teacher on the training: two because the original trial teacher had left the school and two because the original trial teacher was deemed too busy to attend. Therefore, in total, 76 schools accessed the training course. One TAU school chose to send a different teacher on the TCM training, as the original teacher had taken on extra responsibilities and did not have the time to attend that year; however, this replacement teacher attended only two sessions, feeling that the course was too basic for them, and later left the school. We therefore offered the training place to the original teacher the following year, in their final year of involvement in the trial. *Figure 17* shows how many teachers accessed the training and in which year of involvement in the study this occurred.

The process evaluation was based on the 77 teachers who attended the TCM training and the respective 76 schools at which these teachers taught.

Teachers

Five separate FGs were conducted with teachers 2 months after they completed their TCM training.

Thirty-one teachers participated in a FG and 14 took part in an alternative telephone interview (see *Appendix 2* for further details). Three teachers did not participate at all; one was unable to attend owing to a staff bereavement within the school and two were unable to attend as they were on maternity leave. Twenty-nine teachers were not invited to take part in a FG/interview as data saturation had been reached. Characteristics were similar across teachers in each of these groups. There was, however, a higher percentage of teachers who taught KS 2 in the group of teachers who chose to be interviewed rather than attend a FG (see *Appendix 3*).

Headteachers

All cohort 1 intervention and TAU school headteachers ($n = 15$) were invited to participate in an interview immediately after their teacher had completed the TCM training. Of these, 14 responded and 13 completed an interview. All cohort 2 intervention school headteachers ($n = 15$) were also invited to take part in an interview 1 year after their teacher had completed TCM training. Of those invited, 12 headteachers responded and 10 successfully completed an interview.

No other headteachers were invited to participate as data saturation had already been reached.

Group leaders

All group leaders were invited to take part in a FG after all the TCM courses had finished. A total of five group leaders attended the FG. One group leader, who was unable to attend, completed a telephone interview.

Results and discussion

Below are the key themes that emerged for each research question. Additional examples of each theme/subtheme can be found in *Appendix 4*. The source of quotations is identified as originating from teachers, group leaders or headteachers, and is followed by an indication of whether the quotation was obtained during an interview or FG. Interviews and FGs are identified by number so readers are able to get a sense of how many individual participants expressed particular themes. However, care has been taken to ensure that it is not possible to identify any participant or school from the content of the quotation.

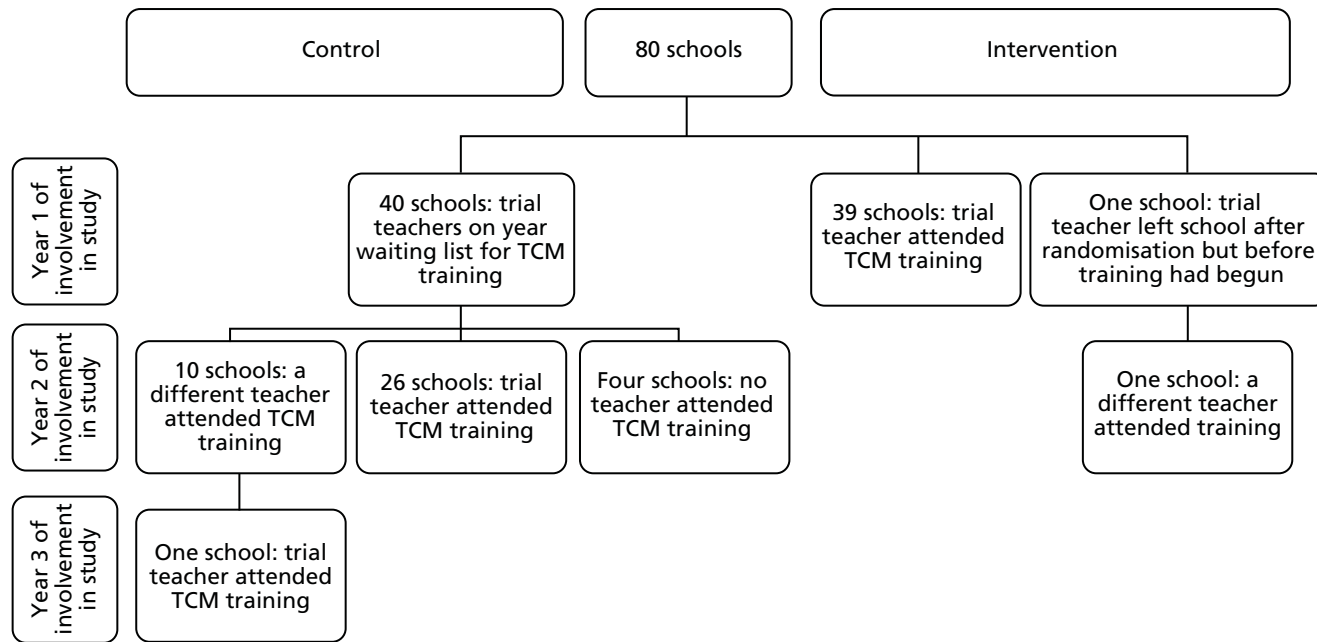


FIGURE 17 Number of teachers who attended the TCM training in each year.

Influences on teachers' learning: what were the key elements of the Teacher Classroom Management training from the teachers' perspective?

Teachers' and group leaders' feedback revealed a number of factors that influenced teachers' learning on the TCM course. Headteachers, however, did not comment on this, although one mentioned that sharing may have been important. Lack of feedback from headteachers may be a result of teachers not discussing this element of TCM with them.

Structure

Teachers appreciated the structure of the course days (i.e. 1 day per month for 6 months), as this allowed for time to practice strategies between sessions:

What's been nice is the gap between the sessions [. . .] it's having the time to go back and work on things, reflect on it and come back together later has been good.

Teacher, FG 2

They also liked that the content 'evolved' as the ethos became more embedded in their practice. They felt that the longer-term nature of the course and the number of sessions gave time for relationships within the group to develop and allowed discussions to deepen:

I think it's nice having so many sessions, 'cos when we all first sat down it's a bit nerve wracking [. . .] you don't know really what to expect, what to say. Then as we got to know each other and it was a really nice group of people so you just relax a little bit more so you are a bit more honest. We got to know each other as the year went on, you can be more open and honest.

Teacher, FG 2

The teachers' appreciation of the 'evolution' of the course was echoed by group leaders who felt that the content was a real strength of the course. However, the depth and breadth of curriculum was experienced as overwhelming at times in terms of the quantity of material that could be included, with group leaders sometimes finding it hard to know what to include and exclude in each session.

Training as a group

Working together and learning as a group emerged as important for teachers and group leaders, with some teachers reporting that this was the most important element of the course (even over course content).

Teachers thought that having teachers from a spread of different schools in different socioeconomic areas and across different year groups facilitated the sharing of ideas and professional practice and gave them a wider perspective from which to reflect on their own practice and context:

We don't get enough chances to talk to other schools about what they do [. . .] it's interesting to do that because sometimes you're very much stuck in your own room. And when you hear about how other people work and what their schools are like [. . .] you do take on some of their ideas.

Teacher, interview 01

Because I work in the lower end of the school, hearing what colleagues further up the school do from other schools, so not just what happens in my school, it's opened up a whole lot of things like 'oh I can do that' or 'colleagues could do that', that's a really positive thing.

Teacher, FG 2

Group leaders also report a mixture of experience, which was positive for the group dynamic, with the more experienced teachers being able to share their experiences with the NQTs and the NQTs bringing their enthusiasm and passion. However, less commonly, some teachers said that this mix made it difficult

for the trainers to cater for everyone's needs. This led them to discuss whether or not it would be preferable to have more targeted courses, for example for NQTs, or for KS 1/KS 2 teachers, although they recognised that this had to be balanced against the advantages of a mixed group.

Teachers reflected that they learnt a lot from group discussions; different teachers interpreted and used strategies in slightly different ways and discussing this helped them to learn from each other. Hearing about how the strategies were working for others also had the potential to motivate some teachers to try them for themselves when they may otherwise have been sceptical.

Teachers liked the size of the group (8–12 participants) and felt that larger groups may have been 'a bit intimidating' and would have limited how well participants got to know each other. Feeling that the group was a 'safe' environment was key for teachers as it meant that they felt able to openly share problems, experiences and ideas. Many teachers commented on the benefits of sharing in a group in which they felt supported and respected:

It was a very safe open environment where everyone's valued.

Teacher, interview 06

For me it was like therapy for teachers [. . .] having the time to come away from the classroom and realise that those little things that really bug you on a day to day basis everyone feels the same and it's OK to get, you know, to feel at times frustrated [. . .] just be reminded of the strategies and ways to deal with it, and that it's OK, was really powerful for me and I went back to school each time for that next sort of few weeks feeling really great [. . .] to have that conversation away from all the stress and the hassle at school and to feel 'actually I am good at my job' and you know that's not about me and that's how I can deal with it.

Teacher, FG 3

Group leaders also noted that one of the things teachers liked best was sharing each other's practices and problems:

They talked a lot [. . .] learning from each other and supporting each other and that felt a very powerful part of the programme.

Group leader, FG

The 'therapeutic' value of being in a group in which they could discuss problems without fear of judgement and realise that the problems are common to others was important for teachers and enabled them to address those issues positively and proactively, increasing their sense of self-efficacy:

It was a chance to be yourself as a teacher but out of the classroom with a bit of time to consider, to think and reflect [. . .] and nobody judging you for the fact that you're not necessarily struggling to cope, but feeling tired, stressed out, pulled in a million different directions and not knowing the right thing to do.

Teacher, FG 1

Being in a group with teachers from different schools contributed to teachers feeling able to reflect honestly on their practice, whereas within their own school they may feel the need to be more reserved.

Group leaders also noticed how openly the teachers were able to talk to each other and this aided their learning. Group leaders believed that teachers enjoyed the chance to take the theory they had learnt each week, try it out in class and report back the following month. The group atmosphere of the course allowed them to talk truthfully about how it had gone and whether or not it was a success.

Group leadership

The way in which the course was delivered made a difference to teachers' experiences of the course. Teachers' comments about the way the course was facilitated differed in how positive they were between cohorts and groups, suggesting that delivery styles may have varied.

There was a widespread belief among teachers that the group leaders' facilitation styles were an important factor in creating the right environment for learning together:

A lot of the course being so useful was the way it was delivered 'cos they were fantastic.

Teacher, FG 3

Because if they felt uncomfortable or awkward or they weren't personable, people wouldn't share in the same way and then you wouldn't get from it what we did.

Teacher, FG 3

Teachers acknowledged that the peer support element of the course was enabled by the group leaders who actively encouraged participants to value and support each other. The teachers fed back that, predominantly, they found the group leaders to be welcoming, friendly, open and supportive:

There was a lot of humour but also they didn't ever make you, they didn't put you in a position where you felt awkward or uncomfortable, you were invited to share but not forced to and I think that is quite important.

Teacher, FG 3

More rarely, a few teachers felt that sometimes the group leaders' approach could be perceived as patronising or that they were not open to other ideas.

Generally, there was a consensus among teachers that the group leaders recognised their expertise as experienced teachers, and this was clearly important. Teachers described how this contributed to an environment of collaboration, which in turn resulted in them feeling that they could be honest without feeling judged.

Teachers commented that the group leaders were good at delivering the course input, while also being flexible in response to the needs of the group. However, there were some exceptions to this, with some teachers reporting that the course content was too fixed and not sufficiently responsive.

Most teachers were positive about the pace of the course and the way that the group leaders achieved a balance between delivering the course material and encouraging teachers' input:

They were fantastic [. . .] and gave us all the time to explore our thinking as well as moving through the course.

Teacher, additional feedback after FG

Teachers reflected that the relaxed delivery style encouraged their participation and receptiveness:

We learnt a lot but it was done in a fun way. I think we all looked forward to the next session. Which made it work.

Teacher, FG 3

Less commonly, there were some teachers who thought that the course was too basic, or did not teach them anything new; however, one teacher reflected that with retrospect they understood why 'the basics' had been emphasised:

The first few sessions I felt were kind of like behaviour management 101 and [I] thought well, been doing this for many moons and I thought a lot of it was very, very basic [. . .] Looking at it at the end, overall I can understand why there was an emphasis on that bit because it all sort of falls into perspective as to what is most important.

Teacher, FG 1

Delivery methods: modelling, rehearsing, using in practice and reviewing

Teachers talked about the benefits of the course adopting a cycle of learning, involving having input on the theory behind a strategy, seeing it modelled, practising it within the group, trying it in their classroom and then coming back to discuss how the implementation went with their colleagues:

Because behaviour policies from schools don't generally talk about all of those little positive things that you do as a teacher, they go right in with the sanctions. So for me I just found it really useful actually just having all of that reiterated and then being able to go into school, put those things into practice, do a bit of trialling, see what worked, see what didn't and then going back and discussing it again was actually really, really beneficial.

Teacher, interview 14

Teachers commented that seeing strategies being modelled helped to deepen their understanding and trying them within the group gave them the confidence to use them in their own classroom:

Using the video clips alongside with [the] training especially seeing it in action cos sometimes it's quite hard to think, especially when you are [looking] at it in the book and you read about it, how you would apply that in a classroom situation. So being able to see it first hand on the screen and being able to have the chance to role-play it within the group, that has been extremely useful.

Teacher, interview 03

Feedback from teachers about the use of role-play as a delivery method did vary. Some teachers found role-play uncomfortable and 'contrived', and others commented that, although it may have felt awkward at first, they could see its value in terms of understanding the child's point of view and rehearsing strategies:

I think they also made you think a lot about the child's perspective 'cos we had to do the role-play, which to us felt really awkward, but it did make you really stop and think about how the child when you are talking to them and are responding to them what they are feeling like, so to have to try and get into their role made you stop and think a little bit more.

Teacher, FG 2

Materials

Teachers' comments about the materials used as part of the TCM training were varied. Each participant was provided with a comprehensive textbook; teachers commented that, although this was interesting and a useful reminder of the session input, the size of it was daunting and they often did not have time to read it. Interestingly, a few teachers suggested that having an official text was useful not necessarily for themselves but for others, helping them to discuss strategies with sceptical parents or to justify their time out of school to senior leaders:

It was really useful though to show senior leadership team, and say look this is the level that this course is going to be working at, so that was very useful, when it was like you got all these sessions out.

Teacher, FG 2

A few teachers talked about using the book as a resource that had extra material and ideas to supplement what they had learnt on the course.

Teachers commented positively that the handouts (given during each session) and 'buzz documents' (flipchart notes from discussions in the session sent out afterwards) served as a useful reminder of what they had learnt and pointed them towards salient points to focus on for the next session.

Watching and discussing video vignettes of teachers managing classroom behaviour is a core element of TCM training and teachers found this modelling of the strategies useful. However, they were critical of some aspects of the vignettes, commenting that they were dated and not culturally relevant as they depicted a small class in a specialist American school rather than a mainstream UK primary school:

It was very American, very American. So something that doesn't quite fit with, you know, sort of British schools the way we are [. . .] I think it might have seemed a little more relevant to us in that respect.

Teacher, FG 2

Some of them were obviously small groups in a very specific type of school, whereas we are all [. . .] class teachers in mainstream schools with quite different set of issues within our classroom really.

Teacher, FG 2

Group leaders saw the videos as a great strength of the course but recognised these limitations. All group leaders felt that the videos should be contextualised at the start of the course, highlighting that, although they are old and some of the classrooms shown are very different from those in the UK, the concepts being demonstrated are valid and the videos provide a great starting point for group discussion about both best and worst practice:

I think the use of the video clips to illustrate the particular principles for each workshop again give it a kind of life so it's not just a dry delivery of PowerPoints [Microsoft Corporation, Redmond, WA, USA] and bullet points but they actually come to life, you can see real children, real teachers, demonstrating or they'll give you examples of a particular technique or strategy that can be used in the classroom.

Group leader, interview

Some teachers also commented that it would be useful if the general materials were electronic, as this would make it easier to use them in their own schools, and that some of the IY TCM classroom materials were expensive and difficult to source.

Time to reflect

One of the most important parts of the course for many of the teachers was the time that it provided, away from the everyday demands and stresses of the classroom to reflect on their practice and talk to colleagues. This was a strong, overarching theme that encompassed many of the elements already described:

I think what's really helped me with the course is having that opportunity to step back, sort of, for the day and reflect and actually think 'oh yes, maybe I am getting a little bit negative with the behaviour or always maybe looking at the wrong thing' and that day having out of class helps you to re-evaluate and think about the good things and you come back and you feel all inspired to try something else.

Teacher, interview 03

Teachers recognised that this gave them the opportunity to re-evaluate their practice. The modelling and rehearsing of strategies helped to give teachers the confidence to try out new approaches in their classroom and the next session would then provide another opportunity to reflect and evaluate their effectiveness.

Barriers to and facilitators of the use of Teacher Classroom Management strategies in the classroom

Teachers described a number of factors that either facilitated or hindered their implementation of TCM strategies within their own classrooms. Some of these related to the course itself and others related to the contextual situation in which the teacher was attempting to implement the strategies. Group leaders and headteachers did not describe any facilitators, although they did touch on a few potential contextual barriers.

Facilitators

Teachers reported that understanding the theory underpinning the strategies gave them a more informed and structured approach to behaviour management. It encouraged them to think more about both what they were doing and why, and also what the child was doing and why:

I've had lots of challenging children in my class over the last few years [. . .] and Behaviour Support have come in or we have worked with ways in school, but this has joined all the dots really. You might get half an hour of Behaviour Support person going 'OK try this, try this, try this'. This for me has been the background to all that they are saying which has made it make sense [. . .] so it has definitely, you know, filled in the gaps that I had.

Teacher, FG 5

Teachers particularly commented on the TCM pyramid (Figure 18) as a useful graphic to illustrate the principles underlying the TCM approach:

It's a nice ethos to think, don't get to the top, try and think of everything you can possibly do before you reach that ultimatum, so it does just make you stop and think about the ways that you can organise your class and the ways that you can respond to the children before you reach that, that top of the pyramid.

Teacher, FG 2

Teachers liked the practical nature of the strategies for managing behaviour and building relationships, and the fact that they were 'easy to use' for the whole class while also meeting individuals' needs. Teachers liked the fact that they had a range of strategies at their fingertips:

We've got so many more tools in our box, and it's just so much easier, you know if one strategy you try doesn't work with a child you've got a plethora of other ones you can have a go at.

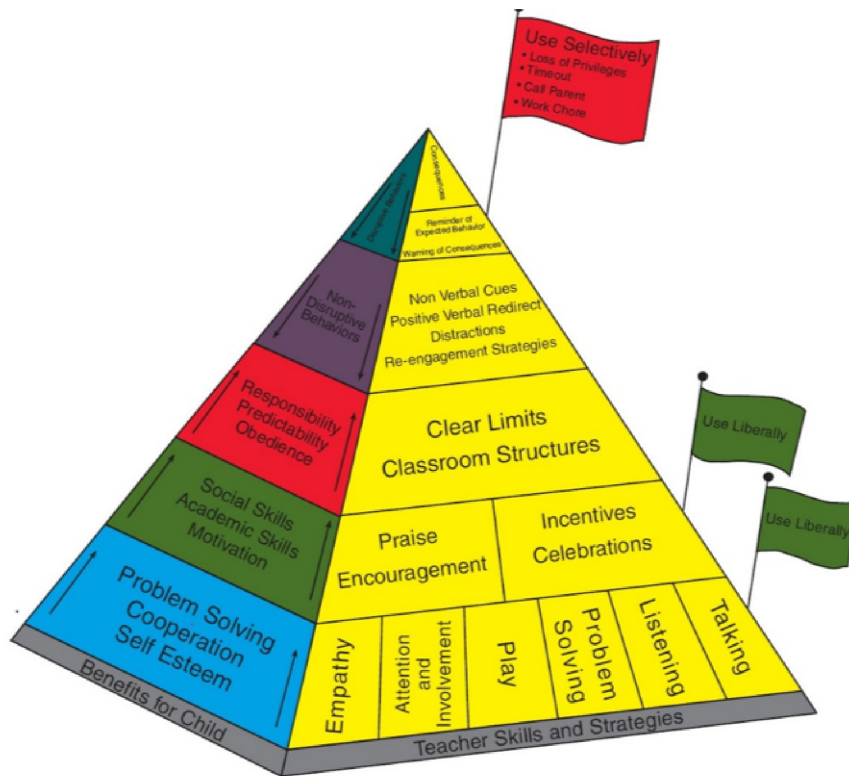
Teacher, FG 1

They also talked about the benefit of learning to adjust and fine-tune strategies in order to suit the context, whether this be the child's needs, the school environment or just the fact that 'no one thing will work forever'.

Teachers liked the fact that, as well as recognising that different strategies may work with different children, there was an acknowledgement in the way the course was taught that different techniques may or may not suit an individual teacher's approach or personality.

Barriers

The factors that teachers, headteachers and group leaders talked about as barriers to their implementation of TCM strategies fell into two categories: (1) limitations related to the content or structure of the course and (2) obstacles within the context in which they were working.



Teaching Pyramid®



FIGURE 18 Teaching Pyramid®. © The Incredible Years® Inc. and Dr Carolyn Webster-Stratton. Reproduced with permission.

Course content

There were two aspects of the TCM training content that some teachers reported did not adequately address their needs. Some teachers expressed impatience with having to wait until the later sessions in the course for information about how to deal with 'the children who were driving us around the twist'. Some teachers wanted more specific or in-depth advice about managing the behaviour that is most challenging and managing the behaviour of those with specific behavioural needs:

If it was sort of to the next level [. . .] 'cos they were obviously the experts in their field and they go in and deal with these particularly difficult situations, I would have liked to have known. I've got a school refuser, who is an extreme, so I would like some guidance on extreme behaviour.

Teacher, FG 2

I'm just not convinced that [the course] has worked for my main offenders if you like, the ones who have ODD [oppositional defiant disorder] and ADHD and all of that kind of thing. I don't think it's really worked for them [. . .] but for general classroom management it definitely has worked.

Teacher, FG 1

One headteacher believed that some of the strategies would not work with particular children and this may have affected the teacher's use of strategies. Although another felt that their teacher had not utilised many of the strategies in school because the behaviour in the classroom was already good and the target child had moved school.

The appropriateness of the TCM training for teachers of older children was also raised as an issue. One teacher thought that the TCM training in general was more geared towards younger children and less appropriate for children in the upper primary age range:

I had one of the older classes on the course so I have a Year 4, and a lot of the other people kind of had a Year 2 or Year 1, foundation, and I understand that obviously a lot of it needs to be geared towards them, but there was little I could take away and use in an upper key stage two classroom.

Teacher, interview 02

Although this view was not generally echoed by other KS 2 teachers, several teachers commented that one particular technique – commentary and social coaching – did not seem to be appropriate for older children.

One final factor in the delivery of the course that teachers mentioned as unhelpful in terms of implementation was the timing of the course within the school year. They noted that it would have been preferable to know some of the strategies in advance of the autumn term so that they could begin the new academic year with structures already in place in their class:

But like next year now, going into September, we'll have the strategies to put in place to start afresh, although we wanted to work hard with our children this year and we've done the best that we can and put those start the rules in January rather than in September but going fresh in September I'll start it from then going up through [. . .] yeah it wasn't in sync with my class [this year] but it will be for next year.

Teacher, FG 1

Context

Many teachers believed that in order to implement TCM strategies fully it would be necessary to have a 'whole-school approach', pointing out the importance of the underlying ethos being shared:

You need everybody on board really don't you because if you're going to be consistent with your behavioural management you need not just you to be doing, you need the other adults that you work with to be doing the same thing.

Teacher, FG 1

It's not something that I can implement within school myself. It'd have to be a whole-school approach.

Teacher, interview 01

A lack of consistency was seen as one of the barriers to implementing TCM at a whole-school level but also within the teacher's own classroom. Some teachers identified a difficulty in overcoming other staff's lack of understanding or misconceptions, and working with what they described as a 'default focus on negatives' and teaching assistants' reluctance to acknowledge positive behaviour. Other teachers talked about the danger that staff who do not understand how a particular strategy is being used (e.g. ignoring) will 'step in' to a situation and hinder the effectiveness of the strategy.

Training teaching assistants and lunchtime assistants was seen as important, as teachers recognised that these staff spend a lot of time with children and that staff using the 'wrong' methods, such as reacting angrily, can result in children not being in the right frame of mind to learn when they come back into the classroom:

They don't understand that they just can't be like that with particular children because it does not get the best out of them [...] it's just going to make them more angry and they're going to end up doing something else wrong which will get into a spiral of negativity.

Teacher, FG 1

Teachers also talked about a lack of time and space as a barrier to implementing strategies as they are 'always racing to fit things in'. Some struggled to balance the need to share their newly gained knowledge with students, colleagues and parents with their lack of time to do this. One teacher talked about discussing the approach with their teaching assistant as having felt 'like I've had to talk about behaviour management non-stop for the past 6 weeks':

I've found myself having to qualify and explain all of these strategies to her on a very regular basis, about the reasons why we do it, and how it can benefit the children, and how she can use them and I've talked about tools through your tool kit and stuff like that so it's kind of imparting that onto her.

Teacher, FG 1

However, teachers also recognised that if a school adopted certain strategies as a universal approach throughout the school then this would, in the long term, save 'time and energy'. One teacher gave the example of the 'show me five' rule being adopted across their school:

It will save us every year having to repeat the same rules in our own interpretive way about however we say them to get children to essentially learn which is 'sit down, carpet, be quiet, look at me, listen', and all you need to say is one or two words, 'high five', and the children no matter what age they are theoretically should know what they are.

Teacher, FG 1

Some teachers mentioned that, although they would like to use the TCM approach, they were aware of being expected to follow their own school's policies or systems, and that sometimes these clashed:

The only thing I haven't really done is probably the time-out side of the programme and that's more because we don't use time-out at our school at all. Because we have a whole-school behaviour policy I couldn't really implement that just in my class.

Teacher, interview 14

Group leaders also noted it was more of a challenge for teachers to implement TCM strategies when they went against school behaviour plans, making it difficult for the teacher. Some teachers further recognised that it can also be more difficult to maintain the motivation to do things differently when this goes against an existing school culture.

In contrast, some teachers described how the senior leadership team had involved the teacher who had been on the course in the development of a new whole-school behaviour plan. Where this was not the case, others talked about finding ways to adapt a TCM approach while still working within school systems, for example by using more strategies before reaching the warning or sanctions stage in the school behaviour policy:

I'm still following my behaviour policy but I've put so much more in place that I don't hit the first step of it yet [...] so for instance ours is name on the board, three dots, headteacher [...] but I'm getting to the point now where I'm putting in so much in before that I'm not needing to put a name on the board [...] because that's personally something I don't agree with [...] but obviously I will follow that because that's my school policy.

Teacher, FG 5

A number of teachers thought that a lack of management support for adopting a new approach hindered their ability to implement and certainly disseminate TCM strategies:

I like this approach but when you work in that kind of environment you also have to go along with school policy [. . .] I have said you know I will try things out because I am on this course and you've put me on it and I'm going to try them but now that I, the course, has finished am I still going to get away with those things?

Teacher, FG 3

One teacher felt that it was difficult to implement some of their learning as the school leadership team had changed since they started the course and were not necessarily supportive. Some teachers had shared their learning with their headteacher or senior leadership team in order to gain approval to make changes. Others suggested that feeding back to their staff teams should be a mandatory part of the course as this would give them an opportunity to disseminate their learning and therefore make it more effective.

Teachers talked about the pressure, especially when their practice is being observed, to either conform to the school's existing approach or be ready to justify the strategies that they were using. Teachers certainly felt that being observed, whether by an internal observer but especially by an Ofsted inspector, would put pressure on them to explicitly justify their use of some strategies:

If you're being observed in-house, actually you have more of a chance to argue why that's the right strategy [. . .] but with Ofsted you have to be seen to be doing things [. . .] it's just making sure that those strategies are very clear.

Teacher, FG 1

In another example, a teacher talked about using an ignoring strategy in response to a child's behaviour and having to verbalise what they were doing so the observer knew that they were ignoring them for a reason.

Group leaders also noted that there may be pressures on teachers in this area. Group leaders felt that teachers can feel that strategies they are using, such as ignoring, may be seen unfavourably by either Ofsted or their headteacher:

Sometimes during most of the courses you deliver, some teachers raise the issue of 'Well we've got Ofsted coming up' or 'If we had a snap inspection and I'm busy trying to ignore little J who is just kind of tapping on the desk and rocking backwards and forward on his chair, this is going to look very bad because I'm supposed to be on top of things.

Group leader, interview

Group leaders were able to offer effective strategies to teachers to help overcome these concerns by suggesting that a behaviour plan should have such strategies written down, thereby showing the headteacher and Ofsted that it is part of an ongoing behaviour strategy. However, whether or not teachers found this effective is unclear.

On a personal level, some teachers voiced a concern that when no longer attending the course the pressures of the prevailing culture in their schools, and the lack of time out of the classroom to reflect, would make it difficult to maintain changes.

What do teachers feel has changed in their approach? What impact has it had on others?

Impact on teachers' practice

Teachers, headteachers and group leaders all noted an impact on teachers' practice after attending the course.

Change of 'mindset'

Many teachers reported that participating in the TCM course had had a significant impact on their teaching practice. Some teachers described this not only in terms of adopting new strategies but also in terms of a major shift in their ethos or approach:

I like the way this course has made me.

Teacher, FG 4

I think one thing I grasped is the idea that we are important, teachers, and how much we do mean to the children and how we can actually make a difference. And I think for me, that's been very strong, actually made me think again about when I was a child and actually the teachers did matter to me and what they thought about me. And it's changed me I think and my relationship towards the children, I take far more interest in them as individuals and far more interest in their personal lives as well.

Teacher, interview 01

One NQT talked about how they had been able to deal with the difficult behaviour of a new child, who had been expelled from previous schools, as a result of the course and how their behaviour management skills had been recognised by the educational psychologist:

My whole mindset has changed. Everything I've learnt at uni[versity], it's not gone out the window but I think my mindset and my practice and the way I deliver and my lessons and my behaviour management has completely changed because of the things we've discussed, the way I've learnt from others here. And the Ed Psych [educational psychologist] said 'There's no way I would have said you were an NQT watching your behaviour management'.

Teacher, FG 5

A child's eye view

Teachers talked about how the course had made them think about things from a child's point of view – how they are feeling and why a behaviour may be happening. The course reminded them that children were at the core of their job, something that can easily be forgotten in the daily demands on their time:

It just gets lost in the level standards, targets, no time, leadership, and everything else that you're doing, you forget that your job is those 30 children, you have to nurture them for a year, help them to grow academically, social, emotionally, everything else. And actually you know whether a sheet is filled in with their targets or not isn't as important as what that child's just gone through on their way into school in the morning.

Teacher, FG 1

Group leaders found that certain teachers changed their views about individuals in their class respectfully, not just viewing the child as naughty but thinking more about why they might behave that way, and developing more empathy towards the child, thereby allowing a better relationship to grow.

Teachers observed that they developed a deeper understanding of the child's point of view and that this greater empathy prompted them to make changes, for example in the systems they used:

[talking about the smiley/sad face system for reward/punishment] I thought that it was a great system but when I sat and thought about it and how the children saw it, it wasn't.

Teacher, FG 3

Some also changed the way in which they spoke or gave instructions to children:

When you say 'Right, tidy up' not assuming that they know what that means [. . .] So there was a session where we kind of walked through exactly what it meant and I've taken pictures of the children and now I've got a slide that I just put up that says 'When we tidy up I need to' [. . .] and it's pictures of our class putting their book in the right box and putting the pencils back in and sitting where they should be and so I found that quite helpful. Particularly with transitions or getting ready for PE [physical education] or at the end of the day.

Teacher, FG 5

This change was also noticed by headteachers, with one headteacher observing that a teacher's language had changed to put more emphasis on being clear in outlining consequences and instructions since attending the TCM training.

Teachers also described a more general impact on their practice as a result of being more empathetic, such as being more patient or not taking things personally. Several teachers spoke about their increased awareness of the separation between their own state of mind and that of the child's, 'thinking about the child the whole time rather than how I'm feeling':

You just need to put aside if you've had a really bad night's sleep or a really bad day or a really bad morning [. . .] and it's not because you've had a bad time so you shouldn't take that out on the child.

Teacher, FG 1

Teachers talked about being reminded that recognising children's social and emotional needs is a prerequisite for creating a positive learning environment:

It's made me realise again that [. . .] we teach people, they are little people and they're going through difficult things just like we are and actually sometimes as a teacher it doesn't matter how their writing is coming along and how this that and the other, that they might just actually need you in another way [. . .] because they can't learn can they? If they're in a poor place emotionally [. . .] they're not in a place to learn so it's so important.

Teacher, FG 4

Building relationships

Teachers are encouraged to nurture positive relationships with children as one of the foundations of the TCM approach, and teachers discussed various strategies that they had introduced to proactively do this. Examples included making time to greet children individually as they come into class in the morning and ask about their evening or weekend, chatting to them about themselves at lunchtime rather than talking about school, introducing a 'compliments circle' and trying to make the class feel like 'family':

I found that in one of the vignettes we saw it was about greeting them in the morning. Some of my children are just sort of shoved into the playground [. . .] so I just make sure that I am making a real conscious effort to say 'morning', 'you alright?', 'you OK today?'. So I tend to say the same thing about 30 times but I think they all really like that. So I do try and do that every day.

Teacher, FG 2

Teachers talked about deliberately strengthening relationships and seeing the impact of those changing relationships on behaviour and learning:

The thing that stands out for me is the developing relationships. You know I have had a couple of challenging children, and over the course of the year taking on board some of the pointers from this course, the relationships have developed a lot, there is a lot more trust, they want to work for me now because I am rewarding them and praising them and seeing the good things that they do.

Teacher, FG 3

One teacher talked about how they had also used this approach in their interactions with their colleagues and it had changed their relationships.

Teachers also talked about recognising that the relationship building was two-way and that there was a benefit to being more open and allowing children to get to know them better:

I think it's remembering that you're human. You're not a superhero. And that children are people, they are little children [. . .] you need to be a human with them.

Teacher, FG 4

Headteachers noted that building better relationships was one of the things their teacher now tries to do and group leaders also noticed that relationships between some teachers and their children had been strengthened as a result of attending the course:

One of the things that most of them do notice and mention is how their relationships have improved with their children. That it seems to be a much happier and much more relaxed situation and that they feel more at ease if you like.

Group leader, interview

Thinking before responding

Some teachers reported that the combination of a strategic approach based on theory and a consideration of the child's point of view had resulted in them responding differently to situations in the classroom. Teachers talked about this affecting not only their planning but also how they respond in the moment – taking time to step back from a situation, think about why it may be happening and then make a choice about what to do, rather than immediately reacting:

I have found it has made a big difference [. . .] it's lots of the little day to day and stopping myself before I say, thinking about what I want to do before I say it and I think that has been one of my biggest changes in what I use from the TCM.

Teacher, FG 3

One teacher described how this approach enabled them to pay more attention to all of the children in the class, not just the more vocal ones:

It's been a case of stepping back as well and just hang on a minute, this doesn't have to be rushed through. So I'll just sit back and I'll watch them and think 'Right OK, who can I praise? Who's really focused? Who is doing a really good job?' And it surprised them the first time I did it because the name is called out [. . .] 'You're doing a fantastic job there C I can see you using those time connectors really well' and it's just like [. . .] shocked. It's the invisible ones that sometimes you forget.

Teacher, FG 5

Positivity

Teachers' feedback seems to suggest that one of the most important things that has changed for them is being more positive in their practice in the classroom. One teacher talked about how the whole class had become 'a much more positive place' as a result, and teachers commonly noticed a shift in the classroom environment:

By being really positive [. . .] the difference in the class has been incredible.

Teacher, FG 1

A big overarching thing I've got from it is walking into a classroom positively, thinking you're going to have a good day, being positive with the children, and if you walk in with that frame of mind it's always a better day.

Teacher, FG 1

Teachers talked about becoming aware of how subtle changes in the way in which they instruct or respond to children can have an impact, in terms of both the children's well-being and their motivation:

It's a nicer environment to be in and it's made me think now just about exactly what I say and how that comes across so it's made me more aware of the things that I say and just to make sure that they are phrased in a positive way and that I'm not always picking up on the negatives and that I am praising them when they're doing things that they should be and that's really helped. I feel happier and I think they do as well.

Teacher, interview 11

Several teachers reported that the way in which this positive approach had helped to prevent negative behaviour and promote a positive learning environment had been noticed by other colleagues:

My student has said that she can see [. . .] that when you're positive with the children, the massive impact it has on how the course of the day or the lesson goes.

Teacher, FG 1

The Head wants me to do a staff meeting based on this so that I can bring some of the strategies in. Because she's been in and she's seen it and she's said the class has changed and it's a really, really lovely class to come into [. . .] she came in and she went 'It's such a nice relaxed class' and I think that's another positive that's come out of this.

Teacher, FG 5

Teachers discussed how it had started to 'just become second nature' to approach their class with a positive attitude and as a result feel more in control.

Headteachers also noticed this change, describing their teachers as being more positive both in themselves and in the way in which they talk and deal with the children in the classroom since attending the TCM training:

I've observed her, I've observed her with a key stage leader and it's all very, very positive and I think it's quite clear that she tries to be really positive with the children all the time. I think that was probably the case before but maybe it's just a bit more accentuated.

Headteacher, interview 19

This was the case even in the absence of any other changes in the teacher's practice, and one headteacher mentioned that their teacher had even been teaching the children about the importance of being positive through the use of 'positivity buckets' and had shared this with other members of staff. Another headteacher reported that this positivity extended to their teacher being more positive about trying new things suggested by the school.

Although this positivity made for a happy classroom environment, two headteachers believed that it was almost excessively positive.

Feeling calmer, more confident and in control

Teachers reported that being able to deal with children's behaviour proactively and having a 'toolkit' to dip in to as needed had led to them feeling calmer:

I'm not so stressed out anymore, things like behaviour it has helped me in that way that I kind of don't let it get me down when things have not quite worked out [. . .] I'll look at it and try something different, rather than beating myself up about it.

Teacher, interview 07

Group leaders also thought that having a plan in place to deal with certain children had brought teachers' stress levels down and enabled them to feel more proactive and in charge:

[The teachers] both said that having a plan and a strategy so that they could deal with the initial irritation of whatever the behaviour was, but knew what they were going to do next and then what they were going to do after that. So they felt scaffolded themselves which brought their stress levels down and then the TAs [teaching assistants] reported that their stress levels had been brought down because they were scaffolded by having a plan in a way that they weren't before so they weren't feeling quite so responsive and emotional, they were actually feeling more proactive and in charge.

Group leader, FG

For group leaders, the feedback they had had led them to believe that teachers had a renewed sense of confidence:

From the feedback at the end it was like we'd turned their entire lives around [. . .] they'd actually been under so much pressure and they were feeling stressed and they were feeling inadequate that to have been given new tools seemed to help their confidence and get it back.

Group leader, FG

Some teachers also discussed how the TCM training had built or restored their confidence in their own behaviour management skills:

It gave me a little bit of self-confidence, that actually I am doing quite a bit of [. . .] what is advised, so that was quite powerful.

Teacher, FG 2

I think if our head said, 'Is there anyone [. . .] prepared to now coordinate [. . .] some sort of behaviour support', before I would have said 'no' but now I would say 'Yes go on then let's look at that' because I feel like I have got the confidence to do that.

Teacher, FG 2

One teacher gave an example of finding it difficult to deal with a child who was constantly calling out in class but was unconvinced that the punitive measures that their colleagues had suggested were the right strategy. After going on the course this teacher felt confident to deal with the situation in their own way, believing that the strategies suggested were not the best idea.

Similarly, headteachers also noted an impact on teacher well-being and particularly that the course had helped the teacher gain confidence/experience. For two headteachers this was demonstrated by teachers' ability to share learning with colleagues when they would not have done so before:

I think it's given her confidence [. . .] that came through with her leading the staff meeting [. . .] she did it very, very well, very confident, and any questions they asked she was able to answer very well.

Headteacher, interview 01

One headteacher also mentioned that the time to reflect had been particularly helpful for their teacher:

No, I think it's been really good for my teacher. It's been really good for my practitioner. It's made him sit and think about what he was doing and the way that he behaves with his children. He's better for it [. . .] I think he felt that he knew it all.

Headteacher, interview 14

Creating a positive cycle through role modelling

As previously described, teachers reported that changes in their own behaviour in the classroom led to changes in children's behaviour and in the well-being of both teachers and children:

When you are looking and spending time and you are really listening to them and being really positive I've found as a person it must have made me better at my job, calmer, happier and the children wanted to please and not be that person who is in time-out.

Teacher, FG 3

Teachers reflected that having a happier teacher seemed to result in a happier class. Teachers reported changes in their own behaviour in class creating a cycle of positive reinforcement; as their approach became calmer and more confident they saw the children's behaviour changing in a similar way. Some teachers also noted that this role-modelling technique worked in a similar way when they were more open about mistakes or challenges with the children; this helped to reduce teacher stress but also helped children to learn about emotional self-regulation and resilience.

Impact on children

Changes in behaviour: children becoming ready to learn

Teachers noticed that changes in their practice had led to changes within the culture of the whole class and within individuals. Some teachers talked about how addressing children's social and emotional needs created a climate in which children were ready to learn:

I have noticed that the one boy [. . .] I didn't think I was getting anywhere with him and it's only been in the last month where his behaviour has changed completely. He is producing an awful lot more work [. . .] we have a home schooling book where parents are getting involved, and there has been a notable change in his attitude to school and also in the amount of work he is doing.

Teacher, FG 2

My PPA [Planning, Preparation and Assessment] teacher [. . .] said she's seen a noticeable difference in my class [. . .] across all of the key stage, she's seen the biggest difference in my class between the September to now in terms of their behaviour and how well they're settled and how hard they work and just generally.

Teacher, FG 1

Teachers noted how shifting the attention that they give to those who are behaving well had also helped children to focus on their work:

I have noticed in my class that those children that mess around and don't do very much work or whatever are actually now producing an awful lot more work because they want the attention and I have withdrawn the attention from them which was 'why haven't you got that done yet, you are always in trouble' and I have given it to these other ones who 'if I want the attention then I need to do the work to get the attention as well' [. . .] it seems so obvious but that kind of turnaround has been really nice to see.

Teacher, FG 2

Headteachers also noted that children were responding to strategies and, as a result, children's behaviour had improved, although in a more general sense. Headteachers gleaned this information from teacher reports (headteachers 09, 10 and 17), observing the impact first hand (headteacher 02), assuming that there has been an impact as fewer children had been sent to them with behavioural issues (headteacher 09) or adopting the strategies themselves and finding them to be successful with the children (headteacher 09).

One headteacher mentioned how there had been additional benefits for children when behaviour improved, describing a classroom that was now 'more able to learn without disruptions, interruptions, etc.' (headteacher, interview 02).

Promoting independence and taking responsibility

Some teachers gave examples of ways in which specific TCM strategies had helped children to become more independent:

It's given the children the power to look around and think 'Oh I know what to do to do this myself', and it's trying to train that little bit of independent thinking and it's been beneficial to the children.

Teacher, FG 2

For one child that came in really struggling [. . .] I gave him the 'acts of responsibility' idea. So his responsibility is to always give out the whiteboard pens, and spot the good behaviour in other children, and it worked a dream and he is such a different child now, and I think that's really been helped by the course.

Teacher, interview 05

Impact on children with behavioural needs

As mentioned previously, teachers gave mixed feedback about the effectiveness of TCM strategies for children with particular behavioural needs. Some teachers believed that using TCM techniques within the classroom complemented the individual support and guidance that some children might be receiving outside the classroom.

Some teachers gave examples of individual children with behavioural needs for whom they had been able to successfully implement specific TCM techniques or formulate an individual behaviour plan as part of the course:

Things like the time-out chair has had a really positive effect for one child who has got severe behavioural difficulties and he now just takes himself to there now [. . .] it's his calming strategy that he's learnt and uses now and so I can feed it up to his next teacher in year one to carry on using that because it's something that he self-soothes with and it really works for him.

Teacher, interview 05

However, some teachers reported instances when the strategies had not worked:

One girl who came in in January I was able to apply a lot of what we have been doing here and I think we kept a lid on her behaviour all the way through to March when it went off. But keeping a lid on it is not what you want to do is it, you want to deal with it rather than just keep a lid on it. But I think if she had been there right from the beginning it might have been a different thing, rather than coming in half way through.

Teacher, FG 2

When we started I had a really difficult child, he's now gone to a special school for behaviour [. . .] but it did make me realise from when I was talking to everyone at the start when he was my focus child, that everything I was trying, I could then say when he did eventually leave that I did try everything, so I didn't sort of give up I tried, literally everything and all the ideas that I got from here I tried. So I think that was really good to know [. . .] I needed that next bit for him which would have been really helpful.

Teacher, FG 2

One headteacher also noted that their teacher thought that the strategies would not have worked for one particularly difficult child in their class.

Although teachers reported that there were reasons why the TCM strategies did not ultimately work, these examples tie in with the feedback from a few teachers (described in the sections above) who suggested that TCM did not always have an impact on the children exhibiting the most challenging behaviour.

The responses of group leaders, however, were less mixed, with most reporting that, for generic supportive processes, the course is very good at supporting those children with additional needs. However, the inclusive structure might not be sufficient for a number of children, so additional strategies may be needed. One group leader commented on the positive impact that the course has had on children with SEN, and this was something noticed by parents as well:

They've reached a kind of almost modus vivendi with these children in that they haven't given up on them, they still have high expectations for them but the children feel more relaxed with them and the parents have reported noticing this as well.

Group leader, interview

Raising self-esteem

Earlier in the chapter we described how teachers noticed increases in their own levels of confidence and self-efficacy that were then reflected in the children. Another strategy that teachers noticed having a direct impact on children's self-esteem was increasing their communication with parents, with a particular focus on reporting positive behaviour:

Another thing we have used is the wow slips and the happy grams to parents all in their contact diaries, just to say so and so worked really hard today [. . .] just by saying we are really proud and sharing that with the parents the child is then 10-foot-tall [sic], that's all they need, they can go away feeling happy and confident.

Teacher, FG 3

Once again, teachers noticed that another outcome of this approach is that it works for all of the children in the class, rather than a smaller number taking the teacher's attention.

No impact

Although most teachers reported an impact on children's behaviour, the majority of headteachers did not, which may be because the impact is not transparent to others, or because headteachers have not had the chance to discuss behaviour with the teacher or spend time with the teacher and children in their class.

Impact on parent-teacher relationships

Developing relationships with parents is advocated by the TCM approach, and teachers mentioned various ways in which they had consciously tried to do this, for example by greeting parents at the door, going out to see them in the playground to pass on positive messages about what their child had done, setting up a home-school book, sending positive letters and making telephone calls. One headteacher also noticed this:

I mean he was thinking that he was always listening and welcoming parents [. . .] but he wasn't so much seeking out opportunities to share positive moments with parents but he feels he does that more now and that parents really appreciate it. And at the end of the day, especially if your child has been previously or in the past been causing problems or having difficulties with learning, that it's nice to be actually told actually he or she is doing really well and trying hard. So he felt he does that a lot more and can feel the sort of appreciation on behalf of the parents when he does that.

Headteacher, interview 22

Some teachers talked about how their efforts to develop better relationships had resulted in parents perceiving the teachers more positively, prompting more positive responses and more discussion:

I have noticed that there has been a much more positive parental response. Because when they see you now they feel they can actually talk to you, more rather than just listening to a litany of things that little Johnny has done wrong today.

Teacher, FG 2

One teacher described how this positive effect had been recognised more widely within the school, and that the whole school was now working towards sending more positive messages to parents, dealing with negatives in a 'gentle' way, and generally encouraging parents to participate more actively in school life:

It's definitely impacted parents which has been a major culture shift, for parents to feel a positive response or message. And actually, our school is driven to work better with parents and open the doors and become more of an open door policy, to encourage children to come in and settle with parents alongside.

Teacher, interview 05

Some teachers had been approached by parents to help them try to tackle difficult behaviour at home using some of the strategies that they had adopted in the classroom. One teacher talked about advising a parent to focus on positive instead of negative behaviours. Others talked about how this new co-operative way of working with parents had the added benefit of not only helping parents but also introducing consistency and therefore reinforcing behaviour changes at school:

Dad initially was very, quite aggressive, now whenever he sees me it's all smiles and 'oh yes', they have implemented the same strategy we have at school at home – the traffic light system, you know, for behaviour – and we have noticed a complete difference in him. He still does the odd thing [...] but there has been a notable change in his attitude to school and also in the amount of work he is doing.

Teacher, FG 2

Another headteacher observed that, in combination with support from the school, the course had helped highlight that the teacher needed to engage more with parents and talk to them positively. The teacher had started to do this, although the headteacher thought that there was still more work to do.

No impact/little impact/unknown impact

Some headteachers reported that the course had had no noticeable impact on their teacher's practice even though the teacher themselves may have felt that it had had an impact:

You know K may say to you it's actually made a great difference to the way she handles her children, I don't know. I can't see a great difference, but she may feel there is a great difference the way she's handling them now. She may feel she's imparted that onto her Key Stage 1 colleagues who she works closely with and that they've taken it on board or she's talked to the preschool practitioner manager and spoke to her about it. There could have been lots of ways it could have impacted. Maybe she has done that and I'm still unaware of it, but you know we shall see.

Headteacher, interview 18

Other headteachers often felt that it was difficult to ascertain whether the course had had an impact or whether change in practice was due to support from the school, increasing teacher experience or a combination of these factors. A few headteachers seemed uncertain about whether or not it had had any impact owing to a lack of feedback from the teacher. Others expressed the view that their teacher had not

gained a great deal extra as they were already doing many things that the course advocated or were already very experienced:

I think she felt it was very much a similar ethos to the ethos she was used to working with. And in that respect I don't think she felt she had to change her practice hugely.

Headteacher, interview 16

I think he felt that it wasn't a lot of things he either hadn't already thought about or wasn't already trying. And I mean he's an experienced teacher anyway and capable and all the rest of it. I don't think there's anything particularly hand on heart he would say that he's changed dramatically. I think he says obviously it does alter your practice because every experience does but nothing like I would say 'Oh he now always does this'.

Headteacher, interview 20

Use of Teacher Classroom Management strategies

The techniques and approaches described below are those that teachers commonly mentioned. Group leaders did not comment on what specific strategies teachers were now using and, although some headteachers mentioned specific strategies that their teacher had been using, many seemed unaware of the strategies teachers had employed, if any.

Specific and proximal praise

Many teachers mentioned specific praise as a key strategy that they were now using regularly with their class and two headteachers also mentioned that this was something their teachers were now using. Teachers were clear about the effectiveness of using specific, targeted praise:

It's that focused praise, 'well done for getting all that work done today, well done for tidying that table up when I asked you to' [. . .] it's not just that general 'oh yes you have been really good today' or 'you have been really nice', it's that very specific reason why you got that particular sticker or that particular reward.

Teacher, FG 2

Some teachers gave examples of particular techniques that they were using to highlight positive behaviour, for example a 'kindness wall' where staff or pupils notice something kind that someone has done, and put their name up and share at the end of the day.

Teachers also talked a lot about using proximal praise, for example reducing disruptive behaviour by giving positive attention to the children who are behaving well rather than focusing on those displaying negative behaviour. Some teachers talked about how, despite finding this difficult at first, they found this to be extremely effective and something that worked quickly:

I found one of the most useful things, in terms of time management, is using the proximal pupil, when you actually use the praise of children around who are doing the right behaviour or what you are looking for and I actually find it so much quicker in bringing the group back than going, for example [. . .] 'Sam turn around, Sam do this' so if you say something like, 'oh Grace is sat beautifully looking at me' and all of a sudden the whole room is turned around. I have found it has made a big difference.

Teacher, FG 3

It also had the added benefit of giving attention to the children who are sometimes ‘invisible’ and encouraging good behaviour in other children:

It's having an impact on other children who suddenly stop and think 'ooh yes I want to be noticed for the right things'.

Teacher, interview 03

Ignoring

In a similar way, lots of teachers were using the ignoring strategy, deliberately not giving attention to low-level disruptive behaviour and teaching children to use the same strategy with their peers. Teachers had used techniques such as role playing with puppets or the idea of ‘using your ignoring muscle’ when someone is annoying you to explain the principle to children and had found it effective:

I use one of the doll puppets in class [. . .] I was actually making a point of how to ignore, what kind of behaviour you are looking to ignore, you are not ignoring your friends just because you are not allowed to talk to them but it's how to sort of show that you are focused without being rude to your friends.

Teacher, FG 3

My children are very good now I'm using it, and ignoring the bad behaviour and knowing they'll get praise for ignoring that bad behaviour.

Teacher, FG 1

Teachers talked about the other children having taken on board the idea of using their ‘ignoring muscle’ to deal with disruptive behaviour from other children, and some teachers had also noticed that teaching assistants (TAs) who had seen it used effectively had also started employing the technique themselves.

Rewards and consequences

Lots of teachers talked about the effectiveness of using reward systems for positive behaviour and one headteacher pointed it out as one of the strategies that their teacher was now using. Examples given by teachers included the use of tokens to help children learn specific behaviours (e.g. ‘using inside voices’), which built up to house points that were counted at the end of the week and resulted in rewards such as DVD time, popcorn, extra playtime, ‘going out to the pirate ship’ or a ‘wheel of fortune’. Other teachers were using rewards such as raffle tickets, stickers or a class dojo (an interactive whiteboard system in which each child has an icon and gains points). Some teachers discussed children ‘taking ownership of their rewards’ by deciding on reward types.

One teacher commented that, although they agreed with recognising/praising good behaviour, they did not feel that material rewards were appropriate, as they can become devalued when overused:

I've stopped using material things as much as I used to, I've actually gone for the more positive comments [. . .] high fives, that sort of thing and that's actually worked better.

Teacher, FG 5

Interestingly, teachers did not believe that the group leaders were supportive of this alternative view, and they reported that sometimes the group leaders’ approaches implied that ‘these are the only ways’, whereas they felt that the ethos of the TCM approach actually recognised teachers’ need to be flexible and respond to the situation:

The one thing that I've taken away from the course as well is that nothing, not one thing, will work forever, so it's about constantly changing.

Teacher, FG 5

Some teachers reported changing the way in which they framed consequences, trying to encourage children to understand the concept of making choices about their behaviour:

I've changed like how I deliver the consequence [. . .] so rather than just say 'You've lost' I say, 'You've now made the choice [. . .] you've chosen to lose your golden time' and although they're only Year 1 I think they're starting to get the hang of that [. . .] I've got quite a particularly difficult class and I think their behaviour has really, really switched.

Teacher, FG 5

Some teachers also talked about realising the importance of trying to give more immediate consequences to actions (rather than a delay such as losing play time later in the week) and making sure that consequences were age appropriate.

Time-out to calm down

The time-out technique advocated by TCM encourages teachers to enable children to use time-out as their space to calm down. Some teachers talked about the value of rethinking time-out and seeing it not as a punishment but as a way of teaching children to recognise that they are feeling angry or upset and to manage the process of calming down; they also talked about the benefits, particularly for children with behavioural needs:

I found the time-out really useful. I was using it already but actually it wasn't [. . .] but then I changed it around and the little boy I use it with has a much clearer focus of why he is having it out, and then when he is having time-out he just then comes straight back in to the lesson and I ask him to carry on, 'You need to be writing your sentences now, you need to be listening', rather than explaining to him why he had time-out. And that has really worked because he has his calm down time and then he just comes back and tries to carry on and that worked really well.

Teacher, FG 3

Some teachers acknowledged that using time-out in this way was different from how they had used it previously and they recognised that some preparatory work had to be put in place for it to be effective.

Some teachers talked about how they had adapted the time-out strategy. One had changed it to be a reward for good behaviour; children could have time-out using the classroom iPads (Apple Inc., Cupertino, CA, USA), for example. Others talked about using the 'calm down thermometer' (a TCM resource) to help children monitor their emotions. For example, one teacher had created a 'cool down area' for children who particularly struggled with becoming angry to use when they needed it. Although this 'took a while to embed', it had become successful:

It's been adapted slightly inasmuch [as] they'll take themselves to it rather than being sent to go and calm down if they need a quiet five minutes, they've got that there in the classroom that they can go to. They've got a timer that they can turn over. And then they've got the faces that we were given with the arrow to move down and then when they get themselves back down to the happy face they can come and join [. . .] there's only about four or five children that really need it but they know it's there they'll take themselves to it quite happily.

Teacher, interview 11

Emotional regulation

Some teachers talked more generally about how the course had made them aware of the need to help children to recognise and deal with their emotions. Teachers had used a range of techniques to address this, including, for example, 'emotion pegs' and mirrors:

I also certainly have a more stronger [sic] access for them to have a knowledge of their own emotions [. . .] I hadn't really considered that nearly as much before [. . .] for example when the children come in

they have a peg and they put their peg on an emotion. And that is also something I've done with PSHE [Personal, Social and Health Education] and focusing on the different emotions and how to cope with them even how to cope with being happy. But there's a lot more emotional intelligence I hope as a result of that [. . .] We have a chair that I now use not just for time-out but also for calming down. It's flooding that world with language. It's flooding that world with empathy as well, that other children experience emotion as well.

Teacher, FG 4

We have actually done that in the nursery and even at 2 [years] they look in the mirror and we've got a thing up with a few words I'm happy, I'm sad, I'm tired, we have kept it quite simple [. . .] it's really nice because we have started to talk about [. . .] if you are happy what does that look like, what does your face look like and how does your body feel, and also, you know, it's OK to be angry it's not a negative thing, it's good to feel angry but you are not allowed to punch someone, it's how you deal with it and giving them the time to look at those strategies because we focus so much on the academic in class, but that's so important.

Teacher, FG 3

Again, teachers also talked about modelling emotional regulation with their own behaviour and how they had found that naming some of their own feelings could help children.

Some teachers talked about how other strategies that they were using to manage behaviour, such as home school books, also helped children to learn about their emotions, helping them talk more about themselves and have open discussions about behaviour.

Behaviour plans

Early in the TCM training, teachers are asked to create behaviour plans focused on particular children with behavioural issues; they then implement and review those plans as the course progresses. Teachers recognised the value of having time to reflect on that particular child's needs and discuss strategies that they could try. A number of teachers talked about how they had successfully used these behaviour plans:

The behaviour plan that we did with one of the children, that was really useful as well because you sit there and really think about what strategies you can put in place and how it supports them and everything and actually it has worked quite a lot [. . .] we don't use it anymore because we have done so well.

Teacher, interview 13

It's thinking about things slightly differently, I think that's what this course has helped me to do. And with the plan, there's a little boy in the class that goes under tables sometimes and he's got learning needs really and I think it's anxiety about the lesson sometimes really and I've tried to put things in place now and he has a daily programme, he goes out 1 hour a week, 1 hour a day with another child and his stress has lessened. He still does sometimes run out of the classroom, it's not totally better, but I think I feel more informed about his needs possibly. So I think the plans that [the group leaders] helped us with, I think that helped.

Teacher, FG 4

Some teachers highlighted the value of behaviour plans in helping to make the approach consistent among different staff who may be working with a child and who they had shared the plan with. One teacher also mentioned that the behaviour plan that they had created for one particular child had allowed them to address behavioural issues that would have not otherwise been addressed, as the child did not necessarily fit the criteria that would normally be used in school.

Although a number of teachers were positive about the impact of using behaviour plans, some gave examples of children with particular needs or contexts in which teachers did not feel that they worked. One teacher realised that focusing on one particular child was not necessarily helpful for that child and that they benefited more from a whole class approach:

Some of the behaviour plans and the charts and all of the things, it's just too fussy [. . .] it was too much for them to take in and they needed more. Actually what I took away from the course was it wasn't the charts and the special things that they necessarily needed, it was the coaching, the positive reinforcement, they actually needed what everybody else needed [. . .] and treating them the same has actually I think helped more than the chart. The one I started off with on the behaviour plans that we made, when I made him one of the charts all he wanted to do was play with it [. . .] so it didn't work. But actually all of the other strategies that I was using with the rest of the class, when I thought hang on a minute maybe he might respond to those he did.

Teacher, FG 5

Another barrier that one teacher identified to making a behaviour plan successful was a lack of support from parents:

Talking about the behaviour plans, I chose a specific boy for the behaviour plan who just was [. . .] who's not awful but he's just constant all of the time. And one of the huge barriers [. . .] was his parents that were, it was their boundaries that were. So I'd make progress with him but then if we did have to employ a sanction he just wouldn't turn up the next day [. . .] sometimes things are outside your control [. . .] following the behaviour plan it didn't seem to have an impact because it was almost like, my voice had been devalued.

Teacher, FG 5

Some teachers also mentioned that because working on behaviour plans continued throughout the course, they were not necessarily always relevant since a child's behaviour had changed, or, conversely, if the child already had SEN then the plan might be a repetition of something that was already in place. In contrast, another teacher commented that with changes in SEN policy, this would not necessarily be the case.

Coaching and commentary

Academic, social and emotional coaching is another strategy taught on the TCM training, including using commentary to acknowledge rather than direct children's actions or behaviour; this is seen as a real strength by group leaders:

I think [coaching] is the most powerful thing that teachers often come back and say 'I was doing it and it actually made a difference'.

Group leader, FG

Coaching was noted by one headteacher as a tool that their teacher was now using. Coaching strategies received a mixed reception from teachers and seemed to work in different ways for different individuals. Some teachers observed that they were more comfortable, or found it more useful, using it in an academic or a social setting, but this was not the same for all teachers.

A number of teachers suggested that coaching as a strategy was useful but more applicable when working with the youngest primary school children:

Positive coaching, both academic and social, that was really, really interesting. And because I work in foundation stage it was absolutely brilliant because it supported my good practice but it just reminded me to do that all the time.

Teacher, interview 05

I have done the social coaching as I work with the youngest children. Whereas before if someone was upset when they came in in the morning I would distract, I would take them away, now just acknowledging how they are feeling, 'oh you must be feeling disappointed, mummy went without a kiss' or whatever, and helping them move through that, rather than 'oh no come and have a story' that has really helped actually.

Teacher, FG 2

We were going through a lot of these social coaching [sic] and actually sitting down with the younger ones which I think is easier than with the older classes.

Teacher, FG 3

One teacher adapted the strategy and used commentary on their own actions as an alternative way to engage children who often do not respond to more direct forms of encouragement to take part in activities.

Comparison of feedback across Teacher Classroom Management training groups

As well as examining the feedback from teachers as a whole, we also considered whether or not there were any differences in teachers' feedback between the different TCM training groups. Post-course feedback was gained from teachers from five different TCM training groups: FGs 1–5. There was one FG per TCM course and a varying number of interviews (see *Appendix 2*).

Feedback appeared to be similar across all TCM groups, with teachers identifying similar influences on learning, barriers to and facilitators of use, impacts and uses of TCM strategies. FG 5, however, tended to be slightly more negative, with teachers mentioning more shortcomings in the TCM content and delivery. In particular, they reported that the feedback they provided as part of the course was not dealt with well, which left them feeling uncomfortable. During each TCM session, teachers are asked to provide written feedback on the session. In the next session, group leaders then go through this feedback and discuss any issues with teachers as a group. Teachers in FG 1 also appeared to be slightly more negative in relation to the delivery of the course. Despite this, however, both groups still reported that they had benefited from the course and reported that the course had had an impact on their teaching life.

Written feedback completed by the teachers

Participant details

All teachers who attended the TCM training ($n = 76$) were asked to complete the TSQ after completing their final TCM session. Sixty-seven of these teachers completed and returned this questionnaire as requested. The number of teachers able to provide data at the later follow-up time points (i.e. 9 months and 21 months) varied as many were no longer part of the STARS trial at this time as their school had completed their 3 years of involvement in the trial. A total of 49 out of 67 completed the TSQ at 9 months post course and 26 of 39 teachers completed the TSQ at 21 months post course.

A total of 74 teachers completed a workshop evaluation form for workshop 1; 71 completed one for workshop 2; 67 completed one for workshop 3; 66 completed one for workshop 4; 62 completed one for workshop 5; and 62 completed one for workshop 6.

Results and discussion

Overall, the majority of teachers rated content, videos, teaching and discussion as either 'helpful' or 'very helpful' (*Figure 19*).

Most teachers thought that the content, teaching and discussion were 'very helpful' (58%, 77% and 83%, respectively) and that the videos were 'helpful' (61%). Videos received more 'not helpful' or 'neutral' ratings than the other three areas (2% and 12%, respectively). Comments from the teachers suggested

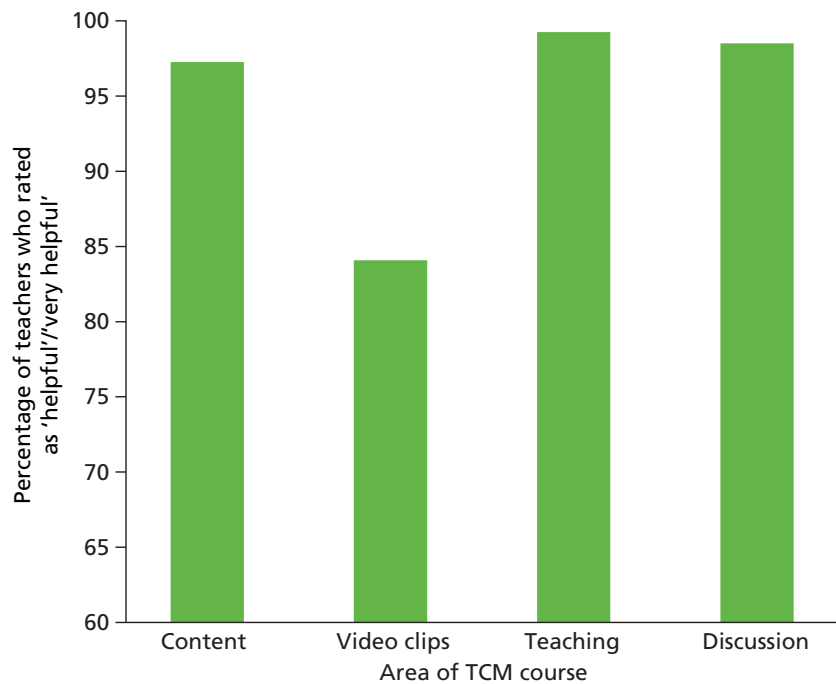


FIGURE 19 Percentage of teachers who found TCM 'helpful' or 'very helpful'. Teacher feedback on course.

that they did find the videos useful to support discussions, but they thought that they were not always culturally relevant to the classrooms in which they teach. This ties in with our findings from the FGs/ interviews.

When studying the teacher feedback by each TCM workshop, teachers rated both the teaching and discussion consistently as either 'helpful' or 'very helpful' across all six workshops. Teachers also rated the content highly across the workshops, although ratings for this dropped slightly in workshops 5 and 6. These two workshops cover topics that could be considered more challenging (see *Table 1* for an overview), which may explain this slight decrease in favourable content ratings. The percentage of teachers who rated the video clips as 'helpful' or 'very helpful' was high across all workshops (although lower than for the other three TCM areas) and appeared to be lower in workshops 2 and 3 (*Figure 20*). Teachers may have thought that the behaviours and strategies covered in these workshops were more 'obvious' and may have felt that they did not need to see as many video clips as a result, valuing other aspects of the workshop more.

When examining feedback by TCM group attended (i.e. TCM groups 1–8), there appears to be greater variation in the percentage of teachers rating each area favourably (*Figure 21*), suggesting that there may have been a difference in how teachers experienced the course depending on what year/group they attended. The greatest variation is found in terms of whether or not the videos were helpful, with them being noticeably less helpful in the final years of the study. This variation in helpfulness is likely to be due to the group dynamic and how well the teachers in that group responded to the video clips, which were often considered outdated and 'Americanised', as discussed previously (see *Materials*).

On average, there was little variation in how useful teachers found, and how frequently they used, the TCM teaching strategies on the TSQ. Immediately post course, teachers reported finding the strategies useful (mean 4.25, SD 0.27) and using them frequently (mean 4.05, SD 0.04). This corresponds with the findings from the qualitative work, which suggested that teachers were employing a range of strategies from the TCM training and finding a number of these strategies useful.

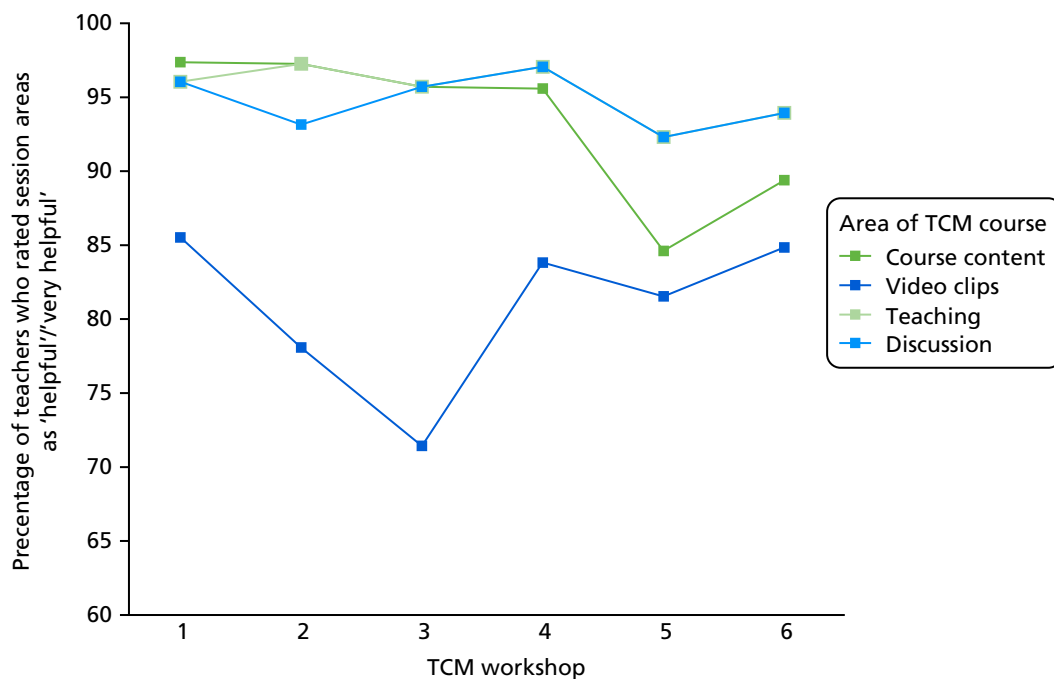


FIGURE 20 Percentage of teachers who rated course content, video clips, teaching and discussion as 'helpful'/'very helpful'. Teacher feedback on course by TCM workshop.

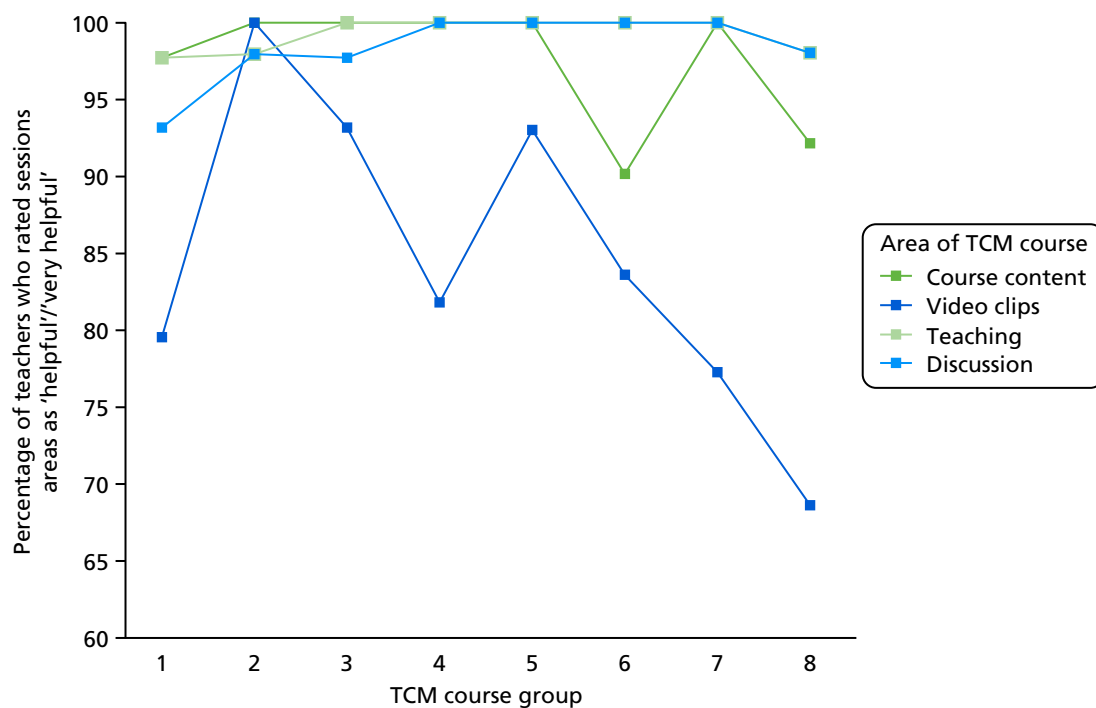


FIGURE 21 Percentage of teachers who rated the course groups as 'helpful' or 'very helpful'. Teacher feedback on course by course group.

Overall discussion

The process evaluation aimed to address four main questions:

1. What are the key elements of the TCM training from a teachers' perspective?
2. What are the barriers to and facilitators of the use of TCM strategies in the classroom?
3. What do teachers believe has changed in their approach and what impact has it had on others?
4. Which TCM strategies have been used/have not been used?

We explored these questions through FGs/interviews with teachers, headteachers and group leaders and also through questionnaires designed to examine teachers' satisfaction with the TCM training.

Overall, qualitative results suggested that teachers mostly enjoyed the TCM training and found it beneficial. This was echoed in the TSQ, which suggested that the majority of teachers found the TCM training either 'helpful' or 'very helpful'.

In relation to our first research question, 'what are the key elements of the TCM training from a teachers' perspective?', qualitative findings suggested that the structure of the course (i.e. being 1 day a month over a 6-month period), being part of a group that came together to discuss practice and learn from one another, having time out of the classroom to 'reflect' on practice, and being able to 'see, practice and implement' the strategies back in the classroom were key elements. The TSQ highlighted that teachers experienced aspects such as the course content, video clips, discussion and the delivery of the TCM training either as 'helpful' or 'very helpful'.

In relation to our second research question, teachers reported limitations in the course content or obstacles within their school settings (e.g. a lack of a 'whole-school' approach, lack of support from colleagues/senior staff members, clashes with school policy, and external bodies such as Ofsted) as their main barriers to the implementation of TCM strategies in the classroom. Although often a widespread view,⁹⁴ only one teacher reported the TCM training as not being as useful for KS 2 children as a whole. Facilitators, however, included understanding the theory behind the strategies (something endorsed by the TCM training), the adaptability/flexibility of the TCM strategies and the fact that the strategies were generally easy to use.

Teachers mentioned using a number of different strategies as a result of the TCM training, including specific and proximal praise, ignoring, rewards and consequences, time-out, emotional regulation and behaviour plans. They also reported attempting to use coaching and commentary, although the feedback on this was more mixed. The TSQ complemented these findings, suggesting that teachers found the strategies promoted by TCM useful and reported using them frequently.

Teachers suggested that the TCM training had affected them in a number of ways, which included changing their mindset; helping them to see things from the child's point of view; helping them to develop better relationships with the children; making them think more before responding; helping them to feel more positive, more confident and more in control; and helping them to create a more positive cycle of behaviour through modelling. These impacts were all extremely positive. They also believed that TCM had affected the children in that they were calmer, more motivated and more ready to learn; however, the views on whether or not the strategies had had an impact on children with behavioural needs were more mixed. A few teachers also reported that TCM had had an impact on their relationships with parents.

Chapter 7 The impact of Teacher Classroom Management training on teachers' well-being and their relationships with work

This chapter reports on the substudy that focused on the potential benefits that TCM training had on the teachers themselves, in particular their sense of well-being and relationship with work.

Introduction

Teaching is commonly acknowledged to be an extremely stressful occupation, with poor and possibly worsening retention in the UK.¹⁰³ The prolonged response to chronic emotional and interpersonal stressors in the workplace can lead to 'burnout', a syndrome of exhaustion, cynicism and reduced professional efficacy, which may underlie poor mental health and low retention among teachers.¹⁰⁴ Teachers often report a lack of training in, and difficulties with, managing disruptive behaviour, which is associated with stress, burnout and exit from the profession.^{11,12,105} Jennings and Greenberg¹⁰⁶ proposed a 'burnout cascade', whereby a teacher's difficulties with behavioural management in the classroom lead to a decreased sense of self-efficacy, which results in negative consequences for their relationship with the children, a more reactive and more negative classroom environment, and increasing distress and burnout for the teacher.

In general, mental ill health is both the largest single cause of disability and the leading cause of sickness absence from work in the UK.¹⁰⁷ The economic impact of poor mental health occurs not just through lost working days and absenteeism but also through the reduced productivity of employees still present at work, known as 'presenteeism'. The education and care sectors have the highest rates of presenteeism of any employment sector, which in turn predicts high levels of absenteeism.¹⁰⁸ Several studies report high rates of depression and anxiety in both primary and secondary school teachers in comparison with the general population.^{109,110} A recent UK study [Wellbeing In Secondary Education (WISE)] of 555 secondary school teachers detected high levels of moderate to severe depression (19%), which were associated with being female, dissatisfaction with work, presenteeism, sickness absence, interpersonal difficulties and low pupil attendance.¹¹⁰

Alongside lost productivity caused by disability, sickness and presenteeism, the staffing of schools is also challenged by the premature loss of experienced teachers from the workforce; mental ill health is the leading cause of exit from the teaching profession. For instance, it was reported as the main factor in the decision to leave of 46% and 37% of Irish and Scottish teachers.^{111,112}

Poor teacher–pupil relationships predict subsequent poor child mental health¹¹³ as well as poor academic attainment.^{18,114} Stressed teachers have more negative classroom management styles and when supported with behaviour management they report reduced emotional difficulties and decreased disruptive behaviour among their pupils.¹¹⁵ Teachers' mental health has also been shown to influence the academic attainment of the children they teach. McLean and Connor¹¹⁶ reported that teachers' depressive symptoms in one term predicted their students' mathematical abilities a term later. Teachers with less-developed classroom management skills have higher reported overall levels of child classroom aggression, peer rejection and exclusion.¹¹⁷ 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.³⁰ There is also an increasing expectation on teachers to identify and manage pupils with mental health needs, which teachers often feel inadequately prepared to do, with consequent effects on teachers' own well-being.^{118,119} Therefore, an intervention that supports teachers to manage disruptive behaviour and promote socioemotional competence has the potential to benefit not only the children they teach but also the teachers themselves.

Previous research has suggested that TCM training improves teachers' perception of their self-efficacy in relation to their classroom management and that such training has the potential to improve well-being and decrease burnout,⁹⁴ but the study lacked the power to detect these changes reliably.

Aims

This chapter examines whether or not TCM training leads to a reduction in teacher-reported 'burnout' and improved self-efficacy and well-being.

Methods

As with the main trial, this study is reported in accordance with CONSORT and TIDieR guidelines.^{120–122} The study design and procedures are presented in full in the published trial protocol,⁵⁷ which was approved by the TSC and DMC. Ethics approval for the conduct of the trial was obtained from the Peninsula College of Medicine and Dentistry Research Ethics Committee (reference number 12/03/141). The class teachers who were invited to attend TCM training were involved in this substudy, and full details about participant identification, consent, masking and randomisation are presented in *Chapter 3*.

Outcomes

Teachers were asked to complete the Everyday Feelings Questionnaire (EFQ), the Teachers' Sense of Efficacy Scale (TSES) and the Maslach Burnout Inventory (MBI) on two occasions: once in September/October (baseline) before training began and then again in the following June, 2 months after training ended (follow-up). The EFQ is a 10-item measure that records how the respondent has felt over the previous 4 weeks. Half of the items focus on well-being and half focus on distress. Items are scored from 0 to 4 for items with distress content and from 4 to 0 for items with well-being content, with a maximum total score of 40 across items; a higher score indicates increased distress. STARS used the shortened 12-item version of the TSES to assess teachers' perceptions of their sense of effectiveness as a teacher on three subscales (each with four items): items cover Student Engagement, Instructional Practice and Classroom Management. Responses are coded 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. The MBI is a 16-item measure that assesses aspects of 'burnout syndrome', which are recorded on 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. A high degree of burnout is reflected in high scores on the Exhaustion and Cynicism subscales and low scores on the Professional Efficacy subscale. A copy of the case report form can be found in *Report Supplementary Material 4*.

Statistical analysis

The main analyses of teachers' outcomes used the intention-to-treat principle to estimate the intervention effect. Linear regression models were fitted to analyse the EFQ, TSES and MBI measures, adjusting for the randomisation balancing factors (urban vs. rural/semi-rural status, KS 1 vs. KS 2 status, percentage of children on free school meals), study cohort (recruited in 2012, 2013 or 2014) and the baseline outcome score. In addition, we adjusted for the following variables when they were prognostic factors for the outcome: school-level IDACI, percentage of children identified as having SEN at the class level, teacher gender, and whether or not the teacher had > 5 years of teaching experience. Analyses were performed using Stata software version 14.2.

Results

We recruited a total of 80 trial teachers; 40 were allocated to attend a TCM training programme and 40 acted as controls (*Figure 22*). Baseline data were collected for all 80 teachers and characteristics were generally balanced between the two arms (*Table 25*). Follow-up data on the EFQ, TSES and MBI were collected for 74 teachers; three TAU and three intervention teachers did not provide these data. TCM training was well attended, with 36 (90%) of the 40 teachers in the intervention arm attending four or more TCM sessions; 23 teachers (58%) attended all six sessions.

Table 26 summarises the comparison between the trial arms at follow-up for the EFQ, TSES and MBI measures. There was little evidence of an intervention effect on any of these outcomes, with self-efficacy, burnout and well-being scores changing very little throughout the year.

Discussion

We found no evidence that TCM training reduced teacher-reported 'burnout' or that it improved self-efficacy and well-being. These results are surprising in the context of the finding of Marlow *et al.*⁹⁴ that there was an improvement in teachers' self-efficacy and are in contrast to the process evaluation findings reported in *Chapter 6*. During FGs and interviews, teachers reported that the TCM training had had a range of impacts on their practice, including changing their mindset, helping them see things from the child's point of view, helping them develop better relationships with the children, making them think more before responding, helping them feel more positive, more confident and more in control, and helping them create a more positive cycle of behaviour through modelling.

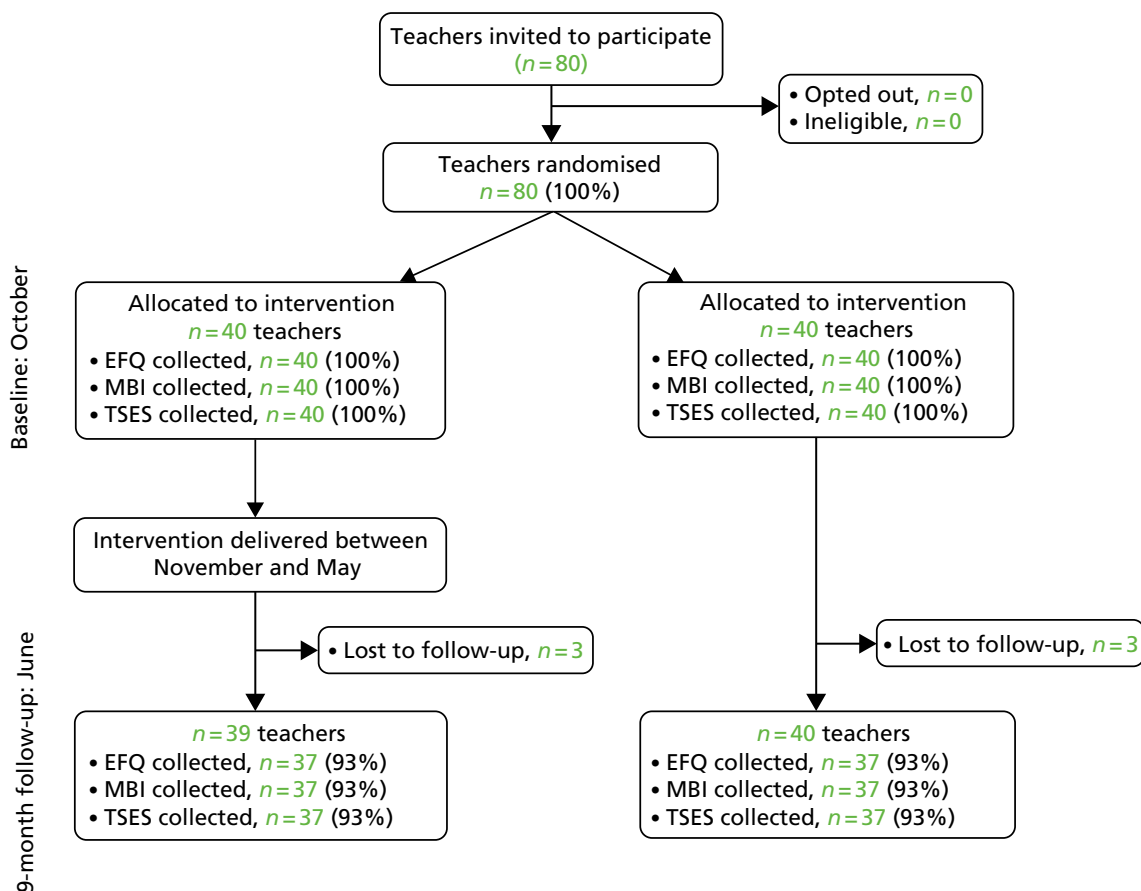


FIGURE 22 Teachers' substudy CONSORT diagram.

TABLE 25 Summary of teacher characteristics at baseline

Characteristic	Intervention (N = 40)	TAU (N = 40)
Length of service (years), mean (SD)	7.5 (6.5)	5.9 (5.9)
Age (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 role, n (%)	4 (10)	2 (5)
Key stage, n (%)		
KS 1	20 (50)	21 (53)
KS 2	20 (50)	19 (47)
Teaching qualification, n (%)		
PGCE	17 (43)	22 (55)
BA, BSc, BEd including QTS	18 (45)	16 (40)
Other	1 (3)	0 (0)
Missing	4 (10)	2 (5)
Whole classroom support, n (%)		
Full time	23 (58)	22 (55)
Part time	17 (43)	15 (38)
One-to-one teaching support, n (%)	18 (45)	22 (55)
Percentage of SEN children in class, mean (SD)	23 (12)	21 (12)
Teacher Self-efficacy Questionnaire, mean (SD)		
Student engagement	6.8 (1.0)	7.1 (1.0)
Instructional practice	6.9 (1.0)	7.2 (0.9)
Classroom management	7.3 (0.9)	7.5 (0.9)
MBI, mean (SD)		
Exhaustion	2.9 (1.4)	2.5 (1.4)
Cynicism	1.2 (1.0)	1.1 (1.0)
Professional efficacy	4.2 (1.0)	4.6 (0.8)
EFQ score, mean (SD)	17.2 (7.0)	13.9 (6.6)

BA, Bachelor of Arts; BEd, Bachelor of Education; BSc, Bachelor of Science; PGCE, Postgraduate Certificate in Education; QTS, Qualified Teacher Status.

Given these reported changes and past work that demonstrates that behaviour management training reduces teacher-reported stress and emotional difficulties,¹¹⁵ we would have expected to see an increase in teacher well-being and a reduction in their reported levels of workplace stress. A possible explanation of why we did not see these changes could be that we were looking to detect a change in the wrong constructs and that TCM's impact is being felt elsewhere by teachers. Teachers voluntarily identified with this possibility and throughout the trial made frequent comments to research staff that, although they were able to identify positive changes as a result of the training, the wider context of their professional life was having a greater impact on their well-being, self-efficacy and burnout. As discussed in *Chapter 6*, teachers report that the biggest impact the training had was on their relationship with the children, but this was not something that we systematically studied.

TABLE 26 Main comparisons of teacher outcomes at the 9-month follow-up (T1)

Outcome	Intervention, <i>n</i> = 37	TAU, <i>n</i> = 37	Unadjusted mean difference	AMD		
	Mean (SD)	Mean (SD)	Estimate	Estimate	95% CI	<i>p</i> -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
MBI						
Exhaustion	2.3 (1.5)	2.5 (1.4)	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
EFQ	15.6 (7.6)	13.6 (6.0)	2.0	−0.2	−2.7 to 2.2	0.85

This substudy was not powered to detect an effect on teachers, so it is still a possibility that a larger sample may reveal intervention effects, but considering our process evaluation the most likely explanation is that we were looking for impacts in the wrong constructs and instead should have been exploring teachers' relationships with the children they teach.

Chapter 8 Comparison of outcomes between children who are taught by teachers in the academic year following Teacher Classroom Management training and those who are taught by teachers while they receive Teacher Classroom Management training

This chapter reports on the substudy that focused on the potential benefits that TCM training had on children taught in the academic year following their teachers' training.

Introduction

During the FGs carried out as part of our process evaluation, teachers repeatedly reported that although the course had been immediately useful to them, they expected that it would have even more of an impact the following year because they would be able to incorporate all the strategies into their planning and fully embed TCM into their practice. This raised a methodological question for us because in the main trial we followed the children into subsequent academic years because we wanted to include as long a follow-up of the children's progress as possible. These comments suggested to us that the children with whom the teachers work in the academic year following their training might actually derive more benefit than the intervention children in the main trial who were taught in the same academic year in which the teacher was trained. The effects of the course might also be increased because the children taught in the following year will be exposed to the intervention from the outset of the academic year as opposed to our trial children, whose teachers did not begin the intervention until November and completed it only 1 month or so before our first follow-up measures were collected.

Aims

In this substudy, we aimed to compare the outcomes of children taught by teachers trained in TCM in the previous year ('Post-TCM') with those of children in the intervention arm of the trial who were taught by recently trained teachers ('Concurrent-TCM'). The former group of children, separate from those children in the main trial, were recruited using a convenience sample of TAU schools that accessed TCM training in the final year of the study. Our hypothesis was that these Post-TCM children might derive greater benefit as the teachers would have had more time to embed what they learnt into their practice and to use it in their planning for the academic year.

Methods

Our final round of TCM training was held in the 2015–16 academic year and involved 21 teachers, all of whom were from TAU schools. All children taught by these TCM-trained teachers in the 2016–17 academic year were eligible for inclusion provided that the class teacher judged that they and their parents had sufficient English-language comprehension to understand recruitment information and complete outcome measures.

Written consent was obtained from the headteacher for the school's participation and from the class teacher for their involvement. Parent information leaflets were sent home with children, and parents were given 2 weeks to 'opt out' their child from the research. In order to opt out their child, parents could

return a pre-prepared letter or contact the research team directly. Verbal assent was obtained from children each time they were asked to complete a questionnaire.

Outcomes

We replicated the outcome measures and time points of each school's first year of participation in the main trial, in that all measures were collected during the first academic term of 2016 and 9 months later (T1) in June 2017. Full details of the measures are provided in *Chapter 2* and summarised in *Figure 23*.

Statistical analysis

Quantitative outcomes were compared between Post-TCM and Concurrent-TCM children using random-effects linear regression models. Binary outcomes were compared using marginal logistic regression models using generalised estimating equations with information sandwich ('robust') estimates of standard error assuming an exchangeable correlation structure. These methods allow for the correlation of children's outcome scores within schools. In addition, the following potential confounders were adjusted for: whether the school was urban or rural/semi-rural; whether the class was KS 1 (Reception to Year 2) or KS 2 (Years 3 or 4); school-level deprivation (whether the percentage of children eligible for free school meals was > 19%, the UK national average in 2012⁴⁷); child gender; and baseline score for the outcome. Analyses were performed using Stata software version 14.2.

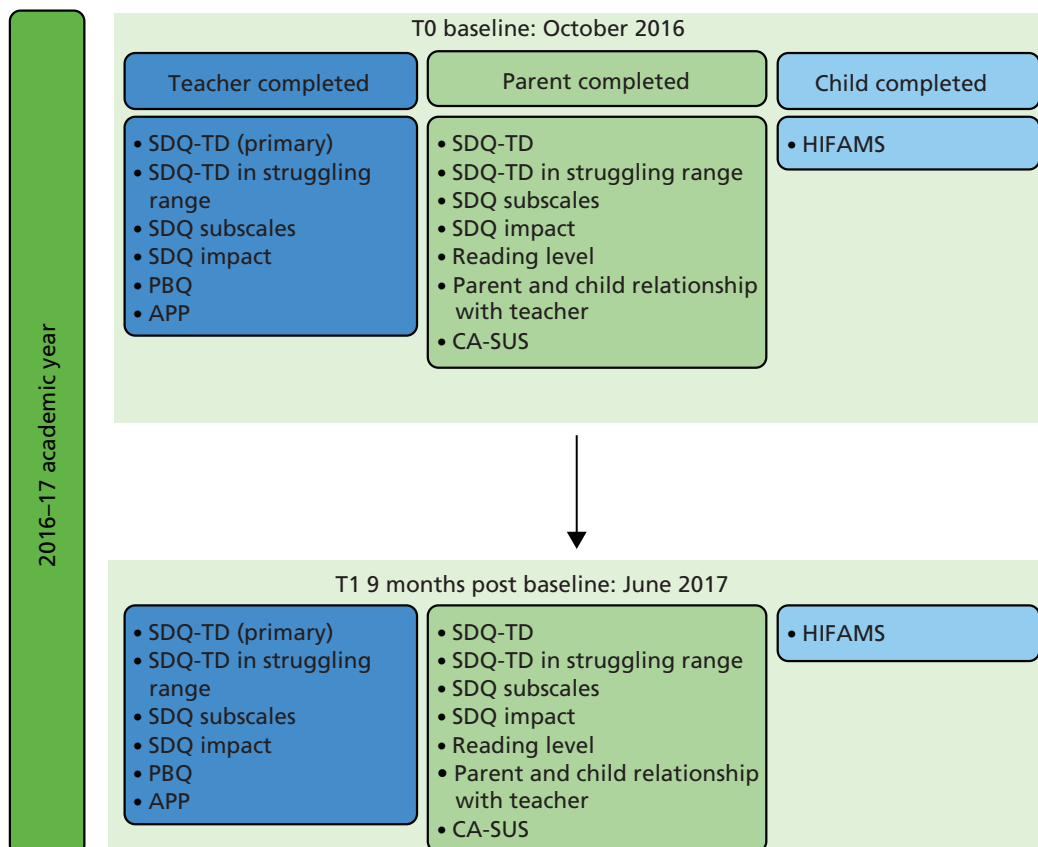


FIGURE 23 Schematic detailing the timing of outcome measures for the Post-TCM substudy.

Results

Fourteen of the 21 schools whose teachers attended TCM training in 2015–16 consented to be part of this study; six schools said that they did not have the time to be involved in this additional work and in one school the teacher who had accessed TCM training had left the school. A total of 397 children were eligible for inclusion in this study; 13 were opted out by a parent, which left a sample size of 384 children. During the year of this additional substudy, two more teachers left the study schools, so follow-up data were not collected for the children in these classes. Our comparison arm (Concurrent-TCM) comprised the main trial's intervention arm, and full details of this sample are provided in *Chapters 2 and 3*. *Figure 24* depicts the flow of participants through the study and *Table 27* compares the baseline characteristics of both arms of the trial.

Table 28 summarises the comparison between the Post-TCM and Concurrent-TCM groups at the 9-month follow-up. There was little evidence that any outcomes differed depending on when their teacher had completed training, with the exception of a reduction in the percentage of children who scored below national expectations in numeracy for the Post-TCM children (23%) versus the Concurrent-TCM children (33%) (OR 0.45, 95% CI 0.23 to 0.89; $p = 0.02$).

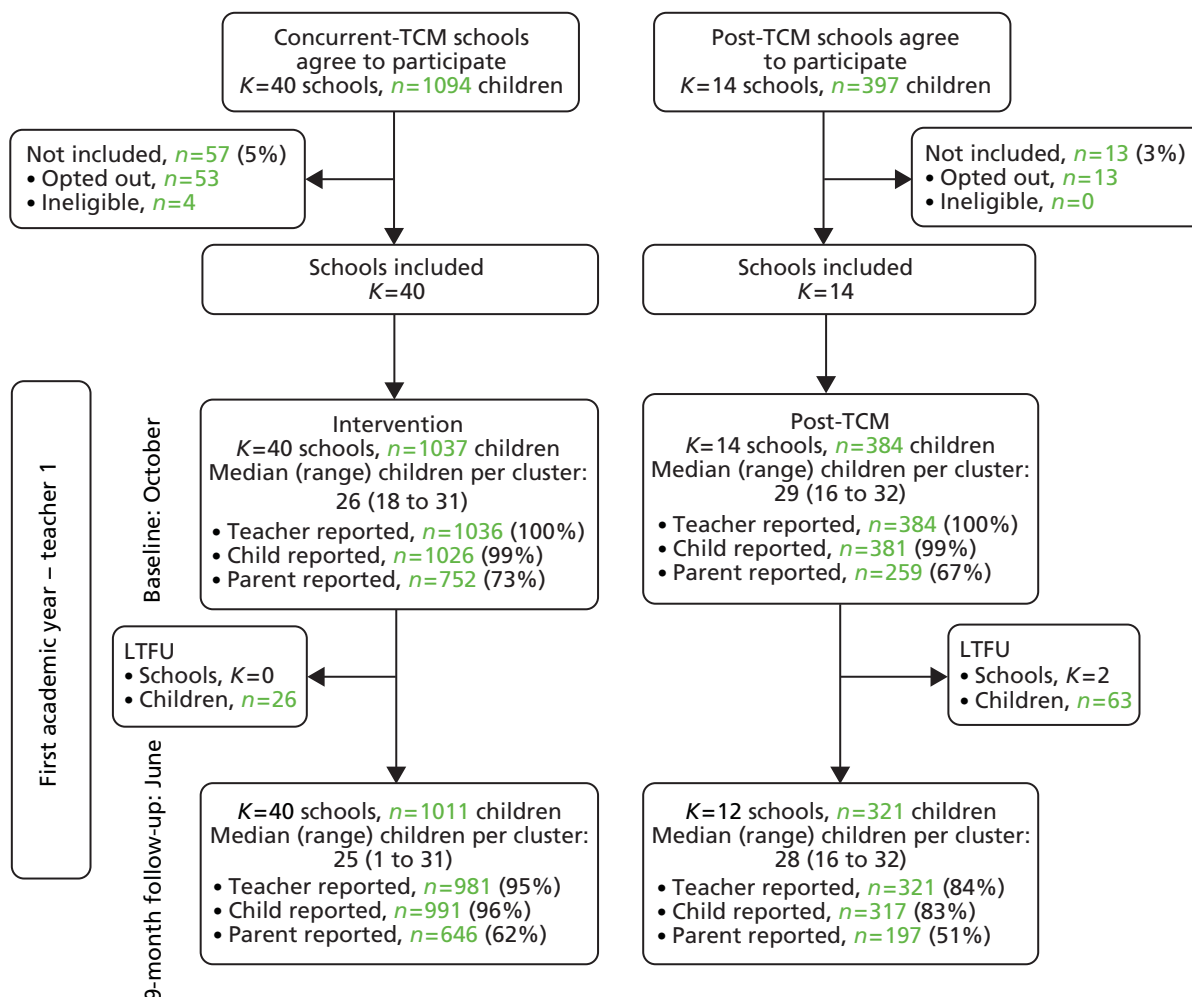


FIGURE 24 Post-TCM substudy flow of participants. *K*, number of schools (clusters); LTFU, lost to follow-up; *n*, number of children.

TABLE 27 Summary of children's characteristics at baseline

Variable	Post-TCM	Concurrent-TCM
School (cluster) characteristics	N_S = 14	N_S = 40
Rural/semi-rural vs. urban school, <i>n</i> (%)		
Urban	7 (50)	22 (55)
Key stage, <i>n</i> (%)		
KS 1	5 (36)	20 (50.0)
KS 2	9 (64)	20 (50.0)
Percentage of children eligible for free school meals, median (IQR)	12 (6–29)	12 (8–24)
IDACI, median (IQR)	0.15 (0.03–0.28)	0.17 (0.08–0.24)
Teacher (cluster) characteristics	N_T = 14	N_T = 40
> 5 years of teaching, <i>n</i> (%)	11 (79)	20 (50)
Age (years), mean (SD)	31.9 (10.0)	34.5 (9.0)
Female, <i>n</i> (%)	14 (100)	32 (80)
NQT, <i>n</i> (%)	0 (0)	2 (5)
Has management position, <i>n</i> (%)	0 (0)	4 (10)
Pupil characteristics	N_P = 384	N_P = 1037
Female, <i>n</i> (%)	175 (46)	483 (46.6)
Age in years at last birthday, mean (SD; range)	7.0 (0.9; 4–8)	6.2 (1.4; 4–9)
	N_P = 241	N_P = 721
Ethnicity, <i>n</i> (%)		
White	233 (97)	689 (95.6)
Black	0 (0)	4 (0.6)
Asian	4 (2)	5 (0.7)
Mixed	4 (2)	20 (2.8)
Other	0 (0)	3 (0.4)
	N_P = 275	N_P = 860
Index of Deprivation Affecting Children Index, median (IQR)	0.13 (0.02 to 0.52)	0.16 (0.08 to 0.64)
	N_P = 260	N_P = 70
Number of children in household, <i>n</i> (%)		
1	32 (12)	125 (16)
2	138 (53)	403 (52)
3	61 (23)	175 (23)
4	23 (9)	45 (6)
≥ 5	6 (2)	22 (3)
	N_P = 258	N_P = 766
Lives in rented housing, <i>n</i> (%)	95 (37)	475 (62)

TABLE 27 Summary of children's characteristics at baseline (continued)

Variable	Post-TCM	Concurrent-TCM
	N_p = 257	N_p = 758
Parent's highest qualification, <i>n</i> (%)		
None	18 (7)	29 (4)
GCSE or equivalent/A level or equivalent	124 (48)	377 (50)
University degree or equivalent and above	115 (45)	352 (46)
	N_p = 384	N_p = 1036
SDQ-TD score (teacher report), mean (SD), raised if ≥ 12	6.5 (6.2)	6.8 (5.6)
SDQ-TD score in struggling range ^a (teacher report), <i>n</i> (%)	75 (20)	206 (20)
SDQ Behaviour score (teacher report), mean (SD), raised if ≥ 3	0.8 (1.4)	0.8 (1.5)
SDQ Emotions score (teacher report), mean (SD), raised if ≥ 3	1.6 (2.2)	1.5 (2.0)
SDQ Overactivity score (teacher report), mean (SD), raised if ≥ 6	3.0 (3.0)	3.3 (3.0)
SDQ Peer Relationships score (teacher report), mean (SD), raised if ≥ 3	1.1 (1.8)	1.2 (1.5)
SDQ Pro-social score (teacher report), mean (SD), low if < 6	7.6 (2.5)	7.3 (2.5)
SDQ Impact score > 0 (teacher report), <i>n</i> (%), raised if ≥ 1	132 (34.4)	395 (38.1)
	N_p = 253 to 259	N_p = 733 to 752
SDQ-TD score (parent report), mean (SD), raised if ≥ 14	7.6 (5.9)	7.8 (6.0)
SDQ-TD score in struggling range ^b (parent report), <i>n</i> (%)	35 (14)	124 (17)
SDQ Behaviour score (parent report), mean (SD), raised if ≥ 3	1.3 (1.5)	1.5 (1.6)
SDQ Emotions score (parent report), mean (SD), raised if ≥ 4	2.0 (2.1)	1.8 (2.0)
SDQ Overactivity score (parent report), mean (SD), raised if ≥ 6	3.2 (2.5)	3.3 (2.6)
SDQ Peer Relationships score (parent report), mean (SD), raised if ≥ 3	1.2 (1.6)	1.2 (1.6)
SDQ Pro-social score (parent report), mean (SD), low if < 6	8.6 (1.8)	8.4 (1.7)
SDQ Impact score > 0 (parent report), <i>n</i> (%), raised if ≥ 1	104 (32)	244 (33)
	N_p = 354	N_p = 1036
PBQ score, mean (SD)	1.5 (2.2)	2.0 (2.4)
	N_p = 379	N_p = 1025
HIFAMS score, mean (SD)	10.8 (2.6)	10.9 (2.5)
	N_p = 258	N_p = 746
Assessment of pupil reading level (parent report)		
Fluent reader, <i>n</i> (%)	129 (50)	320 (42.9)
	N_p = 260	N_p = 753
Assessment of pupil relationship with teacher (parent report)		
Good relationship, <i>n</i> (%)	221 (85)	632 (83.9)
	N_p = 259	N_p = 731
Assessment of parent relationship with teacher (parent report)		
Good relationship, <i>n</i> (%)	166 (64)	465 (63.6)
	N_p = 351	N_p = 1036
Below average on literacy, <i>n</i> (%)	136 (39)	440 (42.5)
Below average on numeracy, <i>n</i> (%)	123 (35)	343 (33.1)

A level, Advanced level; GCSE, General Certificate of Secondary Education; N_p, pupils denominator; N_s, schools denominator; N_t, teachers denominator.

a Struggling defined as scoring ≥ 12 out of 40.

b Struggling defined as scoring ≥ 14 out of 40.

TABLE 28 The comparison between the intervention and post-TCM arms at the 9-month follow-up

Variable	Post-TCM (<i>N</i> = 384)	Concurrent-TCM (<i>N</i> = 981)	Unadjusted, mean difference/OR ^a		95% CI	<i>p</i> -value
	<i>N</i> = 321	<i>N</i> = 981				
SDQ-TD score (teacher report), mean (SD)	6.4 (6.0)	5.6 (5.4)	0.81	0.72	-0.48 to 1.92	0.24
SDQ Behaviour score (teacher report), mean (SD)	0.9 (1.6)	0.7 (1.5)	0.14	0.13	-0.15 to 0.42	0.35
SDQ Emotions score (teacher report), mean (SD)	1.7 (2.2)	1.3 (1.9)	0.41	0.26	-0.21 to 0.75	0.28
SDQ Overactivity score (teacher report), mean (SD)	2.7 (2.9)	2.6 (2.9)	0.10	0.24	-0.22 to 0.70	0.31
SDQ Peer Relationships score (teacher report), mean (SD)	1.0 (1.6)	0.8 (1.4)	0.16	0.17	-0.12 to 0.46	0.26
SDQ Pro-social score (teacher report), mean (SD)	8.0 (2.4)	8.2 (2.3)	-0.22	-0.30	-0.82 to 0.23	0.26
SDQ Impact score, mean (SD)	1.2 (2.1)	1.1 (2.0)	0.10	0.02	-0.31 to 0.34	0.92
	<i>N</i> = 197 to 196	<i>N</i> = 645 to 624				
SDQ-TD score (parent report), mean (SD)	6.9 (5.8)	7.7 (6.4)	-0.96	-0.33	-0.96 to 0.31	0.32
SDQ Behaviour score (parent report), mean (SD)	1.1 (1.5)	1.5 (1.7)	-0.31	-0.04	-0.24 to 0.15	0.66
SDQ Emotions score (parent report), mean (SD)	1.9 (2.1)	2.0 (2.2)	-0.14	-0.15	-0.42 to 0.12	0.29
SDQ Overactivity score (parent report), mean (SD)	2.9 (2.4)	3.3 (2.7)	-0.35	-0.16	-0.44 to 0.12	0.26
SDQ Peer Relationships score (parent report), mean (SD)	1.0 (1.5)	1.2 (1.7)	-0.15	-0.02	-0.22 to 0.19	0.86
SDQ Pro-social score (parent report), mean (SD)	8.7 (1.6)	8.6 (1.7)	0.12	-0.10	-0.31 to 0.11	0.34
SDQ Impact score (parent report), mean (SD)	1.6 (3.0)	1.8 (3.1)	-0.20	0.11	-0.23 to 0.44	0.53
	<i>N</i> = 321	<i>N</i> = 981				
PBQ score, mean (SD)	1.8 (2.3)	1.7 (2.3)	0.13	0.33	-0.08 to 0.73	0.11
	<i>N</i> = 317	<i>N</i> = 991				
HIFAMS score, mean (SD)	10.8 (2.9)	10.8 (2.5)	0.02	0.12	-0.28 to 0.52	0.57
Assessment of pupil reading level (parent report)	<i>N</i> = 194	<i>N</i> = 639				
Fluent reader, <i>n</i> (%)	116 (60)	373 (58)	1.19	0.94	0.59 to 1.52	0.81
Assessment of pupil relationship with teacher (parent report)	<i>N</i> = 197	<i>N</i> = 643				
Good relationship, <i>n</i> (%)	184 (93)	591 (92)	1.22	1.19	0.54 to 2.59	0.67

TABLE 28 The comparison between the intervention and post-TCM arms at the 9-month follow-up (*continued*)

Variable	Post-TCM (N = 384)	Concurrent-TCM (N = 981)	Unadjusted, mean difference/OR ^a	AMD/OR ^b	95% CI	p-value
Assessment of parent relationship with teacher (parent report)	N = 197	N = 646				
Good relationship, n (%)	162 (82)	503 (78)	1.26	1.02	0.64 to 1.64	0.92
	N = 321	N = 953				
Below average on literacy, n (%)	87 (27)	336 (35)	0.73	0.64	0.34 to 1.21	0.17
Below average on numeracy, n (%)	73 (23)	312 (33)	0.63	0.45	0.23 to 0.89	0.02

a Mean difference reported for quantitative outcomes and ORs reported for binary outcomes.

b Analyses adjusted for baseline outcome score, gender, and school level variables (urban vs. non-urban, KS and free school meals percentage).

Within-school (within-cluster) correlation accounted for in the analyses using random-effects linear regression models for quantitative outcomes and marginal logistic regression models using generalised estimating equations with information sandwich ('robust') standard errors assuming an exchangeable correlation structure for binary outcomes.

Discussion

We found no evidence that the impact of TCM training improved children's outcomes for the children taught by a Post-TCM teacher compared with those children whose teacher was currently undertaking the training, except for numerical attainment, which was better for the children being taught by a Post-TCM teacher. However, attainment data in the main study were measured while the KS 1 and KS 2 curricula were undergoing a number of substantial changes,^{123,124} and this substudy took place once these changes had been embedded into practice. It is therefore possible that the improvement seen is the result of improvements across academic cohorts and not directly related to changes in the teachers' schemes of work and classroom practice as a result of TCM training. In addition, given the number of statistical tests undertaken, this finding may just be a chance finding rather than a true difference.

This substudy was limited by the sample of trained teachers and willing schools available to explore the often stated belief that the TCM training would be more effective when the core principles could be embedded in planning. We therefore had limited power to detect differences. We were disappointed not to have been able to recruit more of our TAU schools to this exploratory study, but we are aware that access to training was a major motivator for schools to become involved. This substudy involved yet more disruption to classes after 3 years and four data-collection points, and occurred at a time of pressure due to curriculum change.

Our findings do not suggest that having the full range of TCM strategies available to inform planning from the beginning of the academic year produces additional benefits to the emotional well-being and behaviour of the class. Only a single teacher was trained in each school. Without the ongoing support of a trained peer group, and particularly given the large changes in policy and curriculum during this study, teachers may not have been able to implement the strategies learnt as effectively as they initially thought. It is possible that with ongoing coaching and/or a whole-school approach, teachers' beliefs about the impact of the course on planning might have been borne out. Further study is needed to make a firm conclusion.

Chapter 9 Discussion

In this chapter the findings of all strands of the STARS trial are drawn together, along with the additional work involving the Post-TCM children and the linkage to the NPD.

Effectiveness and cost-effectiveness as a universal child mental health intervention

We detected a small but statistically significant improvement in teacher-reported children's mental health at 9 months (primary outcome: SDQ-TD). Almost all of the 95% CI for the mean difference, however, lies below the assumed minimum clinically important difference (effect size of 0.3 or raw difference of 1.8). Similar effect sizes have been reported on the same measure before and after attendance at CAMHS¹²⁵ and small effects from universal interventions are common, including the Campbell Collaborations systematic review of TCM strategies.^{14,126} The findings provide tentative evidence that TCM may provide teachers with strategies that affect children's mental health and, although small, the effect detected suggests that this intervention could usefully be tested further.

Small effect sizes are not uncommon in prevention studies. The population-level influence of universal interventions will be differentiated across subgroups, so that the intervention promotes health for those who are currently thriving while preventing deterioration or actively treating others who are vulnerable or struggling. Small population effects, therefore, do not necessarily demonstrate a lack of effectiveness for all subgroups.¹⁴ Indeed, our exploratory but planned subgroup analyses suggest that children with poorer mental health derived the most benefit according to teacher report. This is interesting as it echoes the findings from the trial of TCM and child programmes among high-risk children and classrooms, which also suggested that the intervention led to larger improvements among the children with more severe problems.³⁰ Similarly, a recent meta-analysis based on study-level summary data of three TCM trials²⁹ also reported a significantly larger effect on children with worse behaviour, and a similar pattern was reported in a meta-analysis based on individual participants of trials of the parallel IY parent training intervention.⁶ That the STARS trial has produced a similar pattern of effects adds weight to the argument that the small but statistically significant improvement is more than a chance finding, and suggests that TCM is worth further investigation as a universal intervention in the UK context. Finally, a UK-based cluster RCT of the Good Behaviour Game in 77 primary schools also suggested that boys' risk of developing conduct disorder demonstrated the most improvement, although the intervention demonstrated little effect on academic attainment or behaviour overall.¹²⁷

Given the quantitative findings that those in poor initial mental health seem to respond the most, the comments from teachers in the FGs were more mixed in terms of the impact on children who were struggling, although some were describing children with extremely complex needs who ultimately were not able to cope with mainstream educational provision. These comments highlight the tension commonly described by teachers between supporting children with SEN and undertaking their duty to the rest of the class,¹²⁸ and it is important to manage the expectations of universal interventions accordingly.

The effect of TCM on the primary outcome was not maintained at 18 and 30 months, which could mean that TCM has no longer-term impact, or could be a result of the children's reactions to the teaching style of their subsequent teachers who had not accessed TCM training. Given the current financial climate, this intervention was stripped back to the bare minimum for the STARS trial; just one teacher per school was trained. The IY institute currently recommends that teachers are supported during the training by additional coaching,¹²⁹ and this was an added component of Reinke's study.⁴⁰ Indeed, the consensus from the process evaluation was that a lack of the 'whole-school' approach or lack of understanding or support from colleagues interacting with the trial children undermined the effectiveness of the TCM strategies in the classroom. Similarly, the group leaders thought that individual coaching around particularly challenging

situations would be a beneficial addition. Most effective universal programmes report that they employ a whole-school approach,¹⁴ and the findings from this study would suggest that training all school staff, including learning support and playground assistants, to use the same strategies might amplify and sustain any initial impact on children's mental health that training a single teacher might have.

The teachers and group leaders in the process evaluation were arguing for the use of the IY TCM approach throughout the school and the training of a critical mass (if not all) of staff when they used the term 'whole-school approach'. Others would reserve this term for broader programmes that worked at universal, targeted and indicated levels of prevention simultaneously, which often include a mixture of environment, didactic and individualised content.^{130,131} The training of a critical mass of school staff in IY TCM shows promise. A quasi-experimental Norwegian evaluation matched 21 schools in which all Grade 1–3 teachers (children aged 6–8 years) were trained with 22 matched schools that did not receive training.¹³² The comparison of pre and post scores between children in TCM and TAU schools suggested a small improvement in social skills and a very small improvement in behaviour. This study involved the simultaneous training of staff with pre and post measures collected in the same academic year. The teachers in the STARS trial believed that any impact would increase in the following year when they could apply their strategies to planning as well as delivery, and children in the Norwegian study were not exposed to the IY strategies for very long. The effects of a whole-school approach may, therefore, increase over time. A critical mass of staff might also provide longer-term support for each other after the training was completed, which might also enhance any initial impact TCM had, as might the continued provision of coaching after TCM. Future research should test both the coaching and the training of more school staff in the UK context.

The addition of other IY programmes, which were initially conceptualised as a multistrategy approach to children with severe behavioural difficulties³¹ might further enhance the impact of IY TCM. In southern Ireland, such a combined programme demonstrated elevated SDQ scores at school entry and reduced subsequent SDQ scores in a non-randomised evaluation of schools that implemented IY TCM compared with those that did not. TCM was combined with access to the IY parenting programmes and the use of the IY classroom Dina programme for children.¹³³ This report, plus those reporting on other whole-school programmes, provide some important lessons for implementation.^{130,131,133,134} As highlighted by the Gatehouse project (an Australian, multilevel, integrated whole-school approach to well-being in secondary schools),¹³⁰ a clear and broad conceptual framework, plus support for schools to respond to their own context and own data, appears key to successful implementation, along with time and resources for reflective practice and sustainability. A recent US-based cluster RCT of the School-Wide Positive Behaviour Intervention and Support in Elementary schools, which was followed by a related implementation study, demonstrated improved behaviour, punctuality and attendance plus reduced exclusion.^{135,136} The theoretical basis of this programme is very similar to that of IY TCM, with an emphasis on proactive behaviour management and non-punitive responses. Interestingly, effectiveness was correlated with the extent to which schools implemented the different elements of this whole-school approach.

There may be a methodological explanation for the lack of effect at later follow-up. Baseline mental health measures and follow-up after 9 months were completed by the same teacher, whereas these measures were completed by two different teachers at 18 and 30 months. As would be expected, the ICCs were markedly larger for teacher-reported SDQ-TD score (0.12–0.18) than for the corresponding parent-reported score (0.06), which suggests that variability in how teachers score their pupils may contribute noise at the two later follow-ups.

Furthermore, children in the intervention arm were exposed to TCM strategies for a relatively short duration. In the process evaluation, teachers repeatedly discussed their belief that the TCM training would have a greater effect in the subsequent academic year because they could embed the strategy into their planning as well as their teaching, and the children would be exposed to these strategies from the outset

of the academic year. This led us to collect data from the Post-TCM comparison study, which did not provide empirical support for their belief. Recruitment to this substudy was disappointing, with only 67% of eligible schools consenting and a further 10% unable to provide follow-up data owing to the teacher leaving the school. This limited the power available to detect a difference between the Post-TCM sample and the Concurrent-TCM sample as the large CIs (OR 0.72, 95% CI -0.48 to 1.92) demonstrate that it is plausible that there are benefits in either direction. This additional work used a convenience sample of TAU arm schools that had already participated for 3 years in the STARS trial and occurred in an academic year when primary schools were having to grapple with substantial changes in the primary school curriculum.^{123,124} That participation was lower than anticipated is therefore, perhaps, not surprising. This abrupt change in the curriculum and its impact on the school context, teachers and children may have introduced time or cohort effects that we were unable to adjust for in our analysis. We lacked any process evaluation in this small additional study that might have offered alternative explanations or insight into these processes. The Post-TCM study is sadly inconclusive but certainly does not provide evidence that future studies should follow up the teachers' subsequent classes rather than the children they were teaching as they accessed the course, which is the design chosen by all other trials to date. As stated above, it is possible that a whole-school application of TCM may amplify any effect of TCM when one teacher per school is trained, as children could then benefit from the intervention throughout all their remaining school years.

Although TAU children were never taught by a TCM-trained teacher, in the majority of TAU schools there was a TCM-trained teacher at their school during their second and third years of the study. It is therefore possible that TAU children were exposed to elements of TCM through contamination effects, which would reduce the difference between trial and intervention arms. The fact that all schools eventually accessed TCM will, however, have concealed allocation in many schools; interviews with SEN co-ordinators, whose roles relate to behaviour management, revealed that they were often unaware of the school's involvement in the study.¹³⁷ Ideally, future research would not provide delayed training to TAU schools but rather offer some other type of incentive for participation.

The small but sustained effects on disruptive behaviour and inattention/hyperactivity as measured by the teacher SDQ across all three follow-ups are interesting and warrant replication. More than 40% of secondary school children reported in a survey that their classroom was too noisy to work in,¹³⁸ and overactivity/inattentive traits predict poorer academic attainment at General Certificate of Secondary Education (GCSE) level.¹³⁹ Children in poorly managed classrooms observe that disruptive behaviour commands staff attention, which may then increase later disruptive behaviour as well as disengagement from school, with its attendant risks to health and education among quieter, less disruptive pupils. The potentially positive impacts of TCM are, therefore, exciting and should be explored further in future research. The linkage to the NPD demonstrated some interesting differences in relation to exclusion between the trial arms, which suggest that these data may provide a useful vehicle through which to conduct longer-term follow-up of the STARS participants. However, no firm conclusions can be drawn from the NPD data because only a very small number of children were excluded in the STARS trial, which is what would be expected in primary schools.

We failed to detect any influence of TCM on the HIFAMS measure. Some researchers would recommend 8 years of age as the minimum¹⁴⁰ for reliable reporting on standardised measures of mental health, although increasingly researchers are seeking reports from younger children. HIFAMS was developed for the STARS trial and has demonstrated validity and moderate reliability among children as young as 4 years when tested in this and two other samples.⁵² It has now been translated into Italian and Persian and is being used in several education-based studies in the UK. Our findings would not suggest that being taught by a TCM-trained teacher leaves children happier about their school, but, equally, the study was not powered to detect an effect using this measure. It would seem important to try to capture children's experiences directly, because the correlation between parents and children is known to be low.¹⁴¹

We also did not detect an intervention effect on parent-reported measures, but this was not unexpected. Children respond differently in different situations and TCM targets the classroom rather than the home. In addition, a low level of agreement between parents and teachers is common in child mental health research,¹⁴¹ so the lack of effect on parent-reported outcomes does not necessarily indicate a lack of effectiveness in the school context. Parents will have only second-hand knowledge of how their children function in the classroom after all.

In terms of cost-effectiveness, the primary economic evaluation, using SDQ-TD scores, suggested that TCM may be cost-effective compared with TAU at 30 months, particularly at higher levels of willingness to pay for improvements in SDQ-TD scores. These results were not altered in a sensitivity analysis controlling for additional randomisation variables (CC2) but were affected by the impact of missing data (MI), which showed that the probability of TCM being cost-effective was reduced at lower willingness-to-pay thresholds for a unit improvement in total SDQ-TD scores. Although the SDQ is a commonly used measure of mental health in children and adolescents, it is not associated with a willingness-to-pay threshold for decision-making and is therefore of limited use in allocating scarce health-care resources. With very small differences in costs and effects, as in this study, it is not possible to draw firm conclusions about the cost-effectiveness of TCM without knowing society's willingness to pay for improvements in SDQ-TD scores.

Cost-utility analyses using QALYs generated using a mapping algorithm suggest that TCM has a higher probability of being cost-effective than TAU at 9- and 18-months but not at 30-months follow-up, which is consistent with the clinical findings that suggest that early positive findings were not maintained over the longer term.

Small differences in SDQ-TD and mapped utility scores over the trial follow-up made extrapolation beyond the trial period unnecessary, as the baseline beyond which data would be extrapolated from the end of the trial period was almost identical in the two groups. Longer-term analysis would therefore serve only to replicate the within-trial results. Although small clinical differences can result in large changes in health outcomes over longer-term models, there was limited evidence to suggest this in the trial data, particularly given the sharp drop in the probability of TCM being cost-effective using QALYs from approximately 80% at 18 months to 5% at 30 months. It is therefore fair to assume that any projections of costs and effects beyond 30 months would be of questionable validity. Appropriate temporal extrapolation of costs and utilities is reliant not only on clinical outcomes but also on many other long-term processes such as adaptation to particular health states, data on which are not currently available for children and adolescents.

Fidelity of delivery

The reported checklists completed by the group leaders and teachers, along with the monthly supervision from the programme developer, demonstrate that the course was delivered with fidelity to the model. In addition, we demonstrated, in those schools where T-POT observations were carried out, that the training changed the behaviour of teachers back in the classroom. Teachers in the intervention arm employed more praise and, paradoxically, more indirect commands, and child compliance increased. The two other studies that have used the T-POT observational measure^{38,39} both detected a decrease in negative strategies rather than an increase in positive strategies, but Hutchings *et al.*³⁸ focused their observation on key high-risk target children, rather than the whole class, so the findings are not directly comparable. The observed increase in the use of praise mirrors the findings from observations of parent-child dyads in trials of the IY parenting programmes, but we did not detect any reduction in the use of negative strategies that we would have expected.¹⁴² We were resourced only to conduct observations in one-quarter of participating schools, which limited our power to detect changes in the strategies employed, and may have introduced biases if there were systematic differences between participating schools and those who chose not to participate or were not invited to do so. The observers were blind to treatment allocation and the schools were selected on the basis of cohort (cohort 1 or 2) and geography (cohort 2 only; schools in and around Exeter) rather than enthusiasm (or lack thereof) for the course. The T-POT measure was chosen for comparability with other

European trials of IY. Future research should consider observations in a larger proportion of classes, if not all of them. Given the issues that have arisen in previous studies that have focused on target children when those children are absent or have little interaction with the teacher during the observation period, we would recommend observations that focus on the whole class. A review of moderators of the parallel parenting intervention suggests that observations may be more sensitive than questionnaire measures.¹⁴³ However, they are time and resource intensive in terms of ensuring that researchers are sufficiently trained to conduct and record the observation reliably and in organising when they can slot into the school timetable without disruption that might alter teacher behaviour and thus invalidate the observation. As soon as a teacher alters their behaviour to accommodate the requirements of the observation, then, arguably, the observation is unable to be considered valid and the results should not be regarded as any more reliable than questionnaire measures.

Academic attainment

In the process evaluation FGs, teachers reported that children were calmer, more motivated and more ready to learn, as our logic model predicted. These reported experiences did not translate into empirical evidence of an improvement in academic attainment at a universal level. The interaction with baseline mental health, which is difficult to interpret in terms of direction, is interesting and suggests that the impact on attainment is worth studying in more depth. It is intuitively plausible that a more settled classroom would allow accelerated progress for children who were otherwise thriving. Given the crudeness of the measure available to us, we were not surprised that there was no main effect detected on attainment. It would not have been feasible to conduct individual-level attainment tests at four data points for > 2000 children, and others have commented how difficult it is to find appropriate measures of attainment for use in education-based trials.¹⁴⁴ The linkage to the NPD will allow us to explore whether or not the improved attention suggested by the SDQ Hyperactivity scale translates into better academic attainment over time when all these pupils have taken their national exams in Years 6 (2022) and 11 (2027). A population-based study has suggested that each 1-point increase in teacher reported inattention/hyperactivity symptoms at 7 years of age was associated with a 2- to 3-point decrease in GCSE score and a 10–12% increase in the likelihood that the pupil would not obtain five GCSEs at grade C or above, including in maths and English.¹⁴⁵ Arguably, evidence of 'real world' impacts on attendance, exclusion, SEN and qualifications are more useful and powerful measures than standardised tests, although these children will have been exposed to a great deal of other influences during their school careers. The interesting findings from the Head Start Programme,¹⁴⁶ which began in the USA in the mid-1960s, suggest that even a brief pre-school intervention may produce impacts more than a decade later and suggest that these longer-term influences might be present in the STARS sample.

Teacher mental health, burnout and self-efficacy

In the FGs, teachers reported that attendance at the course changed their mindset, increased their focus on children's needs, improved their relationships with pupils and parents, helped them to be more proactive and less reactive, and left them feeling calmer, more confident and more in control. These findings were predicted by our logic model (see *Chapter 1*) and suggest that TCM was experienced as helpful despite teachers' many competing priorities. There is a stark contrast between the highly positive endorsement of the course by these teachers and group leaders and the lack of impact on teacher's mental health, self-efficacy and burnout in the quantitative measures.

Teachers' comments during the process evaluation and anecdotally during the course of data collection (see *Chapter 6*) suggest that although the TCM training was hugely helpful to them in the classroom, the external and internal pressures on primary schools have escalated during the course of the study. Teachers, particularly in the current climate of reduced external support and shrinking budgets, are juggling many competing priorities. Given the impact of poor mental health on academic outcomes,^{16,17} and the regular

and longitudinal contact with children and their families, schools are necessarily a key forum for mental health promotion and prevention, now enshrined in English government policy.¹¹⁹ Our findings suggest that IY TCM was perceived to be useful and relevant to the practice of participating teachers.

Many teachers commented on the pressures they face, which may contribute to the shockingly high and sustained levels of psychological distress that we detected among teachers in this study; 10% scored above a cut-off point that suggests moderate depression at all four data points.¹⁴⁷ These high levels of stress and distress in the wider school environment may also have served to undermine any positive influence of TCM on self-efficacy, particularly in relation to the classroom management subscale, on which we did detect a statistically significant improvement in our uncontrolled feasibility work.⁹⁴ It may also be that our change to the short Teacher Self-Efficacy Scale in an effort to reduce participant burden rendered the measure insensitive, so future studies should consider using the original longer form¹⁴⁸ as we did in our feasibility work. In addition, some teachers reported that what changed most in the classroom was their relationship with the children, so future studies should consider including a measure of the teacher–pupil relationship. The impact of teacher–pupil relationships should not be underestimated; poor teacher–pupil relationships predict poor mental health, particularly behaviour problems, and were also related to poor family function (primary school-age children only) and future exclusion (secondary school-age children only) in a large population-based study of primary school children.¹¹³

The feasibility and acceptability of the Teacher Classroom Management course

The high levels of attendance and positive feedback from teachers and group leaders in our process evaluation suggest that TCM is an acceptable and feasible course for primary school teachers in the UK context. Two-thirds of teachers attended all sessions offered, and 89% attended four or more sessions. Other evaluations of this programme have achieved higher attendance by offering ‘catch up’ sessions to teachers who had to miss a session, which the STARS trial lacked the time and resources to do.³⁰ However, the CACE analysis suggests that increased attendance would not have increased the effectiveness of the intervention substantially, so we lack evidence to endorse the provision of catch up sessions.

The teachers reported that they valued the structure of the course, the opportunity to ‘see, practice and implement’ TCM strategies and the chance to discuss their experiences. The collaborative ethos of the course and the time to share and reflect outside their classrooms were also highly valued. On the whole, teachers rated aspects of the course, such as the course content, video clips, discussion and delivery of the course, as either ‘helpful’ or ‘very helpful’, although, in general, ‘satisfaction’ type questions are often answered very positively after courses and workshops in all fields (see *Chapter 6*).

Teacher-reported facilitators to learning included an understanding of the theory behind the strategies, as endorsed by the TCM training, the adaptability and flexibility of the TCM strategies and the fact that the strategies were generally easy to use. Teachers mentioned that they used a number of different strategies as a result of their attendance at the TCM course, which included specific and proximal praise, ignoring, rewards and consequences, time-out, emotional regulation and behaviour plans. They also reported attempting to use coaching and commentary, although the feedback on this was more mixed, which may indicate a technique less familiar or less appropriate to the UK context, or one that group leaders need to increase the focus on. Limitations to implementation were mainly external to the course, as reported by all stakeholders in our process evaluation, and included a lack of support from colleagues or senior management teams, clashes with school behavioural policy, and fears about the perception of TCM strategies by external bodies such as Ofsted. However, these barriers are potentially tractable if senior managers and school inspectors are brought on board to ensure effective implementation within both schools and teacher training models.

Methodological issues

The internal validity of the STARS trial derives from the high retention of schools, teachers and pupils over 30 months, the delivery of TCM with fidelity by experienced practitioners trained and supervised by the programme developer, independent randomisation and the use of a strongly validated and widely used primary outcome measure. Trial arms were balanced in terms of school characteristics. There were small imbalances in terms of children (more poor readers and 4-year-olds in the intervention arm) and teachers (higher levels of psychological distress, lower self-efficacy, fewer qualified for > 5 years and older in the intervention arm), which may have made it harder to detect an effect on teacher outcomes and academic attainment.

High levels of attendance suggest that teachers valued TCM, and the participating schools were generalisable to the UK population in terms of class size and eligibility for free school meals. The trial benefited from a rigorous cost-effectiveness analysis and a detailed process evaluation.

Like all studies, the STARS trial had some methodological limitations. It was not possible to mask teachers to their allocation, risking response bias, particularly for the 9-month follow-up when intervention teachers reported on behaviour and mental health outcomes following attendance on a course designed to specifically improve these. Thus, our small positive effect at 9 months may be either a chance finding or may be due to teachers' enjoyment of the course and their belief that it was effective, rather than a true indication of effectiveness. However, outcomes at 18 and 30 months were completed by different teachers who did not attend the training and the decrease in both the teacher-reported SDQ Hyperactivity subscale and the PBQ across all follow-ups undermines the argument that the findings can be wholly explained by reporting bias, despite the negative findings on the primary outcome at this time point. Moreover, the blind observations that we were able to undertake did detect changes in teacher behaviour and in child compliance on some strategies that were similar to the findings of previous studies in this area.^{38,39}

Similarly, the size of the research team and the need to develop close and sustained relationships with such a large number of schools, which underpinned our success in recruitment and retention, also meant that only the researchers who conducted observations within the classroom were blind to allocation status. Most measures were completed independently; teachers completed their measures directly on to a database, and parents completed theirs via mail-outs in the book bag or direct to their home. Researchers supported children under the age of 7 years to complete HIFAMS individually, whereas children in KS 2 completed this questionnaire as a whole-class activity, which allows little opportunity for the researchers to influence responses. In addition, we did not detect a difference by trial arm on this measure.

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 their school. This is in fact what occurred with the only school that was lost. In addition, such schools would already have had the involvement of a number of external organisations that 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. The distribution of teacher SDQ-TD scores at baseline, however, was almost identical to that from population surveys of children of this age, which suggests that any selection bias was minimal.⁴⁹

Headteachers are used to considerable autonomy, and it was clear from our feasibility work that any attempt to control the selection of teachers would be a major disincentive to their school's participation in the study. There are two potential biases that might occur from headteachers' selection of teachers to attend the course. If teachers were selected because they struggled with behaviour management, we might overestimate the impact of the intervention, whereas if they were selected because of a particular

interest in socioemotional well-being, we might underestimate the impact if interest correlates with skills, or overestimate the impact if interest correlates with receptiveness. As the 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 4*). Our process evaluation involved interviews with headteachers and suggested a number of reasons for their choice of teacher to nominate, which included NQT status, allocation of a class known to be particularly challenging or a known interest in behaviour management.

Some might query why we detected a difference on the PBQ across all three follow-ups according to teachers but no difference on the Conduct Subscale of the SDQ according to teachers or parents. The former was derived from the longer scale used in the School Effectiveness Literature¹⁴⁹ and so focused on common low-level disruptive behaviours that impede effective teaching and are the direct targets for TCM. It covers talking out of turn, interrupting other pupils, unnecessary non-verbal noisiness, cheeky or rude comments to the teacher, and verbal or physical aggression towards peers. In contrast, the SDQ Conduct Subscale focuses on more antisocial behaviours, such as cheating, lying, stealing, tantrums and disobedience, which are uncommon in this age group in a mainstream school context, are less easily influenced by teachers' general classroom management strategies and are not so closely aligned to the aims of the TCM training. We did not anticipate changes on this subscale, which was the reason for the inclusion of the PBQ.

Finally, there were some potentially interesting background characteristics of the children about whom we lacked information, such as a previous history of poor mental health either in themselves or in their family. There is always a tension between what it is desirable to know about and the risk of overburdening participants, which risks incomplete data.

The contrast between the highly positive reports from teachers in the process evaluation and the tentative evidence of effects from the quantitative findings is interesting and puzzling. FGs set out to invite all teachers who attended the TCM training (although data collection stopped when data saturation was reached); response rates were high and teachers were incentivised to attend, and we also actively encouraged those who could not attend and who had dropped out of the course to discuss the same topics via an interview. It seems unlikely, therefore, that these teachers were particularly enthusiastic and they did not seem to differ systematically from the rest of the sample of participating teachers. Teachers who took part had varying levels of experience, were of varying ages and worked within both KS 1 and KS 2. The FGs and interviews themselves were conducted using a robust, reproducible methodology and a published protocol.¹⁰¹

However, like the main trial, the process evaluation also has limitations. Although every effort was made to ensure that the FGs and interviews were conducted by a researcher who had not acted as the main point of contact for the teacher during their involvement, researchers were not independent of the STARS trial and may have been known to the teachers. This could have led to information bias if teachers felt inhibited in reporting negative results.¹⁵⁰ Given that feedback was consistent across FGs and that interviews were conducted with > 40 teachers, we are confident that we minimised such potential.

Although we successfully obtained data from > 70% of parents at baseline, attrition was marked. Response rates for the CA-SUS were particularly affected by low parent report, with only 48% responding. Although MI of the missing data did not markedly affect the results, any conclusions reached must be considered alongside this low response fraction.

The economic evaluation was limited by the need for mapping to create QALYs. Although NICE recommends the use of mapped health state utility estimates when directly collected data are not available,⁷³ the validity of these mapping methods has not been fully addressed. Mapping can result in loss of information and increased uncertainty and thus requires a closer assessment of the agreement between target and source instruments. It has been argued that current methods for mapping are not known to be conceptually robust⁸⁸ and carry a 'significant risk that may be harmful to population health'.⁸⁹ It is therefore important to consider

results of the mapped cost–utility analysis with greater uncertainty. It is possible that vital information is lost when mapping five subscores of the SDQ (emotional symptoms, conduct problems, inattention/hyperactivity, peer problems and a pro-social scale) to the nine dimensions of the CHU-9D (sad, worried, pain, annoyed, tired, homework or schoolwork, daily routine, activities and sleep), which could therefore lead to erroneous conclusions about cost-effectiveness. Authors of the mapping algorithm used in the present study noted that, although the algorithm performed well in predicting mean group observed utility values, it did not perform well in predicting at individual participant level.⁷⁸ Despite published guidelines and a checklist for the reporting of mapping algorithms,^{90,91} only a few studies have been found to fulfil or partly fulfil all criteria on the checklist.⁹²

It was not possible to calculate QALYs because the CHU-9D is not suitable for the very young children included in this sample, whose ages ranged from 4 to 9 years at recruitment and 6 to 11 years at final follow-up.¹⁵¹ Feasibility work that we conducted of a child version of the EQ-5D suggested that our very young sample struggled to understand the concepts involved and that the measure was not valid or reliable.⁷⁶

The economic evaluation was also limited by failing to consider spillover effects in teachers' quality of life. TCM has been designed to increase the well-being of teachers as well as of students, and future research should attempt to capture these effects by directly administering a preference-based quality-of-life instrument such as the EQ-5D, which is commonly used in adults.⁹³ However, the results presented in *Chapter 7* suggest that there was no evidence that TCM training improved teacher well-being, so the impact of failing to measure teacher utility is likely to be limited.

Future research

The findings of the STARS trial emphasise the importance of the school context to children's mental health and provide tentative evidence that teacher training and continuing professional development might usefully explore how classroom management style may influence children's mental health, behaviour and attainment. Future studies should consider the following questions:

1. How does TCM influence teacher–pupil relationships?
2. If children are taught by teachers who have accessed TCM in 2 (or more) sequential years, is the impact on the mental health of children sustained or even amplified?
3. Does the impact of TCM increase in the year after the teacher has accessed the course as teachers reported it would?
4. What is the influence of TCM training on robust measures academic attainment in the short and long term?
5. Does the addition of the IY parenting course and/or the IY Dina programme for children with particularly challenging behaviour enhance the effectiveness of TCM? Is this cost-effective?
6. Does the addition of individual coaching for teachers outside the TCM sessions amplify the effectiveness of TCM? Is this cost-effective?
7. Would TCM be effective in schools with particularly high levels of behavioural disturbance, such as Alternative Provision for excluded children?
8. Does the addition of individual coaching for teachers during and after the course increase the effectiveness and cost-effectiveness of TCM?
9. Does the level of behavioural challenge that the school is faced with moderate the impact of the course and the effect of adding coaching or parent training?

There are some methodological issues that could also be explored, such as the extent to which blind observations correlate with questionnaire measures and whether or not this differs by informant. The former are expensive and time-consuming and require considerable training and reliability checks to yield reliable data. Implementation studies could address the extent to which schools that appear to be struggling because they lack a substantive headteacher or were judged as failing by Ofsted were able to implement the research

and the intervention. It would be important to identify barriers to and facilitators of including such schools to inform supportive strategies. Both could inform the design of future trials.

There is also scope for some interesting secondary analyses of the existing data, which include the longer-term follow-up of educational outcomes via linkage of the data to the NPD. In addition, we plan to undertake a mixed-methods 'following the thread' analysis to explore the relationship between the strategies adopted by teachers from the TCM training, teacher–pupil relationships and child mental health using both qualitative and quantitative data in order to refine the intervention theory to better inform future research.

Conclusions

The findings of the STARS trial provide strong evidence that TCM is feasible and acceptable in the UK context, and, in those schools in which observations were carried out, led to changes in teachers' behaviour and improved compliance from children in blind observations. There is some evidence that being taught by a teacher who has attended the course may reduce low-level disruption in primary school classrooms and be associated with improved attention and concentration over 30 months. There may also be differential effects according to baseline mental health in terms of the influence of TCM on mental health and academic attainment, but there was no evidence that TCM improved children's happiness at school or teacher mental health, burnout or self-efficacy.

Although there is insufficient evidence to recommend widespread implementation, these findings suggest that further evaluation is warranted. Future research should explore the impact of training more school staff in TCM and focus on the impact on teacher–pupil relationships and academic attainment in addition to mental health, behaviour and concentration. TCM might generate the most impact if employed as a whole-school approach, with or without the provision of additional individual coaching or parental support. These are questions that should be tested empirically. There are also obvious implementation issues to work through, such as the practicalities for schools of simultaneously releasing multiple teachers for training, and the need to have adequately trained and supervised group leaders to ensure fidelity. One solution would be to provide TCM as part of initial teacher training or within the NQT programme.

Acknowledgements

We are very grateful to Carolyn Webster-Stratton and to IY for their input and support in providing the initial training and subsequent supervision of our group leaders. This input ensured that we are confident that the intervention was delivered with fidelity and was of high quality. The IY team were involved in only the training and supervision of group leaders and had no input into the research design, analysis or write up.

We would like to thank Bill Wright, Kate Beard, Kristin Cain and Linda Lardner who, along with Malcolm Fletcher and Will Pritchard, delivered the TCM training. We are grateful to the members of our TSC [Paul Stallard (chairperson), David Glenny, Gail Seymour, Shirley Larkin, Tobit Emmens, Nicola Nathan, Ruth Dixon and Catherine Shotton] and DMC [Paul Ewings (chairperson), Andrew Richards and Siobhan Creanor] for their valuable advice and support during the project. We would like to thank Harriet Hunt for completing the randomisation. We would like to thank Pete Aighton for his help in the early stages of the trial. We would like to thank Elizabeth Nye, Emily Rhodes, Becky Bates and all of the students who have contributed to the study. We wish to acknowledge the Plymouth Association of Primary Heads and Devon Association of Primary Heads for their knowledge and assistance at the beginning of the trial and for assisting with recruitment. Most importantly, we are very grateful to all of the schools, teachers, parents and children for their time in taking part in this trial.

Contributions of authors

Tamsin Ford (Professor of Child and Adolescent Psychiatry, Child Health) as chief investigator was responsible for the original proposal, securing funding, writing the statistical analysis plan and the initial draft of this manuscript, and had overall responsibility for the management of the trial.

Rachel Hayes (Research Fellow, Child Health) was responsible for writing the statistical analysis plan, data collection and cleaning, management of the trial and writing the initial draft of this manuscript.

Sarah Byford (Professor of Health Economics) was responsible for the original proposal, securing funding, writing the statistical analysis plan, the management of the trial, co-ordinating the health economics analysis and writing the initial draft of this manuscript.

Vanessa Edwards (Senior Research Fellow, Child Health) was responsible for the original proposal, securing funding, data collection and the management of the trial.

Malcolm Fletcher (IY Group Leader, Behavioural Consultant) was responsible for the original proposal, securing funding, the management of the trial and delivering the TCM training.

Stuart Logan (Cerebra Professor of Paediatric Epidemiology, Child Health) was responsible for the original proposal, securing funding and the management of the trial.

Brahm Norwich (Professor of Educational Psychology and Special Educational Needs) was responsible for the original proposal, securing funding and the management of the trial.

Will Pritchard (IY Group Leader, Behavioural Consultant) was responsible for the original proposal, securing funding, the management of the trial and delivering the TCM training.

Kate Allen (Associate Research Fellow, Child Health) was responsible for data collection and cleaning, and for writing the initial draft of this manuscript.

Matthew Allwood (Associate Research Fellow, Child Health) was responsible for data collection and cleaning.

Poushali Ganguli (Research Associate, Health Economics) was responsible for co-ordinating the health economics analysis and for writing the initial draft of this manuscript.

Katie Grimes (Associate Research Fellow, Child Health) was responsible for data collection.

Lorraine Hansford (Associate Research Fellow, Child Health) was responsible for data collection and cleaning and for writing the initial draft of this manuscript.

Bryony Longdon (Associate Research Fellow, Child Health) was responsible for data collection and cleaning and for writing the initial draft of this manuscript.

Shelley Norman (Associate Research Fellow, Child Health) was responsible for data collection and cleaning.

Anna Price (Associate Research Fellow, Child Health) was responsible for data collection and cleaning.

Abigail Emma Russell (Research Fellow, Child Health) was responsible for conducting the T-POT analysis and for writing the initial draft of this manuscript.

Obioha C Ukoumunne (Associate Professor in Medical Statistics) was responsible for the original proposal, securing funding, writing the statistical analysis plan, the management of the trial, conducting the statistical analysis and writing the initial draft of this manuscript.

All authors contributed to, and approved, the final manuscript.

Publications

Ford T, Edwards V, Sharkey S, Ukoumunne OC, Byford S, Norwich B, Logan S. 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* 2012;**12**:719.

Marlow R, Hansford L, Edwards V, Ukoumunne OC, Norman S, Ingarfield S, *et al.* Teaching classroom management – a potential public health intervention? *Health Education* 2015;**115**:230–248. <https://doi.org/10.1108/HE-03-2014-0030>

Hansford L, Sharkey S, Edwards V, Ukoumunne O, Byford S, Norwich B, *et al.* Understanding influences on teachers' uptake and use of behaviour management strategies within the STARS trial: process evaluation protocol for a randomised controlled trial. *BMC Public Health* 2015;**15**:119. <https://doi.org/10.1186/s12889-015-1486-y>

Marlow R, Norwich B, Ukoumunne OC, Hansford L, Sharkey S, Ford T. A comparison of teacher assessment (APP) with standardised tests in primary literacy and numeracy (WIAT-II). *Assess Educ* 2014;**21**:412–26.

Nye E, Gardner F, Hansford L, Edwards V, Hayes R, Ford T. Classroom behaviour management strategies in response to problematic behaviours of primary school children with special educational needs: views of special educational needs coordinators. *Emot Behav Difficult* 2015;**21**:1–18.

Allen K, Marlow R, Edwards V, Parker C, Rodgers L, Ukoumunne OC, *et al.* 'How I Feel About My School': The construction and validation of a measure of wellbeing at school for primary school children. *Clin Child Psychol Psych* 2017;**23**:25–41. <https://doi.org/10.1177/1359104516687612>

Price A, Allen K, Ukoumunne OC, Hayes R, Ford T. Examining the psychological and social impact of relative age in primary school children: a cross-sectional survey. *Child Care Health Dev* 2017;**43**:891–8. <https://doi.org/10.1111/cch.12479>

Ford T, Hayes R, Byford S, Edwards V, Fletcher M, Logan S, *et al.* 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 [published online ahead of print July 18 2018]. *Psychol Med* 2018.

Allwood M, Allen K, Price A, Hayes R, Edwards V, Ball S, *et al.* The reliability and validity of the pupil behaviour questionnaire: a child classroom behaviour assessment tool. *Emot Behav Difficult* 2018;**23**:361–71.

Data-sharing statement

The corresponding author had full access to all of the data in the study and final responsibility for the decision to submit for publication. Parents were informed that the research data would be accessed only by the research team and, therefore, supporting data cannot be made openly available. All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review.

References

1. Jenkins A, Ueno A. Classroom disciplinary climate in secondary schools in England: what is the real picture? *Br Educ Res J* 2017;**43**:124–50. <https://doi.org/10.1002/berj.3255>
2. Ford T, Macdiarmid F, Russell AE, Racey D, Goodman R. The predictors of persistent DSM-IV disorders in 3-year follow-ups of the British Child and Adolescent Mental Health Surveys 1999 and 2004. *Psychol Med* 2017;**47**:1126–37. <https://doi.org/10.1017/S0033291716003214>
3. Ford T, Goodman R, Meltzer H. The British Child and Adolescent Mental Health Survey 1999: the prevalence of DSM-IV disorders. *J Am Acad Child Adolesc Psychiatry* 2003;**42**:1203–11. <https://doi.org/10.1097/00004583-200310000-00011>
4. Scott S, Spender Q, Doolan M, Jacobs B, Aspland H. Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice. *BMJ* 2001;**323**:194–8. <https://doi.org/10.1136/bmj.323.7306.194>
5. Achenbach TM, Ivanova MY, Rescorla LA, Turner LV, Althoff RR. Internalizing/externalizing problems: review and recommendations for clinical and research applications. *J Am Acad Child Adolesc Psychiatry* 2016;**55**:647–56. <https://doi.org/10.1016/j.jaac.2016.05.012>
6. Leijten P, Gardner F, Landau S, Harris V, Mann J, Hutchings J, *et al*. Research Review: harnessing the power of individual participant data in a meta-analysis of the benefits and harms of the Incredible Years parenting program. *J Child Psychol Psychiatry* 2018;**59**:99–109. <https://doi.org/10.1111/jcpp.12781>
7. Abikoff H. ADHD psychosocial treatments: generalization reconsidered. *J Atten Disord* 2009;**13**:207–10. <https://doi.org/10.1177/1087054709333385>
8. Franklin C, Kim JS, Beretvas TS, Zhang A, Guz S, Park S, *et al*. The effectiveness of psychosocial interventions delivered by teachers in schools: a systematic review and meta-analysis. *Clin Child Fam Psychol Rev* 2017;**20**:333–50. <https://doi.org/10.1007/s10567-017-0235-4>
9. Collie RJ, Shapka JD, Perry NE. School climate and social–emotional learning: predicting teacher stress, job satisfaction, and teaching efficacy. *J Educ Psychol* 2012;**104**:1189. <https://doi.org/10.1037/a0029356>
10. Beltman S, Mansfield C, Price A. Thriving not just surviving: a review of research on teacher resilience. *Educ Res Rev* 2011;**6**:185–207. <https://doi.org/10.1016/j.edurev.2011.09.001>
11. Brouwers A, Tomic W. A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teach Teach Educ* 2000;**16**:239–53. [https://doi.org/10.1016/S0742-051x\(99\)00057-8](https://doi.org/10.1016/S0742-051x(99)00057-8)
12. Kokkinos CM. Job stressors, personality and burnout in primary school teachers. *Br J Educ Psychol* 2007;**77**:229–43. <https://doi.org/10.1348/000709905X90344>
13. Merrett F, Wheldall K. How do teachers learn to manage classroom behaviour? A study of teachers' opinions about their initial training with special reference to classroom behaviour management. *Educ Stud* 1993;**19**:91–106. <https://doi.org/10.1080/0305569930190106>
14. Greenberg MT, Abenavoli R. Universal interventions: fully exploring their impacts and potential to produce population-level impacts. *J Res Educ Effect* 2017;**10**:40–67. <https://doi.org/10.1080/19345747.2016.1246632>
15. Patalay P, Fink E, Fonagy P, Deighton J. Unpacking the associations between heterogeneous externalising symptom development and academic attainment in middle childhood. *Eur Child Adolesc Psych* 2016;**25**:493–500. <https://doi.org/10.1007/s00787-015-0758-5>

16. Deighton J, Humphrey N, Belsky J, Boehnke J, Vostanis P, Patalay P. Longitudinal pathways between mental health difficulties and academic performance during middle childhood and early adolescence. *Br J Dev Psychol* 2018;**36**:110–26. <https://doi.org/10.1111/bjdp.12218>
17. Panayiotou M, Humphrey N. Mental health difficulties and academic attainment: evidence for gender-specific developmental cascades in middle childhood. *Dev Psychopathol* 2018;**30**:523–38. <https://doi.org/10.1017/S095457941700102X>
18. Korpershoek H, Harms T, Boer H, Kuijk M, Doolaard S. A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Rev Educ Res* 2016;**86**:643–80. <https://doi.org/10.3102/0034654315626799>
19. Webster-Stratton C. *The Incredible Years Series*. Washington, DC: US Department of Justice; 2000.
20. Barrish HH, Saunders M, Wolf MM. Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. *J Appl Behav Anal* 1969;**2**:119–24. <https://doi.org/10.1901/jaba.1969.2-119>
21. Whear R, Thompson-Coon J, Boddy K, Ford T, Racey D, Stein K. The effect of teacher-led interventions on social and emotional behaviour in primary school children: a systematic review. *Br Educ Res J* 2013;**39**:383–420. <https://doi.org/10.1080/01411926.2011.650680>
22. Webster-Stratton C, Reid MJ, Hammond M. Preventing conduct problems, promoting social competence: a parent and teacher training partnership in head start. *J Clin Child Psychol* 2001;**30**:283–302. https://doi.org/10.1207/S15374424JCCP3003_2
23. Webster-Stratton C, Reinke WM, Herman KC, Newcomer LL. The incredible years teacher classroom management training: the methods and principles that support fidelity and training delivery. *Sch Psychol Rev* 2011;**40**:509–29.
24. Patterson GR. *Coercive Family Process*. Eugene, OR: Castalia; 1982.
25. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 1977;**84**:191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
26. Piaget J, Inhelder B. *The Psychology of the Child*. New York, NY: Basic Books; 1962.
27. Beck AT. *Cognitive Therapy and Emotional Disorders*. New York, NY: Meridian; 1976.
28. Pidano AE, Allen AR. The incredible years series: a review of the independent research base. *J Child Fam Stud* 2015;**24**:1898–916. <https://doi.org/10.1007/s10826-014-9991-7>
29. Nye E. *Classroom Behaviour Management to Support Children's Social, Emotional, and Behavioural Development*. Oxford: University of Oxford; 2017.
30. Webster-Stratton C, Jamila Reid M, Stoolmiller M. Preventing conduct problems and improving school readiness: evaluation of the Incredible Years Teacher and Child Training Programs in high-risk schools. *J Child Psychol Psych* 2008;**49**:471–88. <https://doi.org/10.1111/j.1469-7610.2007.01861.x>
31. Webster-Stratton C, Reid MJ, Hammond M. Treating children with early-onset conduct problems: intervention outcomes for parent, child, and teacher training. *J Clin Child Adolesc Psychol* 2004;**33**:105–24. https://doi.org/10.1207/S15374424JCCP3301_11
32. Carlson JS, Tiret HB, Bender SL, Benson L. The influence of group training in the Incredible Years Teacher Classroom Management Program on preschool teachers' classroom management strategies. *J App School Psychol* 2011;**27**:134–54. <https://doi.org/10.1080/15377903.2011.565277>

33. Davenport J, Tansey A. *Outcomes of an Incredible Years Classroom Management Programme with Teachers from Multiple Schools*. Dublin: Trinity College Dublin/National Educational Psychological Service; 2009.
34. Fergusson DM, Horwood LJ, Stanley L. A preliminary evaluation of the Incredible Years teacher programme. *NZ J Psychol* 2013;**42**:51–6.
35. Kirkhaug B, Drugli MB, Handegård BH, Lydersen S, Åsheim M, Fossum S. Does the Incredible Years Teacher Classroom Management Training programme have positive effects for young children exhibiting severe externalizing problems in school?: a quasi-experimental pre-post study. *BMC Psychiatry* 2016;**16**:362. <https://doi.org/10.1186/s12888-016-1077-1>
36. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psych* 1997;**38**:581–6. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
37. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psych* 2001;**40**:1337–45. <https://doi.org/10.1097/00004583-200111000-00015>
38. Hutchings J, Martin-Forbes P, Daley D, Williams ME. A randomized controlled trial of the impact of a teacher classroom management program on the classroom behavior of children with and without behavior problems. *J Sch Psychol* 2013;**51**:571–85. <https://doi.org/10.1016/j.jsp.2013.08.001>
39. Hickey G, McGilloway S, Hyland L, Leckey Y, Kelly P, Bywater T, *et al*. Exploring the effects of a universal classroom management training programme on teacher and child behaviour: a group randomised controlled trial and cost analysis. *J Early Child Res* 2017;**15**:174–94. <https://doi.org/10.1177/1476718X15579747>
40. Reinke WM, Herman KC, Dong N. The Incredible Years Teacher Classroom Management Program: outcomes from a group randomized trial. *Prev Sci* 2018;**19**:1043–54. <https://doi.org/10.1007/s11221-018-0932-3>
41. Murray DW, Rabiner DL, Kuhn L, Pan Y, Sabet RF. Investigating teacher and student effects of the Incredible Years Classroom Management Program in early elementary school. *J Sch Psychol* 2018;**67**:119–33. <https://doi.org/10.1016/j.jsp.2017.10.004>
42. Martin, Daley D, Hutchings J, Jones K, Eames C, Whitaker CJ. The Teacher-Pupil Observation Tool (T-POT): development and testing of a new classroom observation measure. *Sch Psychol Int* 2010;**31**:229–49. <https://doi.org/10.1177/0143034310362040>
43. Pianta RC, La Paro KM, Hamre BK. *Classroom Assessment Scoring System™: Manual K-3*. Baltimore, MA: Brookes Publishing Co.; 2008.
44. Koth CW, Bradshaw CP, Leaf PJ. Teacher observation of classroom adaptation—checklist: development and factor structure. *Meas Eval Couns Dev* 2009;**42**:15–30. <https://doi.org/10.1177/0748175609333560>
45. Gifford-Smith M. *Teacher Social Competence Scale, Fast Track Project Technical Report*. Durham, NC: Duke University; 2000.
46. Eldridge S, Kerry S, Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? *BMJ* 2009;**339**:b4006. <https://doi.org/10.1136/bmj.b4006>
47. Department for Education. *Statistical First Release: Schools, Pupils, and their Characteristics, January 2012*. London: Department for Education; 2012.
48. Raab GM, Butcher I. Balance in cluster randomized trials. *Stat Med* 2001;**20**:351–65. [https://doi.org/10.1002/1097-0258\(20010215\)20:3<351::AID-SIM797>3.0.CO;2-C](https://doi.org/10.1002/1097-0258(20010215)20:3<351::AID-SIM797>3.0.CO;2-C)
49. Goodman A, Goodman R. Population mean scores predict child mental disorder rates: validating SDQ prevalence estimators in Britain. *J Child Psychol Psych* 2011;**52**:100–8. <https://doi.org/10.1111/j.1469-7610.2010.02278.x>

50. Gray J, Richer J. *Classroom Responses to Disruptive Behaviour*. London: Macmillan Education; 1988.
51. Allwood M, Allen K, Price A, Hayes R, Edwards V, Ball S, et al. The validity and reliability of the pupil behaviour questionnaire: a child classroom behaviour assessment tool. *Emot Behav Difficult* 2018;**23**:361–71. <https://doi.org/10.1080/13632752.2018.1478945>
52. Allen K, Marlow R, Edwards V, Parker C, Rodgers L, Ukoumunne OC, et al. 'How I Feel About My School': the construction and validation of a measure of wellbeing at school for primary school children. *Clin Child Psychol Psych* 2018;**23**:25–41. <https://doi.org/10.1177/1359104516687612>
53. Barrett B, Byford S, Chitsabesan P, Kenning C. Mental health provision for young offenders: service use and cost. *Br J Psych* 2006;**188**:541–6. <https://doi.org/10.1192/bjp.bp.105.010108>
54. Byford S, Harrington R, Torgerson D, Kerfoot M, Dyer E, Harrington V, et al. Cost-effectiveness analysis of a home-based social work intervention for children and adolescents who have deliberately poisoned themselves. Results of a randomised controlled trial. *Br J Psych* 1999;**174**:56–62. <https://doi.org/10.1192/bjp.174.1.56>
55. Harrington R, Peters S, Green J, Byford S, Woods J, McGowan R. Randomised comparison of the effectiveness and costs of community and hospital based mental health services for children with behavioural disorders. *BMJ* 2000;**321**:1047–50. <https://doi.org/10.1136/bmj.321.7268.1047>
56. Smith T, Noble M, Noble S, Wright G, McLennan, Plunkett E. *The English Indices of Deprivation 2015: Research Report*. London: Department for Communities and Local Government; 2015.
57. Ford T, Edwards V, Sharkey S, Ukoumunne OC, Byford S, Norwich B, Logan S. 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* 2012;**12**:719. <https://doi.org/10.1186/1471-2458-12-719>
58. Beckett C, Beecham J, Doolan M, Ford T, Kallitsoglou A, Scott S, et al. *Which Type of Parenting Programme Best Improves Child Behaviour and Reading? The Helping Children Achieve Trial DFE-RR261*. London: Department for Education; 2012.
59. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research*: London: Arnold; 2010.
60. Sayal K, Hornsey H, Warren S, MacDiarmid F, Taylor E. Identification of children at risk of Attention Deficit/Hyperactivity Disorder: a school-based intervention. *Soc Psychiatry Psychiatr Epidemiol* 2006;**41**:806–13. <https://doi.org/10.1007/s00127-006-0100-0>
61. SDQ. *Normative SDQ Data from Britain*. URL: www.sdqinfo.org/norms/UKNorms.html (accessed 18 February 2019).
62. Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger P. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Surv Methodol* 2001;**27**:85–95.
63. Dunn G, Bentall R. Modelling treatment-effect heterogeneity in randomized controlled trials of complex interventions (psychological treatments). *Stat Med* 2007;**26**:4719–45. <https://doi.org/10.1002/sim.2891>
64. Great Britain. *Data Protection Act 1998*. London: The Stationery Office; 1998.
65. Wille N, Badia X, Bonsel G, Burström K, Cavrini G, Devlin N, et al. Development of the EQ-5D-Y: a child-friendly version of the EQ-5D. *Qual Life Res* 2010;**19**:875–86. <https://doi.org/10.1007/s11136-010-9648-y>

66. Phillips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al.* Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8**(36). <https://doi.org/10.3310/hta8360>
67. Byford S, Barrett B, Roberts C, Clark A, Edwards V, Smethurst N, Gowers SG. Economic evaluation of a randomised controlled trial for anorexia nervosa in adolescents. *Br J Psych* 2007;**191**:436–40. <https://doi.org/10.1192/bjp.bp.107.036806>
68. Fonagy P, Butler S, Cottrell D, Scott S, Pilling S, Eisler I, *et al.* Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): a pragmatic, randomised controlled, superiority trial. *Lancet Psych* 2018;**5**:119–33. [https://doi.org/10.1016/S2215-0366\(18\)30001-4](https://doi.org/10.1016/S2215-0366(18)30001-4)
69. Department of Health and Social Care. *NHS Reference Costs 2014–15*. London: Department of Health and Social Care; 2015.
70. Curtis L. *Unit Costs of Health and Social Care*. Canterbury: Personal Social Services Research Unit, University of Kent; 2015.
71. Royal Pharmaceutical Society of Great Britain. *British National Formulary*. BMJ Group and RPS Publishing; 2016. URL: www.medicinescomplete.com/mc/bnf/current/ (accessed 18 February 2019).
72. Office for National Statistics (ONS). *Annual Survey of Hours and Earnings: 2015*. Newport: ONS; 2015. URL: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworking/hours/adhocs/004906annualsurveyofhoursandearnings (accessed 18 February 2019).
73. National Institute for Health and Clinical Excellence (NICE). *Guide to the Methods of Technology Appraisal 2013*. London: NICE; 2013. URL: www.nice.org.uk/process/pmg9/chapter/foreword (accessed 18 February 2019).
74. Netten A, Knight J, Dennett J, Cooley R, Slight A. *A 'Ready Reckoner' for Staff Costs in the NHS*. Personal Social Services Research Unit, University of Kent Canterbury, Kent; 1998.
75. Barmby P. Improving teacher recruitment and retention: the importance of workload and pupil behaviour. *Educ Res* 2006;**48**:247–65. <https://doi.org/10.1080/00131880600732314>
76. Thompson HL, Reville M-C, Price A, Reynolds L, Rodgers L, Ford T. The Quality of Life Scale for Children (QoL-C). *J Child Services* 2014;**9**:4–17. <https://doi.org/10.1108/JCS-05-2013-0019>
77. Stevens KJ. Working with children to develop dimensions for a preference-based, generic, pediatric, health-related quality-of-life measure. *Qual Health Res* 2010;**20**:340–51. <https://doi.org/10.1177/1049732309358328>
78. Furber G, Segal L, Leach M, Cocks J. Mapping scores from the Strengths and Difficulties Questionnaire (SDQ) to preference-based utility values. *Qual Life Res* 2014;**23**:403–11. <https://doi.org/10.1007/s11136-013-0494-6>
79. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;**14**:487–96. <https://doi.org/10.1002/hec.944>
80. Efron B, Tibshirani R. *An Introduction to the Bootstrap*. New York, NY: Chapman Hall; 1993. <https://doi.org/10.1007/978-1-4899-4541-9>
81. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Stat Med* 2000;**19**:3219–36. [https://doi.org/10.1002/1097-0258\(20001215\)19:23<3219::AID-SIM623>3.0.CO;2-P](https://doi.org/10.1002/1097-0258(20001215)19:23<3219::AID-SIM623>3.0.CO;2-P)
82. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *BMJ* 2000;**320**:1197–200. <https://doi.org/10.1136/bmj.320.7243.1197>

83. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. *Br J Psych* 2005;**187**:106–8. <https://doi.org/10.1192/bjp.187.2.106>
84. Weichle T, Hynes DM, Durazo-Arvizu R, Tarlov E, Zhang Q. Impact of alternative approaches to assess outlying and influential observations on health care costs. *Springerplus* 2013;**2**:614. <https://doi.org/10.1186/2193-1801-2-614>
85. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;**18**:341–64. [https://doi.org/10.1016/S0167-6296\(98\)00039-3](https://doi.org/10.1016/S0167-6296(98)00039-3)
86. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence (NICE). *Lancet* 2002;**360**:711–15. [https://doi.org/10.1016/S0140-6736\(02\)09832-X](https://doi.org/10.1016/S0140-6736(02)09832-X)
87. Kwon J, Kim SW, Ungar WJ, Tsiplova K, Madan J, Petrou S. A systematic review and meta-analysis of childhood health utilities. *Med Decis Mak* 2017;**38**:277–305. <https://doi.org/10.1177/0272989x17732990>
88. Round J, Hawton A. Statistical alchemy: conceptual validity and mapping to generate health state utility values. *Pharmacoecon Open* 2017;**1**:233–9. <https://doi.org/10.1007/s41669-017-0027-2>
89. McCabe C, Edlin R, Meads D, Brown C, Kharroubi S. Constructing indirect utility models: some observations on the principles and practice of mapping to obtain health state utilities. *Pharmacoeconomics* 2013;**31**:635–41. <https://doi.org/10.1007/s40273-013-0071-4>
90. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, Gray A. The MAPS reporting statement for studies mapping onto generic preference-based outcome measures: explanation and elaboration. *Pharmacoeconomics* 2015;**33**:993–1011. <https://doi.org/10.1007/s40273-015-0312-9>
91. Petrou S, Rivero-Arias O, Dakin H, Longworth L, Oppe M, Froud R, Gray A. Preferred reporting items for studies mapping onto preference-based outcome measures: the MAPS statement. *Int J Technol Assess Health Care* 2015;**31**:230–5. <https://doi.org/10.1017/S0266462315000379>
92. Dakin H, Abel L, Burns R, Yang Y. Review and critical appraisal of studies mapping from quality of life or clinical measures to EQ-5D: an online database and application of the MAPS statement. *Health Qual Life Outcomes* 2018;**16**:31. <https://doi.org/10.1186/s12955-018-0857-3>
93. Brooks R. EuroQol: the current state of play. *Health Policy* 1996;**37**:53–72. [https://doi.org/10.1016/0168-8510\(96\)00822-6](https://doi.org/10.1016/0168-8510(96)00822-6)
94. Marlow R, Hansford L, Edwards V, Ukoumunne O, Norman S, Ingarfield S, et al. Teaching classroom management – a potential public health intervention? *Health Educ* 2015;**115**:230–48. <https://doi.org/10.1108/HE-03-2014-0030>
95. Martin PA. *From Small Acorns: The Positive Impact of Adopting Simple Teacher Classroom Management Strategies on Global Classroom Behaviour and Teacher–Pupil Relationships*. PhD thesis. Bangor: University of Bangor; 2009.
96. Robinson EA, Eyberg SM. The dyadic parent–child interaction coding system: standardization and validation. *J Consult Clin Psychol* 1981;**49**:245–50. <https://doi.org/10.1037/0022-006X.49.2.245>
97. Tapp J, Wehby J, Ellis D. A multiple option observation system for experimental studies: MOOSES. *Behavior Res Method Inst Comp* 1995;**27**:25–31. <https://doi.org/10.3758/BF03203616>
98. Webster-Stratton C, Reid MJ. Strengthening social and emotional competence in young children – the foundation for early school readiness and success: incredible years classroom social skills and problem-solving curriculum. *Infants Young Child* 2004;**17**:96–113. <https://doi.org/10.1097/00001163-200404000-00002>

99. Hutchings J, Daley D, Jones K, Martin P, Bywater T, Gwyn R. Early results from developing and researching the Webster-Stratton Incredible Years Teacher Classroom Management Training Programme in North West Wales. *J Child Serv* 2007;**2**:15–26. <https://doi.org/10.1108/17466660200700023>
100. Reinke WM, Herman KC, Stormont M, Newcomer L, David K. Illustrating the multiple facets and levels of fidelity of implementation to a teacher classroom management intervention. *Adm Policy Ment Health* 2013;**40**:494–506. <https://doi.org/10.1007/s10488-013-0496-2>
101. Hansford L, Sharkey S, Edwards V, Ukoumunne O, Byford S, Norwich B, et al. Understanding influences on teachers' uptake and use of behaviour management strategies within the STARS trial: process evaluation protocol for a randomised controlled trial. *BMC Public Health* 2015;**15**:119. <https://doi.org/10.1186/s12889-015-1486-y>
102. Srivastava A, Thomson SB. Framework analysis: a qualitative methodology for applied policy research. *J Admin Govern* 2009;**4**:8.
103. Weale S. Four in 10 new teachers quit within a year. *The Guardian*, 31 March 2015. URL: www.theguardian.com/education/2015/mar/31/four-in-10-new-teachers-quit-within-a-year (accessed 18 February 2019).
104. Maslach C, Schaufeli WB, Leiter MP. Job burnout. *Annu Rev Psychol* 2001;**52**:397–422. <https://doi.org/10.1146/annurev.psych.52.1.397>
105. Greene RW, Beszterczey SK, Katzenstein T, Park K, Goring J. Are students with ADHD more stressful to teach? Patterns of teacher stress in an elementary school sample. *J Emot Behav Disord* 2002;**10**:79–89. <https://doi.org/10.1177/10634266020100020201>
106. Jennings PA, Greenberg MT. The Prosocial Classroom: teacher social and emotional competence in relation to student and classroom outcomes. *Rev Educ Res* 2009;**79**:491–525. <https://doi.org/10.3102/0034654308325693>
107. Chief Medical Officer. *Public Mental Health Priorities: Investing in the Evidence*. London: Department of Health and Social Care; 2014.
108. Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. *J Epidemiol Commun Health* 2000;**54**:502–9. <https://doi.org/10.1136/jech.54.7.502>
109. Hinz A, Zenger M, Brahler E, Spitzer S, Scheuch K, Seibt R. Effort–reward imbalance and mental health problems in 1074 German teachers, compared with those in the general population. *Stress Health* 2016;**32**:224–30. <https://doi.org/10.1002/smi.2596>
110. Kidger J, Brockman R, Tilling K, Campbell R, Ford T, Araya R, et al. Teachers' wellbeing and depressive symptoms, and associated risk factors: a large cross sectional study in English secondary schools. *J Affect Disord* 2016;**192**:76–82. <https://doi.org/10.1016/j.jad.2015.11.054>
111. Maguire M, O'Connell T. Ill-health retirement of schoolteachers in the Republic of Ireland. *Occup Med* 2007;**57**:191–3. <https://doi.org/10.1093/occmed/kqm001>
112. Brown J, Gilmour WH, Macdonald EB. Ill health retirement in Scottish teachers: process, outcomes and re-employment. *Int Arch Occup Environ Health* 2006;**79**:433–40. <https://doi.org/10.1007/s00420-005-0060-9>
113. Lang IA, Marlow R, Goodman R, Meltzer H, Ford T. Influence of problematic child–teacher relationships on future psychiatric disorder: population survey with 3-year follow-up. *Br J Psychiatry* 2013;**202**:336–41. <https://doi.org/10.1192/bjp.bp.112.120741>
114. Cadima J, Doumen S, Verschueren K, Buyse E. Child engagement in the transition to school: contributions of self-regulation, teacher–child relationships and classroom climate. *Early Childhood Res Q* 2015;**32**:1–12. <https://doi.org/10.1016/j.ecresq.2015.01.008>

115. McGilloway S, Hyland L, Mháille G, Lodge A, O'Neill D, Kelly P, et al. *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)*. Dublin: Archways; 2011.
116. McLean L, McDonald Connor C. Depressive symptoms in third-grade teachers: relations to classroom quality and student achievement. *Child Dev* 2015;**86**:945–54. <https://doi.org/10.1111/cdev.12344>
117. Kellam SG, Ling X, Merisca R, Brown CH, Ialongo N. The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Dev Psychopathol* 1998;**10**:165–85. <https://doi.org/10.1017/S0954579498001564>
118. Rothì DM, Leavey G, Best R. On the front-line: teachers as active observers of pupils' mental health. *Teach Teach Educ* 2008;**24**:1217–31. <https://doi.org/10.1016/j.tate.2007.09.011>
119. Department of Health and Social Care and Department for Education. *Transforming Children and Young People's Mental Health Provision: A Green Paper*. London: HMSO; 2017.
120. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;**345**:e5661. <https://doi.org/10.1136/bmj.e5661>
121. Schulz KF, Moher D, Altman DG. CONSORT 2010 comments. *Lancet* 2010;**376**:1222–3. [https://doi.org/10.1016/S0140-6736\(10\)61879-X](https://doi.org/10.1016/S0140-6736(10)61879-X)
122. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**:g1687. <https://doi.org/10.1136/bmj.g1687>
123. Alexander R. Neither national nor a curriculum? *Forum* 2012;**54**:369–84. <https://doi.org/10.2304/forum.2012.54.3.369>
124. House of Commons Library. *The School Curriculum in England, Briefing Paper Number 06798, 16 April 2018*. London: House of Commons Library; 2017. URL: www.parliament.uk/commons-library/intranet.parliament.uk/commons-library (accessed 18 February 2019).
125. Fugard AJB, Stapley E, Ford T, Law D, Wolpert M, York A. Analysing and reporting UK CAMHS outcomes: an application of funnel plots. *Child Adoles Mental Health* 2015;**20**:155–62. <https://doi.org/10.1111/camh.12086>
126. Oliver R, Wehby JH, Reschly D. *Teacher Classroom Management Practices: Effects on Disruptive or Aggressive Student Behaviour*. Oslo: The Campbell Collaboration; 2011. <https://doi.org/10.4073/csr.2011.4>
127. Humphrey N, Hennessey A, Ashworth E, Frearson K, Black L, Petersen K, et al. *Good Behaviour Game. Evaluation Report and Executive Summary. July 2018*. London: Education Endowment Foundation; 2018.
128. Gwernan-Jones R, Moore DA, Cooper P, Russell AE, Richardson M, Rogers M, et al. A systematic review and synthesis of qualitative research: the influence of school context on symptoms of attention deficit hyperactivity disorder. *Emot Behav Difficult* 2016;**21**:83–100. <https://doi.org/10.1080/13632752.2015.1120055>
129. Reinke WM, Stormont M, Webster-Stratton C, Newcomer LL, Herman KC. The incredible years teacher classroom management program: using coaching to support generalization to real-world classroom settings. *Psychol Sch* 2012;**49**:416–28. <https://doi.org/10.1002/Pits.21608>
130. Bond L, Butler H. The Gatehouse Project: A Multi-Level Integrated Approach to Promoting Wellbeing in Schools. In Killoran A, Kelly M, editors. *Evidence-based Public Health*. 1st edn. Oxford: Oxford University Press; 2009. <https://doi.org/10.1093/acprof:oso/9780199563623.003.016>

131. Murphy JM, Abel MR, Hoover S, Jellinek M, Fazel M. Scope, scale, and dose of the world's largest school-based mental health programs. *Harv Rev Psychiatry* 2017;**25**:218–28. <https://doi.org/10.1097/HRP.000000000000149>
132. Aasheim M, Reedtz C, Handegård B, Martinussen M, Mørch W-T. Evaluation of the Incredible Years teacher classroom management program in a regular Norwegian school setting [published online ahead of print May 29 2018]. *Scand J Educ Res* 2018. <https://doi.org/10.1080/00313831.2018.1466357>
133. Morgan M, Espey K. *Whole-School Implementation of Incredible Years: An Action Research Study: youngballymun*. 2012. URL: www.youngballymun.org/fileadmin/user_upload/pdf/Full_report_Final__web_ready_.pdf (accessed 18 February 2019).
134. Pearson M, Chilton R, Wyatt K, Abraham C, Ford T, Woods HB, Anderson R. Implementing health promotion programmes in schools: a realist systematic review of research and experience in the United Kingdom. *Implement Sci* 2015;**10**:149. <https://doi.org/10.1186/s13012-015-0338-6>
135. Ward B, Gersten R. Research into practice. A randomized evaluation of the safe and civil schools model for positive behavioral interventions and supports at elementary schools in a large urban school district. *Sch Psychol Rev* 2013;**42**:317–33.
136. Smolkowski K, Strycker L, Ward B. Scale-up of safe and civil schools' model for school-wide positive behavioral interventions and supports. *Psychol Sch* 2016;**53**:339–58. <https://doi.org/10.1002/pits.21908>
137. Nye E, Gardner F, Hansford L, Edwards V, Hayes R, Ford T. Classroom behaviour management strategies in response to problematic behaviours of primary school children with special educational needs: views of special educational needs coordinators. *Emot Behav Difficult* 2016;**21**:43–60. <https://doi.org/10.1080/13632752.2015.1120048>
138. Wilson S, Benton T, Scott E, Kendall L. *London Challenge: Survey of Pupils and Teachers 2006*. Research Report RR823. Nottingham: National Foundation for Educational Research; 2007.
139. Stergiakouli E, Martin J, Hamshere ML, Heron J, St Pourcain B, Timpson NJ, et al. Association between polygenic risk scores for attention-deficit hyperactivity disorder and educational and cognitive outcomes in the general population. *Int J Epidemiol* 2017;**46**:421–8. <https://doi.org/10.1093/ije/dyw216>
140. Schwab-Stone ME, Shaffer D, Dulcan MK, Jensen PS, Fisher P, Bird HR, et al. Criterion validity of the NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3). *J Am Acad Child Adolesc Psych* 1996;**35**:878–88. <https://doi.org/10.1097/00004583-199607000-00013>
141. Collishaw S, Goodman R, Ford T, Rabe-Hesketh S, Pickles A. How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? *J Child Psychol Psych* 2009;**50**:571–80. <https://doi.org/10.1111/j.1469-7610.2008.02026.x>
142. Posthumus JA, Raaijmakers MA, Maassen GH, van Engeland H, Matthys W. Sustained effects of incredible years as a preventive intervention in preschool children with conduct problems. *J Abnorm Child Psychol* 2012;**40**:487–500. <https://doi.org/10.1007/s10802-011-9580-9>
143. Menting AT, Orobio de Castro B, Matthys W. Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: a meta-analytic review. *Clin Psychol Rev* 2013;**33**:901–13. <https://doi.org/10.1016/j.cpr.2013.07.006>
144. Anders J, Brown C, Ehren M, Greany T, Nelson R, Heal J, et al. *Evaluation of Complex Whole-School Interventions: Methodological and Practical Considerations: A Report for the Education Endowment Foundation*. London: Education Endowment Foundation; 2017.

145. Sayal K, Washbrook E, Propper C. Childhood behavior problems and academic outcomes in adolescence: longitudinal population-based study. *J Am Acad Child Adolesc Psych* 2015;**54**:360–8.e2. <https://doi.org/10.1016/j.jaac.2015.02.007>
146. Garces E, Thomas D, Currie J. Longer-term effects of head start. *Am Econ Rev* 2002;**92**:999–1012. <https://doi.org/10.1257/00028280260344560>
147. Titheradge D, Hayes R, Longdon B, Allen K, Price A, Hansford L, *et al*. Psychological distress amongst primary school teachers: a comparison with clinical and population samples. *Public Health* 2019;**166**:53–6. <https://doi.org/10.1016/j.puhe.2018.09.022>
148. Gibson S, Dembo MH. Teacher efficacy: a construct validation. *J Educ Psychol* 1984;**76**:569–82. <https://doi.org/10.1037/0022-0663.76.4.569>
149. Department of Education and Science and the Welsh Office. *Discipline in Schools. Report of the Committee of Enquiry Chaired by Lord Elton*. London: HMSO; 1989.
150. Kitzinger J. Qualitative research: introducing focus groups. *BMJ* 1995;**311**:299–302. <https://doi.org/10.1136/bmj.311.7000.299>
151. Stevens K. Valuation of the Child Health Utility 9D Index. *PharmacoEconomics* 2012;**30**:729–47. <https://doi.org/10.2165/11599120-000000000-00000>

Appendix 1 Topic guides used in the process evaluation

Supporting Teachers And childRen in Schools Pilot Phase – Post Course and Research Process

Topic Guide: Focus Group Interview for intervention teachers to take place June 2013 (C01)

[More detailed interview guides will be developed to assist researchers and maintain consistency during data collection, to include: introductions, equipment checks, parameters and ground rules, brief explanation of interview process, roles and timing]

Topics

1. TCM course :delivery, content and relevance
 - a. Views on how the course has been run generally
 - b. Any particularly positive/negative aspects?
 - c. Which aspects of the content of the TCM course do they think were most valuable to help manage children in the classroom? Prompts? for example:
 - Theoretical
 - Practical skills
 - Materials in course
 - Social or peer support – shared burden
 - Style of delivery
 - Mode of delivery (e.g. video clips)

Prompt to give example of how this is manifest in what they do in their practice in the classrooms
 - d. Where do they think TCM strategies fit in with their general teaching practice?
 - e. Were any of the sessions more useful than others? If so, in what way?
 - f. Views on refresher/top-ups for TCM skills: have they any suggestions about when and how to run – prompt – couple of classroom visits/whole revision days?
2. Context: School and teachers:
 - a. Do they think that the course would be useful for other teachers, and if so when and in what ways? e.g., in general training, in newly qualified teacher year, or at what point - 2 yr/5yr/10yr experienced?
 - b. Transference of TCM principles within and beyond school
 - c. Have they discussed TCM principles with colleagues? If so, how received?
 - d. Any other similar programmes in their school / in their area that they are aware of – how does TCM compare – prompt e.g. complement, supplement, clash - ask specifically about Thrive

3. The research study/process:
 - a. Arrangements for attending TCM course:
 - o Feasible within school day, ease of attending, things that would make it easier and barriers in attending
 - o What could be improved?
 - o What worked and what did not work
 - o Any suggestions as to how the STARS team could alter/improve arrangements of the TCM course itself?
 - b. Recruitment: For example – how they were asked about/get selected to go on TCM course? Any thoughts on the consent process?
 - c. Questionnaires:
 - Views in general, ease of completion, any difficulties, irrelevance to job, etc
 - Any important areas that we have not covered?
4. Any other comments on course or study?

Sum up/confidentiality reminder

Supporting Teachers And childRen in Schools

Main trial – C01 teachers who leave TCM course

Topic Guide and questions: Teachers telephone interviews

Ask if speaking with.....

- Introduce self and why ringing. *Hello. My name is and I'm a researcher with the University of Exeter Medical School, ringing as arranged to talk with you about the STARS study and the Teacher Classroom Management course. Is this time still convenient?*
- Explain purpose of the interview: *The aim of the interview is to hear your views and get your feedback on the TCM course content and how the study has been run. This will be really helpful to us in running future courses and research. You are the expert here and you'll be able to provide us with important information. So questions will be focused on TCM course and the research, but there will also be time at the end if you want to raise anything.*
- Explain what will happen during the interview – format (questions, wanting to hear their experiences and views about STARS); maximum half an hour; no right or wrong answers
- Explain strict confidentiality, have ethical approval, would like to use their feedback in future research, right not to answer/withdraw.
- Remember to Check understanding – but not too often!
- Any questions from them at this stage?
- Remind about tape recorder (might they hear anything?) and check ok?
- Also that you will probably take notes to help with interviewing eg. coming back to anything
- Some of the questions may not be relevant depending on the individual teacher's circumstance – use only the relevant ones

1. Now, lets get going on our discussion. Can we start with general feedback on the TCM course if thats ok? I'm interested in how you think the course has been run.

- a) How has it been run generally? Overall?
- b) Any particularly positive/negative aspects?
- c) Which aspects of the content of the TCM course do you think were most valuable to help manage children in the classroom? Prompts? – for example:
- a. Theoretical
 - b. Practical skills
 - c. Materials in course (note for me – US/UK)
 - d. Social or peer support – shared burden
 - e. Style of delivery
 - f. Mode of delivery (eg video clips)
- Prompt – examples?*

- d) How do you think TCM fits in with your general teaching practice? For example, Have you been able to apply TCM principles in the classroom?
How easy? Hard? Feasible? Barriers/facilitators?
Encourage them to give Examples
- e) Were any of the sessions more useful than others? If so, in what way?

2. Now can we talk about TCM in the context of your school:

- a) Do you think the course would be useful for other teachers, and if so when and in what ways? – in general training, in Newly Qualified teacher year, or at what point - 2 yr/5yr/10 yr experienced?
- b) Have you been able to discuss TCM principles with colleagues in your schools? If so
 - which professional groups?
 - Whats their views? How received?
 - Used by them?
- c) Are they aware of any similar programmes in their schools / areas?
 - a. Used?
 - b. Compare to TCM? How? Strengths/weaknesses, complement, supplement, clash?
Prompt – ask specifically about Thrive?

3. Thinking about the research study –

- a) How have the arrangements been for attending the TCM course?
 - Feasibility within the school day, ease of attending, things that would make it easier and barriers in attending
 - What could be improved?
 - What worked and what didn't work
 - Any suggestions as to how the STARS team could alter/improve arrangements for the TCM course itself
- b) How did recruitment go? How were you asked about/get selected to go on TCM course? Any thoughts on the consent process?

4. Any other comments on the course or the study?

Thanks

Sum up

Remind re confidentiality

Very helpful

Supporting Teachers And childRen in Schools **Main Phase - Impact of course one year on**

Topic Guide: Telephone Interviews for intervention Head- teachers to take place March/April (2015 C02 & 2016 C03) **following previous year's course**

Audio File No..... Date.....Time.....

- Ask if speaking with.....
- Introduce self and why ringing: *Hello. My name is and I'm a researcher with the University of Exeter Medical School, ringing as arranged to talk with you about the STARS study and the Teacher Classroom Management course. Is this time still convenient?*
- Explain: *the aim of the forthcoming telephone interview is to explore your experiences and views about a teacher from your school attending the Teacher Classroom Management (TCM) course and to discuss any impact you feel that this might have had on their teaching practice and/or on the school*
- Explain what will happen during the interview – format (questions, wanting to hear their experiences and views about STARS both positive and negative); maximum half an hour; no right or wrong answers
- Explain strict confidentiality, have ethical approval, would like to use their feedback in future research, right not to answer/withdraw.

Remember to check understanding – but not too often!

- Any questions from them at this stage?
 - Remind about tape recorder and check ok?
 - Also that you will probably take notes to help with interviewing eg. coming back to anything
-

School

1. Ok lets start: I'd like to get some **background details** about your school first if that's ok?
 - a) Can I check some details please about your type of school and locality? (urban, rural, community school, church school, academy)
 - a. Confirm -Type of school and locality– urban, rural, community school, chu,
 - b) Have there been any particular changes to your school since you signed up to the STARS study? (e.g. flagship or outstanding school, special measures, change in leadership)
2. Any general reflections on the previous year for your school and teachers since the TCM training?

Thank you. We'll move on now to talk about the Teacher Classroom Management (TCM) training course.

So reflecting over the last year, since [NAME OF TEACHER] finished their TCM training course, it would be useful to get your views and observations on a few areas:

3. Have you noticed any changes in the type and range of external support the school has received or asked for in relation to managing children's behaviour in the classroom?

If so in what way? From whom?

4. Have you observed, or had feedback about, any changes in the ways in which [TEACHERS NAME] talks about or manages children's behaviour in the classroom?

If so, what has changed? Can you give examples?

5. Would you say there have been any changes in the ways in which other members of staff in the school (e.g. other teachers or classroom support staff) talk about or manage children's behaviour in the classroom?

If so, what/who? Any examples?

6. How have teachers specifically, and the school more generally, integrated TCM principles into the classroom environment?

7. Are there any new programmes in your school / area for managing children's behaviour in the classroom?

If so, how does TCM work alongside them?

Prompt e.g. complement, supplement, clash - ask specifically about Thrive

8. Have you noticed any changes to the teacher/parent relationships in the school, specifically relating to management of children's behaviour in the classroom?

9. Has there been any other impact from the school being involved in a research project? (On teacher, school, Head)

Prompts – positive or negative impacts, access to additional training/funding, organisational change, impact on children and/or parents, relationships with external teachers/schools/organisations?

10. Any there any other comments you'd like to make or feedback you would like to give us?

Supporting Teachers And childRen in Schools

Topic guide for Group Leaders focus group

[A more detailed guide will be developed to assist researchers and maintain consistency during data collection, to include: introductions, equipment checks, parameters and ground rules, brief explanation of focus group process, roles and timing]

Course

1. Strengths / weaknesses within the TCM programme - suggestions and views? Which aspects of the content of the TCM course do you think are most valuable to help manage children in the classroom? Are there aspects of the course that you feel are not valuable?

Prompt: Theory, practical skills, materials

2. What (from TCM) are the benefits for students with SEN who have special behavioural needs and where might the training fall short?

Prompt: How do you think pupils with SEN who also exhibit problematic behaviours responded to TCM strategies? How might teachers need to adapt these management strategies when working with students with SEN who have additional behavioural needs?

3. Delivery methods – what works well and not so well?

Prompt: Style of delivery, mode of delivery (e.g. video clips)

4. Do you think any of the sessions (or parts of sessions) were more useful for teachers than others? If so, in what way?
5. Any comments about how social / peer support might impact on teachers' experience of TCM?

Prompt: Group size, peer group dynamics, similar levels of experience?

Use and Impact of TCM

6. What influences how receptive teachers are to the course?
 - Is it more useful depending on level of experience / time since qualification?
 - Do you think teacher's type of qualification makes a difference to how useful TCM is (i.e. B.Ed. or PGCE)?
 - Method of recruitment / reason for taking part?
7. Other influences on the uptake and use of TCM strategies by teachers?
8. How does teachers' practice change as a result of TCM?

9. Are you aware of any impact on
- Children
 - Parents
 - The wider school environment that teachers are working within

School context

10. How does school context influence dissemination and wider uptake within schools?
11. What do you think about the benefits of training a group of teachers from one school (to embed the approach in the whole school) vs. the benefits of training a group of teachers from different schools (being able to talk freely and openly away from their own school)?
12. What other similar programmes in schools are you aware of and how does TCM compare?

Prompt: Strengths / weaknesses, complement, supplement, clash? Thrive?

13. What do you think is the impact of OFSTED (or other observers / reviewers) on teacher's uptake and use of TCM strategies?

Future of TCM

14. Do you think top up/refresher sessions would be useful for teachers? If so in what format?
15. How would you take the work forward?
16. How useful have you found the supervision process? Any other comments about training/support/accreditation for group leaders?
17. Any other comments from the perspective of a group leader?

Impact of research

18. Any other impact from being involved in research? For themselves or their organisation? (resources, funding, training, support, links with schools / other organisations)

Sum Up

Reminder re. confidentiality

Appendix 2 Number of teachers attending focus groups/interviews

Year	FG	Number of teachers who attended	Number of telephone interviews	Number of teachers who did not attend
2012/13	1	6	2	1
2013/14	2	5	3	2
	3	7	2	0
2014/15	4	8	1	0
	5	5	6	0

Appendix 3 Participant characteristics

Participant characteristics	FG	Interview	No FG or interview	Not invited
Female, <i>n</i> (%)	24 (77)	10 (77)	3 (100)	27 (93)
Age, mean (SD)	34.6 (10.3)	30.3 (6.3)	38.7 (4.5)	31.1 (9.1)
Number in full-time employment, <i>n</i> (%)	29 (94)	13 (100)	2 (67)	28 (97)
Number teaching KS 1, <i>n</i> (%)	20 (65)	4 (31)	2 (67)	17 (59)
Number teaching KS 2, <i>n</i> (%)	11 (35)	9 (69)	1 (33)	12 (41)
SEN children in class, mean % (SD)	22.0 (13.2)	24.7 (13.5) ^a	34.5 (9.0)	18.7 (10.3) ^b
Full-time classroom support, <i>n</i> (%)	19 (61)	7 (54)	3 (100)	16 (55)
Teaching years, mean (SD)	6.8 (5.9)	6.2 (6.3)	14.7 (4.9)	4.3 (4.7)
Leadership roles, <i>n</i> (%)	3 (10)	0 (0)	1 (33)	0 (0)
Number of NQTs, <i>n</i> (%)	1 (3)	1 (8)	0 (0)	1 (3)
Number whose initial qualification was a PGCE, <i>n</i> (%)	15 (48)	4 (31)	1 (33)	16 (55)
Number whose initial qualification was other, <i>n</i> (%)	14 (45)	8 (62)	1 (33)	11 (38)

PGCE, Postgraduate Certificate in Education.

^a One teacher did not send information about the percentage of children with SEN.

^b We do not have information about the percentage of children with SEN for the nine replacement teachers.

Appendix 4 Additional examples of each theme/subtheme

Theme	Description	Additional examples
1. Influences on teachers' learning: what were the key elements of the TCM training from the teachers' perspectives		
Structure of course	Content of course real strength	<i>... kind of a bottom up approach, so you start off with the building relationships and having a predictable classroom routine and structure and using praise and incentives before you get to the kind of hardnosed stuff of sanctions and discipline</i> <i>Group leader, interview</i>
	Depth and breadth experienced by group leaders as overwhelming at times	<i>Right, just how much of this stuff do I actually need to make sure I get across? Can I afford to leave certain points out? Can I do a bit of cutting here, a bit of pasting there?</i> <i>Group leader, interview</i>
Training as a group	Spread of schools in different socioeconomic areas and year groups important	<i>There are different pockets of schools with different social issues going on. That's really useful, because my day [would] probably be quite different to A's day ... that's quite important as well because there is an understanding of what those children are coming in with could be very different depending on where you are</i> <i>Teacher, FG2</i>
		<i>In terms of whether it is better for NQTs, I think a mix works. Because downtrodden teachers have been in it a long time, they've got a lot to teach and a lot to give to young enthusiastic whippersnappers who are coming in and vice versa, it's not a 'these people have this and these people don't'</i> <i>Group leader, FG</i>
	Hearing about how others used or interpreted strategies was helpful for learning	<i>... people's interpretation of what they heard could be slightly different so they tried something that you tried but they tried it in a slightly different way, and therefore you could 'oh I didn't actually think of it like that, I thought this' and you learned from each other all the time</i> <i>Teacher, FG2</i>
	Hearing about how strategies were working for others was useful for motivation	<i>... sometimes you may have been not wanting to try the different strategies but hearing everyone else and how they got on with it it's like 'OK yeah actually we'll do this now in my classroom because it's worked for everyone else this may work for me'</i> <i>Teacher, FG2</i>
	Teachers liked sharing each other's practices and problems	<i>Something that the teachers highlight in their feedback many times is that it gives them a chance to share good practice with colleagues and also to share difficulties or worries as well in terms of classroom management in general and individual children's difficulties, this gives them an outside viewpoint, another perspective from practitioners who are in the trench at the moment</i> <i>Group leader, interview</i>
	Being able to discuss problems without fear or judgement was important	<i>... keeping yourself sane and having that opportunity to come out of the class and talk to other teachers ... I think that has been very positive</i> <i>Teacher, interview 3</i>
	Being in a group with teachers from different schools helped teachers reflect honestly on practice	<i>You generally are just talking to people in your school and so you're all in that same situation but it's all a little too close for comfort sometimes, and actually this has kind of provided the opportunity to talk to people elsewhere and actually reflect well no, this is not going so well I need to look at what I can do to improve that</i> <i>Teacher, FG1</i>

Theme	Description	Additional examples
	Talking about theory with the group and reporting back on how it went in the classroom with group was useful	<i>It's the theory into practice bit so they're coming back having taken it from the theory and they're either saying 'Do you know what I tried this and it's starting to work or it isn't' or they go 'Oh my God why did you think that was a good idea?' and it's two ends of the spectrum sometimes and they put it out there and again because there's no expert in the room and it's the colleagues people start exploring that together. So it feels safe and the safer it feels the more that people are open and honest</i> <i>Group leader, FG</i>
Group leadership	Group leaders were welcoming, friendly, open and supportive	<i>They acknowledged your worth as well . . . the way they came over was that you are good at what you do . . . that was great you know, because it bolstered you</i> <i>Teacher, FG3</i>
	Group leaders were perceived as patronising or not open to other ideas	<i>I think they were teaching 'These are the ways and these are the only ways' sometimes, just sometimes</i> <i>Teacher, FG5</i>
	Group leaders good at delivering course but also being flexible	<i>They did sort of let us lead it as well as giving us structure</i> <i>Teacher, FG2</i>
	Content too fixed/not responsive enough	<i>We could take it in our own direction and pull out things we wanted to pull out</i> <i>Teacher, FG2</i>
Delivery methods	Role-play helped with learning	<i>That's probably one negative output I would say about it is the fact that it is scripted and you can't deviate from that. As professionals I would like us to be able to have that ability to do it, we have to do it in the classroom . . . and it would have been nice if we had that in here</i> <i>Teacher, FG5</i>
		<i>I also found one thing that really hit me personally was the practicing. We talked lots about building relationships and eye contact and turning your body and things like that. We were made to practice and somebody had their back to us and it wasn't very comfortable and I went back into my classroom and actively thought to myself 'Right I need to make sure that I am looking at people' and I thought that I was looking at people a lot more than I was . . . So I'm making myself turn . . . to actually talk to the child and they are responding in a much more positive way . . . They're not tugging at me or hitting me or tapping me or repeating my name over and over again because they know that when they say my name I will turn around and listen to them . . . so that's been massive in my class</i> <i>Teacher, FG4</i>
Materials	Book useful for extra material and ideas	<i>I'd say that's one of the biggest things I've taken from it, is what I expect from them and what they expect of me. It's understanding that they have expectations too. Because you don't think that as a teacher do you?</i> <i>Teacher, FG4</i>
	Video vignettes dated and not culturally relevant but are a real strength	<i>I've referred to the book quite fair bit in the past four weeks because I had a particularly difficult child join my class after being expelled from her last three schools . . . So I flipped through the book to try and tried to read into any strategies that might work and so far the ones that I've used from the book that I've put in place have worked really well . . . I wouldn't have known about that if I hadn't read the book</i> <i>Teacher, FG5</i>
		<i>Another downside which the teachers themselves do mention is that and I think it's one of the things that often gets mentioned as a downside is all of the video clips, the vignettes. They find some of them rather dated and sometimes find it hard just to kind of take a step back and look for and identify the teaching principles that we're hoping they'll pick up . . . and some of the informality maybe of the practice in the States, they find it's a little bit of a shock to the system sometimes</i> <i>Group leader, interview</i>

Theme	Description	Additional examples
		<i>I think the video things are really strange because I think that's a real strength in terms of people seeing stuff but it's quite weird that because also people initially said that in previous groups they found that quite difficult and sometimes they talked about it being American and not being in the UK and people aren't able to make those kind of links</i> Group leader, FG
Time to reflect	Time to step back and reflect was deemed one of the most important aspects of course	<i>It's one of those things that you don't normally get to do just without the children, get to sit with another group of teachers and talk about what you do in class and hear what they do in class and think 'oh that's a better idea I might do that as well' or whatever . . . it gives you that opportunity to reflect on what you do, which you don't often in the heat of the day in the classroom, you sort of just try and get to the end of the day sometimes</i> Teacher, FG2
2. Barriers to and facilitators of influencing the use of TCM strategies in the classroom		
Facilitators	Understanding theory was important as a facilitator	<i>I think it has helped me to realise . . . the reasons why they do the things that they do and that looking a bit deeper into why they do it . . . So finding the kind of the root of the problem . . . And taking into consideration that other factors that could be involved in it</i> Teacher, interview 7
		<i>. . . really understanding where those strategies come from as opposed to just going 'Oh yeah we'll give stickers because that's nice . . . It's really going through why you're saying what you're saying and what effect that that's having so it's much more thought through</i> Teacher, FG4
	IY pyramid was a useful graphic	<i>In terms of theory, you said about the pyramid, sharing that with other members of staff across the school made a big difference because they did, they all spotted and went 'wow that's a long way to go before you get to sanctions' . . . we had realised that we had got stuck into this trap of 'we expect you to do the right thing but if you don't this is what happens' . . . but there is nowhere in the school that told the children what the right thing was or how to do it</i> Teacher, FG3
	Being able to fine tune strategies and adapt strategies was useful	<i>'It started with one thing and if it didn't seem to be quite working, it was slightly adjusting what you were doing, rather than chucking that out and bringing another thing in to deal with it, and it was then by that fine tuning that by the end, I found that I was maybe not always getting it right but almost anticipating situations before they happen to you, which obviously made things a lot easier</i> Teacher, FG2
		<i>'I really liked the message that was being given . . . it was nice to actually have it all pieced back together, if you do this first, and then if that doesn't work try this and then it was nice to have it step by step you knew where to go if it didn't work</i> Teacher, interview 11
		<i>It's nice to have a change and have a bank of those in mind if something doesn't work or becomes stale it's good to be able to have other things up your sleeve</i> Teacher, interview 10
	Not all strategies were suitable for everyone and this was OK	<i>I can't do puppets, I've tried, I really have, I find it really uncomfortable, I just can't do it. And that's not a problem with the course, we've all realised it's OK to go 'No I just can't do that' . . . Actually it just becomes a joke because I look ridiculous, I feel ridiculous whereas other people said 'I really enjoyed that' but other things I've said I've done people have gone 'Oh I could never do that', but it's understanding that's OK</i> Teacher, FG4

Theme	Description	Additional examples
Barriers	More in-depth advice about managing most challenging behaviour/those with specific behavioural needs needed	<p><i>That impulsive child who is very likely to, as one of my children did last week, stick a pencil through my interactive whiteboard because she wanted me to switch it off, so we could do something else, so it's those people that would be good to do. I think we have touched on it</i></p> <p style="text-align: right;"><i>Teacher, FG2</i></p> <p><i>It's kind of like how does STARS [the course] affect like individual specifics like ADHD or whatever. Or is it just a general if you do this with your class, most people will be on board but there will be the odd one or two who regardless of whatever you do will always be . . .</i></p> <p style="text-align: right;"><i>Teacher, FG1</i></p>
	Lack of challenging behaviour in class so not able to use techniques	<p><i>I think some of the issues she had maybe were that behaviour isn't as much of a problem in our school so she wasn't able to use some of the techniques and strategies because actually apart from a couple of children the behaviour was relatively good</i></p> <p style="text-align: right;"><i>Headteacher, interview 12</i></p>
	Commentary not appropriate for older children	<p><i>The only thing I haven't really used is the like the commentary . . . I think as they're a little bit older it's not, I don't feel it was as effective . . . I've used it once and it wasn't quite, I don't know . . . I don't think it quite worked as well as it would if they were younger children</i></p> <p style="text-align: right;"><i>Teacher, interview 7</i></p>
	Timing of course not ideal	<p><i>I had to redo my rules after like in January when actually we should be setting the rules in September but it was just a different way of doing it . . . it just would've been nice if while we had the new class at the start of the process going through</i></p> <p style="text-align: right;"><i>Teacher, FG1</i></p>
	Need for whole-school approach	<p><i>It's about the philosophy, it's an ethos and that's why I think as a whole school it's important. Although everyone has different teaching styles and brings different strengths to that job, this . . . is kind of an umbrella over all, that can not necessarily dictate how we behave, but structure or help inform how we might behave with the children</i></p> <p style="text-align: right;"><i>Teacher, FG1</i></p>
	Lack of space and time	<p><i>Space I think for me and my particular classroom because it's a tiny Edwardian classroom, can be a problem . . . I also sometimes think with the number of children you have, there's 30 in a class, it's quite different, difficult to do the is it the conferencing, or the talking to the children about how they're working together. . . That sometimes is finding time, I think that's one of the thing. Space and time have been difficult. Practical things</i></p> <p style="text-align: right;"><i>Teacher, interview 1</i></p>
	Difficult to implement strategies when they go against school behaviour plans/policies	<p><i>Some of the challenges that we've had are that teachers go back and reward systems aren't in place in school and won't be allowed, we don't use stickers or we don't have wheels of fortune because that's not what we do. So those are the challenges</i></p> <p style="text-align: right;"><i>Group leader, FG</i></p>
	Difficult to maintain motivation to do things	<p><i>. . . because they've got energy and they need to run around and their behaviour will just be worse if that doesn't happen so you really need to think through what's the reason for that? What is the reason for keeping them in at play? And it is reminding yourself because you do slip back into it. Because it's hard when everybody in your school does that . . . that's what I found really difficult especially when it's been 3 or 4 weeks since I've been here and that is a common thing in schools is miss playtime. And I found that quite difficult to keep the momentum of this up</i></p> <p style="text-align: right;"><i>Teacher, FG4</i></p>

Theme	Description	Additional examples
	Working with school to overcome barriers	<i>Our school already used the rewards system it was just refreshing that rewards system with the children I think is good. One thing . . . was they lose a privilege instantly, we don't follow that system we use our golden time system . . . But I feel what we've learnt from this study is actually more beneficial and actually I've been having conversations with SLT [senior leadership team] to see if we can change that to do the agreement of our school policy and I think what you guys have been teaching is a much more effective way so I'm hoping that we'll be able to implement that</i> <i>Teacher, interview 12</i>
	Pressure of being observed	<i>Verbalise it . . . like 'I know this looks like Jake's not paying attention at the moment, that's fine, we're going to give him a few minutes, then I'm sure he'll come back and join us'. So whoever's observing knows you're ignoring him for a reason . . . that you've chosen to do that</i> <i>Teacher, FG1</i> <i>We've had to try and encourage them that so long as it's part of the plan, your classroom discipline plan, you've got it documented and it's there or if it's for an individual child that you're using a strategy that might seem a little bit kind of lax in Ofsted speak, if it's part of the documented behaviour plan you've got and it's been agreed to by the SENCo or the child's parents then you should be okay by that</i> <i>Group leader, interview</i>
	No time to reflect once teachers are no longer on the course	<i>I think it's been useful having the monthly sessions because I feel for me you sort of need a little time out to reflect on how things have worked and then it's quite refreshing to then go back and you've got a bit more motivation. Whereas I feel like recently because I don't know whether I sort of psychologically feel like the course has come to an end and so I haven't had that time to come and sort of offload so I feel like you have to consciously be positive, it's not something that just comes naturally, I have to consciously think 'Right I'm going to' so it takes a lot of effort . . . So I'm quite sad that it's over</i> <i>Teacher, FG4</i>

3. What do teachers feel has changed in their approach? What impact has it had on others?

Change of mindset	Course has changed teachers' mindsets/ethos	<i>I think the course itself has changed us</i> <i>Teacher, FG5</i>
A child's eye view	Reminded that children are the core of their job	<i>I mean obviously as a teacher you do have to say sometimes 'This is not acceptable' . . . but it's tapping into the child's emotions, what they're going through and what their experiences are, where they're coming from, why those behaviours are happening and I think this course has been so useful in that in how to deal with certain children and to think 'It's OK, I don't have to shout at this moment'. I think in the past I might have thought, 'That child is under the table, get out of my class', now I wouldn't even dream of doing that I would go under with them 'What's the matter?' and then if that child's in high stress it can be dealt with in a different way and I think this course has really helped me to think 'Yes there are other ways of thinking'</i> <i>Teacher, FG4</i> <i>Having to role-play . . . you really thought about what a child would be hearing . . . you might think that you are being really clear and that child should interpret whatever you said the way you said it, but in actual fact they misinterpreted it entirely. So rather than getting cross because they're not doing it as you asked them to, it might be purely because they didn't understand what you asked them to do . . . and sort of getting you to think well to yourself actually, I will try and explain it a different way or I will go about it a different way rather than just saying 'why haven't you done that yet?' . . . So you are not this angry person that controls the classroom</i> <i>Teacher, FG2</i>

Theme	Description	Additional examples
		<p><i>It's made me realise again that . . . we teach people, they are little people and they're going through difficult things just like we are and actually sometimes as a teacher it doesn't matter how their writing is coming along and how this that and the other, that they might just actually need you in another way . . . Because they can't learn can they? If they're in a poor place emotionally . . . they're not in a place to learn so it's so important'</i></p> <p style="text-align: right;"><i>Teacher, FG4</i></p>
	Changed views on individuals in the class	<p><i>She was really stuck with that particular focused child and she saw him as a really naughty boy and as a problem . . . But after she said it's really changed, 'I've really changed my view, I've realised that I wasn't liking him and that first of all I needed to like him' and so he was the focus and she was adding and building on that behaviour plan</i></p> <p style="text-align: right;"><i>Group leader, FG</i></p>
	Changing practice now think more about the child	<p><i>You have got to create a level playing field so have a homework club so that everyone gets the opportunity or else you are penalising people for something that's beyond their control. That's really worked nicely</i></p> <p style="text-align: right;"><i>Teacher, FG3</i></p>
	Being more clear with language	<p><i>. . . as opposed to 'No don't do that'. Whereas before it was 'You shouldn't be doing that, you're breaking a school rule' now it's 'You have a choice, you can continue to do that which you know is against XYZ and this will be the result or you could do this and this will be the result'. It's just language of choice</i></p> <p style="text-align: right;"><i>Headteacher, interview 14</i></p>
Building relationships	Used strategies with colleagues too and this had a positive impact	<p><i>I've been giving my colleagues lots of positive compliments so that if anything does go wrong or I need to talk about anything with them slightly more negative that it's going to be all right and they're not going to take it to heart . . . And actually they've all completely changed with me</i></p> <p style="text-align: right;"><i>Teacher, FG4</i></p>
	Benefit of developing better relationships with children	<p><i>I think personally unintentionally I was becoming a bit aloof from my children and there was a bit of a distance between us so from all that I sort of changed the way I approached it all and I spent more time trying to build positive relationships with them by play. Because I work with Year 6 classes it was a bit trickier but nevertheless I still tried to find time to play with them or do things they want or talk about things that they're interested in. I'm also giving away some of my things I get up to as well just to get them to know a bit more about me</i></p> <p style="text-align: right;"><i>Teacher, interview 09</i></p>
Thinking before responding		<p><i>You've got to think about why they're doing it and still going through the same things rather than jumping down their throats because they've done something little, actually making sure you're thinking 'Well what should I be doing?'</i></p> <p style="text-align: right;"><i>Teacher, FG1</i></p>
Positivity	Positive attitude and language has led to a happier class	<p><i>It definitely has more impact and it leads to you know a happier classroom, the kids' self-confidence is up, they are more willing to do things and try really hard because they know if they're doing what you've asked them to do they're going to get the praise, they're going to get the rewards</i></p> <p style="text-align: right;"><i>Teacher, interview 6</i></p>
	New, more positive approach noticed by colleagues	<p><i>My TA commented on after I came back and every session, she was like so 'what're you thinking of trying this time then?' you know and she's been really positive about implementing it all as well and joining in on the next wave of interventions or the next wave of new practice that we're going to try and made comment on she's seen a change in the children since</i></p> <p style="text-align: right;"><i>Teacher, FG1</i></p>

Theme	Description	Additional examples
		<p><i>When I was observed recently the teacher said 'you're very much in command, you're very relaxed and the children respond to that' and I think 'yeah, I think they do'</i></p> <p style="text-align: right;"><i>Teacher, FG1</i></p>
	Teaching children to be more positive too	<p><i>Yeah one of the things he did . . . there was a particular book about filling your buckets with positivity and positive things. And that business whereby if somebody smiles at you and is nice to you in the morning it adds to your life, but if somebody says something unpleasant to you or about you it takes from your life. And explaining to children by actually saying unkind things to other people, they're actually stealing other people's pleasure in life and actually not helping themselves</i></p> <p style="text-align: right;"><i>Headteacher, interview 14</i></p>
	Although more positive, two headteachers felt now excessively positive	<p><i>I know she's using positive praise, lots of positive praise. I think possibly too much</i></p> <p style="text-align: right;"><i>Headteacher, interview 1</i></p> <p><i>I think for example when we did some work on marking and feedback to children, he probably got a bit too much in the positive but his comments and his academic guidance let's say for the learners are very positive and very much building up their confidence and his class are very engaged in learning, there's a great attitude to learning through that. What we're now working on is, OK great to be positive but now we need to look at next steps and how to challenge children and how do you move them forward in their learning</i></p> <p style="text-align: right;"><i>Headteacher, interview 22</i></p>
Calmer more confident and in control	Teachers more confident in ability	<p><i>It's made me kind of have confidence to listen to my colleagues but also think 'I think you're wrong' . . . and before as quite a new teacher I've always thought 'Oh maybe they know because they've been teaching for 25 years so they probably know better than me' . . . and actually sometimes having the confidence to go 'I'm going to do it my way actually' and trying it</i></p> <p style="text-align: right;"><i>Teacher, FG4</i></p> <p><i>They feel generally more confident and more kind of relaxed. The problems don't go away but I think they feel more able to be with them and keep them in some kind of perspective that we've got a plan, we've got a structure, we can go back to the behaviour plan, we can go back to the functional analysis of behaviour and look at it . . .'</i></p> <p style="text-align: right;"><i>Group leader, interview</i></p> <p><i>Yeah she would make comments at the staff meetings that were a direct result of the course. She'd be bringing things back and saying things that I don't think she would have done otherwise</i></p> <p style="text-align: right;"><i>Headteacher, interview 21</i></p> <p><i>'I think that T's gained, give her her due she's such a hard worker and she's embraced it, I think that as long as it follows what a school believes in . . . With T, it's definitely built T's confidence</i></p> <p style="text-align: right;"><i>Headteacher, interview 10</i></p>
	Helping to make 'teaching life good'	<p><i>There's a system which is working and making her teaching life good</i></p> <p style="text-align: right;"><i>Headteacher, interview 9</i></p>
Creating positive cycle through role modelling		<p><i>And the other thing is just to be myself, because sometimes you feel like you should be this certain person and you've got these good things about you that you're not revealing or showing. So just being myself and just letting them see who I am, that I make mistakes, that I can have a laugh, that I'm not always on their case you know . . . I think that has really helped the class. Yeah, there were some pointers about not being so uptight and if you make a mistake, making them aware that we all make mistakes and so on one hand it's showing that you are human but also reassuring people about resilience and things like that if children make mistakes and not to let that hold them back too much. And that's really helped me</i></p> <p style="text-align: right;"><i>Teacher, FG3</i></p>

Theme	Description	Additional examples
<i>Impact on children</i>		
Changes in behaviour: children becoming ready to learn		No additional examples
Promoting independence and taking responsibility		No additional examples
Impact on children with behavioural needs	Complements other support and guidance some children are already receiving	<i>This supports those children in the classroom . . . you might have children in your class that have THRIVE so they are going out for their individual kind of emotional kind of time but then it's about how we support them in class . . . It's about how you can make their day with you kind of happy and stress free and I think that's kind of where this has been most supportive</i> <i>Teacher, FG3</i>
	Worked well for children with behavioural needs	<i>I use a lot of the modelling with him sort of on our own, we spent time doing it, but we use like the 'I can' statements and the photos of him doing things . . . making it really clear simple steps for each thing, because when he first started 'right go and put your things away' it was like a 10-minute wandering around . . . he found instructions very difficult so when we said 'oh, you can say to yourself "I can put my coat up"' . . . he goes off and hangs it up. So we have used a lot of the 'making it about him' so that he gets that the instruction is for him</i> <i>Teacher, FG3</i>
	Strategies not sufficient for certain children with behavioural needs	<i>The teacher would go to some of the training sessions and come back and say this isn't how this child can be managed, this isn't actually going to work for him</i> <i>Headteacher, interview 23</i>
	Good for general supportive processes	<i>In terms of the generic supportive processes and things like that it's cracking, but for each of those children who are presenting certain behaviours the underlying need may also need additional support from specialists teams</i> <i>Group leader, FG</i>
Raising self-esteem	Positive impact for the whole class	<i>. . . so instead of having the bottom end and the top end we have much more of an even spread</i> <i>Teacher, FG3</i>
No impact		No additional examples
<i>Impact on parents</i>		
Parent relationships	Parents perceive teachers more positively	<i>Parents have said they're thrilled to have good news rather than always having to be told about maybe not so good choices in the classroom</i> <i>Teacher, interview 3</i>
		<i>. . . so the parents actually now see me in a much more positive light . . .</i> <i>Teacher, interview 5</i>
	Parents now approaching teachers for help with behaviour management – more co-operative way of working	<i>[The boy's mother]'s like 'I'm always telling him to stop' and I'm like 'Yes but you're telling him what he shouldn't be doing so you're making him do it more . . . instead of doing that tell him the positive, what you do want to see and not what you don't want to see all the time because actually he just wants attention, he's just doing the wrong things'</i> <i>Teacher, FG1</i>
No impact		No additional examples

Theme	Description	Additional examples
4. Use of TCM strategies		
Specific and proximal praise	Found difficult at first, but really effective	<i>That's quite challenging to do that as well, 'cos you find yourself kind of almost kind of 'well done, well done, well done, well done' . . . But . . . I'm really trying to pick out the positive children who are doing the right thing, and . . . it does have a positive effect on the others. You were saying about the right ones are getting the attention, and that seems fair, but it did take a while to get used to doing that rather than reverting to what I would have before – picked up on the ones that weren't doing things. So it's kind of like changing your mindset</i> Teacher, FG2
	Added benefit of giving attention to children who are sometimes 'invisible'	<i>I found that that was praising the children for doing the right thing. And the children who you would normally have a go at because they were doing something wrong, ignoring that behaviour entirely and focusing completely on all the children who were doing the right thing, because they were getting the praise for a change rather than the others getting the attention. And I think that that was something that I do every day now</i> Teacher, FG2
Ignoring	Ignoring as a useful strategy	<i>They don't want to be ignored . . . they want you</i> Teacher, FG3 <i>The praise has been the biggest thing in the classroom as well and just ignoring the minor disruptions and silly behaviour really, you focus more on the positives</i> Teacher, FG3
	TAs also taking on ignoring strategy	<i>I hear my TA talking about ignoring muscles all the time</i> Teacher, FG2 <i>My TA's got a son at [another] school, which is like a behavioural school, and I know she's used a few strategies with him at home, I think she's used the ignore muscle because he gets quite wound up by other people so she's going to try and get him to use his ignore muscle. I think it's quite good that she'd seen it work in the classroom and then she's taken it home</i> Teacher, FG1
	Rewards and consequences	<i>I took the pictures and laminated them and that was like 'I'm a friend' or 'I can help' or I can whatever, and I had the little names and they stuck their names on the posters when they'd done something and then we sat down as a group and talked about who's done what to help . . . The kids loved it yeah, they really liked that and it was very effective</i> Teacher, FG1
Time-out	Need to put in preparatory work before using	<i>I thought I know why you have to plan it out now 'cos I went to use it and it didn't work so I was like that's sort of my own fault, I didn't give the background that it needed</i> Teacher, FG3
Emotional regulation	Modelling emotional regulation is also helpful	<i>Sometimes you say actually 'I am really frustrated' and this is what it looks like, so if you're feeling like that you can label it as frustrated, and using grown up words as well I think that's something that we were saying about. Happy or sad, happy or sad and using 'I feel exhausted', 'I feel irritated', and actually using those with the children from an early age I think has helped some of them</i> Teacher, FG3
	Other strategies also helped with emotional regulation	<i>I've got a lot more books, home school books, I've probably got about seven in my class now and I thought 'This is going to be ridiculous' but I do them throughout the day with the child so it gives me that discussion with the child on their behaviour. It helps them with their emotional intelligence, they're able to talk a bit more about themselves</i> Teacher, FG4

Theme	Description	Additional examples
Behaviour plans	Behaviour plans successfully implemented	<i>I had one child who is SEN and he's Asperger's . . . I certainly needed to do something because our relationship wasn't right and I had ideas from the SEN Coordinator but it was really kind of formalising it . . . I probably would have done stuff before but now I knew that I was going to make a plan so it's slightly different, the way you think about it. I'm going to make a plan, I'm going to be proactive and do step A, B and C</i> <i>Teacher, FG4</i>
	Useful in making approach consistent when shared with other staff working with child	<i>I spent time with her because she's new to being in the TA role and I'm trying to model as much as possible how I'm doing it and have it literally as much depth as possible written down in this behaviour plan saying 'This is what we do' and try to explain to her why</i> <i>Teacher, FG4</i>
	Repetition of things already used	<i>I've used the individual behaviour plans, they've gone into their IEP [Individual Education Plan] if they've got them so they're used and they're shared with the TAs, they all know the strategies, they're consistent</i> <i>Teacher, interview 11</i>
	Helped address issues that may not have otherwise been addressed as child may not have met criteria	<i>Yeah I mean we've got behaviour plans in school anyway so I was pretty much just copying what I had in my SEN file for him</i> <i>Teacher, FG5</i>
	Changes in SEN policy mean that behaviour plans are important	<i>It was no one [who] would ever appear on a Thrive [programme], well might but not to get any kind of intervention . . . Whereas I've put in a plan of action for this specific child . . . who is very able . . . And [in] another class might not have got anything . . . They're not SEN, they don't come under Thrive</i> <i>Teacher, FG4</i>
Coaching and commentary	Academic/social coaching difficult to feel comfortable doing	<i>Now I guess you have IEPs [Individual Education Plans], SEN children have Individual Education Plans so you would probably be doing it in that way but our school has taken a lot of those IEPs away</i> <i>Teacher, FG4</i>
	Adapting and using on self too	<i>I get emotional coaching, I think that works well here and there . . . but if you are building a town or something and I'm saying 'Oh you've picked this and you're doing that' mine just look at me a bit like 'What are you talking about yeah? Why are you just commenting on what I'm doing?' . . . We all said it goes against everything you're taught to do. You should be saying 'Do you think you could do this?' and you can't do any of that. And I know a lot of us found that not, not useful that's the wrong way of putting it, but probably reception is probably the only place I could see it working</i> <i>Teacher, FG4</i>

A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care

Published by the NIHR Journals Library