



Health Technology Assessment

Volume 28 • Issue 39 • August 2024

ISSN 2046-4924

Impact of Social Stories on social and emotional health of autism spectrum primary school children: the ASSSIST2 RCT with economic evaluation

Barry Wright, Kerry Jane Bell, Jane E Blackwell, Catarina Teige, Laura Mandefield, Han I Wang, Charlie Welch, Arabella Scantlebury, Jude Watson, Dean McMillan, Emma Standley, Leah Attwell, Hayley Carrick, Amelia Taylor, Olivia Taylor, Rachel Hodgkinson, Hannah Edwards, Hannah Pearson, Steve Parrott, David Marshall, Danielle Varley, Rebecca Hargate, Ann McLaren and Catherine Elizabeth Hewitt



Impact of Social Stories on social and emotional health of autism spectrum primary school children: the ASSSIST2 RCT with economic evaluation

Barry Wright¹, Kerry Jane Bell¹, Jane E Blackwell¹, Catarina Teige², Laura Mandefield¹, Han I Wang¹, Charlie Welch¹, Arabella Scantlebury¹, Jude Watson¹, Dean McMillan¹, Emma Standley¹, Leah Attwell², Hayley Carrick², Amelia Taylor², Olivia Taylor¹, Rachel Hodgkinson², Hannah Edwards², Hannah Pearson², Steve Parrott¹, David Marshall¹, Danielle Varley¹, Rebecca Hargate², Ann McLaren³ and Catherine Elizabeth Hewitt^{1*}

¹University of York, York, UK

²Leeds and York Partnership NHS Foundation Trust, Yorkshire, UK

³Patient and public involvement representative

*Corresponding author

Published August 2024
DOI: 10.3310/JBTM8017

This report should be referenced as follows:

Wright B, Bell KJ, Blackwell JE, Teige C, Mandefield L, Wang HI, *et al.* Impact of Social Stories on social and emotional health of autism spectrum primary school children: the ASSSIST2 RCT with economic evaluation. *Health Technol Assess* 2024;**28**(39). <https://doi.org/10.3310/JBTM8017>

Health Technology Assessment

ISSN 2046-4924 (Online)

Impact factor: 3.6

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

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 HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal

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

Reviews in *Health Technology Assessment* 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.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This article

The research reported in this issue of the journal was funded by the HTA programme as award number 16/111/91. The contractual start date was in June 2018. The draft manuscript began editorial review in June 2022 and was accepted for publication in May 2023. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care 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, the HTA 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, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Copyright © 2024 Wright *et al.* This work was produced by Wright *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).

Abstract

Impact of Social Stories on social and emotional health of autism spectrum primary school children: the ASSIST2 RCT with economic evaluation

Barry Wright¹, Kerry Jane Bell¹, Jane E Blackwell¹, Catarina Teige², Laura Mandefield¹, Han I Wang¹, Charlie Welch¹, Arabella Scantlebury¹, Jude Watson¹, Dean McMillan¹, Emma Standley¹, Leah Attwell², Hayley Carrick², Amelia Taylor², Olivia Taylor¹, Rachel Hodgkinson², Hannah Edwards², Hannah Pearson², Steve Parrott¹, David Marshall¹, Danielle Varley¹, Rebecca Hargate², Ann McLaren³ and Catherine Elizabeth Hewitt^{1*}

¹University of York, York, UK

²Leeds and York Partnership NHS Foundation Trust, Yorkshire, UK

³Patient and public involvement representative

*Corresponding author catherine.hewitt@york.ac.uk

Background: Differences in the way autistic children experience the world can contribute to anxiety and stress. Carol Gray's Social Stories™ are a highly personalised intervention to support children by providing social information about specific situations in an individual story.

Objectives: This randomised controlled trial aimed to establish whether Social Stories are clinically effective and cost-effective in improving social responsiveness and social and emotional health in children on the autism spectrum in schools.

Design: A multisite pragmatic cluster randomised controlled trial comparing Social Stories with care as usual.

Setting: Eighty-seven schools (clusters) across Yorkshire and the Humber.

Participants: Two hundred and forty-nine children were randomised via a bespoke system hosted at York Trials Unit (129 Social Stories and 120 care as usual). Recruitment was completed in May 2021. Participants were children aged 4–11 years with a diagnosis of autism, alongside teachers, interventionists and caregivers. Recruitment was via schools, NHS trusts, support groups and local publicity.

Intervention: The intervention included training for educational professionals and caregivers covering psychoeducation and implementation of Social Stories. Stories were written around contextualised goals around the child's need for social information. Interventionists read the Social Story™ with the child at least six times over 4 weeks during school.

Main outcome measure: The primary outcome was the Social Responsiveness Scale-2 completed by teachers at 6 months (the primary end point), which measures social awareness, cognition, communication and behaviour. Data were collected from caregivers and educational professionals at 6 weeks and 6 months through questionnaires. Blinding of participants was not possible.

Results: At 6 months, the estimated difference in expected teacher-reported Social Responsiveness Scale-2 T-score (the primary end point) was -1.61 (95% confidence interval -4.18 to 0.96 , $p = 0.220$), slightly favouring the intervention group. The estimated differences for the parent-reported secondary outcomes at 6 months were small and generally favoured the control group except the measure of children's quality-adjusted life-year ($+ 0.001$, 95% confidence interval -0.032 to 0.035) and parental stress (-1.49 , 95% confidence interval -5.43 to 2.46 , $p = 0.460$), which favoured the intervention group. Children in the intervention group met their individual goals more frequently than children who received usual care alone (0.97 confidence interval 0.21 to 1.73 , $p = 0.012$). The intervention is likely to save small costs ($-\pounds 191$ per child, 95% confidence interval -767.7 to 337.7) and maintain a similar quality of life compared to usual care. The probability of Social Stories being a preferred option is 75% if the society is willing to pay $\pounds 20,000$ per quality-adjusted life-year gained. Limitations include considerable disruptions during the coronavirus disease 2019 pandemic.

Conclusion: Social Stories are used in schools and represent a low-cost intervention. There is no clinically evident impact on social responsiveness, anxiety and/or depression, parental stress or general health. Benefits were observed for specific behavioural goals as assessed by the teacher, and Social Stories may serve as a useful tool for facilitating dialogue between children and school staff to address specific behavioural challenges. Usage should be at the school's discretion.

Future work: Given the uncertainty of the results in light of coronavirus disease 2019, further work to establish the impact of Social Stories is merited.

Trial registration: This trial is registered as ISRCTN11634810.

Funding: This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 16/111/91) and is published in full in *Health Technology Assessment*; Vol. 28, No. 39. See the NIHR Funding and Awards website for further award information.

Contents

List of tables	xi
List of figures	xv
List of supplementary material	xvii
List of abbreviations	xix
Plain language summary	xxi
Scientific summary	xxiii
Chapter 1 Introduction	1
Background	1
<i>Autism spectrum condition</i>	1
<i>Supporting children on the autism spectrum in primary school</i>	1
<i>Social Stories™ intervention</i>	2
Rationale for research, aims and objectives	4
<i>Rationale</i>	4
<i>Study aims</i>	4
Chapter 2 Methods	5
Main trial methods	5
<i>Trial design</i>	5
<i>Patient and public involvement</i>	5
<i>Regulatory approvals and research governance</i>	5
<i>Trial registration</i>	6
<i>Setting</i>	6
<i>Recruitment</i>	6
<i>Consenting participants</i>	7
<i>Sample size</i>	8
<i>Randomisation</i>	9
<i>Blinding</i>	9
<i>Group allocation</i>	9
<i>Participant follow-up</i>	11
<i>Outcomes</i>	11
<i>Adverse event data</i>	13
<i>Participant withdrawal</i>	13
<i>Trial completion and exit</i>	13
<i>Data management</i>	13
<i>Data analysis</i>	14
<i>Data collection schedule</i>	15
<i>Coronavirus pandemic impact analyses</i>	15
<i>Subgroup analysis</i>	15
<i>Secondary analyses</i>	16
<i>Trial oversight</i>	16
<i>Internal pilot</i>	16
Process evaluation methods	17
<i>Research questions</i>	17

CONTENTS

<i>Process evaluation methods</i>	17
<i>Qualitative data</i>	17
<i>Design</i>	19
<i>Sampling and participants</i>	19
<i>Recruitment and consent</i>	19
<i>Data collection</i>	20
<i>Process evaluation analysis</i>	20
Economic evaluation methods	21
<i>Effectiveness</i>	21
<i>Cost of the Social Stories intervention</i>	21
<i>Costs of service use</i>	21
<i>Descriptive analysis</i>	22
<i>Missing data</i>	22
<i>Primary economic analysis</i>	22
<i>Sensitivity analyses</i>	23
Summary of changes to the protocol	23
<i>Inclusion criteria</i>	23
<i>Baseline data collection</i>	23
<i>Recruitment</i>	23
<i>Follow-up procedures</i>	23
<i>Consenting procedures</i>	24
<i>Qualitative component</i>	24
<i>Randomisation procedures</i>	24
Chapter 3 Clinical effectiveness results	25
Recruitment and retention	25
<i>Recruitment</i>	25
<i>Retention</i>	26
Baseline data	27
<i>School-level baseline data</i>	27
<i>Participant-level baseline data</i>	30
<i>Parent, teacher and interventionist baseline demographics</i>	35
Intervention delivery	39
Primary outcome analyses	39
<i>Descriptive analyses</i>	39
<i>Primary analysis</i>	41
Subgroup analyses	42
<i>Teacher treatment preference</i>	42
<i>Diagnosed mental health or psychological problems</i>	43
<i>Diagnosed cognitive problems and/or learning difficulties</i>	44
<i>Complier-average causal effect estimation</i>	44
<i>Missing outcome data</i>	46
<i>Data collection timelines</i>	48
<i>Coronavirus disease impact analyses</i>	51
Secondary outcomes	53
<i>Teacher-reported outcomes</i>	53
<i>Goal-based outcome scores</i>	53
<i>Parent/carer-reported outcomes</i>	55
<i>Revised Children's Anxiety and Depression Scale total score</i>	55
<i>Parental Stress Index total stress score</i>	55
<i>European Quality of Life-5 Dimension Youth Questionnaire general health visual analogue scale</i>	56
Adverse events and safeguarding	56

Chapter 4 Economic evaluation	59
Aims	59
Availability of quality-of-life and cost data	59
Multiple imputation	59
Baseline characteristics	60
Costs	60
<i>Intervention costs</i>	61
<i>Service use and costs</i>	61
Quality of life	64
Primary analysis	64
Sensitivity analysis	66
Summary	67
Chapter 5 Process evaluation	69
Objectives	69
<i>What were the main objectives of the process evaluation?</i>	69
Defining usual practice	69
<i>What is usual care for autistic children in schools in Yorkshire and the Humber?</i>	69
Description of interventionists	70
<i>What were the profiles of the interventionists delivering Social Stories?</i>	70
Participant responsiveness	70
<i>How do participants respond to or engage with the Social Stories intervention?</i>	70
Fidelity: adherence, exposure and dose	73
<i>Did the children on the autism spectrum receive the intervention as intended?</i>	73
Suitability of Social Stories for children on the autism spectrum	75
<i>Which goals are most appropriate for Social Stories?</i>	75
Quality of delivery	76
<i>Which factors of delivering Social Stories were particularly important?</i>	76
Comprehensiveness of the intervention	80
<i>Unique features of Carol Gray Social Stories as facilitators to intervention delivery</i>	80
<i>Navigating the goal-setting process</i>	80
Training model	81
<i>Profile of trainers</i>	82
<i>The training model</i>	82
<i>Participants' views on training</i>	84
Difficulties associated with delivering a school-based intervention	87
<i>What were the challenges faced in delivering Social Stories?</i>	87
Programme differentiation	89
<i>What are the considerations to make when implementing Social Stories in the future?</i>	89
Summary	89
Chapter 6 Discussion	91
Summary of key findings	91
<i>Trial population</i>	91
<i>Primary outcome</i>	91
<i>Secondary outcomes</i>	92
<i>Cost-effectiveness</i>	92
<i>Process evaluation</i>	92
Comparison with previous studies	93
<i>Comparison with current evidence base</i>	93
Strengths and limitations	94
Equality, diversity and inclusion	96

CONTENTS

Implications for practice	97
Recommendations for research	97
Conclusion	98
Additional information	99
References	103
Appendix 1 Additional tables	107
Appendix 2 Unit cost tables	117

List of tables

TABLE 1 Original sample size assumptions and possible updates to assumptions and their implication on power	8
TABLE 2 Process evaluation methods overview	18
TABLE 3 Available eligibility data as reported in the screening CRF	25
TABLE 4 Reasons for pre-randomisation withdrawals	25
TABLE 5 School cluster sizes (as randomised) by allocation and overall	26
TABLE 6 Post-randomisation withdrawals by allocation	28
TABLE 7 Completeness of teacher- and parent/carer-completed data	29
TABLE 8 Brief baseline characteristics (stratification factors and cluster sizes) of randomised schools	30
TABLE 9 Randomisation stratification factors (participant-level summaries)	31
TABLE 10 Parent-reported demographics of randomised participants	31
TABLE 11 Baseline scores for teacher- and parent-reported outcomes and treatment preferences (range of possible scores are noted in parentheses for each outcome measure)	33
TABLE 12 Demographics of parents/carers associated with the 249 randomised participants	36
TABLE 13 Demographics of teachers associated with randomised participants at baseline	37
TABLE 14 Demographics of interventionists associated with randomised schools	38
TABLE 15 Intervention delivery by allocation	40
TABLE 16 Time between randomisation and first and last Social Stories sessions (participants that had ≥ 1 session only)	40
TABLE 17 Teacher-reported SRS-2 total T-scores by allocation and overall	41
TABLE 18 Treatment effect estimates at 6 weeks and 6 months post randomisation from the planned primary analysis model	41
TABLE 19 Subgroup summaries by allocation and overall	42
TABLE 20 Treatment effect estimates at 6 weeks and 6 months by teacher-reported baseline treatment preference	43

TABLE 21 Social Stories sessions delivered/attended by allocation (dichotomised at 6 in accordance with the definition and estimation of the CACE)	44
TABLE 22 Complier-average causal effect estimates at 6 weeks and 6 months	45
TABLE 23 Primary outcome data summaries by availability of compliance data (intervention group participants with outcome data available for at least one follow-up)	45
TABLE 24 Complier-average causal effect estimates at 6 weeks and 6 months assuming intervention group participants with missing compliance status received ≥ 6 sessions (i.e. were compliers)	46
TABLE 25 Completeness of teacher- and parent-reported outcome data by allocation	47
TABLE 26 Estimated treatment effects at 6 weeks and 6 months under various departures from MAR	49
TABLE 27 Teacher week 6 data collection timelines	49
TABLE 28 Teacher month 6 data collection timelines	50
TABLE 29 Treatment effect estimates at 6 weeks and 6 months post randomisation obtained from the primary analysis model fitted using only data collected within the specified windows at 6 weeks and 6 months	51
TABLE 30 Teacher-completed follow-up data collection timelines	51
TABLE 31 Treatment effect estimates at 6 weeks and 6 months by different sensitivity analyses	52
TABLE 32 Summary of the treatment effects of the secondary outcomes	53
TABLE 33 Details of types of behavioural goals set at baseline (participants with available goal-setting data only)	54
TABLE 34 Summary of reported serious and non-serious AEs by allocation	57
TABLE 35 Details of reported serious and non-serious AEs	57
TABLE 36 Availability of QoL and cost data (before imputation)	59
TABLE 37 Key baseline characteristics by trial arm	60
TABLE 38 Intervention costs by trial arm	61
TABLE 39 Average service use by the trial arm (complete case, $n = 112$)	62
TABLE 40 Average costs of service use in 6 months by trial arm	63
TABLE 41 Utility scores at baseline and month 6 by trial arm	64

TABLE 42 Base-case cost-effectiveness results (from the societal perspective)	65
TABLE 43 Sensitivity analyses	67
TABLE 44 Areas for consideration when using and implementing social stories in the future	90
TABLE 45 Progression criteria (Stop/Go)	107
TABLE 46 Components of intervention fidelity as described by Hasson	107
TABLE 47 Interviews conducted with parents, teachers and interventionists	108
TABLE 48 Missing primary outcome data patterns by allocation	109
TABLE 49 Key baseline characteristics by inclusion in the primary analysis	109
TABLE 50 Teacher-reported SRS-2 total raw scores by allocation	111
TABLE 51 Teacher-reported scores for the goal-based outcome measure by time point and allocation	111
TABLE 52 Treatment effect estimates from post hoc semiparametric analysis of the week 6 teacher-reported goal-based outcome scores	113
TABLE 53 Treatment effect estimates from post hoc semiparametric analysis of the month 6 teacher-reported goal-based outcome scores	113
TABLE 54 Parent/carer-reported SRS-2 total T-scores and raw scores by time point and allocation	113
TABLE 55 Parent-reported RCADS total scores by time point and allocation	115
TABLE 56 Parental stress index total stress scores by time point and allocation	115
TABLE 57 European Quality of Life-5 Dimension Youth Questionnaire general health VAS scores (parent-reported) by time point and allocation	116
TABLE 58 Treatment effect estimates at 6 weeks and 6 months for the post hoc semiparametric analysis of the EQ-5D-Y general health VAS score (parent-reported)	116
TABLE 59 The interventions received by time period	117
TABLE 60 General health community services	118
TABLE 61 Mental health-related, community-based services	118
TABLE 62 Hospital-based services	118
TABLE 63 Medications	119

LIST OF TABLES

TABLE 64 School-based health services	120
TABLE 65 Education services/support	120
TABLE 66 Private expenses	120
TABLE 67 Productivity cost	121
TABLE 68 European Quality of Life-5 Dimension Youth Questionnaire responses by trial arms by data collection time points	121

List of figures

FIGURE 1 Recruitment and teacher completed data collection	27
FIGURE 2 Recruitment and parent/carer completed data collection	28
FIGURE 3 Base-case cost-effectiveness plan of Social Stories compared with usual care (from the societal perspective)	65
FIGURE 4 Cost-effectiveness acceptability curve of Social Stories compared with usual care	66

List of supplementary material

- Report Supplementary Material 1** Information sheets
- Report Supplementary Material 2** Consent forms
- Report Supplementary Material 3** Qualitative information sheets
- Report Supplementary Material 4** Qualitative consent forms
- Report Supplementary Material 5** Qualitative topic guides
- Report Supplementary Material 6** Survey
- Report Supplementary Material 7** Additional figures

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/JBTM8017>).

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

AE	adverse event	NIHR	National Institute for Health and Care Research
ASC	autism spectrum condition		
ASSIST-2	Autism Spectrum Social Stories™ In Schools Trial 2	PCA	Prescription Cost Analysis
		PIC	participant identification centre
CACE	complier-average causal effect		
CAMHS	Child and Adolescent Mental Health Services	PIS	participant information sheet
		PPI	patient and public involvement
CEAC	cost-effectiveness acceptability curve	PSI	Parental Stress Index, Fourth Edition
CI	confidence interval	PTSD	post-traumatic stress disorder
CONSORT	Consolidated Standards of Reporting Trials	QALY	quality-adjusted life-year
		QoL	quality of life
CRF	case report form	RA	research assistant
CYP	children and young people	RCADS	Revised Children's Anxiety and Depression Scale
DMEC	Data Monitoring and Ethics Committee		
		RCT	randomised controlled trial
EOI	Expression of Interest	REC	Research Ethics Committee
EQ-5D-Y	EuroQol-5 Dimensions-Youth Questionnaire	SAE	serious adverse event
		SAP	statistical analysis plan
GP	general practitioner	SD	standard deviation
HRA	Health Research Authority	SENCO	Special Educational Needs Coordinator
HRG	Healthcare Resource Group		
HTA	Health Technology Assessment	SOP	standard operating procedure
ICC	Intracluster Correlation Coefficient	SRS-2	Social Responsiveness Scale, Second Edition
		TA	teaching assistant
ITT	intention to treat	TMG	Trial Management Group
LA	Local Authority	TSC	Trial Steering Committee
LYPFT	Leeds and York Partnership Foundation Trust	VAS	visual analogue scale
MAR	missing at random	WTP	willingness to pay
NICE	National Institute for Health and Care Excellence	YTU	York Trials Unit

Plain language summary

Autism affects the way children experience the world, and some children find social situations stressful. We wanted to know whether Social Stories™, developed by Carol Gray, helped children with their social skills and behaviour in school and whether they offered value for money.

A randomised controlled trial design was used, which gave schools an equal chance of being asked to deliver Social Stories or to continue providing care as usual.

Two hundred and forty-nine children from 87 schools took part and we trained school staff and parents to write and deliver Social Stories. We agreed with teachers and parents, what each child needed help with and wrote stories with this in mind. Trained staff read the Social Story with the child at least six times over 4 weeks. Follow-up information was collected from parents and school staff at the start of the study, after 6 weeks and 6 months.

After 6 months, teachers completed a questionnaire called the Social Responsiveness Scale-2 which measures the child's social skills.

Using these measures, the results suggest that Social Stories do not lead to any significant changes in social skills, mental health, parent stress, general health or quality of life but children in schools allocated to Social Stories met their goal more frequently and incurred less costs than children who did not. Parents and educational professionals found the Social Stories intervention and training beneficial.

Based on our findings, Social Stories do not appear to improve general social skills in primary-aged autistic children. Benefits were observed for specific goals, and school-based costs were reduced.

Scientific summary

Sections of this summary have been adapted from the trial protocol by Wright *et al.* (Wright B, Teige C, Watson J, Hodkinson R, Marshall D, Varley D, *et al.* Autism Spectrum Social Stories In Schools Trial 2 (ASSIST2): study protocol for a randomised controlled trial analysing clinical and cost effectiveness of Social Stories™ in primary schools. *BMC Psychol* 2020;8:60. <https://doi.org/10.1186/s40359-020-00427-z>). This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article unless otherwise stated.

Background

Children on the autism spectrum experience a higher prevalence of mental health problems than typically developing children, including anxiety and low mood. Many children on the autism spectrum struggle to manage social anxiety and feelings of frustration, which can lead to behaviours that challenge them. The International Society for Autism Research (INSAR) has highlighted that more research evaluating early interventions for children on the autism spectrum is needed to ensure practitioners and policy-makers have robust data on intervention effectiveness and implementation to secure optimal outcomes for children. One intervention that attempts to alleviate social difficulties while not being intrusive, time-consuming or requiring extensive involvement from experts is Carol Gray's Social Stories™. Social Stories are a highly personalised intervention aiming to share accurate, meaningful information about a particular goal or topic that the child needs help with in a positive and reassuring way. Social Stories can be written and delivered by both parents and professionals in a range of settings and represent a less time-consuming and intrusive intervention than alternatives.

Previous studies examining the use of Social Stories have yielded mostly positive results but have largely been single-case studies with a lack of evidence from randomised controlled trials (RCTs). Despite the lack of rigorous evidence, numerous schools and families are already accessing Social Stories training and delivering the intervention with children and therefore a fully powered trial is timely. Schools have limited resources and limited access to specialist practitioner interventions, and therefore it is important that interventions such as Social Stories undergo robust evaluation. If they are found to be clinically effective and cost-effective, they can be delivered within schools on a day-to-day basis.

Objectives

The aim was to assess whether Social Stories alongside care as usual is clinically effective and cost-effective in improving child social impairment, reducing anxiety and improving social and emotional health in children on the autism spectrum in primary and special educational needs (SEN) schools when compared with care as usual alone.

Primary objective

The primary objective of the study was to establish whether Social Stories can improve social responsiveness in children on the autism spectrum in primary schools across Yorkshire and the Humber, when compared to children who have received care as usual only. Social responsiveness can be broadly defined as social awareness, social cognition, social communication, social motivation and mannerisms.

Secondary objectives

The secondary objectives of this trial were:

1. to investigate whether Social Stories can reduce behaviours that challenge children on the autism spectrum in primary schools
2. to investigate whether Social Stories can improve social and emotional health in children on the autism spectrum in primary schools
3. to assess the cost effectiveness of Social Stories
4. to examine the effects of Social Stories delivered in the classroom on general measures of health-related quality of life
5. to examine whether Social Stories improve classroom attendance
6. to assess sustainability of Social Stories in an educational setting across a 6-month period
7. to examine any changes in parental stress
8. to examine any associations between treatment preference and outcomes
9. to examine how elements of session delivery (e.g. session frequency, length and any associated problems/adverse events) are associated with outcomes.

Methods

This trial was a multisite pragmatic cluster RCT comparing Social Stories and care as usual with a control group receiving care as usual alone. Care as usual is defined as the existing support routinely provided for a child with autism spectrum condition (ASC) from educational and health services such as specialist autism teacher teams, mental health teams or other associated professionals. The trial included an internal pilot, economic evaluation and a nested process evaluation.

Setting

The Social Stories intervention was primarily delivered within educational settings by educational professionals. Educational settings included both mainstream primary schools and SEN schools. Parents/caregivers were invited to receive Social Stories training and had the option of also delivering the intervention within the home.

Participants

Inclusion criteria

- The child was aged 4–11 years at the time of recruitment.
- The child attended a participating primary or SEN school within Yorkshire and the Humber.
- The child has a clinical diagnosis of ASC and daily challenging behaviour.
- Parents/guardians of the child were able to self-complete the English language outcome measures (with assistance if necessary).

Exclusion criteria

- The school had used Social Stories for any pupil in the current or preceding school term.
- The child or interventionist teacher had taken part in the previous Social Stories feasibility study (ASSIST-1). Schools that have taken part were not excluded.

Where families were recruited but the child's school was unwilling or unable to participate, families were notified by a telephone call and/or an e-mail.

Where children were confirmed eligible and consent had been taken from all relevant parties, baseline questionnaires were distributed to parents/carers and educational professionals. A school was

considered 'ready to randomise' once all consent forms and baseline questionnaires had been received by the trial team.

Interventions

Children in the intervention arm of the trial received the Social Stories intervention in addition to their care as usual. The Social Stories intervention was delivered by a trained educational professional (the interventionist) who was employed within each school allocated to the intervention arm. The interventionist varied between the schools [e.g. a teacher, teaching assistant (TA) or Special Educational Needs Coordinator (SENCO)] but was most typically a TA. A core aspect of the intervention was to first agree a goal around which the story would be set. The goal was typically a behavioural challenge the child was struggling with, for example, sharing with their teacher how they were feeling. This goal was agreed during a collaborative 'goal-setting meeting' attended by the child's teacher, a parent/caregiver, a member of the research team and sometimes the child's TA. Occasionally, the children themselves were able to feed into the goal-setting process if deemed appropriate. The intervention is designed to provide social information to the child, and so the goal sought to reflect this. In this way, many goals sought to reduce the child's anxiety or frustration by equipping them with information or providing reassurance around appropriate behaviours within a given circumstance. We assessed goal attainment in terms of how frequently a child was able to implement the desired behaviours, for example, if a goal for a particular child was to use calming strategies when they were upset, we asked teachers to rate how often a child was able to do this.

Interventionists were trained by members of the research team, who had received training via a cascade model overseen by child psychiatrist, Professor Barry Wright. The training of interventionists included some psychoeducation around the differences in children on the autism spectrum to facilitate understanding around what factors may be driving the observed behaviours associated with the behavioural goal. Training also provided key information on the design and implementation of Social Stories, with materials based on those developed in the preceding feasibility study with the support of Carol Gray and a Social Stories manual produced by Professor Barry Wright and a Clinical Psychologist with expertise in autism. During the training session, interventionists constructed a Social Story. Parents/guardians were also invited to attend these sessions. Following training, all Social Stories were assessed against a fidelity checklist by a member of the research team to ensure they conformed to the 10 established criteria central to Carol Gray's Social Stories. They were then delivered to children in the intervention arm by the interventionist at least six times over a 4-week period.

Sample size

The primary outcome was that the teacher completed the Social Responsiveness Scale, Second Edition (SRS-2) T-score at 6 months. Within the pilot data, outcomes were measured at 6 and 16 weeks. The correlation between baseline and 6 weeks [$r = 0.67$, 95% confidence interval (CI) 0.44 to 0.80] was lower than that at 16 weeks ($r = 0.83$, 95% CI 0.68 to 0.91) for the pilot data. To be conservative, the lower 95% confidence limit was chosen for the lowest correlation between baseline and follow-up that we observed within our pilot data ($r = 0.44$). Assuming a difference of 3 points, standard deviation (SD) = 7, 5% alpha, 90% power, average cluster size 1.35, intraclass correlation coefficient (ICC) = 0.34, correlation = 0.44 and 25% attrition, a total sample size of 278 was required. Support was gained from the trial steering committee and funder to stop recruitment with 249 children, which we modelled would still retain at least 80% power under a variety of different scenarios.

Randomisation

Randomisation was completed by unblinded members of the trial team via a bespoke trial management system. Stratified blocked randomisation was used to allocate school clusters, with randomly varying block sizes (4, 6 and 8), stratification by school type (SEN school or mainstream school) and the number of participating children within a school (≤ 5 or > 5 participating children). There were no stipulations regarding the minimum number of participants per school cluster (i.e. any school with ≥ 1 eligible/consented child with available baseline data was eligible for randomisation).

Main outcome measure

The primary outcome of the trial was the SRS-2 at 6 months post randomisation, as reported by the child's teacher. This was also collected at 6 weeks post randomisation. The SRS-2 identifies the presence and severity of social difficulty within the autism spectrum condition and consists of 65 questions. For each question, the person completing the form picks a score from 1 to 4 (1 = not true, 2 = sometimes true, 3 = often true, 4 = almost always true) that best describes the child's behaviour. A T-score is calculated based on the sex of the child and the person completing the form (teacher).

Economic evaluation

The primary analysis for the economic evaluation was a within-trial cost-utility analysis conducted from a societal perspective. Combining costs and quality-adjusted life-years (QALYs), an incremental cost-effectiveness ratio (ICER) of cost per QALY was calculated and evaluated against the willingness-to-pay threshold of £20,000–30,000 per QALY gained to assess the cost effectiveness of Social Stories compared to usual care. Regression models on an intention-to-treat basis were used to compare mean costs and QALYs. To take uncertainty into consideration, a non-parametric bootstrap resampling method was used to produce CIs around the cost and QALY differences and ICER. The following sensitivity analyses were conducted to test assumptions made in the primary analysis: a complete case analysis; a sensitivity analysis from the UK NHS and personal social services (PSS) perspective; and a sensitivity analysis from the joint perspectives of the NHS/PSS and education perspectives to representative a global public sector perspective.

Process evaluation

The process evaluation was cross-sectional and longitudinal, encompassing all aspects of the Social Stories intervention. The aim of the process evaluation was to assess the fidelity of the programme, consider the views of various stakeholders and identify barriers and facilitators to successful implementation. We aimed to achieve this through a combination of data collection techniques, including interviews, focus groups, questionnaires, surveys and diaries (session logs).

Statistical analyses

For the primary outcome, differences in expected SRS-2 scores were estimated using a linear mixed-effect covariance pattern model with both post-randomisation time points (6 weeks and 6 months) included as outcomes and fixed effects for treatment group, time point and their interaction. Further fixed effects were included for the following cluster and participant-level baseline covariates: school SEN status (binary, SEN/non-SEN), number of consented children attending school (binary, ≤ 5 / > 5), baseline score (linear term), age at randomisation (linear term) and sex (binary, female/male). Dependence between participants within a cluster was modelled using school cluster-level random intercepts, and dependence between repeated measurements within participants was modelled using an unstructured correlation matrix for the residual errors. Similar models were used to analyse the secondary outcomes.

Results

The primary analysis included all 249 randomised children. After 6 months, a reduction of 1.61 points was found on the SRS-2 in children on the autism spectrum in the intervention group (95% CI –4.18 to 0.96; $p = 0.220$). No statistically significant differences were found in overall symptoms of anxiety and/or depression, parental stress or general health. Children in the intervention group met their individual goals more frequently than children who received usual care alone, and this difference was statistically

significant (0.97, CI 0.21 to 1.73; $p = 0.012$). The primary analysis was conducted under the principles of intention to treat. Coronavirus disease 2019 (COVID-19) impacted upon delivery in schools and school routines, and only 62.8% of the sample is known to have received the intervention as per protocol. Sensitivity analyses suggested that there may be a dose effect with improved clinical effects in those receiving the intervention as per protocol compared to those with fewer sessions. The economic evaluation showed that, compared to usual care, Social Stories slightly decreased the service use costs over the 6-month period by £191 (95% CI -337.7 to 767.7) per child and maintained similar QALYs. If society is willing to pay £20,000 for extra QALY gained, then the probability of Social Stories being a dominant and preferred option is 75%. The results of both primary and sensitivity analyses, which considered costs derived from various perspectives, are consistent but limited to the given data within the short study time frame, number of missing data and disruption of COVID-19. The qualitative results suggest parents and educational professionals found the Social Stories training and intervention beneficial. Limitations include considerable disruptions during the COVID-19 pandemic.

Conclusion

We found no impact of Social Stories on autistic children's overall social responsiveness skills. There was some evidence that Social Stories are well supported by parents and teachers, and it may be effective at addressing a wider set of individual goals and could reduce costs. Based on the evidence generated through this trial, we cannot recommend Social Stories for the purposes of improving social skills, anxiety and/or depression, parental stress, general health and quality of life in autistic children. The COVID-19 pandemic had a large impact on schools during this trial, and sensitivity analyses suggested that poor compliance may have reduced the effectiveness of the intervention and that effects were better in those receiving Social Stories as per protocol. We did not find any negative effects, and Social Stories are already frequently used in schools to support autistic children and represent a low-cost and potentially cost-saving intervention. Despite limited evident impact on global social skills, based on the data elicited through the process evaluation, it appears that Social Stories may serve as a useful tool for facilitating dialogue between children and school staff to better understand the needs of autistic children, and usage should be at the school's discretion.

Trial registration

This trial was registered as ISRCTN11634810.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment programme (NIHR award ref: 16/111/91) and is published in full in *Health Technology Assessment*; Vol. 28, No. 39. See the NIHR Funding and Awards website for further award information.

Chapter 1 Introduction

Materials throughout this chapter have been adapted from the trial protocol by Wright *et al.*¹ This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution and reproduction in any medium provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article unless otherwise stated.

Background

Autism spectrum condition

Around 1 in 57 (1.76%) children in England are on the autism spectrum.² Autism spectrum condition (ASC) is a lifelong neurodevelopmental condition characterised by differences in social communication and interaction and restricted, repetitive patterns of behaviour and interests.³ Autism is also associated with strengths, such as some individuals having excellent attention and memory for detail and a strong drive to detect patterns.⁴ Throughout this report, we will refer to 'autistic children' or 'children on the autism spectrum' to reflect the general preference for identity-first language rather than person-first language in the autistic community⁵ and amongst our public involvement representatives.

Children on the autism spectrum experience a higher prevalence of mental health problems than typically developing children, including anxiety and low mood.⁶ Many children on the autism spectrum experience anxiety in social settings and feelings of frustration.⁷ The International Society for Autism Research (INSAR) has recently highlighted that more research evaluating early interventions for autistic children is needed to ensure practitioners and policy-makers have robust data on intervention effectiveness and implementation to secure optimal outcomes for children on the autism spectrum.⁸

Supporting children on the autism spectrum in primary school

Children on the autism spectrum often struggle socially and academically in educational settings, which may be due to differences in their experiences of social situations, differences in social communication and experiences of anxiety.⁹ The educational environment can present many challenges for a child on the autism spectrum,¹⁰ such as busy classroom environments, frequent changes in school routines and difficulties navigating peer relationships. Given these challenges, the school environment can be difficult for autistic children to follow and engage with.⁹ Autistic children can also have difficulty understanding the point of view, thoughts or feelings of someone else and have difficulties with the social use of language which can make forming and maintaining peer relationships challenging.¹¹ Children on the autism spectrum may struggle to manage their compulsions, sensory sensitivities and preoccupations within the school routine if reasonable adjustments are not made to account for these differences. Educational professionals face many demands on their time and may not be able to focus enough on the child's needs for them to achieve their full educational potential.

Little is known about the best ways to support children on the autism spectrum¹² in primary school, and more information is needed about what training is required for staff to feel confident supporting these children. Current interventions designed to support children on the autism spectrum can be time-consuming, needing involvement of outside experts. For example, specialist social skills support is delivered within Child and Adolescent Mental Health Services (CAMHS). It is important that educational professionals recognise the social challenges of children with autism,¹¹ and there is a need for low-cost, child-friendly and evidence-based interventions that can be delivered in community settings to support children on the autism spectrum with the aim of potentially reducing the need for specialist referral elsewhere. A systematic review of school-based social skills interventions for autistic children suggested that these can be effective in improving social outcomes for students on the autism spectrum.¹¹

However, further randomised controlled trials (RCTs) are needed to further understand which interventions are most clinically effective and cost-effective.¹²

Social Stories™ intervention

One intervention that attempts to alleviate these social difficulties while not being intrusive, time-consuming or requiring extensive involvement of outside experts is Carol Gray's Social Stories.⁹ Social Stories are a highly personalised intervention aiming to share accurate, meaningful information about a particular goal or topic that the child needs help with in a positive and reassuring way.⁹ Social Stories can be written and delivered by both parents and professionals in a range of settings and represent a less time-consuming and intrusive intervention than other more intensive interventions. They are typically better suited to primary-aged children (age 4–11 years),¹³ though they are also sometimes used with older children.

Social Stories are short stories which describe a social situation or skill to help children on the autism spectrum to understand a situation applicable to the child more easily. They write a child into their own personal story about themselves to help them learn new social information. They are commonly used to enable children to understand socially expected behaviours and norms. Social Stories are defined by 10 criteria which guide story development as detailed below.¹⁴

- #1 **The Social Story Goal.** Authors follow a defined process to share accurate information using content, format and voice that is descriptive, meaningful and physically, socially and emotionally safe for the Audience.
- #2 **Two-Step Discovery.** Authors gather information to (1) improve their understanding of the Audience in relation to a situation, skill or concept and (2) identify the topic and focus of each Story/Article. At least 50% of all Social Stories applaud achievements.
- #3 **Three Parts and a Title.** A Social Story/Article has a title and introduction that clearly identifies the topic, a body that adds detail and a conclusion that reinforces and summarises the information.
- #4 **FOURmat.** The Social Story format is tailored to the individual abilities, attention span, learning style and – whenever possible – talents and/or interests of its Audience.
- #5 **Five Factors Define Voice and Vocabulary.** A Social Story/Article has a patient and supportive 'voice' and vocabulary that is defined by five factors. These factors are: (1) First- or Third-Person Perspective; (2) Past, Present and/or Future Tense; (3) Positive and Patient Tone; (4) Literal Accuracy and (5) Accurate Meaning.
- #6 **Six Questions Guide Story Development.** A Social Story answers relevant 'wh' questions that describe the context, including place (WHERE), time-related information (WHEN), relevant people (WHO), important cues (WHAT), basic activities, behaviours or statements (HOW) and the reasons or rationale behind them (WHY).
- #7 **Seven is About Sentences.** A Social Story is comprised of Descriptive Sentences, as well as optional Coaching Sentences. Descriptive Sentences accurately describe relevant aspects of context, including external and internal factors, while adhering to all applicable Social Story Criteria.
- #8 **A GR-EIGHT Formula.** One Formula ensures that every Social Story describes more than directs.
- #9 **Nine to Refine.** A story draft is always reviewed and revised if necessary to ensure that it meets all defining Social Story criteria.
- #10 **Ten Guides to Implementation.** The Ten Guides to Implementation ensure that the Goal that guides Story/Article development is also evident in its use. They are: (1) Plan for Comprehension; (2) Plan Story Support; (3) Plan Story Review; (4) Plan a Positive Introduction; (5) Monitor; (6) Organise the Stories; (7) Mix and Match to Build Concepts; (8) Story Reruns and Sequels to Tie Past, Present and Future; (9) Recycle Instruction into Applause; and (10) Stay Current on Social Story Research and Updates.

Their original designer, Carol Gray, believes these criteria are the hallmark of their success. The criteria guide story development to ensure an overall patient supportive quality and relevant content that is descriptive, meaningful and safe for the audience. These stories are written in a specific way using a

variety of defined sentence types and a formula for the ratio of sentences in a social story. Using this formula and framework suggested by Carol Gray, Social Stories can be a flexible intervention that can be individualised to different social situations.

The imperative to tailor these stories is particularly important for helping a child on the autism spectrum, as every autistic child has different strengths and weaknesses.⁴ When deciding on the goal for a Social Story, it is important to consider: 'What is the challenge or problem facing the child?', 'What do we know about why this is happening from the child's perspective?', 'What does the child and young person need help with?' and 'What positive outcome would we like to see for the child?'. Individuals are asked to consider the child's individual needs, including their current known empathy/socio-emotional skills, their ability to 'get the gist' of social situations, their communication and imagination skills and how they engage with their sensory environment. Understanding the child's preoccupations, routines and how stress, anxiety or frustration emerge in various situations is also important so that the child's needs and perspectives can be considered within a Social Story. The goal that is set for a Social Story should be positive for the child and specific, measurable, achievable, realistic and time-limited (also known as a SMART goal). Social Stories can also be used to coach adults around the child about how to respond to an autistic child in certain situations and can help adults understand the child's perspective so that they are better able to support the child.

Previous case studies have suggested that following exposure to tailored Social Stories, children on the autism spectrum have shown improvements in mealtime skills,¹⁵ making independent choices and playing appropriately,¹⁶ reducing anxiety,¹⁷ supporting improved communication,¹⁸ managing and reducing frustrations¹⁸ and reducing behaviours that challenge.¹⁹⁻²¹ Three systematic reviews have examined the effectiveness of Social Stories for supporting children on the autism spectrum.²²⁻²⁴ As reported in Marshall *et al.*,¹³ these reviews indicate an overall positive effect of Social Stories on a number of social and behavioural outcomes in individual case studies. Wright *et al.*'s²⁴ systematic review included any study that used a standardised, numerical measure of outcomes but also included non-standardised, numerical measures. For single-case designs, studies had to report repeated measurements of the target behaviour to be eligible for inclusion. Outcomes explored in the reviewed literature included social abilities and awareness, communication, restricted behaviours, life skills, emotional development and sustained attention. In addition to case studies, Wright *et al.*'s²⁴ systematic review included seven between-group studies, four of which were RCTs.²⁵⁻²⁸ The interventions included in these randomised studies were often delivered over a much shorter time frame than was typically used in single-case designs, and the interventions were standard ones, rather than individually tailored to each child, meaning they did not fulfil Carol Gray's criteria.²⁴ The authors highlighted that the included literature was vulnerable to selection and reporting bias and was largely US-based.²⁴

Based on the recommendation from the systematic reviews²²⁻²⁴ that further well-designed large-scale RCTs were needed in this area, Wright *et al.*¹³ conducted a feasibility RCT to assess the feasibility of recruiting children, parents and educational professionals to a trial of Social Stories and assess retention, the appropriateness of outcome measures and any barriers to delivering training and the intervention in mainstream primary and secondary schools. This trial focused on writing and delivering individualised Social Stories that followed the Carol Gray criteria within the school setting.¹³ The results suggested that a future trial would be feasible to conduct and could inform the policy and practice of using Social Stories in primary schools. It was not recommended to include secondary schools in the full trial, as qualitative work revealed that the intervention was not viewed as appropriate for secondary-aged children due to being too simplistic and more difficult to implement. The feasibility results were in line with previous research that suggested that it is possible to train tier-one professionals, for example, teachers and teaching assistants (TAs), to develop and use Social Stories tailored to a child.²⁹ As previously described, studies examining the use of Social Stories have yielded mostly positive results but have largely been single-case studies. Despite the lack of rigorous evidence, numerous schools and families are already accessing Social Stories training and delivering the intervention with children, and therefore a fully powered RCT is timely. Schools have limited resources and limited access to specialist

practitioner interventions, and therefore it is important that interventions such as Social Stories undergo robust evaluation. If they are found to be clinically effective and cost-effective, they can be delivered within schools on a day-to-day basis.

Rationale for research, aims and objectives

Rationale

Individuals on the autism spectrum require varying levels of support from different services (such as the NHS and charitable organisations), as the condition can have a widespread and persistent impact on quality of life (QoL), relationships, employment and standards of living.³⁰ An evaluation of the economic cost of autism in the UK³⁰ revealed that the annual costs of supporting children on the autism spectrum were estimated to be £2.7 billion each year, and for adults, these costs were £25 billion each year. Given this significant cost and limited funding available for specialist support in schools³¹ and for other services, there is a need for low-cost, child-friendly and evidence-based interventions that can be delivered in community settings such as schools to support children on the autism spectrum, with the aim of potentially reducing the need for specialist referral elsewhere.

The current work aims to address this gap by exploring whether one promising intervention (Social Stories) could be effective in supporting children on the autism spectrum in primary school. This RCT follows a feasibility study,¹³ which explored the acceptability of running a trial examining Social Story use with 50 children across 37 primary and secondary mainstream schools. This demonstrated a high degree of acceptability with young people, families and schools. This main trial now seeks to assess the clinical and cost effectiveness of Social Stories, addressing the lack of fully powered RCTs in this area.

Study aims

As outlined in the study protocol,¹ the aim of this trial was to assess whether Social Stories alongside care as usual is clinically effective and cost-effective in improving child social impairment, reducing behaviours that challenge and improving social and emotional health in children on the autism spectrum in primary schools when compared with care as usual alone.

Primary objective

The primary objective of the study was to establish whether Social Stories can improve social responsiveness in children on the autism spectrum in primary schools across Yorkshire and the Humber when compared to children who have received care as usual only.

Secondary objectives

The secondary objectives of this trial were:

1. to investigate whether Social Stories can reduce behaviours that challenge children on the autism spectrum in primary schools
2. to investigate whether Social Stories can improve social and emotional health in children on the autism spectrum in primary schools
3. to assess the cost-effectiveness of Social Stories
4. to examine the effects of Social Stories delivered in the classroom on general measures of health-related QoL
5. to examine whether Social Stories improve classroom attendance
6. to assess the sustainability of Social Stories in an educational setting across a 6-month period
7. to examine any changes in parental stress
8. to examine any associations between treatment preference and outcomes
9. to examine how elements of session delivery [e.g. session frequency, length and any associated problems/adverse events (AEs)] are associated with outcomes.

Chapter 2 Methods

Materials throughout this chapter have been adapted from the trial protocol by Wright *et al.*¹ This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article unless otherwise stated.

Main trial methods

Trial design

This trial was a multisite pragmatic cluster RCT comparing Social Stories and care as usual with a control group receiving care as usual alone. Care as usual being defined as the existing support routinely provided for a child on the autism spectrum from educational and health services such as specialist autism teacher teams, mental health teams or other associated professionals. The trial included an internal pilot, an economic evaluation and a nested process evaluation. The full trial protocol is published online via the National Institute for Health and Care Research (NIHR) Journals Library (<https://fundingawards.nihr.ac.uk/award/16/111/91>).

Patient and public involvement

The development of the Autism Spectrum Social Stories In Schools Trial 2 (ASSIST-2) trial was supported by three core patient and public involvement (PPI) members, a parent of an autistic child, a specialist teacher and a member of the National Autistic Society. Two of these representatives were named co-applicants for the funding awarded from the NIHR to deliver the trial, and all were involved in the trial from its conception. In addition to these core members, a local youth council, which also worked with members of the trial team on a previous feasibility study of Social Stories,²⁴ reviewed and contributed to the design of the present trial. This included shaping information leaflets and giving advice on young people's preferences regarding the research study.

Furthermore, the study team utilised findings from the qualitative interviews completed during the previous feasibility study which included young people with autism, teachers and parents/carers.²⁴ The findings from these interviews informed the use of questionnaires, the presentation and content of materials such as information sheets and consent forms, the logicising of intervention delivery and the best means of contacting participants. These interviews, in conjunction with our PPI representatives, also informed the training manual used during this study.⁹ Advice was also provided from parents, teachers and educational psychologists regarding more general aspects of trial delivery within education settings. Both the Trial Steering Committee (TSC) and Trial Management Group (TMG) included PPI representatives who were able to contribute to discussions around the trial throughout its implementation.

A PPI event was held at the end of the trial to disseminate the key findings to representatives from key sectors, specifically educational settings, parents, charities, mental health services and the local authority. Advice and guidance were sought with regard to dissemination activities, including the appropriateness of language and terminology around autism as well as key considerations for any roll-out and future research avenues.

Regulatory approvals and research governance

Initially, the trial opened as a non-NHS trial, recruiting schools and families directly with no NHS involvement. Ethics approval for the trial was obtained from the Health Sciences Research Governance Committee embedded within the Department of Health Sciences at the University of York on 2 July 2018.

At a later stage of the research, ethical approval was sought and awarded from the Health Research Authority (HRA) on 24 July 2019 and the North East – York Research Ethics Committee (REC) on 23 July 2019 (REC reference number 19/NE/0237) to allow opening of an additional recruitment stream via the NHS. Substantial and non-substantial amendments to approve changes to the protocol and study documentation were submitted to the REC, HRA, the Department of Health Sciences Research Governance Committee and each NHS site's research and development office as required during this trial. The trial sponsor was Leeds and York Partnership NHS Foundation Trust.

Trial registration

The trial was assigned the International Standard Randomised Controlled Trial Number (ISRCTN) ISRCTN11634810 on 23 April 2019.

Setting

The Social Stories intervention was delivered within educational settings by educational professionals. Educational settings included both mainstream primary schools and Special Educational Needs (SEN) schools. Parents/caregivers were invited to receive Social Stories training and were encouraged to utilise Social Stories within the home environment, though this was not considered part of the intervention. Parents were specifically advised not to deliver the story written for the purposes of the trial at home as the trial sought to evaluate the intervention through a school-based delivery model. No data were collected on the use of Social Stories at home.

Recruitment

Recruitment was on a rolling basis through the recruitment period, with schools being randomised as and when all pre-randomisation criteria had been fulfilled, that is, all relevant parties had provided consent and all baseline data had been collected. No additional children could be recruited from a school once it had been randomised. Participants were recruited via the following recruitment pathways.

Recruitment from schools

Primary schools across Yorkshire and the Humber were contacted with information about the trial via e-mail or post. Meetings were then set up with school head teachers to establish willingness at the school level to support the trial and undertake the necessary trial activities. Where schools agreed to participate, they were asked to distribute information sheets (see [Report Supplementary Material 1](#)) and Expression of Interest (EOI) forms to parents/guardians of eligible children. Parents/guardians who returned an EOI form or consented for school staff to pass on their details were contacted by a member of the research team to discuss the trial, answer any questions and arrange collection of informed consent from parents/guardians and, where possible, assent from the child. Consent was also obtained from education professionals taking part in the trial (see [Report Supplementary Material 2](#)).

Recruitment via the National Health Service

In total, seven CAHMS teams from five NHS Trusts within the Yorkshire and Humber area were opened to act as a participant identification centre (PIC). PICs were required to screen their patient case lists to identify potential children who were likely to meet the inclusion criteria for the trial. They then distributed recruitment packs prepared by the York Trials Unit (YTU) containing information sheets about the trial and EOI forms to be returned to the trial team. A member of the trial team then contacted families via their preferred method to confirm eligibility and request contact details for the child's education setting. Recruitment then followed the procedure set out for schools, as it was necessary for the child's school to agree to take part before the family could be formally consented.

Recruitment from local community groups

Local community groups within the catchment area, for example, AWARE, a parent-run family support group, were contacted with information about the trial, and researchers distributed study information to interested parents/guardians. On receipt of an EOI, recruitment followed the procedure outlined for schools.

Liaising with local authority professionals

The trial protocol outlined that relevant local authority professionals, for example, educational specialists in autism, could be contacted and asked to disseminate study information and EOI forms to the parents/guardians of potentially eligible children. On receipt of an EOI, recruitment could then follow the procedure outlined for schools. However, this recruitment pathway was not implemented due to the success of the other channels; hence, the additional burden this would place on local authorities was deemed unnecessary.

Recruitment from local publicity

Some families heard about the trial from a range of sources, for example, publicity on Twitter (@COMICResearchUK), and contacted the research team directly. In this instance, the team confirmed the child's eligibility and sent out study information. It was made clear that participation was dependent on the child's school taking part, and recruitment continued in line with the procedure outlined for schools.

Consenting participants

Participation in the ASSSIST-2 trial was entirely voluntary. There were three types of participants involved in the trial, namely, educational professionals, parents/caregivers and the children themselves. The consent process varied slightly depending on whether participants joined the trial prior to or during the coronavirus disease 2019 (COVID-19) pandemic.

Prior to the COVID-19 pandemic, consent was typically obtained in person during face-to-face visits to the child's school by a trained research assistant (RA). Information sheets were always distributed to the relevant parties at least 24 hours prior to consent being taken. During the COVID-19 pandemic, consent forms were posted to participants and an appointment was scheduled for a RA to go through the form with the participants either over the telephone or via videoconferencing.

Educational professionals and parents/caregivers provided their own consent. Due to the age of the children, parents/caregivers provided consent on behalf of the children, though verbal assent was gained if possible. During the consent process, participants could opt in to being contacted to participate in the qualitative interviews for the process evaluation. They could also opt in to be sent details about future relevant research.

Participant eligibility

All children and schools were screened based on the following inclusion and exclusion criteria. After an EOI to participate was received, screening took place via two telephone calls led by a RA, one with a parent/carer and one with the child's teacher. Screening questionnaires were completed during screening calls.

Inclusion criteria

- The child was aged 4–11 years at the time of recruitment.
- The child attended a participating primary or SEN school within Yorkshire and the Humber.
- The child has a clinical diagnosis of ASC as confirmed by the parent/carer during a screening call and has daily challenging behaviour as confirmed during a screening call with the child's teacher.
- Parents/guardians of the child were able to self-complete the English language outcome measures (with assistance if necessary).

Exclusion criteria

- The school had used Social Stories for any pupil in the current or preceding school term.
- The child or interventionist teacher had taken part in the previous Social Stories feasibility study (ASSSIST). Schools that have taken part were not excluded.

Where families were identified to be eligible to participate but the child's school was unwilling or unable to participate, families were notified by telephone or e-mail.

Where children were confirmed eligible and consent had been taken from all relevant parties, baseline questionnaires were distributed to parents/carers and educational professionals. A school was considered 'ready to randomise' once all consent forms and baseline questionnaires had been received by the trial team.

Sample size

The primary end point was that the teacher completed the Social Responsiveness Scale, Second Edition (SRS-2) T-score at 6 months. Within the pilot data, outcomes were measured at 6 and 16 weeks. The correlation between baseline and 6 weeks [$r = 0.67$, 95% confidence interval (CI) 0.44 to 0.80] was lower than that at 16 weeks ($r = 0.83$, 95% CI 0.68 to 0.91) for the pilot data. To be conservative, the lower 95% confidence limit was chosen for the lowest correlation between baseline and follow-up that we observed within our pilot data ($r = 0.44$). Assuming a difference of 3 points, standard deviation (SD) = 7, 5% alpha, 90% power, average cluster size 1.35, ICC = 0.34, correlation = 0.44 and 25% attrition, a total sample size of 278 was required.

Given the COVID-19 pandemic and subsequent closure of schools during this period, recruitment fell behind target. Recruitment began in mid-November 2018, and at the time of the first UK lockdown on 23 March 2020, 194 children (across 67 schools) had been screened, of which 134 (51 schools) had been randomised (53.8% of the final sample). The trial status was appraised by the TMG in October 2020, and it was decided that recruitment should end on 31 May 2021 with a total of 249 participants recruited. This was informed by the data collected to date and the impact on power was modelled. Keeping all sample size assumptions the same as the original outlined above, 249 participants would reduce the power from 90% to 86%. We also knew that the original assumption of the average cluster size of 1.35 was too low and this should be increased to 2.86 (the average school size), which would lead to a further reduction in power. As we had recent data to provide more accurate estimates of correlation between baseline and 6-month SRS-2 T-scores and attrition, we updated the sample size scenarios as outlined in [Table 1](#). The estimate for correlation increased from 0.44 to 0.54, and attrition decreased from 25% to 23%. Unfortunately, we did not have enough data to get reliable estimates of the ICC; however, we had some information which may inform our choice of ICC further than that observed in the pilot trial. The ICC of the teacher SRS-2 T-scores at baseline was 0.09 (95% CI 0.00 to 0.24). Scenario 1 in [Table 1](#) shows that 249 participants would provide 89% power, assuming the reduced estimate of the ICC (0.09) and other updated estimates of correlation (0.54), attrition (23%) and average cluster size (2.87). As the updated estimate of the ICC was not based on 6-month follow-up data, it is possible that our true ICC may be greater than this. For further assurance, we looked at the highest value of ICC we could have while still providing 80% power, and this would still be achieved with an ICC of 0.3, which is over three times > 0.09 and is greater than the upper confidence limit of the baseline ICC of 0.24. Evidence suggests that an ICC of 0.2 would be plausible in an educational setting,³² which

TABLE 1 Original sample size assumptions and possible updates to assumptions and their implication on power

Scenario	Cluster size ^a	ICC	Correlation	Attrition	Delta	Power (%)	<i>n</i>
Original	1.35	0.34	0.44	25%	0.43	90	278
Updated recruitment	1.35	0.34	0.44	25%	0.43	86	249
Updated based on trial estimates on 27 May 2021							
Scenario 1	2.86	0.09	0.54	23%	0.43	89	249
Scenario 2	2.86	0.30	0.54	23%	0.43	80	249

^a Average cluster size.

gives us some confidence that a value of 0.3 would be conservative. Both updated scenarios provide at least 80% power, which gave us some assurance that ending recruitment with 249 participants would mean we had an adequate level of power even if the ICC were greater than expected. This change to the sample size was agreed by the trial oversight committees and funder.

Randomisation

Participants were entered on to a bespoke trial management system developed at YTU if they fulfilled the eligibility criteria, and their parent/caregiver and school staff provided written consent to take part. This system was used to record key identifiable data required to facilitate trial participation and data collection, as well as the various dates (e.g. baseline completion, randomisation, etc.) used to schedule and manage follow-up data collection. Randomisation was clustered at the school level; hence, all pupils within a school were allocated the same trial arm, Social Stories or usual care. Randomisation occurred after consent/assent and valid baseline data had been obtained from all participating families and educational professionals within a school. The randomisation process was completed by unblinded members of the trial team via the trial management system. Stratified blocked randomisation was used to allocate school clusters, with randomly varying block sizes (4, 6 and 8), and stratification by school type (SEN school or mainstream school) and the number of participating children within a school (≤ 5 or > 5 participating children). There were no stipulations regarding the minimum number of participants per school cluster (i.e. any school with ≥ 1 eligible/consented child with available baseline data was eligible for randomisation). Prior to COVID-19, schools were notified of their study allocation via telephone, e-mail and postal letter with parents/guardians receiving only a letter. Post COVID-19, all correspondence regarding allocation was via e-mail or telephone.

Blinding

Owing to the nature of the intervention, all participant types were unblinded to allocation. Research assistants collecting outcome data and the main trial statistician were blinded to study allocation until final data analysis. The Data Monitoring and Ethics Committee (DMEC) had access to unblinded data throughout the study. Any instances of unblinding were recorded using a bespoke form. In such instances, a substitute blinded RA collected participant data wherever possible.

Group allocation

Intervention

Children in the intervention arm of the trial received Social Stories in addition to their care and education as usual. The Social Stories intervention was delivered by a trained educational professional (the interventionist) who was employed within each school allocated to the intervention arm. The interventionist varied between the schools [e.g. a teacher, TA or Special Educational Needs Coordinator (SENCO)] but was most typically a TA. A core aspect of the intervention was to first agree a behavioural goal around which the story would be set. The goal was typically a behavioural challenge the child was struggling with, for example, sharing with their teacher how they were feeling. This goal was agreed during a collaborative 'goal-setting meeting' attended by the child's teacher, a parent/caregiver, a member of the research team and sometimes the child's TA. Occasionally, the children themselves were able to feed into the goal-setting process if deemed appropriate. The goal-setting meeting lasted around 15 minutes. Parents/carers were asked if they had a preference for the goal beforehand via e-mail or during baseline. The parents' preferred goal was discussed during the meeting. If teachers and parents/carers disagreed on the goal, ultimately the teachers' preferred goal was used given that the intervention is school-based and intended to support school-based challenges. Parents/carers were invited to attend training so that they could write their own stories at their discretion.

The intervention is designed to provide social information to the child, and so the goal sought to reflect this. In this way, many goals sought to reduce the child's anxiety or frustration by equipping them with information or providing reassurance around appropriate behaviours within a given circumstance.

METHODS

We assessed goal attainment in terms of how frequently a child was able to implement the desired behaviours, for example, if a goal for a particular child was to use calming strategies when they were upset, we asked teachers to rate how often a child was able to do this.

In this trial, stories were developed collaboratively during a training session attended by the child's interventionist (often a TA) and facilitated by a member of the trial team. Sometimes the child's class teacher was also able to attend this session. Parents/carers were invited to attend the session with the view of providing contextual information around the child's behaviour so that they could write their own stories at home if they so wished, though any such stories did not form part of the trial and no data were collected with regard to home use. Schools were provided with a resource pack containing worksheets designed to help them develop the stories. Included in this resource pack was a copy of a handbook⁹ designed during the feasibility trial.²⁴

Initially, all training was delivered in person; however, due to restrictions imposed as a consequence of COVID-19, schools randomised after 23 March 2020 were trained online. In-person training typically consisted of a 3-hour training session comprising psychoeducational training around the differences associated with autism, training around the trial procedures (e.g. completion of questionnaires and session logs) and the collaborative writing of the story.

Those who were trained online were required to complete more directed independent learning by watching an online version of the training presentation and spending some time thinking about the child/children who were participating in the trial. They then received a training session online via videoconferencing, which reiterated key components of the training presentation, provided information about trial procedures and the collaborative writing of the story. These online sessions typically took around 1 hour per story.

Stories were around eight pages in length and were required to adhere to the 10 Carol Gray criteria as outlined in [Background](#). A fidelity checklist was used to assess conformity to the criteria and was completed by one of the trainers on the research team. This checklist was developed during the feasibility trial and is documented within the Social Stories manual given to schools.

One of the Carol Gray criteria centres on making refinements to the story if needed. The intervention affords some level of flexibility, and schools were free to make changes to the story if circumstances changed, which meant content within the story was no longer accurate or relevant or if content was found to be upsetting to the child. Schools were also free to adapt the way the story was delivered, for example, if a child needed to move around, they could place pages of the book around a room so the child could move between them.

Schools were asked to read the story at least six times over a 4-week period with the child and to record each reading within a session log.

Usual care

Participants allocated to the control arm of the trial received care and education as usual. Schools in the control arm were asked to continue delivering any other support but to refrain from delivering any Social Stories for the duration of their trial involvement. Given that there is considerable variation in the level of support needed by children on the autism spectrum, we sought to define what comprises 'usual care' through the process evaluation. Teachers were asked at both baseline and 6-month follow-up what interventions, including Social Stories, children had received in the previous 6 months. They were also asked to describe any classroom-based support. This is described in [Chapter 5](#).

Participant follow-up

All educational professionals (teachers and interventionists) and parents were followed up approximately 6 weeks and 6 months post randomisation. Interventionists were also asked to return session logs documenting their intervention (story-reading) sessions with the children. Follow-up data were collected in several ways.

Educational professionals

Educational professionals were primarily followed up using postal questionnaires. This approach offered the most flexibility, with educational professionals being able to complete questionnaires at their convenience. However, some completed questionnaires over the telephone or via videoconferencing with a RA, depending on their preference. Interventionists were provided with a stamped address envelope to return session logs to the trial team at the end of their 6-week intervention period.

Parents/carers

Prior to the COVID-19-related lockdown in March 2020, all parent/carer follow-ups were completed in person, either at home or within school, by a RA. After March 2020, parent/carer data were collected through a combination of postal questionnaires and telephone/videoconferencing by RAs, depending on parent/carer preference.

In all cases where data were not received within 2 weeks of the due date, RAs made four attempts to contact parents/carers to request data. This approach included a combination of telephone and e-mail contacts. At the 6-week follow-up point, if data could not be collected after four contacts, no further attempts were made, but the parents/carers remained in the trial and were sent their 6-month questionnaires at the usual time interval. Where data could not be collected after four contacts at the 6-month time point, participants were considered lost to follow-up.

Outcomes

Primary outcome

The primary outcome of the trial was the teacher-reported SRS-2, with the primary end point being the measurement of this outcome obtained at 6 months post randomisation. This teacher was not the interventionist. This outcome was also collected at 6 weeks post randomisation. The SRS-2 identifies the presence and severity of social difficulty within the autism spectrum³³ and consists of 65 questions. For each question, the person completing the form selects a score from 1 to 4 (1 = not true, 2 = sometimes true, 3 = often true, 4 = almost always true) that best describes the child's behaviour. The responses to these 65 items are transformed by subtracting 1 from the available item scores, and 17 of the items are reverse scored so that all non-missing items have an integer score between 0 and 3, where higher scores indicate greater social difficulties. If there are strictly fewer than eight missing responses (across all 65 items), then the transformed and reversed item scores are used to generate five raw subscale scores by summing various mutually exclusive subsets of items: Social Awareness (0, 24), Social Cognition (0, 36), Social Communication (0, 66), Social Motivation (0, 33) and Restricted Interests/Repetitive Behaviour (0, 36). If there are eight or more missing items, then these scores are not calculated. These five raw subscale scores are then summed to generate the SRS-2 total raw scores (0, 195), with higher scores indicating greater social difficulties.

The available SRS-2 total raw scores are used to generate a T-score based on the sex of the child and the person completing the form (teacher). These T-scores are integers in the range (38, 90) (higher scores indicate greater difficulties) and are based on scores obtained for a nationally representative standardisation sample by the developers of the instrument. The total raw scores are mapped to T-scores so that the mean T-score in the standardisation sample is 50, with a SD of 10. The developers of the SRS-2 provide the following guidance regarding the interpretation of the total T-scores.

SRS-2 total T-score (T) range	Guideline interpretation
$T \leq 59$	Within normal limits
$60 \leq T \leq 65$	Mild range
$66 \leq T \leq 75$	Moderate range
$T \geq 76$	Severe range

Secondary outcomes

All secondary outcomes are listed below (grouped by respondent) and were collected at baseline, 6 weeks and 6 months post randomisation unless otherwise specified.

Parent questionnaires:

1 SRS-2.³³

- Scoring and interpretation essentially identical to the primary outcome (Teacher-reported SRS-2).

2 Demographic information pertaining to the child and the parent – *baseline only*.

3 Parenting Stress Index short form.³⁴

- Thirty-six Likert items with five levels of response (5 = Strongly agree, 4 = Agree, 3 = Not sure, 2 = Disagree, 1 = Strongly disagree).
- Used to derive three subscale scores [Parental Distress (12, 60), Parent-Child Dysfunctional Interaction score (12, 60) and Difficult Child score (12, 60)], no missing items permitted, higher scores indicate greater distress, dysfunctionality or difficulties.
- Subscale scores summed to obtain the Total Stress Score (36, 180) (higher scores indicating greater stress) which is mapped to a Total Stress percentile (0, 100) (higher scores indicating greater stress).
- Guidelines for interpretation of Total Stress percentile (p); $p \leq 80$ = Typical stress, $81 \leq p \leq 89$ = High stress, $p \geq 90$ = Clinically significant stress.

4 The EuroQol-5 Dimensions-Youth Questionnaire (EQ-5D-Y) (3L proxy version).³⁵

- Five items (Mobility, Self-care, Usual activities, Pain/discomfort, Anxiety/depression), with three levels of response.
- Used to derive health state and utilities for economic analyses.

5 Revised Children's Anxiety and Depression Scale (RCADS) short form.³⁶

- Forty-seven Likert items with four levels of response (0 = Never, 1 = Sometimes, 2 = Often, 3 = Always) used to derive six subscale scores [Social Phobia (0, 27), Panic Disorder (0, 27), Major depression (0, 30), Separation Anxiety (0, 21), Generalised Anxiety (0, 18), Obsessive – Compulsive (0, 18)], up to two missing items per subscale permitted (missing items pro-rated using non-missing subscale scores), higher scores indicate presence of more severe symptoms.
- Subscale scores summed to obtain the RCADS total score (0, 141), where higher scores indicate greater symptoms of anxiety and depression.

6 Bespoke resource use questionnaire, capturing healthcare and non-health resource use of participants and parents/carers – *baseline and 6 months only*.

7 Bespoke treatment preference questionnaire – *baseline only*.

- Single visual analogue scale (VAS) (0, 100), where 0 indicates strong preference for usual care/support, 100 indicates strong preference for Social Stories and 50 indicates indifference/no preference.

Associated teacher/TA questionnaires:

1. SRS-2.³³
2. A goal-based outcome measure (adapted from the Child Outcomes Research Consortium).³⁷
 - Single integer rating (0, 10) of the frequency that the child is meeting the goal set for them (0 = None of the time, 5 = Half of the time, 10 = All of the time).
3. Bespoke resource use questionnaire – *baseline and 6 months only*.
4. Bespoke treatment preference questionnaire – *baseline only*.
5. Bespoke resource use questionnaire determining current school care/education plan interventions – *baseline and 6 months only*.
6. Teacher demographics.
 - Age (years), Current professional role(s), Time spent working with children/young people (years), Self-reported knowledge and experience of working with children with autistic spectrum disorder (ASD) (Limited, Moderate, Sound, In-depth).

Interventionist Teacher/TA Questionnaires:

1. Bespoke Social Story session log – *used after each Social Story session*.
2. A bespoke sustainability questionnaire – *6 weeks and 6 months only*.
3. Interventionist demographics.

Adverse event data

Although unlikely due to the nature of the intervention, possible harm as a result of the study was monitored according to YTU standard operation procedures (SOPs). Adverse events reported by individuals participating in the study were recorded using a bespoke Adverse Events Recording Form and assessed for seriousness. It was outlined in the trial protocol¹ that any AE will be recorded as a serious adverse event (SAE) if it results in death, is life-threatening, prolongs or requires hospitalisation or results in disability or incapacity. All AEs were reported to the Chief Investigator, Professor Barry Wright, a child psychiatrist. The protocol specified that any SAEs related to the study should be reported to the study Sponsor, DMEC and TSC. The TSC and DMEC committees reviewed AE data at approximately 6-monthly intervals throughout the trial.

Participant withdrawal

Parents/carers and/or educational professionals could withdraw from the trial at any point during the course of the study. Parents/carers also had the option of fully withdrawing their child from the trial. Unless the parent/carer opted to fully withdraw the child from the trial, follow-up of the other participants associated with the child continued as usual where possible. If a participant indicated that they wanted to withdraw from the study, a member of the research team completed a change of circumstance form, and the new participant status was recorded on the trial management system.

Trial completion and exit

Participants completed the trial once they had completed the 6-month follow-up period post randomisation. At this stage, the educational professionals and parents/caregivers allocated to the usual care arm were offered Social Stories training. No further follow-up was required post training; however, support in story writing was available on demand. Participants (of any level) were considered to have exited the trial if they had withdrawn or lost to follow-up at the 6-month follow-up point. Parents/carers were offered a Love2Shop voucher to the value of £20 upon completion of the trial. These were delivered via post. Schools were offered £50 per participating child, paid in cash.

Data management

All information collected during the study was kept strictly confidential and stored on a secure password-protected server located at the University of York for the purposes of assisting in follow-ups

during the study. All paper documents were stored securely, initially at the sponsor's office (Leeds and York Partnership NHS Foundation Trust (LYPFT), throughout data collection prior to transferring the documents to the University of York, which occurred on a monthly basis.

Case Report Forms (CRFs) were initially checked for errors by the research team, and any queries were raised immediately with participants. CRFs were then logged on the YU's bespoke data management system and scanned using Cardiff Teleform. Original data sheets were securely stored at YU. All data were collected and retained in accordance with the Data Protection Act 2018, the General Data Protection Regulation 2018 and YU SOPs. All data will be archived for 10 years following the end of the study and then securely destroyed.

Data analysis

All outcomes were analysed after the trial had ended according to a pre-specified statistical analysis plan (SAP). Analyses were conducted using Stata version 17,³⁸ following the principles of intention to treat (ITT), with available outcome data analysed according to the participants' allocated group regardless of protocol deviations or non-compliance (unless otherwise stated). All statistical tests are two-sided tests of a point null hypothesis, and reported interval estimates are based on two-sided 95% CIs. The trial was designed and reported according to the Consolidated Standards of Reporting Trials (CONSORT) guidelines.³⁹ Participant withdrawals (number, type and timing) and follow-up response rates are presented overall and by allocation. Compliance data, as measured through session logs, are also summarised. All participant baseline data are summarised by trial arm and overall, with no formal statistical comparisons being undertaken. All continuous/quantitative data are reported in terms of their mean, SD, median, interquartile range, minimum and maximum, with categorical data reported in terms of frequencies and percentages.

Primary analysis

The primary end point was the teacher-reported SRS-2 total T-score at 6 months post randomisation, with an earlier measurement of this outcome being collected at 6 weeks post randomisation. The derivation and analysis of the primary outcome (at each time point) were checked by a second statistician. Differences in expected score (Intervention – Control) were estimated using a linear mixed-effect covariance pattern model with both post-randomisation time points (6 weeks and 6 months) included as outcomes and fixed effects for treatment group, time point and their interaction. Further fixed effects were included for the following cluster and participant-level baseline covariates: school SEN status (binary, SEN/non-SEN), number of consented children attending school (binary, ≤ 5 / > 5), baseline score (linear term), age at randomisation (linear term) and sex (binary, female/male). Dependence between participants within clusters was modelled using school cluster-level random intercepts (Gaussian with expected value 0 and variance estimated using the data), and dependence between repeated measurements within participants was modelled using an unstructured correlation matrix for the residual errors (distinct terms for the three residual error variance/covariance parameters). The model was fitted using restricted maximum likelihood estimation. No small sample degrees of freedom corrections were implemented when calculating the reported interval estimates or test statistics.

Missing data

We undertook analyses to assess the sensitivity of the results of the primary analysis to departures from the missing at random (MAR) assumption underpinning this analysis.⁴⁰ In particular, we used a delta-based sensitivity analysis (implemented via a pattern mixture modelling approach) to obtain point and interval estimates of the treatment effect, assuming the missing outcome data exhibit various systematic departures from MAR.^{41,42}

Adherence

We undertook analyses to estimate the complier-average causal effect (CACE) at 6 weeks and 6 months post randomisation. For the purposes of estimating CACE, participants allocated to the intervention group were defined as compliers if they were administered at least six Social Stories sessions. For

identification, we assume there are no 'defiers' (i.e. participants that would receive at least six Social Stories sessions if and only if they were allocated to control) and that the usual assumptions required for causal inference apply (i.e. exchangeability, consistency, positivity and no interference). We also assume that randomisation is a valid instrument for treatment received (where receipt of treatment is defined as having/attending at least six Social Stories sessions) and that the effect of treatment received on outcome is not modified by any observed or unobserved confounders of this pathway. We use two-stage least squares estimators to estimate the CACE at 6 months and 6 weeks (with randomisation being used as the instrument in the first-stage regression). We adjusted for the same baseline covariates in the first and second stages as in the primary analysis and modelled dependence between participants within the same cluster using a school cluster random intercept.

Data collection schedule

To examine the possible impact of mistimed data collection (e.g. due to delays resulting from school holidays), analyses were conducted using only teacher-reported SRS-2 data that were collected within the data collection windows specified in the SAP, namely 2/+ 4 weeks for the 6-week follow-up and -4/+ 8 weeks for the 6-month follow-up.

Coronavirus pandemic impact analyses

Due to the impact of the COVID-19 pandemic (e.g. nationwide lockdowns and school closures), certain aspects of the trial were not delivered as intended. In particular, numerous participants had their follow-up data collection schedule disrupted (e.g. some month 6 follow-ups scheduled during summer 2020 could not be completed until September due to school closures) or received Social Stories from interventionists who were trained remotely (online) rather than in person. A number of pre-specified analyses were undertaken to investigate the possible impact of these disruptions on the effectiveness of the intervention.

To assess whether delays in the completion of the primary end point were associated with variation in treatment effect, we refitted the primary analysis model with an extra term denoting delayed outcome completion (binary – defined as the month 6 follow-up being completed more than 8 weeks after the planned follow-up date) and additional terms for all of the two- and three-way interactions between this variable, allocation and time point.

To investigate more broadly the extent to which disruption of follow-up was associated with variation in treatment effects, we refitted the primary analysis model with an additional term denoting whether or not the participant was due their month 6 follow-up during the school closures in the first UK lockdown (23 March 2020 to 3 September 2020) and additional terms for all of the two- and three-way interactions between this variable, allocation and time point.

Finally, to investigate the extent to which disruption of follow-up and online intervention delivery training were associated with variation in treatment effects, we refitted the primary analysis model, including an additional term denoting whether all available follow-up data for the primary outcome (at 6 weeks and 6 months) were provided/completed before/after the date of the start of the first UK lockdown (23 March 2020) and additional terms for all of the two- and three-way interactions between this variable, allocation and time point.

Subgroup analysis

Subgroup analyses were performed to explore the potential modifying effect of the following subgroups:

- The teacher's preferred randomised group collected at baseline. Treatment preference was elicited using a VAS with a range (0, 100), where 0 indicates strong preference for usual care/support, 100 indicates strong preference for Social Stories and 50 indicates indifference/no preference. The participants were assigned to three preference subgroups based on the preference score indicated: (0, 50) = prefers usual care, (50) = no preference and (50, 100) = prefers Social Stories.

- Whether the child has been diagnosed with cognitive/intellectual problems or a learning disability as collected on the comorbidities CRF at baseline (Binary – Yes/No).
- Whether the child has been diagnosed with mental health or psychological problems as collected on the comorbidities CRF at baseline (Binary – Yes/No).

For each of the three baseline variables considered, the primary analysis model was augmented with additional terms for the main effect of the relevant subgroup, the two-way interactions between subgroup and allocation and subgroup and time point and the three-way interaction between subgroup, allocation and time point. The fitted models were used to obtain treatment effect estimates at 6 weeks and 6 months by subgroup, together with appropriate 95% CIs and *p*-values. As recommended by the literature, the subgroup analyses were restricted to the primary analysis, and subgroups were defined by baseline data, that is, data that are not dependent on the intervention.⁴³

Secondary analyses

The goal-based outcome score (6 weeks and 6 months), SRS-2 total score completed by parents (6 weeks and 6 months), RCADS total score (6 weeks and 6 months), EQ-5D-Y proxy VAS (6 weeks and 6 months) and Parental Stress Index (PSI) total stress score (6 weeks and 6 months) were analysed similarly to the primary outcome (i.e. using linear mixed-effect models with the same fixed and random effects and unstructured residual error covariance structure), but with the baseline score for the relevant secondary outcome included in the linear predictor in place of the baseline score for the primary outcome. The fitted models were used to estimate the between-group difference in expected score (for the relevant outcome) at 6 weeks and 6 months post randomisation together with appropriate 95% CIs and *p*-values.

Trial oversight

The conduct of the study was governed by three oversight committees:

- a TSC
- an independent DMEC
- a TMG.

These committees functioned in accordance with YTU SOPs. The DMEC and TSC were both independent from the sponsor. The TSC consisted of an independent chair, an independent subject specialist, an independent clinical academic, an independent statistician and a PPI representative. The DMEC consisted of an independent chair, an independent statistician and another independent member experienced in research with children and families. The TSC and DMEC met approximately every 6 months from the start of the trial. The TMG comprised co-applicants, members of the trial team (including the data manager), PPI representatives and the trial managers. Co-applicants and trial team members were invited as required, depending on their roles.

Internal pilot

The first 10 months of the trial following the start of recruitment served as the internal pilot. Clear stop/go criteria were established at the outset of the trial with the view of reviewing progress against these criteria before proceeding to the full trial. The stop/go criteria centred on the feasibility of recruitment (target 110 children), retention (target 44 final follow-ups complete) and safety outcomes (see [Appendix 1, Table 45](#)).

At the end of the 10-month period, the trial team reported to the oversight committees, the TSC and DMEC, who then critically reviewed the feasibility of continuing recruitment. The progress against the stop/go criteria was as follows.

Recruitment

A total of 35 schools accounting for 100 children were randomised between 25 February 2019 (date of first randomisation) and 25 December 2019 falling just short of the 110-recruitment target.

Retention

In total, 49 participants had reached the 6-month follow-up point at the end of the pilot phase. Of these, 37 teacher 6-month follow-ups had been returned in addition to 36 parent follow-ups; this amounted to approximately a 75% response rate for the primary outcome.

Safety

Four AEs were reported to the trial team during the course of the pilot period, all of which were deemed unrelated or unlikely to be related to Social Stories; hence, no safety concerns were raised.

Process evaluation methods

The process evaluation was cross-sectional and longitudinal, encompassing all aspects of the Social Stories intervention. The aim of the process evaluation was to assess the fidelity of the programme, consider the views of various stakeholders and identify barriers and facilitators to successful implementation.⁴⁴ We aimed to achieve this through a combination of data collection techniques, including interviews, focus groups, questionnaires, surveys (see [Report Supplementary Material 6](#)) and diaries (session logs).

Research questions

The research questions for the process evaluation were:

1. What does baseline practice in participating schools look like (in the control and intervention arms) in terms of interventions to support children on the autism spectrum in school?
2. To what extent do the schools and interventionists implementing the Social Stories intervention adhere to the intended delivery model (e.g. sessions delivered six times over 4 weeks)?
 - a. What variability in implementation exists? Are there any barriers or adaptations?
 - b. How well have the Social Stories been delivered, and how well have children on the autism spectrum and school staff engaged with it?
3. What are the views of specified stakeholders (interventionists, teachers and parents) about the implementation and effectiveness of the intervention during the trial period?
4. To explore the views of RAs and trainers about the goal-setting process as part of the Social Stories intervention and to describe any challenges faced.
5. To what extent do schools continue to deliver the intervention after the trial period?
 - a. What are the barriers and facilitators to continuing to deliver the intervention?

Process evaluation methods

A full overview in relation to each of the process evaluation objectives is presented in [Table 2](#).

Qualitative data

The Social Stories intervention is complex, as not only does it involve several interacting components (training, goal-setting, writing and delivering Social Stories), but it also requires tailoring to meet the needs of each child, setting and behaviour. The extent to which interventions are delivered as intended is a particular challenge when designing and evaluating pragmatic RCTs of complex interventions.⁴⁵⁻⁴⁷ There is a wealth of evidence that has aimed to define the key components of intervention fidelity.⁴⁸ The Hasson framework⁴⁴ is an updated version of a conceptual framework, which was originally developed by Carroll *et al.*⁴⁴ as a result of a critical review that synthesised existing research on intervention fidelity. Hasson proposes nine elements, as a way of conceptualising intervention fidelity: adherence, exposure or dose, quality of delivery, participant responsiveness, programme differentiation, intervention complexity, facilitation strategies, recruitment and context (see [Appendix 1, Table 46](#)). The Hasson framework⁴⁴ has since been used to explore how qualitative findings can be used to understand

TABLE 2 Process evaluation methods overview

Method of data collection	Who or what?	N planned	Research questions
Semistructured interviews conducted via videoconference or telephone at 3 or 6 months post randomisation	Parents/carers who attended Social Stories training	Approximately 20 in total	What are the views of parents about the implementation and effectiveness of the intervention during the trial period?
Semistructured interviews conducted via videoconference or telephone at 3 or 6 months post randomisation	Interventionists who attended Social Stories training		What are the views of interventionists about the implementation and effectiveness of the intervention during the trial period?
Semistructured interviews conducted via videoconference or telephone at 3 or 6 months post randomisation	Teachers who may have attended Social Stories training and were involved in completing outcome measures about the participating children		What are the views of teachers about the implementation and effectiveness of the intervention during the trial period?
Focus groups facilitated by independent researcher at the end of the data collection period	Social Stories trainers	1	What were the trainers' experiences of delivering Social Stories training? What did they think about the training they had received?
Focus groups facilitated by independent researcher at the end of the data collection period	RAs involved in setting Social Stories goals with teachers and parents/carers in person (prior to COVID-19)	1	What are the views of RAs about the goal-setting process as part of the Social Stories intervention? What challenges were faced?
Focus groups facilitated by independent researcher at the end of the data collection period	RAs involved in setting Social Stories goals with teachers and parents/carers during online meetings (post COVID-19)	1	What are the views of RAs about the goal-setting process as part of the Social Stories intervention? What challenges were faced?
Questionnaires completed by the teachers and interventionists at baseline, 6 weeks and 6 months post randomisation	Teachers and interventionists	All participating teachers and interventionists	What does baseline practice in participating schools look like (control and programme) in terms of interventions to support children on the autism spectrum in school? To what extent do schools continue to deliver the intervention after the trial period?
Session logs completed by interventionists during the intervention period (4 weeks)	Interventionists	All participating interventionists	To what extent do the schools and interventionists implementing the Social Stories intervention adhere to the intended delivery model? How well have the Social Stories been delivered, and how well have children on the autism spectrum and school staff engaged with it?
Online survey – following completion of the 6-month follow-up	Parents/carers	All participating parents who consented to additional research	What are the views of parents about the implementation and effectiveness of the intervention during the trial period?
Online survey – following completion of the 6-month follow-up	Interventionists	All participating interventionists who consented to additional research	All research questions addressed
Online survey – following completion of the 6-month follow-up	Teachers	All participating teachers who consented to additional research	All research questions addressed

intervention fidelity within a NIHR-funded pragmatic RCT that evaluated a multifaceted podiatry intervention trial.^{46,49} In this process evaluation, we have adopted a similar approach, and by using the Hasson framework⁴⁴ to underpin our analysis, we have addressed the aims of our qualitative study and provided an in-depth exploration of the key factors that affected the fidelity of the Social Stories intervention during the ASSSIST-2 trial. This approach has also enabled us to identify the key areas to consider prior to any potential future rollouts of the intervention.

The decision to use Hasson's components of intervention fidelity as an analytical framework was based on early familiarisation with our qualitative data set during the early stages of data collection and analysis, where it became clear that many of the initial themes we had identified and indeed our research questions related to fidelity. This decision was also a pragmatic one due to the number of researchers involved in analysis, and we needed a framework which could be used consistently, broadly and flexibly (i.e. allowing new themes and/or subthemes to be identified) by the group, while at the same time allowing for an in-depth, deductive analysis of our qualitative data based on the overarching constructs proposed by Hasson to be undertaken.

Design

The qualitative study consisted of (1) interviews with interventionists who had received Social Stories training and delivered the intervention, (2) interviews with teachers who may have attended the training and were involved in completing outcome measures about the participating children, (3) interviews with parents who attended training and (4) focus groups with trainers to explore their experiences of delivering training and with the study RAs to explore their experiences of setting goals for the Social Stories with teachers and parents.

Sampling and participants

In the protocol,¹ we specified that we would undertake semistructured interviews with a minimum total of 20 participants and would purposively select individuals from the cohort of trainers, interventionists and associated teachers who have been involved in the trial, as well as parents who attended training. To achieve maximum variation, we proposed selecting participants according to age, gender, time in profession and experience with ASC. The COVID-19 pandemic meant that we had to move from an in-person to a virtual training model. We therefore adapted our sampling frame and recruitment strategy to ensure that we selected participants according to whether they had received training online or in person, at both early and late stages of the transition to the online training model. To ensure further variation in our sampling, we also selected participants that represented schools and areas that were diverse geographically and in terms of the student population (e.g. school size and type) and setting (rural, urban). Data collection continued until a varied sample of teachers, interventionists and parents had been obtained⁵⁰ (see [Appendix 1, Table 47](#)). This was assessed through monthly qualitative team meetings held between the RAs, the trial coordinator and lead qualitative researcher.

In total, 30 participants took part in the qualitative study, with 21 interviews (five teachers, seven interventionists and nine parents) and three focus groups (nine research team members) conducted. The 21 participants represented 15 schools throughout Yorkshire and the Humber. Schools were diverse in size (9 schools > 200 pupils, 6 schools < 200 pupils) and rurality (seven urban, eight rural). Three focus groups were conducted: one with trainers who delivered training ($n = 4$), one with RAs who were involved in goal-setting with parents and teachers when they could visit participants in person (prior to the COVID-19 pandemic) ($n = 2$) and one with RAs who were involved in goal-setting with parents and teachers during online meetings (during the COVID-19 pandemic) ($n = 4$). One RA participated in two focus groups as they had become unblinded to help deliver the Social Stories training.

Recruitment and consent

All ASSSIST-2 trial participants (interventionists, teachers, parents/carers) who expressed an interest in participating in a qualitative interview on the consent form and who had been involved in the intervention were eligible for participation in the qualitative study. Up until April 2021, eligible

participants were approached to take part in the qualitative study after they had returned all 6-month follow-up data. This was to avoid the RAs who were conducting the interviews becoming unblinded. After April 2021, one RA became unblinded and was able to invite participants to take part in the qualitative study 3 months after randomisation (between the 6-week and 6-month follow-up points). Trial participants were recruited via e-mail which contained a participant information sheet (PIS) (see [Report Supplementary Material 1](#)) and consent form (see [Report Supplementary Material 2](#)). Recruitment e-mails also informed trial participants that a RA would contact them via telephone to find out if they would be willing to take part in the qualitative study. Once consent forms had been completed by participants and received by the research team, a date and time for the interview were arranged. If the interview took place via videoconference, the participants could choose whether or not they appeared on video and whether the video was recorded. For focus groups, RAs and trainers were contacted directly by the lead qualitative researcher via e-mail and were asked if they were willing to take part in a focus group. Prior to all interviews and focus groups, informed consent was obtained, and participants were assured anonymity and confidentiality and were given the opportunity to ask questions.

Data collection

All focus groups and interviews were semistructured and were conducted via the videoconferencing software Zoom or via telephone (depending on the participants' preference) between August 2020 and December 2021. Interviews and focus groups ranged from 10 minutes to 1 hour 20 minutes in length. Four separate topic guides were developed for the interviews, and two separate topic guides were developed for the focus groups. Topic guides were used flexibly to ensure key topics were discussed, while allowing for probing by interviewers and for participants to voice issues they considered important (see [Report Supplementary Material 5](#)). Areas explored during interviews and focus groups broadly covered participants' views and experiences of goal-setting, training, trial participation and the intervention. Interviews with interventionists also explored participants' experiences of delivering the Social Stories intervention with the children. Interviews were conducted by six trained RAs. All had previous experience of working with children or individuals on the autism spectrum and had received Social Stories training from the study team. Given that the RAs were instrumental to the delivery of the trial and were involved in setting goals for participants, focus groups with RAs and trainers were conducted by a qualitative researcher who was not involved in the design of the ASSSIST-2 trial or its delivery (AS). Arabella Scantlebury is an experienced qualitative methodologist with expertise in using qualitative research alongside RCTs. She has no prior experience of Social Stories and joined the ASSSIST-2 team in October 2020. Analysis was primarily undertaken by the RAs and trial co-ordinator and overseen by AS. Arabella Scantlebury led the analysis of focus group data. No key differences were identified in the data collected or analysed by any of the researchers.

Process evaluation analysis

All interviews and focus groups were audio-recorded and transcribed verbatim prior to analysis. Initially, we developed a coding framework, which was informed by our topic guides and Hasson's conceptual framework of intervention fidelity.⁴⁴ This involved generating a 'list' of themes and subthemes that were mapped onto the nine components of Hasson's conceptual framework. For instance, participant responsiveness (how participants respond to or engage with the intervention) included data relating to previous experiences of Social Stories, motivations to take part in the trial and views of the intervention. Data were then analysed according to this framework for thematic content, with our coding framework adapted as data collection and analysis progressed, reflecting our growing familiarity and understanding of our data. Lastly, to aid readability and to give a comprehensive overview of intervention fidelity during the ASSSIST-2 trial, quantitative data pertaining to intervention fidelity that was collected through questionnaires and screening logs were integrated and presented alongside qualitative data under the components of intervention fidelity as described by Hasson in the process evaluation results chapter (see [Chapter 5](#)).

Data analysis was undertaken alongside data collection, with monthly qualitative team meetings used to discuss and refine our coding framework and facilitate the interpretation of our data. All participants

were allocated unique personal IDs to protect anonymity and confidentiality. Due to the number of researchers involved in data analysis who worked across different institutions, data were managed using Microsoft Word and Excel.

Economic evaluation methods

The health economics component of ASSSIST 2 was a cost-utility analysis, measuring the incremental cost-effectiveness ratio (ICER) of the Social Stories intervention over and above the control arm.

Effectiveness

Effectiveness of the health economics analysis was measured using EQ-5D-Y to calculate quality-adjusted life-years (QALYs). The EQ-5D-Y (3L proxy version)³⁵ is a five-item generic preference-based measure of health-related QoL that allows a proxy person to complete the measure for the participant. The EQ-5D-Y was chosen for the following reasons. Firstly, the EQ-5D-Y has been shown to be a reliable and valid health-related QoL instrument for use in children and young people (CYP)⁵¹ and with children on the autism spectrum or with related conditions, such as attention deficit hyperactivity disorder (ADHD), speech/language disorder⁵² and functional disability.⁵³ Secondly, the EQ-5D-Y is commonly used in CYP with psychosocial problems, including our previously successfully delivered Social Stories feasibility study – ASSSIST.¹² Finally, the EQ-5D is the instrument suggested by the National Institute for Health and Care Excellence (NICE) for cost-utility analysis.⁵⁴ The measure allows for the comparison between groups across a wide range of comparable dimensions and, thus, can be used to calculate ICERs compared with the national willingness-to-pay (WTP) threshold of £20,000–30,000 per QALY gained.

In EQ-5D-Y, health of CYP can be classified on a 3-point scale (1, 'no problems'; 2, 'some problems' and 3, 'a lot of problems') over five dimensions: mobility (walking about), looking after myself, doing unusual activities, having pain or discomfort and feeling worried, sad or unhappy. Additionally, the EQ-5D-Y includes a VAS where participants can indicate their overall health status from 0 (worst imaginable state) to 100 (best imaginable state). All questions refer to the health state 'today'. The measured utilities at baseline and follow-up were further joined through time, and the area under the curve approach was used to calculate QALYs for further cost-utility analyses.⁵⁵

Cost of the Social Stories intervention

Costs of the Social Stories intervention included the cost of training and the cost of delivering the intervention. Training costs were measured by the time spent by the trainers. Costs per minute of time, based on salary costs, were allocated, and any travel costs, costs of preparation time and the cost of materials used for the training were also recorded, whereas costs associated with delivering the Social Stories intervention were collected using a bottom-up estimation of the time spent by professionals to plan and conduct sessions and to undertake any additional work. Tailored questionnaires for collecting the costs of training and delivering Social Stories were developed by the research team based on our LEGO®-based therapy trial – I-SOCIALISE⁵⁶ to capture the resource use in both Social Stories (trial) and usual care (control) arms.

Costs of service use

Service use information was collected using the tailored questionnaires (completed by the parent/guardian) to capture data on the use of the following services:

- community health services, including appointments with general practitioner (GP), nurse, child development centre, walk-in centre, social worker and family support worker
- mental health services, including appointments with psychiatrist, psychologist, CAMHS therapist and mental health nurse

METHODS

- hospital services, including outpatient visits, inpatient admissions, accidents, emergency visits and urgent care centre visits
- school-based interventions/support provided by teachers, school nurses or school staff
- other, including medication and informal parental/guardian care (productivity cost).

Service use was multiplied by unit costs to arrive at a total cost in each arm. Unit costs of health and social service use were obtained from two public sources: the National Cost Collection 2019–20⁵⁷ and the Unit Costs of Health and Social Care report produced by the Personal Social Services Research Unit (PSSRU).⁵⁸ The costs of medication were based on the most recent version of the Prescription Cost Analysis – England,⁵⁹ and the costs of teachers were estimated based on the data reported by the Department for Education.⁶⁰ Privately paid services were estimated based on market prices, and the productivity costs were valued according to national average wage rates. [Appendix 2, Tables 60–67](#), presents the summary of key unit costs used in the study.

Descriptive analysis

Total costs, including the intervention cost and service use costs, and QALYs were compared between the intervention and control groups at baseline and 6-month follow-up. Unadjusted costs and QALYs were presented using appropriate descriptive analyses.

Missing data

Based on the completed service use questionnaires, data were deemed as missing when all sections under a particular item were left blank. For instance, where one question had subquestions, all were deemed as missing when all were left blank. If one of these questions is answered, then the others will be assumed zero. For EQ-5D-Y, due to the structure of the questionnaire, the whole section is considered missing if any of the five questions is not answered. The missing data in service use and EQ-5D-Y were imputed using Rubin's multiple imputation method.⁶¹

Primary economic analysis

All the outcomes were analysed after the trial ended according to a pre-specified health economic analysis plan (HEAP). The primary analysis was a within-trial cost-utility analysis conducted from a societal perspective. Combining costs and QALYs, an incremental ICER of cost per QALY was calculated and against the WTP threshold of £20,000–30,000 per QALY gained to assess the cost-effectiveness of Social Stories compared to usual care in children with ASC. Discounting was not applied due to the short-term nature of the trial.

To adjust any imbalanced characteristics at baseline, regression models on an ITT basis were used to compare mean costs and QALYs. The regression analyses were controlled for age, gender, stratification factors (SEN status stratum: non-SEN or SEN and number of participants stratum: ≤ 5 or > 5), parent-completed SRS-2 score, baseline utility and baseline costs from the societal perspective. The specification of the model followed the approach recommended by Glick *et al.*,⁵⁵ which considers the distribution of the dependent variable as well as any correlation between the cost and QALY outcomes. The regression coefficients on treatment then represented the difference in mean cost and mean QALYs between groups.

To take uncertainty into consideration, a non-parametric bootstrap resampling method on the basis of 5000 iterations was used to produce CIs around the cost and QALY differences and ICER. The number of iterations was chosen because it was considered to be sufficient to generate robust estimates of standard errors⁶² and is widely used in trial-based cost-effectiveness analyses for mental health studies.^{63,64}

The outcomes of bootstrapped iterations were presented in the conventional form of a cost-effectiveness plane (CE plane) and a cost-effectiveness acceptability curve (CEAC). The CEAC presents the probability of the intervention being cost-effective over a range of WTP thresholds per QALY.⁶⁵

The higher the probability, the more likely it is that the treatment is cost-effective at the particular WTP threshold.

Sensitivity analyses

The following sensitivity analyses were conducted to test assumptions made in the primary analysis:

- To assess the impact of the missing data, a sensitivity analysis was carried out using the complete case analysis (CCA), whereby results are analysed only for those participants who had both the completed cost and outcome data at each time point.
- A sensitivity analysis was conducted based on the costs from the UK NHS and personal social services (PSS) perspective.
- A sensitivity analysis was conducted based on the costs from the NHS/PSS and education perspective.

Summary of changes to the protocol

Several changes were made to the protocol and study procedures during the course of the trial. These are described in full here.

Inclusion criteria

In the original trial design, we planned to only include primary schools in the trial, given that the evidence indicated that the intervention is best suited to this population. However, this study involved both mainstream and SEN schools. As SEN schools often have a different structure than mainstream schools, we identified that there may be occasions where a SEN school participant moved from year 6 to year 7 without there being a change to any practical or environmental circumstances likely to influence intervention delivery (i.e. they will be in the same school, same class and have the same teachers). As part of a substantial amendment approved by North East – York REC on 2 January 2020, we specified that if a participant in this situation wishes to be part of the trial, we would consent them to the trial as if they were to remain a primary school student throughout their participation in the research.

Baseline data collection

Due to the structure of the academic year and school holidays, as well as other unexpected delays, we found that there was sometimes a time lag between data being collected and randomisation. As part of a substantial amendment approved by North East – York REC on 1 July 2020, we amended the baseline data collection procedure to allow for baseline data to be recollected if significant time delays were observed between the date of data collection and randomisation so that no data were collected more than 3 months prior to the date of randomisation.

Recruitment

As part of a substantial amendment approved by North East – York REC on 1 July 2020, we were granted approval to expand our recruitment strategy to include social media. Following this approval, information about the study, including eligibility criteria, could be posted on the research teams' social media accounts, specifically including The Child Oriented Mental Health Intervention Centre Twitter account (@COMICResearchUK), Facebook (COMIC Research) and the website (www.comic.org.uk). A further change to recruitment was later made via a non-substantial submitted on 17 August 2020. The purpose of this was to extend the pre-specified recruitment window to reflect an extension to the trial granted by the funder in light of the COVID-19 pandemic.

Follow-up procedures

Two changes were made to the follow-up procedures during the course of the trial. Firstly, clarifications were made to the follow-up procedure to formally document that where CRFs were not returned within 2 weeks of their due date, researchers would contact participants by e-mail, telephone and/or text

(depending on participant contact preferences) a maximum of four times. For the 6-week follow-up, if no response was received after four attempted contacts, no further action was taken. For the 6-month follow-up, participants were considered lost to follow-up after four attempted contacts. These clarifications were accepted by North East – York REC on 1 July 2020.

Secondly, as part of the same amendment, due to the COVID-19 pandemic and associated social distancing guidelines, we received approval to amend the follow-up procedures from researcher-based data collection to allow for postal data collection. This meant that parents could choose whether to receive a postal questionnaire or to continue with researcher-supported data collection (via their preferred contact method of telephone or videoconferencing).

Consenting procedures

As a consequence of the COVID-19 pandemic and associated social distancing guidance, we were required to update our consenting procedures which were previously all undertaken in person. We amended our processes to allow for researchers to take consent remotely via telephone or videoconferencing. Participants were sent consent forms via post to complete during telephone/video calls, and these were returned to the trial team. This change was approved as part of a substantial amendment by North East – York REC on 1 July 2020.

Qualitative component

Revisions were made to the approved interview and focus group schedules, and a survey was introduced as part of the qualitative enquiry. These reflected only minor changes and were approved by North East – York REC on 1 July 2020. Final version copies of all associated paperwork from the qualitative work can be found in [Report Supplementary Materials 3–6](#).

Randomisation procedures

Prior to the COVID-19 pandemic, the trial employed very strict timelines surrounding the randomisation process. Specifically, interventionists needed to be trained within 2 weeks post randomisation and following this, interventionists had a 4-week period to deliver the intervention before the 6-week CRF was due. This timeline meant that we had to book provisional training dates for all schools and then cancel those for schools allocated to usual care. Following the school closures associated with the pandemic and the limited time with which to randomise schools without needing to extend the trial into another school year, we proposed amending the randomisation procedure so that follow-ups were triggered by the date that schools were trained and could start delivering the intervention rather than by date of randomisation. This allowed us to randomise all schools and then schedule training only for those schools allocated to the intervention arm. This change was reviewed by all the trial oversight committees (TSC, DMEC, TMG) as well as the funder and was submitted as a non-substantial amendment on 22 February 2021.

Chapter 3 Clinical effectiveness results

Recruitment and retention

Recruitment

Between November 2018 and May 2021, 295 children across 98 schools were screened for trial eligibility. Of these, 268 (90.8%) were found eligible according to the protocol-specified eligibility criteria, 26 (8.8%) were ineligible and 1 (0.3%) had unclear/unknown eligibility status (due to missing data on age). Of the 295 children screened, 292 (99.0%) had screening data available for analysis. A summary of the available eligibility data is given in [Table 3](#). The other three children were consented and randomised and are therefore assumed to have been eligible. These participants are included in all subsequent analyses. Of the 26 children that were deemed ineligible, the reasons for ineligibility were (possibly more than one per participant): not attending a school in Yorkshire or the Humber ($n = 1$), the parent/carer reporting that they did not feel the child had behavioural problems at school ($n = 16$), took part in the ASSSIST feasibility study ($n = 1$) and use of Social Stories by the school in the current or preceding school term ($n = 15$).

Initial consent to trial participation was obtained for all 268 children who were eligible. Of these, 45 (16.8%) subsequently withdrew (or were withdrawn) from trial participation prior to being randomised. Hence, 223 children were eligible, consented and were not withdrawn prior to randomisation. Brief details for the pre-randomisation withdrawals are given in [Table 4](#).

TABLE 3 Available eligibility data as reported in the screening CRF (boldface indicates the response required for eligibility)

Eligibility criteria	Yes, <i>n</i> (%)	No, <i>n</i> (%)	Unknown, <i>n</i> (%)
Aged 4–11 years	291 (99.7)	0 (0.0)	1 (0.3)
Attends participating schools in Yorkshire and the Humber	291 (99.7)	1 (0.3)	0 (0.0)
Has an ASC diagnosis	292 (100.0)	0 (0.0)	0 (0.0)
Parent/carer feels the child has behavioural problems in school	276 (94.5)	16 (5.5)	0 (0.0)
Parent/carer is able to complete English language outcome measures	292 (100.0)	0 (0.0)	0 (0.0)
Child took part in ASSSIST feasibility study	1 (0.3)	291 (99.7)	0 (0.0)
School has used Social Stories in current or preceding school term	15 (5.1)	277 (94.9)	0 (0)

TABLE 4 Reasons for pre-randomisation withdrawals

Reason	<i>n</i> (%)
School lacking capacity/unable to facilitate trial	14 (31.1)
School and/or parent no longer wanted to be involved/withdrew consent	10 (22.2)
Child moved schools before randomisation could be completed	9 (20.0)
School and/or parent not responding/uncontactable	8 (17.8)
Child reached 12 + years of age before randomisation could be completed	3 (6.7)
Baseline measures not received prior to the end of recruitment period	1 (2.2)

Between February 2019 and May 2021, 87 schools were randomised (44 to the intervention group and 43 to the control group), with a total of 249 children being randomised (129 to the intervention group and 120 to the control group). This figure includes the 223 eligible/consented children referred to above, but also 26 children that were ineligible according to the protocol-specified eligibility criteria (see above). Details of the school cluster sizes (at randomisation) by allocation and overall are given in [Table 5](#). Plots of the cumulative number of clusters and participants randomised over the recruitment period are given in [Figures 1](#) and [2](#) in the [Report Supplementary Material 7](#). Overviews of the screening/recruitment activity (including baseline data collection) are given in [Figures 1](#) and [2](#).

Retention

Following randomisation, five (2.0%) participants were formally withdrawn from either certain aspects of the planned follow-up data collection (e.g. just parent-completed data collection) or all follow-up data collection (i.e. full withdrawal). A summary of post-randomisation withdrawals by allocation is given in [Table 6](#). The completeness of both teacher- and parent-reported data collection (randomised participants only) is given in [Table 7](#). The flow of the participants through the trial is illustrated in [Figures 1](#) and [2](#) (teacher-completed data collection in [Figure 1](#) and parent/carer-completed data collection in [Figure 2](#)).

From [Table 7](#) and [Figure 1](#), it is apparent that the proportion of participants retained for the purposes of teacher-completed follow-up was reasonably stable across the two time points (approximately 85% at both time points). However, it is clear from the figures in [Table 7](#) that it was generally not the same group of participants missing teacher-completed data at each time point. For example, 238 (95.6%) participants had at least one available teacher-completed SRS-2 score (the primary outcome), meaning only 4.4% of participants had no follow-up data for the primary outcome, although 15.3% were missing primary end-point data (i.e. were missing primary outcome data at 6 months). There were some differences in the completeness of teacher-completed follow-up between the randomised groups. For example, at both week 6 and month 6, approximately 80% of participants allocated to the intervention group had teacher-completed SRS-2 data, compared with approximately 90% of participants allocated to the control. This difference is driven primarily by cluster-level loss to follow-up/non-response. For example, at month 6, 10 (22.7%) whole school clusters were lost to follow-up in the intervention group, compared with 5 (11.6%) in the control group.

For the parent/carer-completed data (see [Table 7](#) and [Figure 2](#)), the proportion of participants retained diminished slightly between the week 6 and month 6 follow-ups (approximately 78% at week 6 vs. 74% at month 6). However, loss to follow-up was again not monotonic, meaning a larger proportion (approximately 86%) had parent-completed data for at least one follow-up time point. The completeness of the parent-completed follow-up data differed somewhat between randomised groups at week 6, with the proportion providing data at this time point being 5–10% lower in the intervention group than in the control group, although at month 6 there was little difference between groups. In addition, the proportion of participants with a baseline measurement and at least one follow-up measurement was similar for each of the parent-reported outcomes.

TABLE 5 School cluster sizes (as randomised) by allocation and overall

	Control (N = 43)	Intervention (N = 44)	Total (N = 87)
School cluster size			
N	43	44	87
Mean (SD)	2.8 (1.9)	2.9 (2.0)	2.9 (2.0)
Median (Q1, Q3)	2.0 (1.0, 4.0)	2.0 (2.0, 4.0)	2.0 (1.0, 4.0)
Minimum, maximum	1.0, 9.0	1.0, 11.0	1.0, 11.0

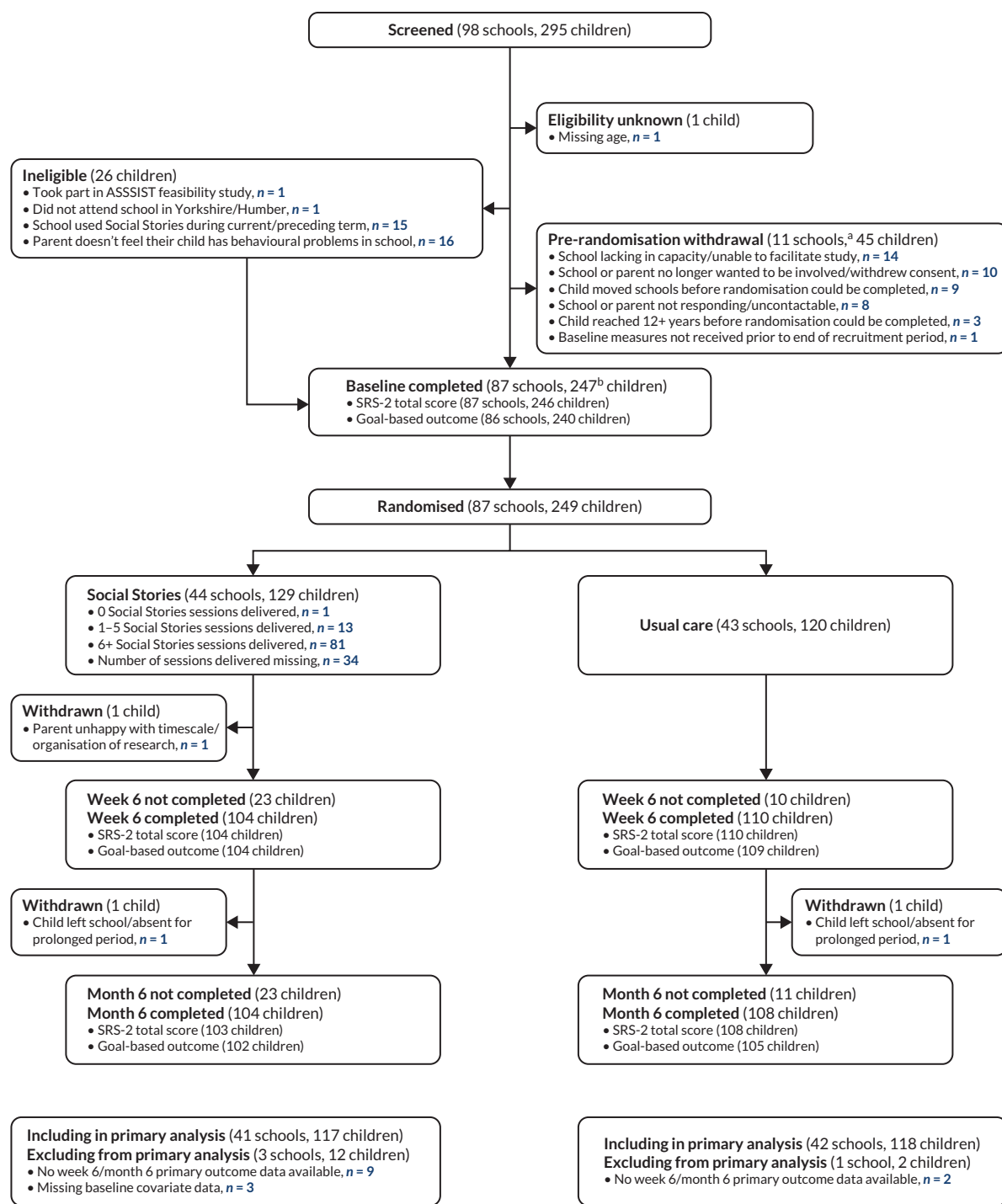


FIGURE 1 Recruitment and teacher completed data collection. a, Eleven schools where all participants withdrew prior to randomisation; b, two participants with no teacher-completed baseline data.

Baseline data

School-level baseline data

Of the 98 schools that provided screening and/or baseline data for at least one participant (or potential participant), 87 (88.8%) were randomised (44 schools to intervention and 43 to control). Randomisation was stratified by school SEN status (whether the school was mainstream or whether the school was a specific specialist provision setting, that is, non-SEN or SEN) and number of participants (≤ 5 or > 5). Summaries of the distribution of these stratification factors within randomised groups are given

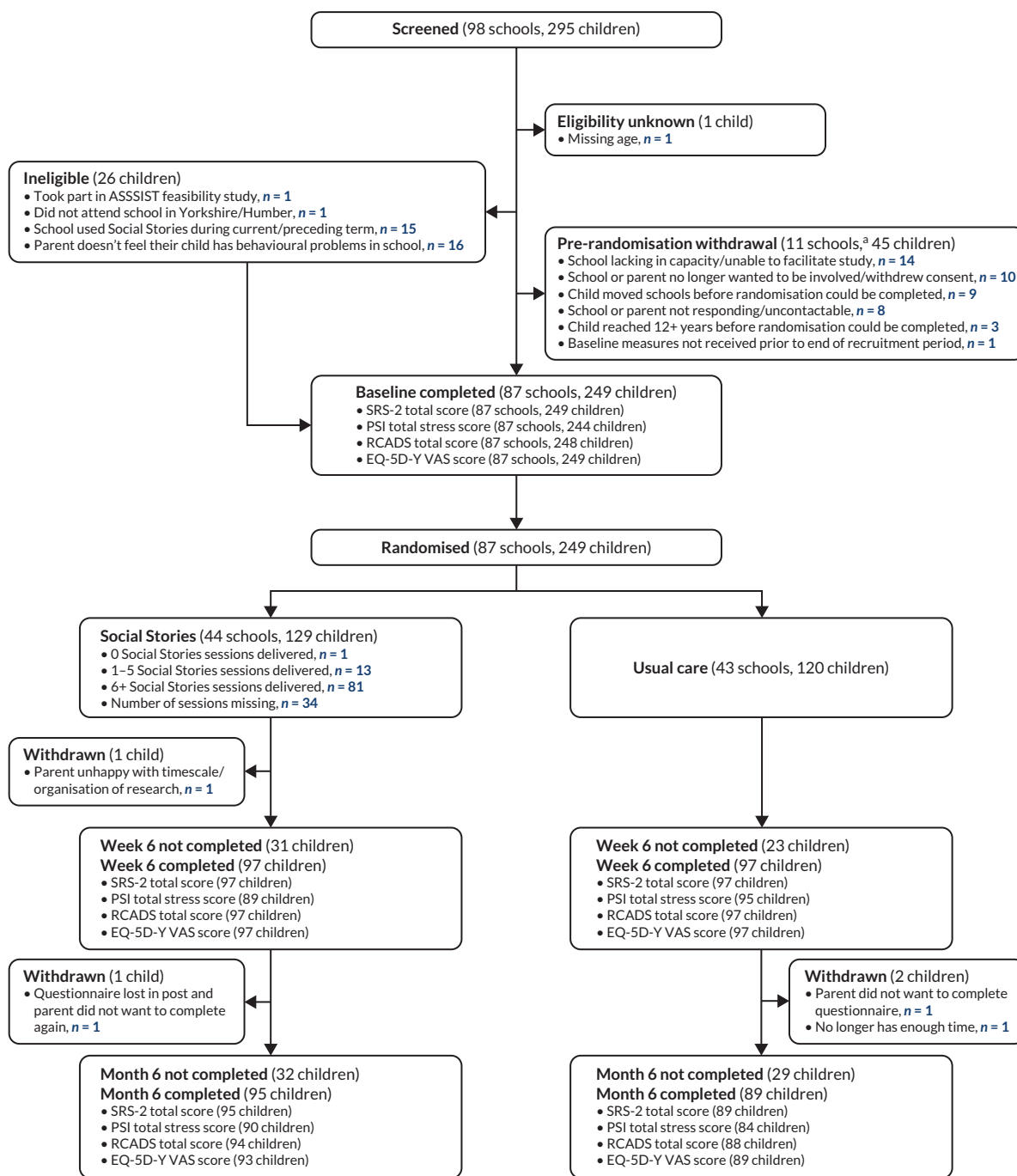


FIGURE 2 Recruitment and parent/carer completed data collection. a Eleven schools where all participants withdrew prior to randomisation.

TABLE 6 Post-randomisation withdrawals by allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Withdrawal from parent/child follow-up	1 (50.0)	1 (33.3)	2 (40.0)
Withdrawal from intervention and teacher follow-up	0 (0.0)	1 (33.3)	1 (20.0)
Full withdrawal	1 (50.0)	1 (33.3)	2 (40.0)

TABLE 7 Completeness of teacher- and parent/carer-completed data

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Teacher completed			
SRS-2 total T-score, n (%)			
Baseline	120 (100.0)	126 (97.7)	246 (98.8)
Week 6	110 (91.7)	104 (80.6)	214 (85.9)
Month 6	108 (90.0)	103 (79.8)	211 (84.7)
Baseline + ≥ 1 follow-up completed	118 (98.3)	117 (90.7)	235 (94.4)
Goal-based outcome, n (%)			
Baseline	114 (95.0)	126 (97.7)	240 (96.4)
Week 6	109 (90.8)	104 (80.6)	213 (85.5)
Month 6	105 (87.5)	102 (79.1)	207 (83.1)
Baseline + ≥ 1 follow-up completed	111 (92.5)	118 (91.5)	229 (92.0)
Parent/carer completed			
SRS-2 total T-score, n (%)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	89 (74.2)	95 (73.6)	184 (73.9)
Baseline + ≥ 1 follow-up completed	103 (85.8)	112 (86.8)	215 (86.3)
PSI total stress score, n (%)			
Baseline	117 (97.5)	127 (98.4)	244 (98.0)
Week 6	95 (79.2)	89 (69.0)	184 (73.9)
Month 6	84 (70.0)	90 (69.8)	174 (69.9)
Baseline + ≥ 1 follow-up completed	100 (83.3)	107 (82.9)	207 (83.1)
RCADS total score, n (%)			
Baseline	120 (100.0)	128 (99.2)	248 (99.6)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	88 (73.3)	94 (72.9)	182 (73.1)
Baseline + ≥ 1 follow-up completed	103 (85.8)	110 (85.3)	213 (85.5)
EQ-5D-YVAS, n (%)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	89 (74.2)	93 (72.1)	182 (73.1)
Baseline + ≥ 1 follow-up completed	103 (85.8)	112 (86.8)	215 (86.3)

in [Table 8](#), together with details regarding the distribution of school cluster sizes. It is evident from [Table 8](#) that the stratification (and blocking) of the randomisation sequence achieved its intended goal of ensuring that the cluster allocations were (approximately) orthogonal with respect to the chosen stratification factors. The very slightly higher mean cluster size and additional allocation in the intervention group resulted in a slight imbalance in the number of participants in each group (129 in

TABLE 8 Brief baseline characteristics (stratification factors and cluster sizes) of randomised schools

	Control (N = 43)	Intervention (N = 44)	Total (N = 87)
School SEN status stratum, n (%)			
Non-SEN	39 (90.7)	40 (90.9)	79 (90.8)
SEN	4 (9.3)	4 (9.1)	8 (9.2)
Number of participants stratum, n (%)			
≤ 5	33 (76.7)	35 (79.5)	68 (78.2)
> 5	10 (23.3)	9 (20.5)	19 (21.8)
Randomisation stratum, n (%)			
Non-SEN and ≤ 5	30 (69.8)	32 (72.7)	62 (71.3)
SEN and ≤ 5	3 (7.0)	3 (6.8)	6 (6.9)
Non-SEN and > 5	9 (20.9)	8 (18.2)	17 (19.5)
SEN and > 5	1 (2.3)	1 (2.3)	2 (2.3)
Cluster size (as randomised)			
N	43	44	87
Mean (SD)	2.8 (1.9)	2.9 (2.0)	2.9 (2.0)
Median (Q1, Q3)	2.0 (1.0, 4.0)	2.0 (2.0, 4.0)	2.0 (1.0, 4.0)
Minimum, maximum	1.0, 9.0	1.0, 11.0	1.0, 11.0

the intervention group and 120 in the control group); however, this imbalance would not be expected to have any material consequences for the analyses undertaken.

Participant-level baseline data

There were 266 children with available teacher-reported baseline data and 276 children with available parent-reported baseline data. However, 19 (7.1%) of the 266 children with available teacher-reported baseline data and 27 (9.8%) of the 276 children with available parent-reported baseline data were withdrawn prior to randomisation. Hence, 247 (99.2%) of the 249 randomised participants had available teacher-reported baseline data and 249 (100%) had available parent-reported baseline data. Participant-level summaries of the randomisation stratification factors are given (by allocation and overall) in [Table 9](#). Parent-reported demographics of the randomised participants are given (by allocation and overall) in [Table 10](#). Participant-level summaries of the baseline scores for the teacher- and parent-completed outcome measures are given (by allocation and overall) in [Table 11](#). Plots illustrating the distribution of the observed baseline teacher-reported SRS-2 total raw and T-scores are given in [Figures 3](#) and [4](#), respectively, in the [Report Supplementary Material 7](#).

The figures in [Table 9](#) show the distribution of participants across the four randomisation strata was reasonably similar across the two groups, although naturally there are some departures from perfect balance at the individual level owing to the cluster randomised design. Again, this would not be expected to have any material influence on the analyses undertaken. The figures reported in [Table 10](#) show that the participant-level demographic factors collected were well completed and that the distributions of these were similar across the groups as randomised. From [Table 11](#), it is evident that the raw scores for the teacher-reported SRS-2 were slightly higher ('worse') in the intervention group than in the control group at baseline, and that this translated into a small difference in the T-scores (the primary outcome). However, in both cases, the discrepancies are well within what would be expected

TABLE 9 Randomisation stratification factors (participant-level summaries)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
School SEN status stratum, n (%)			
Non-SEN	111 (92.5)	116 (89.9)	227 (91.2)
SEN	9 (7.5)	13 (10.1)	22 (8.8)
Number of participants stratum, n (%)			
≤ 5	70 (58.3)	80 (62.0)	150 (60.2)
> 5	50 (41.7)	49 (38.0)	99 (39.8)
Randomisation stratum, n (%)			
Non-SEN and ≤ 5	66 (55.0)	71 (55.0)	137 (55.0)
SEN and ≤ 5	4 (3.3)	9 (7.0)	13 (5.2)
Non-SEN and > 5	45 (37.5)	45 (34.9)	90 (36.1)
SEN and > 5	5 (4.2)	4 (3.1)	9 (3.6)

TABLE 10 Parent-reported demographics of randomised participants

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Age (years)			
N	120	129	249
Mean (SD)	8.6 (1.8)	8.5 (1.7)	8.5 (1.7)
Median (Q1, Q3)	8.9 (7.2, 10.0)	8.7 (7.2, 9.8)	8.7 (7.2, 9.9)
Minimum, maximum	4.9, 11.4	4.5, 11.9	4.5, 11.9
Sex, n (%)			
Male	90 (75.0)	95 (73.6)	185 (74.3)
Female	30 (25.0)	34 (26.4)	64 (25.7)
Ethnicity, n (%)			
White – British	103 (85.8)	107 (82.9)	210 (84.3)
White – Irish	1 (0.8)	0 (0.0)	1 (0.4)
White – GRT	0 (0.0)	1 (0.8)	1 (0.4)
White – other	2 (1.7)	6 (4.7)	8 (3.2)
Black – African	1 (0.8)	1 (0.8)	2 (0.8)
Asian – Indian	2 (1.7)	0 (0.0)	2 (0.8)
Asian – Pakistani	1 (0.8)	5 (3.9)	6 (2.4)
Asian – Bangladeshi	1 (0.8)	0 (0.0)	1 (0.4)
Asian – Chinese	0 (0.0)	1 (0.8)	1 (0.4)
Asian – other	0 (0.0)	2 (1.6)	2 (0.8)

continued

TABLE 10 Parent-reported demographics of randomised participants (*continued*)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Mixed – white and Black Caribbean	1 (0.8)	2 (1.6)	3 (1.2)
Mixed – white and Asian	2 (1.7)	2 (1.6)	4 (1.6)
Mixed – other	3 (2.5)	2 (1.6)	5 (2.0)
Other – other	1 (0.8)	0 (0.0)	1 (0.4)
Prefer not to say	2 (1.7)	0 (0.0)	2 (0.8)
ASC diagnosis confirmed by ^a			
Psychiatrist	8 (6.7)	19 (14.7)	27 (10.8)
Clinical psychologist	50 (41.7)	49 (38.0)	99 (39.8)
Speech and language therapist	30 (25.0)	36 (27.9)	66 (26.5)
Educational psychologist	19 (15.8)	18 (14.0)	37 (14.9)
Paediatrician	64 (53.3)	71 (55.0)	135 (54.2)
Other	16 (13.3)	15 (11.6)	31 (12.4)
ASC diagnosis part of multidisciplinary assessment?, n (%)			
Yes	116 (96.7)	119 (92.2)	235 (94.4)
No	3 (2.5)	10 (7.8)	13 (5.2)
Missing	1 (0.8)	0 (0.0)	1 (0.4)
Child's age at the time of ASC diagnosis (years)			
N	120	129	249
Mean (SD)	5.5 (2.2)	5.2 (2.1)	5.3 (2.1)
Median (Q1, Q3)	5.0 (4.0, 7.0)	5.0 (4.0, 7.0)	5.0 (4.0, 7.0)
Minimum, maximum	2.0, 10.0	2.0, 10.0	2.0, 10.0
Child receiving CAMHS support?, n (%)			
Yes	10 (8.3)	11 (8.5)	21 (8.4)
No	110 (91.7)	118 (91.5)	228 (91.6)
Comorbidities ^a			
Physical health problems	31 (25.8)	33 (25.6)	64 (25.7)
Mental health or psychological problems	16 (13.3)	15 (11.6)	31 (12.4)
Developmental problems or learning difficulties	18 (15.0)	17 (13.2)	35 (14.1)
Cognitive problems or learning disability	4 (3.3)	8 (6.2)	12 (4.8)
Genetic or chromosomal problems	5 (4.2)	2 (1.6)	7 (2.8)

GRT: Gypsy, Roma and Traveller.

^a Possibly more than one category per individual.

TABLE 11 Baseline scores for teacher- and parent-reported outcomes and treatment preferences (range of possible scores are noted in parentheses for each outcome measure)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Teacher SRS-2 total raw score ^a (0–195)			
N	120	126	246
Mean (SD)	90.4 (28.7)	94.5 (31.8)	92.5 (30.3)
Median (Q1, Q3)	89.5 (72.0, 113.0)	98.0 (69.0, 120.0)	94.0 (69.0, 115.0)
Minimum, maximum	22.0, 162.0	26.0, 165.0	22.0, 165.0
Teacher SRS-2 total T-score ^a (38–90)			
N	120	126	246
Mean (SD)	71.2 (9.9)	72.9 (11.8)	72.1 (10.9)
Median (Q1, Q3)	72.0 (64.5, 79.0)	73.5 (62.0, 82.0)	73.0 (63.0, 80.0)
Minimum, maximum	48.0, 90.0	47.0, 90.0	47.0, 90.0
Parent SRS-2 total raw score ^a (0–195)			
N	120	129	249
Mean (SD)	116.5 (24.0)	114.4 (23.6)	115.4 (23.8)
Median (Q1, Q3)	118.0 (102.0, 132.0)	119.0 (101.0, 129.0)	118.0 (102.0, 130.0)
Minimum, maximum	50.0, 170.0	34.0, 161.0	34.0, 170.0
Parent SRS-2 total T-score ^a (37–100)			
N	120	129	249
Mean (SD)	82.5 (8.1)	82.0 (8.4)	82.3 (8.2)
Median (Q1, Q3)	84.0 (78.0, 90.0)	84.0 (78.0, 89.0)	84.0 (78.0, 90.0)
Minimum, maximum	57.0, 90.0	50.0, 90.0	50.0, 90.0
PSI total stress score ^a (36–180)			
N	117	127	244
Mean (SD)	103.2 (20.9)	104.9 (19.5)	104.1 (20.1)
Median (Q1, Q3)	104.0 (90.0, 117.0)	104.0 (94.0, 119.0)	104.0 (91.0, 118.0)
Minimum, maximum	47.0, 147.0	52.0, 147.0	47.0, 147.0
PSI Parental Distress Score ^a (12–60)			
N	119	128	247
Mean (SD)	32.6 (9.5)	32.7 (9.3)	32.6 (9.4)
Median (Q1, Q3)	32.0 (26.0, 39.0)	32.5 (26.0, 40.0)	32.0 (26.0, 39.0)
Minimum, maximum	13.0, 53.0	12.0, 55.0	12.0, 55.0
PSI Parent/Child Dysfunction Score ^a (12–60)			
N	120	128	248
Mean (SD)	31.6 (7.1)	33.1 (6.9)	32.4 (7.1)
Median (Q1, Q3)	31.5 (27.0, 36.0)	34.0 (28.5, 38.0)	33.0 (28.0, 37.0)
Minimum, maximum	14.0, 47.0	13.0, 51.0	13.0, 51.0

continued

TABLE 11 Baseline scores for teacher- and parent-reported outcomes and treatment preferences (range of possible scores are noted in parentheses for each outcome measure) (*continued*)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
PSI Difficult Child Score ^a (12–60)			
N	118	127	245
Mean (SD)	39.1 (8.1)	39.1 (6.9)	39.1 (7.5)
Median (Q1, Q3)	40.0 (34.0, 45.0)	39.0 (34.0, 44.0)	39.0 (34.0, 44.0)
Minimum, maximum	14.0, 56.0	21.0, 55.0	14.0, 56.0
RCADS total score ^a (0–141)			
N	120	128	248
Mean (SD)	45.7 (23.3)	44.7 (21.6)	45.2 (22.4)
Median (Q1, Q3)	45.0 (27.6, 59.5)	41.0 (28.5, 59.0)	43.5 (28.5, 59.0)
Minimum, maximum	2.0, 102.0	5.0, 100.0	2.0, 102.0
RCADS Social Phobia Score ^a (0–27)			
N	120	129	249
Mean (SD)	10.7 (6.9)	10.3 (7.0)	10.5 (7.0)
Median (Q1, Q3)	11.0 (5.0, 16.0)	10.0 (4.0, 15.0)	10.0 (5.0, 16.0)
Minimum, maximum	0.0, 27.0	0.0, 26.0	0.0, 27.0
RCADS Panic Disorder Score ^a (0–27)			
N	120	128	248
Mean (SD)	5.0 (4.4)	5.1 (4.2)	5.1 (4.3)
Median (Q1, Q3)	4.0 (2.0, 7.5)	4.0 (2.0, 8.0)	4.0 (2.0, 8.0)
Minimum, maximum	0.0, 20.0	0.0, 18.0	0.0, 20.0
RCADS Separation Anxiety Score ^a (0–21)			
N	120	129	249
Mean (SD)	9.1 (5.1)	8.7 (4.6)	8.9 (4.9)
Median (Q1, Q3)	9.0 (5.0, 13.0)	8.0 (5.8, 12.0)	9.0 (5.0, 12.0)
Minimum, maximum	0.0, 20.0	0.0, 21.0	0.0, 21.0
RCADS Generalised Anxiety Score ^a (0–18)			
N	120	128	248
Mean (SD)	7.2 (4.5)	6.4 (4.1)	6.8 (4.3)
Median (Q1, Q3)	6.0 (4.0, 10.0)	6.0 (3.0, 9.0)	6.0 (4.0, 9.5)
Minimum, maximum	0.0, 18.0	0.0, 16.0	0.0, 18.0
RCADS Obsessive – Compulsive Score ^a (0–18)			
N	120	128	248
Mean (SD)	4.6 (3.2)	4.7 (3.4)	4.6 (3.3)
Median (Q1, Q3)	4.0 (2.0, 7.0)	4.0 (2.0, 6.5)	4.0 (2.0, 7.0)
Minimum, maximum	0.0, 13.0	0.0, 15.6	0.0, 15.6

TABLE 11 Baseline scores for teacher- and parent-reported outcomes and treatment preferences (range of possible scores are noted in parentheses for each outcome measure) (*continued*)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
RCADS Major Depression Score ^a (0–30)			
N	120	129	249
Mean (SD)	9.2 (4.9)	9.3 (4.8)	9.3 (4.8)
Median (Q1, Q3)	9.0 (6.0, 13.0)	10.0 (5.0, 12.2)	9.0 (6.0, 13.0)
Minimum, maximum	0.0, 21.0	1.0, 21.0	0.0, 21.0
EQ-5D-Y VAS ^b (0–100)			
N	120	129	249
Mean (SD)	79.0 (19.0)	80.6 (18.8)	79.8 (18.9)
Median (Q1, Q3)	80.0 (70.0, 95.0)	90.0 (70.0, 95.0)	85.0 (70.0, 95.0)
Minimum, maximum	30.0, 100.0	25.0, 100.0	25.0, 100.0
Teacher treatment preference ^c (0–100)			
N	116	126	242
Mean (SD)	80.2 (22.8)	79.7 (22.0)	80.0 (22.3)
Median (Q1, Q3)	90.0 (60.0, 100.0)	87.5 (50.0, 100.0)	90.0 (50.0, 100.0)
Minimum, maximum	0.0, 100.0	25.0, 100.0	0.0, 100.0
Parent treatment preference ^c (0–100)			
N	120	129	249
Mean (SD)	85.3 (21.2)	83.3 (21.0)	84.3 (21.1)
Median (Q1, Q3)	100.0 (75.0, 100.0)	98.0 (70.0, 100.0)	100.0 (70.0, 100.0)
Minimum, maximum	20.0, 100.0	25.0, 100.0	20.0, 100.0

a Higher score indicates 'worse' outcome.

b Higher score indicates 'better' outcome.

c Higher scores indicate preference for Social Stories.

due to random error alone. This, coupled with the fact that randomisation was only completed after all baseline data were collected, suggests the observed differences are not the result of any systematic (i.e. non-random) processes or biases. Likewise, minor differences between groups are evident for the other teacher- and parent-reported secondary outcomes, but none are any more extreme than would be expected from purely random variation.

Parent, teacher and interventionist baseline demographics

Demographic data were also collected from participant's parents/carers, teachers and interventionists associated with the participating children at baseline. Demographic data for the parents/carers of the 249 randomised participants are given in [Table 12](#). Demographic data were available for 226 teachers associated with children that went on to be randomised and 184 interventionists associated with schools that went on to be randomised. These demographic data are summarised in [Tables 13](#) and [14](#), respectively.

TABLE 12 Demographics of parents/carers associated with the 249 randomised participants

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Relationship to participant, <i>n</i> (%)			
Parent	115 (95.8)	127 (98.4)	242 (97.2)
Other with parental responsibility	5 (4.2)	2 (1.6)	7 (2.8)
Sex, <i>n</i> (%)			
Male	6 (5.0)	12 (9.3)	18 (7.2)
Female	113 (94.2)	117 (90.7)	230 (92.4)
Prefer not to say	1 (0.8)	0 (0.0)	1 (0.4)
Age (years)			
<i>N</i>	120	129	249
Mean (SD)	38.3 (6.7)	39.2 (7.6)	38.8 (7.2)
Median (Q1, Q3)	38.0 (34.0, 43.0)	38.0 (34.0, 44.0)	38.0 (34.0, 43.0)
Minimum, maximum	25.0, 64.0	24.0, 66.0	24.0, 66.0
Ethnicity, <i>n</i> (%)			
White – British	106 (88.3)	110 (85.3)	216 (86.7)
White – Irish	1 (0.8)	0 (0.0)	1 (0.4)
White – other	4 (3.3)	3 (2.3)	7 (2.8)
Black – British	0 (0.0)	1 (0.8)	1 (0.4)
Black – African	1 (0.8)	1 (0.8)	2 (0.8)
Asian – Indian	2 (1.7)	0 (0.0)	2 (0.8)
Asian – Pakistani	1 (0.8)	6 (4.7)	7 (2.8)
Asian – Bangladeshi	1 (0.8)	0 (0.0)	1 (0.4)
Asian – Chinese	0 (0.0)	1 (0.8)	1 (0.4)
Asian – other	0 (0.0)	1 (0.8)	1 (0.4)
Mixed – white and Black Caribbean	1 (0.8)	1 (0.8)	2 (0.8)
Mixed – white and Asian	1 (0.8)	2 (1.6)	3 (1.2)
Mixed – other	0 (0.0)	1 (0.8)	1 (0.4)
Other – Arab	0 (0.0)	1 (0.8)	1 (0.4)
Other – other	0 (0.0)	1 (0.8)	1 (0.4)
Prefer not to say	2 (1.7)	0 (0.0)	2 (0.8)
Relationship status, <i>n</i> (%)			
Single	17 (14.2)	22 (17.1)	39 (15.7)
Married/civil partnership/cohabiting	85 (70.8)	88 (68.2)	173 (69.5)
Divorced/separated	17 (14.2)	18 (14.0)	35 (14.1)
Widowed	1 (0.8)	1 (0.8)	2 (0.8)

TABLE 12 Demographics of parents/carers associated with the 249 randomised participants (continued)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Highest level of education, n (%)			
Primary or less	1 (0.8)	1 (0.8)	2 (0.8)
Secondary school	18 (15.0)	20 (15.5)	38 (15.3)
Further education	49 (40.8)	54 (41.9)	103 (41.4)
Higher education	51 (42.5)	50 (38.8)	101 (40.6)
Other	0 (0.0)	1 (0.8)	1 (0.4)
Missing	1 (0.8)	3 (2.3)	4 (1.6)
Main occupation, n (%)			
Employed	69 (57.5)	75 (58.1)	144 (57.8)
Full-time parent/carer	40 (33.3)	49 (38.0)	89 (35.7)
Volunteer/between jobs	1 (0.8)	0 (0.0)	1 (0.4)
Student/trainee	3 (2.5)	2 (1.6)	5 (2.0)
Retired	2 (1.7)	2 (1.6)	4 (1.6)
Unemployed	5 (4.2)	0 (0.0)	5 (2.0)
Missing	0 (0.0)	1 (0.8)	1 (0.4)
Ever received support/therapy regarding child's ASC diagnosis, n (%)			
Yes	50 (41.7)	54 (41.9)	104 (41.8)
No	70 (58.3)	75 (58.1)	145 (58.2)

TABLE 13 Demographics of teachers associated with randomised participants at baseline

	Control (N = 118)	Intervention (N = 108)	Total (N = 226)
Sex, n (%)			
Male	25 (21.2)	18 (16.7)	43 (19.0)
Female	92 (78.0)	88 (81.5)	180 (79.6)
Missing	1 (0.8)	2 (1.9)	3 (1.3)
Age (years)			
N	117	108	225
Mean (SD)	36.5 (9.4)	37.4 (9.7)	36.9 (9.5)
Median (Q1, Q3)	35.0 (29.0, 45.0)	35.5 (29.5, 45.0)	35.0 (29.0, 45.0)
Minimum, maximum	21.0, 57.0	22.0, 67.0	21.0, 67.0
Time working with children/young people (years)			
N	117	108	225
Mean (SD)	13.7 (8.3)	13.1 (8.3)	13.4 (8.3)
Median (Q1, Q3)	12.0 (7.0, 20.0)	11.0 (7.5, 16.5)	12.0 (7.0, 18.0)
Minimum, maximum	1.0, 36.0	1.0, 45.0	1.0, 45.0

continued

TABLE 13 Demographics of teachers associated with randomised participants at baseline (continued)

	Control (N = 118)	Intervention (N = 108)	Total (N = 226)
Knowledge and experience of working with ASC, n (%)			
Limited	9 (7.6)	8 (7.4)	17 (7.5)
Moderate	47 (39.8)	60 (55.6)	107 (47.3)
Sound	48 (40.7)	32 (29.6)	80 (35.4)
In depth	14 (11.9)	8 (7.4)	22 (9.7)
Current professional role/occupation ^a			
Teacher/educator	101 (85.6)	99 (91.7)	200 (88.5)
TA	4 (3.4)	2 (1.9)	6 (2.7)
High-level TA	2 (1.7)	0 (0.0)	2 (0.9)
Learning mentor	2 (1.7)	1 (0.9)	3 (1.3)
SENCO	15 (12.7)	10 (9.3)	25 (11.1)
Emotional Literacy Support Assistant	0 (0.0)	0 (0.0)	0 (0.0)
Specialist autism teacher for local authority	2 (1.7)	1 (0.9)	3 (1.3)
Headteacher	0 (0.0)	0 (0.0)	0 (0.0)
Senior staff member	19 (16.1)	20 (18.5)	39 (17.3)
Other	5 (4.2)	4 (3.7)	9 (4.0)

^a Possibly more than one category per individual.

TABLE 14 Demographics of interventionists associated with randomised schools

	Control (N = 76)	Intervention (N = 108)	Total (N = 184)
Sex, n (%)			
Male	4 (5.3)	6 (5.6)	10 (5.4)
Female	70 (92.1)	102 (94.4)	172 (93.5)
Missing	2 (2.6)	0 (0.0)	2 (1.1)
Age (years)			
N	75	107	182
Mean (SD)	43.3 (10.8)	42.1 (11.3)	42.6 (11.1)
Median (Q1, Q3)	45.0 (35.0, 52.0)	43.0 (35.0, 51.0)	45.0 (35.0, 51.0)
Minimum, maximum	22.0, 62.0	20.0, 64.0	20.0, 64.0
Time working with children/young people (years)			
N	75	106	181
Mean (SD)	14.4 (7.8)	12.7 (8.6)	13.4 (8.3)
Median (Q1, Q3)	14.0 (8.0, 20.0)	11.0 (5.0, 19.0)	13.0 (7.0, 19.0)
Minimum, maximum	0.8, 35.0	1.0, 44.0	0.8, 44.0

TABLE 14 Demographics of interventionists associated with randomised schools (continued)

	Control (N = 76)	Intervention (N = 108)	Total (N = 184)
Knowledge and experience of working with ASC, n (%)			
Limited	7 (9.2)	24 (22.2)	31 (16.8)
Moderate	27 (35.5)	39 (36.1)	66 (35.9)
Sound	35 (46.1)	37 (34.3)	72 (39.1)
In depth	7 (9.2)	8 (7.4)	15 (8.2)
Current professional role/occupation ^a			
Teacher/educator	1 (1.3)	1 (0.9)	2 (1.1)
TA	64 (84.2)	90 (83.3)	154 (83.7)
High-level TA	11 (14.5)	8 (7.4)	19 (10.3)
Learning mentor	3 (3.9)	6 (5.6)	9 (4.9)
SENCO	1 (1.3)	1 (0.9)	2 (1.1)
Emotional Literacy Support Assistant	2 (2.6)	1 (0.9)	3 (1.6)
Specialist autism teacher for local authority	1 (1.3)	1 (0.9)	2 (1.1)
Headteacher	0 (0.0)	0 (0.0)	0 (0.0)
Senior staff member	0 (0.0)	0 (0.0)	0 (0.0)
Other	5 (6.6)	6 (5.6)	11 (6.0)

a Possibly more than one category per individual.

Intervention delivery

Of the 44 schools and 129 participants allocated to the Social Stories group, 95 (73.6%) participants [across 36 (81.8%) schools] had available intervention delivery data. Summaries of the number of Social Stories sessions attended/administered and the time frame in which these were conducted are given in [Tables 15](#) and [16](#) for the 95 participants with available session log data. The remaining 34 (26.4%) participants allocated to Social Stories were missing session log data and are therefore missing details of the number and timing of any sessions that were delivered.

Primary outcome analyses

Descriptive analyses

The SRS-2 was collected from participating children's teachers at baseline (prior to randomisation) and 6 weeks and 6 months post randomisation. The responses to the SRS-2 are used to derive a total score that is then converted to a T-score lying between 38 and 90 (inclusive). Of the 249 children randomised, 247 (99.2%) had available teacher-completed SRS-2 data, and for all but one of these participants, a valid total score could be derived. Hence, 246 (98.8%) participants had observed primary outcome data at baseline. Of the 249 randomised participants, 214 (85.9%) had observed primary outcome data at 6 weeks post randomisation, and 211 (84.7%) had observed primary outcome data at 6 months post randomisation (the primary end point). In total, 235 (94.4%) participants had observed primary outcome data at baseline and for at least one post-randomisation time point. Descriptive summaries of the primary outcome data at the three time points are given by allocation and overall in [Table 17](#). Plots of the distribution of the primary outcome at each time point are given in [Figures 5–7](#) in [Report Supplementary Material 7](#), stratified by randomised group.

TABLE 15 Intervention delivery by allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Number of Social Stories sessions delivered			
N	120	95	215
Mean (SD)	0.0 (0.0)	6.3 (1.9)	2.8 (3.4)
Median (Q1, Q3)	0.0 (0.0, 0.0)	6.0 (6.0, 7.0)	0.0 (0.0, 6.0)
Minimum, maximum	0.0, 0.0	0.0, 10.0	0.0, 10.0
Number of Social Stories sessions delivered, n (%)			
0	120 (100.0)	1 (0.8)	121 (48.6)
1	0 (0.0)	1 (0.8)	1 (0.4)
2	0 (0.0)	3 (2.3)	3 (1.2)
3	0 (0.0)	1 (0.8)	1 (0.4)
4	0 (0.0)	2 (1.6)	2 (0.8)
5	0 (0.0)	6 (4.7)	6 (2.4)
6	0 (0.0)	51 (39.5)	51 (20.5)
7	0 (0.0)	14 (10.9)	14 (5.6)
8	0 (0.0)	3 (2.3)	3 (1.2)
9	0 (0.0)	3 (2.3)	3 (1.2)
10	0 (0.0)	10 (7.8)	10 (4.0)
Missing	0 (0.0)	34 (26.4)	34 (13.7)
Number of Social Stories sessions delivered, n (%)			
Attended 0 sessions	120 (100.0)	1 (0.8)	121 (48.6)
Attended 1–5 sessions	0 (0.0)	13 (10.1)	13 (5.2)
Attended 6 + sessions	0 (0.0)	81 (62.8)	81 (32.5)
Missing	0 (0.0)	34 (26.4)	34 (13.7)

TABLE 16 Time between randomisation and first and last Social Stories sessions (participants that had ≥ 1 session only)

		N = 94
Time to first Social Stories session (weeks)		
Mean		3.4 (1.3)
Median (Q1, Q3)		3.4 (2.7, 3.9)
Minimum, maximum		0.6, 9.3
Time to last Social Stories session (weeks)		
Mean		6.9 (2.8)
Median (Q1, Q3)		6.3 (5.3, 8.0)
Minimum, maximum		1.7, 16.0

TABLE 17 Teacher-reported SRS-2 total T-scores by allocation and overall

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Baseline^a			
N	120	126	246
Mean (SD)	71.2 (9.9)	72.9 (11.8)	72.1 (10.9)
Median (Q1, Q3)	72.0 (64.5, 79.0)	73.5 (62.0, 82.0)	73.0 (63.0, 80.0)
Minimum, maximum	48.0, 90.0	47.0, 90.0	47.0, 90.0
Week 6^a			
N	110	104	214
Mean (SD)	70.5 (11.4)	71.2 (11.8)	70.8 (11.5)
Median (Q1, Q3)	71.0 (62.0, 79.0)	72.0 (63.0, 80.5)	71.0 (63.0, 80.0)
Minimum, maximum	44.0, 90.0	46.0, 90.0	44.0, 90.0
Month 6^a			
N	108	103	211
Mean (SD)	70.6 (10.2)	69.9 (11.6)	70.3 (10.9)
Median (Q1, Q3)	71.0 (64.0, 78.0)	70.0 (61.0, 78.0)	71.0 (63.0, 78.0)
Minimum, maximum	44.0, 90.0	45.0, 90.0	44.0, 90.0

^a Score between 38 and 90, where higher scores indicate 'worse' outcome.

Primary analysis

Of the 87 schools (249 children) randomised, 83 (95.4%) schools [238 (95.6%) children] had at least one available follow-up measurement. However, three of these participants were missing a baseline score (due to there being ≥ 8 missing items for their baseline teacher-reported SRS-2 assessment, meaning a baseline T-score could not be derived). In the absence of pre-specified procedures for addressing missing baseline covariate data, these three participants were excluded from the primary analysis model. Hence, the primary analysis model included 235 (94.4%) participants across 83 school clusters. Point and 95% CI estimates for the differences in expected teacher-reported SRS-2 total T-score at 6 weeks and 6 months from the fitted model are given in [Table 18](#) and are illustrated in [Figure 8](#) in [Report Supplementary Material 7](#). Basic diagnostic plots for the fitted primary analysis model are given in [Figures 9](#) and [10](#) in [Report Supplementary Material 7](#).

Conditional on all of the assumptions used to calculate them, the point estimates in [Table 18](#) suggest the data are most compatible with the hypotheses positing that allocation to Social Stories causes small (around half the target difference used for planning) reductions in teacher-reported SRS-2 scores at 6 weeks and 6 months following randomisation. However, the *p*-values in [Table 18](#) for the tests of

TABLE 18 Treatment effect estimates at 6 weeks and 6 months post randomisation from the planned primary analysis model

	Estimated difference (intervention – control) in expected value (95% CI) ^a	<i>p</i> -value ^a
Week 6	-1.14 (-3.35 to 1.06)	0.310
Month 6 (primary end point)	-1.61 (-4.18 to 0.96)	0.220

^a Based on delta method standard errors.

$H_0 : \delta = 0$ at each time point suggest that the observed differences would not be particularly surprising if allocation truly has no effect on these outcomes. Hence, the observed data are quite compatible with allocation to Social Stories, not causing any reduction in scores at 6 weeks and 6 months compared with control (again contingent on the manifold assumptions that the various estimates and test statistics are based on). The residual plots in [Figures 12 and 13](#) in [Report Supplementary Material 7](#) suggest departures from the key distributional assumptions of the fitted model are relatively limited and are unlikely to have any material effect on the substantive conclusions of the primary analysis. For example, removal of the extreme and potentially influential participant identified in [Figure 9](#) results in fairly modest reductions of the estimated treatment effects at 6 weeks and 6 months [6 weeks -0.89 (95% CI -3.00 to 1.21); 6 months -1.59 (95% CI -4.15 to 0.96)], but the substantive conclusions remain relatively unaffected.

Subgroup analyses

Teacher treatment preference

Teachers reported their treatment preference at baseline using a 0–100 VAS, where 0 indicates strong preference for usual care (for the relevant child) and 100 indicates strong preference for Social Stories (for the relevant child). For the purposes of the subgroup analyses reported here, the preference score was used to classify children into three categories: (1) Prefers usual care if < 50 ; (2) No preference if $= 50$ and (3) Prefers Social Stories if > 50 . Summaries of the raw scores and the derived categories are given in [Table 19](#). The distribution of observed preference scores is illustrated in [Figure 11](#) in [Report Supplementary Material 7](#).

TABLE 19 Subgroup summaries by allocation and overall

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Treatment preference score ^a (0–100)			
N	116	126	242
Mean (SD)	80.2 (22.8)	79.7 (22.0)	80.0 (22.3)
Median (Q1, Q3)	90.0 (60.0, 100.0)	87.5 (50.0, 100.0)	90.0 (50.0, 100.0)
Minimum, maximum	0.0, 100.0	25.0, 100.0	0.0, 100.0
Treatment preference category, ^b n (%)			
Prefers usual care	3 (2.5)	1 (0.8)	4 (1.6)
No preference	25 (20.8)	34 (26.4)	59 (23.7)
Prefers Social Stories	88 (73.3)	91 (70.5)	179 (71.9)
Missing	4 (3.3)	3 (2.3)	7 (2.8)
Diagnosed mental health or psychological problem, n (%)			
Yes	16 (13.3)	15 (11.6)	31 (12.4)
No	86 (71.7)	98 (76.0)	184 (73.9)
Missing	18 (15.0)	16 (12.4)	34 (13.7)
Diagnosed cognitive problems/learning difficulties, n (%)			
Yes	19 (15.8)	23 (17.8)	42 (16.9)
No	83 (69.2)	90 (69.8)	173 (69.5)
Missing	18 (15.0)	16 (12.4)	34 (13.7)

a Higher scores indicate preference for Social Stories.

b Prefers usual care if score < 50 , No preference if score $= 50$ and prefers Social Stories otherwise.

Of the 249 participants that were randomised, 230 (92.4%) had observed primary outcome data (week 6 and/or month 6), complete baseline covariate data and complete baseline treatment preference data. To estimate treatment effects at 6 weeks and 6 months by treatment preference category (Prefers usual care, No preference, Prefers Social Stories), the primary analysis model was augmented with additional terms for the main effect of preference subgroup, the two-way interactions between preference and allocation and preference and time point and the three-way interaction between preference, allocation and time point.

The treatment effect estimates, 95% CIs and *p*-values for the 6 time points by subgroup strata are given in [Table 20](#). While the point estimates do show some variation across these strata, this apparent variation is not any more than might be expected, even if there truly is no variation in treatment effect across these strata. The *p*-value for the likelihood ratio test of the full model (main effects and two/three-way interactions for allocation, time point and treatment preference) and the constrained model (main effects for allocation, time point and treatment preference and two-way interactions between allocation and time point and treatment preference and time point) is 0.75. Hence, overall, there is little evidence for any interaction between treatment preference and allocation. However, given the sparsity of the data, the absence of evidence for any interaction provides relatively little information concerning the presence/absence of treatment effect heterogeneity across treatment preference strata.

Diagnosed mental health or psychological problems

Parents of participating children reported whether their child had any diagnosed mental health and/or psychological problems at the time of the baseline data collection. A summary of the binary (Yes/No) responses is given in [Table 19](#). The majority of participants (73.9%) had no diagnosed mental health

TABLE 20 Treatment effect estimates at 6 weeks and 6 months by teacher-reported baseline treatment preference

Time	Subgroup	Estimated difference (intervention – control) in expected value (95% CI) ^a	<i>p</i> -value ^a
Treatment preference			
Week 6	Prefers usual care	4.25 (–13.13 to 21.62)	0.632
	No preference	–0.00 (–4.39 to 4.38)	0.998
	Prefers Social Stories	–1.82 (–4.33 to 0.69)	0.155
Month 6	Prefers usual care	–1.75 (–21.70 to 18.19)	0.863
	No preference	–2.89 (–7.89 to 2.11)	0.257
	Prefers Social Stories	–1.37 (–4.23 to 1.50)	0.350
Diagnosed mental health/psychological problem			
Week 6	Yes	–2.43 (–8.09 to 3.24)	0.401
	No	–1.08 (–3.71 to 1.56)	0.424
Month 6	Yes	–3.17 (–10.06 to 3.72)	0.367
	No	–1.07 (–4.08 to 1.93)	0.483
Diagnosed cognitive problem or learning difficulty			
Week 6	Yes	–4.26 (–9.58 to 1.06)	0.117
	No	–0.60 (–3.28 to 2.09)	0.662
Month 6	Yes	–0.52 (–6.60 to 5.57)	0.868
	No	–1.79 (–4.90 to 1.32)	0.260

^a Based on delta method standard errors.

problem, but a substantial minority (12.4%) had a diagnosis reported. Furthermore, 34 participants (13.7%) were missing information on mental health and/or psychological problem diagnoses. Of the 249 children randomised, 238 (95.6%) had observed primary outcome data at week 6 and/or month 6. Of these 238 children, 33 (13.9%) were missing information on mental health and/or psychological diagnoses, and three (1.3%) were missing other baseline covariate data. Hence, 202 participants had observed primary outcome data (week 6 and/or month 6), complete baseline covariate data and complete baseline mental health/psychological diagnoses data. The treatment effect estimates, 95% CIs and *p*-values are given by subgroup and time point in [Table 20](#). These appear to be relatively stable across subgroups and time points, with a *p*-value for the relevant likelihood ratio test of interaction of 0.82. However, the 95% CIs contain a reasonably large range of values due to relatively sparse data in some strata. Hence, the observed data are not entirely at odds with hypotheses positing moderate treatment effect heterogeneity by mental health diagnosis subgroup.

Diagnosed cognitive problems and/or learning difficulties

Parents of participating children reported whether their child had any diagnosed cognitive problems and/or learning difficulties at the time of [Table 19](#). The majority of participants (69.5%) had no diagnosed cognitive problem/learning difficulty, but a substantial minority (16.9%) had a diagnosis reported. Furthermore, 34 participants (13.7%) were missing information on diagnosed cognitive problems/learning difficulties.

The treatment effect estimates, 95% CIs and *p*-values are given in [Table 20](#). There is some evidence that treatment effect at 6 weeks is larger in the cognitive problem/learning difficulty subgroup than in the subgroup without any cognitive problem/learning difficulty diagnoses. However, this apparent pattern disappears and is even reversed at 6 months, and the *p*-value for the relevant test of interaction is 0.27, suggesting this apparent variation in treatment effect would be relatively unsurprising even if there were truly no treatment effect heterogeneity present. However, the 95% CIs do contain a reasonably large range of values due to relatively sparse data in some strata. Hence, potentially important variations by subgroup cannot be completely ruled out.

Complier-average causal effect estimation

Of the 129 participants allocated to the intervention, 95 (73.6%) had available session log data, meaning 34 (26.4%) had no available session log data. Due to the clustered randomisation, participants allocated to the control group were assumed to have not received any Social Stories sessions. Social Stories delivery/attendance is reported by allocation in [Table 21](#).

Of the 249 participants randomised, 214 (85.9%) had available primary outcome data at 6 weeks post randomisation and 211 (84.7%) had available primary outcome data at 6 months post randomisation. Of the 214 participants with available 6-week data, 23 (10.7%) were missing compliance status (allocated to the intervention but no session log data available) and 2 (0.9%) had complete compliance data but were missing a baseline score. Similarly, of the 211 participants with available 6-month outcome data, 24 (11.4%) were missing compliance status (allocated to the intervention but no session log data available) and 3 (1.4%) had complete compliance data but missing baseline score. Hence, 189 participants were included in the analysis to estimate the CACE at 6 weeks (79 intervention, 110 control) and 184 in the

TABLE 21 Social Stories sessions delivered/attended by allocation (dichotomised at 6 in accordance with the definition and estimation of the CACE)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Number of Social Stories sessions reported (binary), <i>n</i> (%)			
Attended 0–5 sessions	120 (100.0)	14 (10.9)	134 (53.8)
Attended 6 + sessions	0 (0.0)	81 (62.8)	81 (32.5)
Missing	0 (0.0)	34 (26.4)	34 (13.7)

analysis to estimate the CACE at 6 months (76 intervention, 108 control). Point estimates and 95% CIs for the CACE at 6 weeks and 6 months are reported in [Table 22](#), together with *p*-values for a two-sided test of the null hypothesis that treatment receipt has no effect on outcomes.

The estimates of the CACE in [Table 22](#) suggest the data are quite compatible with the hypothesis that Social Stories cause a modest reduction in SRS-2 score at 6 weeks and 6 months for participants in the complier principal stratum (i.e. the latent subgroup of participants that would receive ≥ 6 Social Stories sessions if they were offered it). For example, the interval estimate and *p*-value for the CACE at 6 months suggest that the observed data would be reasonably surprising if the true effect of treatment was zero for participants in the complier principal stratum.

However, some caution is required in interpreting these estimates because a substantial number of intervention group participants have been excluded from the estimation due to having missing compliance data. It is quite plausible that those in the intervention group who are missing compliance data (and therefore excluded from the estimation of the CACE estimands) are not a random subset of the intervention group participants. In particular, they may be a subset of participants with poorer prognosis (treated or untreated) compared with those that have available compliance data. The summaries in [Table 23](#) provide some evidence that this is the case. Hence, exclusion of these intervention participants is likely to have resulted in the intervention and control groups that are no

TABLE 22 Complier-average causal effect estimates at 6 weeks and 6 months

	Control (N = 120)	Intervention (N = 129)	CACE estimate (95% CI)	<i>p</i> -value
Week 6	110 (91.7%)	79 (61.2%)	-2.35 (-5.44 to 0.73)	0.134
Month 6	108 (90.0%)	76 (58.9%)	-3.37 (-6.65 to -0.10)	0.043

TABLE 23 Primary outcome data summaries by availability of compliance data (intervention group participants with outcome data available for at least one follow-up)

	Compliance data missing (N = 28)	Compliance data not missing (N = 92)	Total (N = 120)
Baseline SRS-2 total T-score			
<i>N</i>	28	89	117
Mean (SD)	74.8 (10.1)	72.9 (12.2)	73.4 (11.7)
Median (Q1, Q3)	77.5 (67.5, 82.5)	74.0 (62.0, 83.0)	75.0 (62.0, 83.0)
Minimum, maximum	57.0, 90.0	47.0, 90.0	47.0, 90.0
Week 6 SRS-2 total T-score			
<i>N</i>	23	81	104
Mean (SD)	74.1 (10.7)	70.3 (12.0)	71.2 (11.8)
Median (Q1, Q3)	77.0 (66.0, 81.0)	71.0 (62.0, 80.0)	72.0 (63.0, 80.5)
Minimum, maximum	51.0, 90.0	46.0, 90.0	46.0, 90.0
Month 6 SRS-2 total T-score			
<i>N</i>	24	79	103
Mean (SD)	74.6 (10.9)	68.5 (11.5)	69.9 (11.6)
Median (Q1, Q3)	74.5 (70.0, 82.5)	67.0 (61.0, 77.0)	70.0 (61.0, 78.0)
Minimum, maximum	49.0, 90.0	45.0, 90.0	45.0, 90.0

longer exchangeable with respect to their potential outcomes, with the control group having poorer prognosis (on average) than the intervention group.

To assess the possible influence of the missing compliance data, two additional post hoc estimates of the CACE (at both 6 weeks and 6 months) were obtained. The first was undertaken assuming participants that were missing compliance data received < 6 Social Stories sessions (i.e. were not compliers), and a second assuming these participants received ≥ 6 sessions (i.e. were compliers). The results of these analyses are reported in [Table 24](#). Clearly, these estimates have their limitations (e.g. uncertainty about the missing imputed values not adequately accounted for) and should also be interpreted with caution. However, they do suggest that the magnitudes of the CACE estimates presented in [Table 22](#) are at least partly explained by selection bias arising from exclusion of intervention participants with missing compliance data. That said, the results in [Table 24](#) again suggest the data are most compatible with hypotheses positing modest positive effects of treatment in the complier principal stratum (although hypotheses positing no effect or small negative effects are also reasonably compatible with the observed data).

Missing outcome data

The number of available/complete responses to the various teacher- and parent-reported outcomes is given in [Table 25](#). For both the teacher- and parent-reported outcomes, the response rates were generally somewhat lower in the intervention group than in the control group.

The different patterns of primary outcome data missingness are reported by allocation in [Appendix 1, Table 48](#). From this, it is evident that a considerably higher proportion of the control group had complete baseline and follow-up data for the primary outcome (65.9% of the intervention group had scores for the primary outcome at all three time points vs. 83.3% of the control group). Similarly, the proportion of participants included in the primary analysis (i.e. all those with primary outcome data available for at least one time point and complete baseline covariate data) was around 8% higher in the control group than in the intervention group (90.7% in the intervention group vs. 98.3% in the control group). Key baseline characteristics are reported by inclusion in/exclusion from the primary analysis in [Appendix 1, Table 49](#). Given the relatively small number of participants excluded from the primary analysis, it is difficult to discern the extent to which these participants differ from those that were included. However, there are no immediately obvious patterns suggesting that those excluded differed greatly at baseline from those that were included.

The planned primary analysis assumes that the missing outcome data are MAR conditional on the observed baseline and outcome data included in the model. We believe this assumption to be reasonably plausible in this context given that missingness of primary outcome data is likely to be driven by teacher and/or school-level factors rather than being driven by the outcomes for these patients being better/worse than would be expected given their observed baseline/outcome data. However, to assess

TABLE 24 Complier-average causal effect estimates at 6 weeks and 6 months assuming intervention group participants with missing compliance status received ≥ 6 sessions (i.e. were compliers)

	Control (N = 120) (%)	Intervention (N = 129) (%)	CACE estimate (95% CI)	p-value
Assuming intervention group with missing compliance status were compliers				
Week 6	110 (91.7)	102 (79.1)	-1.63 (-4.27 to 1.00)	0.225
Month 6	108 (90.0)	100 (77.5)	-1.87 (-4.68 to 0.93)	0.190
Assuming intervention group with missing compliance status were not compliers				
Week 6	110 (91.7)	102 (79.1)	-2.25 (-5.85 to 1.36)	0.222
Month 6	108 (90.0)	100 (77.5)	-2.59 (-6.27 to 1.08)	0.167

TABLE 25 Completeness of teacher- and parent-reported outcome data by allocation

	Control (N = 120) (%)	Intervention (N = 129) (%)	Total (N = 249) (%)
Teacher			
SRS-2 total raw score (teacher-reported)			
Baseline	120 (100.0)	126 (97.7)	246 (98.8)
Week 6	110 (91.7)	104 (80.6)	214 (85.9)
Month 6	108 (90.0)	103 (79.8)	211 (84.7)
SRS-2 total T-score (teacher-reported)			
Baseline	120 (100.0)	126 (97.7)	246 (98.8)
Week 6	110 (91.7)	104 (80.6)	214 (85.9)
Month 6	108 (90.0)	103 (79.8)	211 (84.7)
Goal-based outcome score (teacher-reported)			
Baseline	114 (95.0)	126 (97.7)	240 (96.4)
Week 6	109 (90.8)	104 (80.6)	213 (85.5)
Month 6	105 (87.5)	102 (79.1)	207 (83.1)
Parent			
SRS-2 total raw score (parent-reported)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	89 (74.2)	95 (73.6)	184 (73.9)
SRS-2 total T-score (parent-reported)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	89 (74.2)	95 (73.6)	184 (73.9)
PSI total score (parent-reported)			
Baseline	117 (97.5)	127 (98.4)	244 (98.0)
Week 6	95 (79.2)	89 (69.0)	184 (73.9)
Month 6	84 (70.0)	90 (69.8)	174 (69.9)
RCADS total score (parent-reported)			
Baseline	120 (100.0)	128 (99.2)	248 (99.6)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	88 (73.3)	94 (72.9)	182 (73.1)
EQ-5D-Y mobility (parent-reported)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	96 (80.0)	97 (75.2)	193 (77.5)
Month 6	89 (74.2)	94 (72.9)	183 (73.5)

continued

TABLE 25 Completeness of teacher- and parent-reported outcome data by allocation (*continued*)

	Control (N = 120) (%)	Intervention (N = 129) (%)	Total (N = 249) (%)
EQ-5D-Y washing/dressing (parent-reported)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	96 (74.4)	193 (77.5)
Month 6	89 (74.2)	95 (73.6)	184 (73.9)
EQ-5D-Y usual activities (parent-reported)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	89 (74.2)	95 (73.6)	184 (73.9)
EQ-5D-Y pain/discomfort (parent-reported)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	89 (74.2)	94 (72.9)	183 (73.5)
EQ-5D-Y worried/sad/unhappy (parent-reported)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	89 (74.2)	95 (73.6)	184 (73.9)
EQ-5D-Y VAS (parent-reported)			
Baseline	120 (100.0)	129 (100.0)	249 (100.0)
Week 6	97 (80.8)	97 (75.2)	194 (77.9)
Month 6	89 (74.2)	93 (72.1)	182 (73.1)

the robustness of the estimates obtained for the primary analysis to departures from MAR, we consider treatment effect estimates obtained assuming the missing outcome data exhibit various systematic departures from MAR. To do this, we performed a delta-based sensitivity analysis using a pattern mixture model, where the departures from MAR occur in each group individually (with MAR assumed for the other group) and both arms together. Treatment effect estimates for $\delta = -5$ [i.e. expected value of the unobserved scores is 5 points less ('better') than the expected value of the observed scores conditional on the covariates in the substantive model] to $\delta = 5$ [i.e. expected value of the unobserved scores is 5 points higher ('worse') than the expected value of the observed scores conditional on the covariates in the substantive model] in increments of 2.5 are provided in [Table 26](#). The estimates for 6 weeks and 6 months are also illustrated in [Figures 12](#) and [13](#) in [Report Supplementary Material 7](#), respectively. The estimates for the 6-month outcome data in [Table 26](#) show that the estimated treatment effect is reasonably stable across the range of departures from MAR considered. For instance, the point estimate of the treatment effect has the same sign across all of the missing not at random scenarios considered. Similar comments apply to the 6-week estimates.

Data collection timelines

Descriptive summaries of the timing of follow-up data collection are given by allocation in [Tables 27](#) and [28](#) for the 6-week and 6-month outcome data, respectively. The distribution of the observed follow-up times is illustrated in [Figures 14](#) and [15](#) in [Report Supplementary Material 7](#) for the 6-week and 6-month follow-ups, respectively.

TABLE 26 Estimated treatment effects at 6 weeks and 6 months under various departures from MAR

δ	Treatment effect (95% CI)		
	Social Stories only	Usual care only	Both groups
6 weeks			
-5	-2.44 (-4.93 to 0.06)	-1.01 (-3.45 to 1.44)	-2.01 (-4.52 to 0.50)
-2.5	-1.93 (-4.38 to 0.51)	-1.22 (-3.65 to 1.21)	-1.72 (-4.17 to 0.73)
0	-1.43 (-3.86 to 0.99)	-1.43 (-3.86 to 0.99)	-1.43 (-3.86 to 0.99)
2.5	-0.93 (-3.37 to 1.51)	-1.64 (-4.07 to 0.79)	-1.14 (-3.59 to 1.31)
5	-0.43 (-2.92 to 2.07)	-1.85 (-4.30 to 0.59)	-0.85 (-3.36 to 1.66)
6 months			
-5	-2.98 (-5.34 to -0.62)	-1.46 (-3.80 to 0.87)	-2.48 (-4.86 to -0.09)
-2.5	-2.47 (-4.79 to -0.16)	-1.72 (-4.03 to 0.60)	-2.22 (-4.55 to 0.10)
0	-1.97 (-4.27 to 0.34)	-1.97 (-4.27 to 0.34)	-1.97 (-4.27 to 0.34)
2.5	-1.46 (-3.78 to 0.86)	-2.22 (-4.53 to 0.09)	-1.71 (-4.04 to 0.61)
5	-0.95 (-3.31 to 1.41)	-2.47 (-4.81 to -0.14)	-1.46 (-3.84 to 0.93)

TABLE 27 Teacher week 6 data collection timelines

	Control (N = 110)	Intervention (N = 105)	Total (N = 215)
Time between randomisation and teacher week 6 follow-up (weeks)			
N	110	105	215
Mean (SD)	9.5 (6.3)	9.8 (5.5)	9.6 (5.9)
Median (Q1, Q3)	6.9 (6.0, 9.4)	8.0 (6.3, 10.4)	7.3 (6.1, 10.1)
Minimum, maximum	4.0, 31.4	4.4, 27.9	4.0, 31.4
Teacher week 6 follow-up inside protocol specified window ^a			
No	21 (19.1)	33 (31.4)	54 (25.1)
Yes	89 (80.9)	72 (68.6)	161 (74.9)
Teacher week 6 SRS-2 total T-scores for participants followed up inside protocol specified window ^a			
N	89	72	161
Mean (SD)	70.1 (12.1)	70.9 (12.7)	70.4 (12.4)
Median (Q1, Q3)	71.0 (61.0, 79.0)	72.5 (62.0, 81.0)	71.0 (61.0, 80.0)
Minimum, maximum	44.0, 90.0	46.0, 90.0	44.0, 90.0
Teacher week 6 SRS-2 total T-scores for participants followed up outside protocol specified window ^a			
N	21	32	53
Mean (SD)	72.2 (7.2)	71.8 (9.5)	71.9 (8.6)
Median (Q1, Q3)	71.0 (68.0, 79.0)	71.5 (64.0, 79.0)	71.0 (67.0, 79.0)
Minimum, maximum	60.0, 88.0	51.0, 90.0	51.0, 90.0

a 6 weeks -2/+4 weeks.

TABLE 28 Teacher month 6 data collection timelines

	Control (N = 108)	Intervention (N = 104)	Total (N = 212)
Time between randomisation and teacher month 6 follow-up (weeks)			
N	108	104	212
Mean (SD)	31.8 (6.5)	32.1 (9.7)	32.0 (8.2)
Median (Q1, Q3)	30.7 (26.4, 34.3)	28.5 (26.4, 33.0)	29.0 (26.4, 34.3)
Minimum, maximum	24.3, 60.7	24.0, 67.3	24.0, 67.3
Teacher month 6 follow-up within protocol specified window ^a			
No	29 (26.9)	25 (24.0)	54 (25.5)
Yes	79 (73.1)	79 (76.0)	158 (74.5)
Teacher month 6 SRS-2 total T-scores for participants followed up inside protocol specified window ^a			
N	79	78	157
Mean (SD)	70.1 (10.7)	69.5 (11.4)	69.8 (11.0)
Median (Q1, Q3)	71.0 (64.0, 76.0)	68.5 (61.0, 77.0)	70.0 (63.0, 77.0)
Minimum, maximum	44.0, 90.0	45.0, 90.0	44.0, 90.0
Teacher month 6 SRS-2 total T-scores for participants followed up outside protocol specified window ^a			
N	29	25	54
Mean (SD)	71.8 (8.7)	71.4 (12.5)	71.6 (10.5)
Median (Q1, Q3)	72.0 (66.0, 78.0)	72.0 (62.0, 82.0)	72.0 (62.0, 79.0)
Minimum, maximum	53.0, 87.0	49.0, 90.0	49.0, 90.0

^a 6 months -4/+ 8 weeks.

From [Table 27](#) it is apparent that while half of the participants with available week 6 outcome data had these data collected between 6 and 10 weeks post randomisation, approximately 25% were collected at more than 10 weeks post randomisation, and approximately 10% at more than 20 weeks post randomisation. Delays in follow-up data collection at 6 weeks were slightly greater on average in the intervention arm, with a considerably higher proportion of week 6 follow-ups in the intervention arm being collected outside of the pre-specified window for completion (31.4% in the intervention group and 19.1% in the control group). Similarly, approximately half of the teacher follow-ups at 6 months took place between 26 and 34 weeks (in line with the 6 months + 8 weeks window specified), with around 25% being more than 34 weeks after randomisation and around 10% being more than 40 weeks after randomisation. In contrast with the week 6 follow-ups, the timing of the month 6 follow-ups was generally more similar between groups, with similar mean and median follow-up times and similar proportions taking place within the protocol-specified windows (24.0% in the intervention group and 26.9% in the control group). For both time points, there are seemingly no dramatic differences in outcome scores between those followed up within window and those followed up outside.

To examine the possible impact of mistimed data collection (e.g. due to delays resulting from school holidays), analyses were conducted using only primary outcome data that were collected within the data collection windows specified in the SAP, namely -2/+ 4 weeks for the 6-week follow-up and -4/+ 8 weeks for the 6-month follow-up. Treatment effect estimates at 6 weeks and 6 months from this analysis are reported in [Table 29](#). The estimates in [Table 29](#) are broadly similar to those obtained for the primary analysis (i.e. small point estimate in favour of the intervention with a range of small-to-moderate positive treatment effects or small negative treatment effects being compatible with the data), and the

TABLE 29 Treatment effect estimates at 6 weeks and 6 months post randomisation obtained from the primary analysis model fitted using only data collected within the specified windows at 6 weeks and 6 months

	Estimated difference (intervention – control) in expected value (95% CI) ^a	p-value ^a
Week 6	-1.77 (-4.30 to 0.77)	0.172
Month 6	-0.82 (-3.83 to 2.20)	0.596

a Based on delta method standard errors.

substantive conclusions are relatively unaffected. However, the point estimate at 6 weeks moved slightly further from zero, and the point estimate at 6 months moved slightly closer to zero, suggesting that any treatment effect wanes somewhat over time (in contrast to the primary analysis, which suggested a slight increase in treatment effect over time).

Coronavirus disease impact analyses

A crude overview of the timing of teacher-reported SRS-2 data collection is given in [Figure 16](#) in [Report Supplementary Material 7](#). From this, it is evident that numerous participants had follow-ups at 6 weeks delayed by 6 weeks or more, with some participants (particularly those randomised just prior to the pandemic) having week 6 follow-ups completed at around 6 months post randomisation. Similar delays are also evident for the 6-month follow-ups. Details regarding the timing for teacher follow-up data collection are provided by allocation in [Table 30](#). This table clearly shows that while most participants had their week 6 data provided within -2/+ 4 weeks of the planned follow-up date and month 6 data provided within -4/+ 8 weeks of the planned follow-up date, there is a substantial minority whose follow-ups were severely delayed (approximately 12% and 14% at week 6 and month 6, respectively).

To assess whether delays in the completion of the primary outcome (at month 6) were associated with variation in treatment effect, we refitted the primary analysis model with an extra term denoting delayed outcome completion (binary – defined as the month 6 follow-up being completed more than 8 weeks after the planned follow-up date) and additional terms for all of the two- and three-way interactions between this variable, allocation and time point. The estimates from this augmented model are given in [Table 31](#). From this table, there is some weak evidence that delayed primary outcome ascertainment/collection was associated with larger treatment effects at both 6 weeks and 6 months, particularly at the later time point.

TABLE 30 Teacher-completed follow-up data collection timelines

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Teacher week 6 follow-up completion, n (%)			
Week 6 not completed	10 (8.3)	24 (18.6)	34 (13.7)
Followed up at (4, 10) weeks	89 (74.2)	72 (55.8)	161 (64.7)
Followed up at (10, 12) weeks	2 (1.7)	16 (12.4)	18 (7.2)
Followed up at (12, 14) weeks	4 (3.3)	3 (2.3)	7 (2.8)
Followed up at > 14 weeks	15 (12.5)	14 (10.9)	29 (11.6)
Teacher month 6 follow-up completion, n (%)			
Month 6 not completed	12 (10.0)	25 (19.4)	37 (14.9)
Followed up at (22, 34) weeks	79 (65.8)	79 (61.2)	158 (63.5)
Followed up at (34, 38) weeks	10 (8.3)	9 (7.0)	19 (7.6)
Followed up at > 38 weeks	19 (15.8)	16 (12.4)	35 (14.1)

TABLE 31 Treatment effect estimates at 6 weeks and 6 months by different sensitivity analyses

Timing	Response	Estimated difference (intervention – control) in expected value (95% CI) ^a	p-value ^a
6 months follow-up completed more than 8 weeks after planned follow-up date			
Week 6	No	-0.69 (-3.12 to 1.75)	0.580
	Yes	-2.68 (-7.39 to 2.03)	0.264
Month 6	No	-0.34 (-3.27 to 2.59)	0.819
	Yes	-5.41 (-10.45 to -0.37)	0.035
6 months follow-up due during 23 March 2020 to 3 September 2020			
Week 6	No	-2.48 (-4.96 to 0.00)	0.050
	Yes	2.17 (-1.85 to 6.18)	0.291
Month 6	No	-2.19 (-5.08 to 0.71)	0.139
	Yes	-0.35 (-5.41 to 4.72)	0.893
Follow-up (6-week or 6-month) completed after 23 March 2020			
Week 6	No	-0.31 (-4.70 to 4.08)	0.890
	Yes	-1.35 (-3.95 to 1.25)	0.310
Month 6	No	-5.07 (-10.49 to 0.34)	0.066
	Yes	-0.63 (-3.56 to 2.31)	0.676

^a Based on delta method standard errors.

To investigate the extent to which disruption of follow-up due to the pandemic was associated with variation in treatment effects, we refitted the primary analysis model with an additional term denoting whether or not the participant was due for their month 6 follow-up during the school closures in the first UK lockdown (23 March 2020 to 3 September 2020) and additional terms for all of the two- and three-way interactions between this variable, allocation and time point. The estimates from this augmented model are given in [Table 31](#). From this table, there is some weak evidence that disruption to follow-up due to school closures is associated with variation in treatment effect at both 6 weeks and 6 months, particularly at the earlier time point. However, substantial uncertainty is evident (due to the relatively small effective sample size for the analyses undertaken/reported), and the data are also reasonably compatible with hypotheses that posit essentially zero effect at 6 weeks and 6 months in both strata.

To investigate the extent to which disruption of follow-up and online intervention delivery training were associated with variation in treatment effects, we refitted the primary analysis model, including an additional term denoting whether all available follow-up data for the primary outcome were provided/completed before/after the date of the start of the first UK lockdown (23 March 2020) and additional terms for all of the two- and three-way interactions between this variable, allocation and time point. The estimates from this augmented model are given in [Table 31](#). From this table, there is some weak evidence that the effectiveness of Social Stories was greater at the month 6 time point among those who were followed up prior to the beginning of lockdown (little variation in treatment effect was apparent at week 6). However, the estimates for all between-group contrasts are quite imprecise (particularly for the pre-lockdown follow-up group), and a range of treatment effects across the different strata are plausible given the observed data (including the absence of any important treatment effects).

Secondary outcomes

Teacher-reported outcomes

Social Responsiveness Scale-2 total raw scores

Descriptive summaries of total raw scores derived from the teacher-completed SRS-2 responses and the distributions of these scores are given by allocation in [Appendix 1, Table 50](#). These scores were analysed using a similar approach to the analysis of the primary outcome (i.e. the teacher-reported SRS-2 total T-scores), except that the raw scores were conditioned in place of the T-scores. The estimated treatment effects at 6 weeks and 6 months are given in [Table 32](#). Model assumptions were checked following similar approaches as used for the primary analysis. These diagnostics showed no severe departures from the distributional assumptions of the planned analysis model.

The results in [Table 32](#) for the SRS-2 total raw score are broadly comparable with the results for the primary outcome, although the estimated treatment effect appears to be more consistent over time. The point and interval estimates indicate the data are most compatible with hypotheses positing that the intervention causes modest reductions in the expected severity of social impairment associated with participants' autism. However, the upper limits of the reported intervals suggest the data are also reasonably compatible with hypotheses positing minor increases in expected scores due to the intervention. In addition, the reported *p*-values for the tests of point null hypotheses at 6 months and 6 weeks indicate the observed data would not be terribly surprising if there really were no effect of allocation on expected score at these time points.

Goal-based outcome scores

To provide some context for the scores and analyses reported below, details of the types of goals set in consultation with the participants' teachers at baseline are given in [Table 33](#). These figures show that across both groups, the most common types of goals set related to prosocial behaviour and communication and understanding/processing emotions, as would be expected given the characteristics

TABLE 32 Summary of the treatment effects of the secondary outcomes

Outcome	Timing	Estimate (95% CI) ^a	<i>p</i> -value ^a
SRS-2 total raw score (teacher)	Week 6	-3.37 (-9.41 to 2.67)	0.274
	Month 6	-3.32 (-10.28 to 3.63)	0.349
Goal-based outcome (teacher)	Week 6	0.84 (0.14 to 1.54)	0.018
	Month 6	0.97 (0.21 to 1.73)	0.012
SRS-2 total T-score (parent)	Week 6	0.43 (-0.83 to 1.70)	0.504
	Month 6	0.35 (-1.26 to 1.97)	0.668
SRS-2 total raw (parent)	Week 6	2.37 (-1.41 to 6.16)	0.219
	Month 6	1.21 (-3.65 to 6.08)	0.625
RCADS total score (parent)	Week 6	1.10 (-1.81 to 4.02)	0.458
	Month 6	2.35 (-1.37 to 6.06)	0.215
PSI total stress score	Week 6	-1.42 (-4.92 to 2.09)	0.428
	Month 6	-1.49 (-5.43 to 2.46)	0.460
EQ-5D-Y VAS (parent)	Week 6	-0.63 (-4.93 to 3.66)	0.772
	Month 6	-1.92 (-6.70 to 2.86)	0.432

a Based on delta method standard errors.

TABLE 33 Details of types of behavioural goals set at baseline (participants with available goal-setting data only)

	Control (N = 107)	Intervention (N = 128)	Total (N = 235)
Types of goal set ^a			
Preparation for transitions, change or new experiences	14 (13.1)	17 (13.3)	31 (13.2)
Recognising positive self attributes, self-esteem	12 (11.2)	10 (7.8)	22 (9.4)
Information about life skills	29 (27.1)	27 (21.1)	56 (23.8)
Information for keeping safe	7 (6.5)	5 (3.9)	12 (5.1)
Understanding routines	18 (16.8)	25 (19.5)	43 (18.3)
Understanding social behaviour or social communication	42 (39.3)	57 (44.5)	99 (42.1)
Understanding emotions	39 (36.4)	39 (30.5)	78 (33.2)
Understanding attitudes/perspectives/points of view	19 (17.8)	16 (12.5)	35 (14.9)

^a Possibly more than one goal type per participant.

of the sample. Goals relating to understanding routines and dealing with transitions or changes in routines were also reasonably common and fairly evenly distributed between groups. Descriptive summaries of the teacher-reported goal-based outcome measures at each time point and the distribution of these responses are given by allocation in [Table 51](#) in [Appendix 1](#).

The teacher-reported goal-based outcome scores were analysed using a similar approach to the analysis of the primary outcome, except that the baseline scores for this outcome were conditioned on in place of the baseline score for the primary outcome. The estimated treatment effects at 6 weeks and 6 months are given in [Table 32](#). Model assumptions were checked following similar approaches as used for the primary analysis. Despite the discrete nature of the outcome, these diagnostics did not identify severe departures from the distributional assumptions of the fitted model. However, the analysis of this outcome as a continuous outcome assumes that the gaps between adjacent values are all equal (i.e. that a shift between 0 and 1 reflects the same improvement in the underlying construct being measured as a shift between 9 and 10). This assumption is certainly doubtful, placing some limitations on the interpretation of the treatment effects estimated from this model. We therefore also conducted post hoc ordinal analyses of these data using a mixed-effect proportional odds model, with each time point modelled separately using the same fixed and random effects as before. For both models, there was little evidence that the effects of allocation deviated substantially from the proportional odds assumption. The estimated odds ratios for allocation at 6 weeks and 6 months were 1.97 (95% CI 1.05 to 3.70) and 2.08 (95% CI 1.13 to 3.84), respectively. To aid interpretation, we used the fitted model to derive absolute and relative differences in the probability of having a score strictly > 5 (goal met more than half of the time) at both the 6-week and 6-month time points, conditional on various covariate patterns *X*. These are reported in [Tables 52](#) and [53](#) in [Appendix 1](#).

The results in [Table 32](#) suggest allocation to the intervention causes a small-to-moderate increase in the frequency that teachers perceived the goal set at baseline to have been met. The *p*-values indicate considerable discrepancy between the data and the test hypotheses, suggesting the observed data would be reasonably surprising if the entire set of assumptions used to compute the reported *p*-values were true (including the assumption that allocation to Social Stories had no effect on this outcome). Similarly, the estimates in [Tables 52](#) and [53](#) in [Appendix 1](#) suggest an overall shift towards higher scores (i.e. goals being met more of the time) in the intervention group, although some of the lower confidence limits suggest the data are somewhat compatible with hypotheses positing relatively small effects of allocation or none. Furthermore, the apparent treatment effects could be at least partly explained by

the non-blinded nature of the outcome assessment. Overall, these results provide some weak evidence that allocation to Social Stories increases the frequency that children meet specific behavioural goals at 6 weeks and 6 months post randomisation.

Parent/carer-reported outcomes

Social Responsiveness Scale-2 total scores

Descriptive summaries of the parent/carer-reported SRS-2 total T-scores and raw scores at each time point are reported by allocation in [Table 54](#) in [Appendix 1](#). These scores were analysed using a similar approach to the analysis of the primary outcome, except that the parent/carer-reported baseline scores were conditioned on in place of the teacher-reported scores. The estimated treatment effects at 6 weeks and 6 months are given in [Table 32](#) for the T-scores and raw scores. Model assumptions were checked following similar approaches as used for the primary analysis. These diagnostics showed no severe departures from the distributional assumptions of the planned analysis models.

The estimates in [Table 32](#) suggest allocation had relatively little impact on the expected severity of social impairment as perceived by the parents/carers of the participants. However, in contrast with the teacher-reported SRS-2 outcomes, the data are seemingly most compatible with hypotheses positing the intervention causes slight worsening of symptoms compared with care as usual, although hypotheses positing no effect or very slight improvements are also reasonably compatible with the observed data. Overall, these data provide little evidence of any important treatment effects (positive or negative) at either time point for the parent-reported SRS-2.

Revised Children's Anxiety and Depression Scale total score

Descriptive summaries of the RCADS total scores (parent-reported) at each time point are reported by allocation in [Table 55](#) in [Appendix 1](#). These scores were analysed using a similar approach to the analysis of the primary outcome, except that the baseline RCADS scores were conditioned on in place of the baseline measurement of the primary outcome. The estimated treatment effects at 6 weeks and 6 months are given in [Table 32](#). Model assumptions were checked following similar approaches as used for the primary analysis. These diagnostics showed no severe departures from the distributional assumptions of the fitted model.

The estimates in [Table 32](#) suggest allocation to Social Stories has a relatively limited impact on overall symptoms of anxiety and/or depression. The point estimates suggest the data are most compatible with hypotheses positing small negative effects of allocation to Social Stories (i.e. that allocation of Social Stories causes a slight increase in expected levels of anxiety/depression compared with allocation to care as usual). However, the interval estimates and *p*-values suggest that the observed data would not be hugely unexpected if allocation to Social Stories truly had no impact on this outcome at all or even reduced levels of anxiety and depression.

Parental Stress Index total stress score

Descriptive summaries of the PSI total stress scores (parent-reported) at each time point are reported by allocation in [Table 56](#) in [Appendix 1](#). These scores were analysed using a similar approach to the analysis of the primary outcome, except that the baseline PSI total stress scores were conditioned on instead of the baseline measurements of the primary outcome. The estimated treatment effects at 6 weeks and 6 months are given in [Table 32](#). Model assumptions were checked following similar approaches as used for the primary analysis. These diagnostics showed no severe departures from the distributional assumptions of the fitted model.

The estimates in [Table 32](#) suggest allocation to Social Stories has a relatively limited impact on overall stress due to problematic parent-child interaction and behaviours. The point estimates suggest the data are most compatible with hypotheses positing small positive effects of allocation to Social Stories

(i.e. that allocation of Social Stories causes a slight decrease in expected levels of parental stress compared with allocation to care as usual). However, the interval estimates and *p*-values suggest that the observed differences at both time points would not be hugely unexpected if allocation to Social Stories truly had no impact on this outcome at all or even increased levels of parental stress.

European Quality of Life-5 Dimension Youth Questionnaire general health visual analogue scale

Descriptive summaries of the EQ-5D-Y general health VAS scores (parent-reported) at each time point are reported by allocation in [Table 57](#) in [Appendix 1](#). These scores were analysed using a similar approach to the analysis of the primary outcome, except that the baseline EQ-5D-Y VAS scores were conditioned instead of the baseline measurements of the primary outcome. The estimated treatment effects at 6 weeks and 6 months from the planned analysis model are given in [Table 32](#). Model assumptions were checked following similar approaches as used for the primary analysis. These diagnostics revealed some departures from the distributional assumptions of the planned analysis model. In particular, ceiling effects were evident, resulting in heteroscedasticity and substantial departures from conditional normality. Hence, further semiparametric analyses that avoid some of the more restrictive parametric assumptions of the planned analysis model were conducted.

Mixed-effect ordinal logistic regression was used to model the EQ-5D-Y VAS outcome scores at week 6 and month 6 separately, with the same fixed and random effects as previously. The fitted models were used to estimate the difference in expected EQ-5D-Y VAS score between groups at week 6 and month 6, conditional on attendance at a non-SEN school with fewer than six participants, male gender and age and baseline score equal to the relevant sample means. The results of this post hoc analysis are reported in [Table 58](#) in [Appendix 1](#).

The estimates suggest allocation to Social Stories has a limited effect on general health (as proxy reported by participants' parents/carers) at 6 weeks and 6 months post randomisation. The estimates from the planned analysis model suggest that the data are most compatible with hypotheses positing the intervention causes a very slight reduction in expected score compared with the control. However, the data are also seemingly quite compatible with hypotheses positing no treatment effect and even small positive treatment effects in favour of Social Stories. Despite the reversal of direction at 6 weeks, the estimates from the post hoc analysis also suggest allocation to Social Stories has a limited effect on general health at 6 weeks and 6 months post randomisation. Overall, there is limited evidence of any clinically relevant effects of allocation to Social Stories on general health (as perceived by parents/carers).

Adverse events and safeguarding

During the course of the trial, two serious and four non-SAEs were reported. These six AEs occurred in six individuals (i.e. at most one event per individual) and are summarised by allocation in [Table 34](#). Details of the AEs reported are given in [Table 35](#). Despite the slightly higher number of AEs reported in the intervention group, only one was deemed probably related to the intervention, and there was limited evidence of any clear variation in incidence between groups. With the one Story where the incident was related, it highlighted the importance of using careful wording when writing a Social Story, but it was also reported that the Social Story enabled some helpful further discussions related to helping the child with managing their feelings and helping the teacher understanding the child better. However, by themselves, the data are much too sparse to draw any strong conclusions regarding the possible impact of Social Stories on the incidence and/or severity of AEs in this population.

TABLE 34 Summary of reported serious and non-serious AEs by allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
AEs reported			
Non-SAEs	1	3	4
SAEs	1	1	2
None	118	125	243
Participants with ≥ 1 AE reported			
≥ 1 non-SAE	1	3	4
≥ 1 SAE	1	1	2
No AEs reported	118	125	243
Relatedness			
Probably related	0 (0.0)	1 (0.8)	1 (0.4)
Unlikely to be related	0 (0.0)	2 (1.6)	2 (0.8)
Unrelated	2 (1.7)	1 (0.8)	3 (1.2)
None reported	118 (98.3)	125 (96.9)	243 (97.6)

TABLE 35 Details of reported serious and non-serious AEs

Event	Allocation	Description	Seriousness	Relatedness	Expectedness
1	Social Stories	Information from session record log: 'They covered their ears when I started reading the sentence about being angry and shout'. 'They are never proud of their work and start to bang their head on the wall slowly – lasted a few seconds'. 'T.A. had to speak with them that it is ok to find writing hard and if they need help they could ask. They don't need to hit their head on wall'	Non-serious	Probably related	Expected
2	Social Stories	Child was out for the day ... They ... tripped and fell down steps. They broke their collar bone, cut their head open and grazed their body. Child has dyspraxia and weakness in their left side	Non-serious	Unrelated	-
3	Social Stories	Whilst reading <i>participant name</i> the social story, because they were in the wrong place they continued to run away. I followed them and continued to read. They didn't want to be in the classroom	Non-serious	Unlikely to be related	Expected
4	Social Stories	Child is no longer attending school. They were withdrawn in June after struggling with bullying. They also made attempts on their lives and had to go to A&E. This included tying part of their lunch box around their neck. They have been diagnosed with PTSD severe anxiety and attachment disorder. They have since received support from CAMHS, a psychiatrist and a family support worker. They are doing much better and due to start a specialist school for autism in January	Serious	Unlikely to be related	Unexpected

continued

TABLE 35 Details of reported serious and non-serious AEs (*continued*)

Event	Allocation	Description	Seriousness	Relatedness	Expectedness
5	Usual care	The participant went to A&E ... after having prolonged nose bleeds at school, which they could not stop. They went to A&E, by which time the bleeding had stopped, and they were given cream by the doctor. They have difficulties with nose bleeds more generally	Non-serious	Unrelated	-
6	Usual care	Child broke their left arm falling off a day bed at home. They slipped whilst trying to climb from the bottom part of the bed to the top. They broke their humerus, were taken to A&E and had surgery. They spent the night in hospital. During surgery, the surgeons caught a nerve. They currently can't grip with their left hand, although the arm is healed. They have had to change the hand they write with, and it has affected their schoolwork. Schools are being supportive	Serious	Unrelated	-

A&E, accident and emergency; PTSD, post-traumatic stress disorder.

Chapter 4 Economic evaluation

Aims

The primary aim of the economic evaluation was to assess the cost effectiveness of Social Stories versus usual care from the societal perspective, which included costs from the educational system, private out-of-pocket expenses and parental productivity costs reported in the trial. To take uncertainty into consideration, a set of sensitivity analyses (including an evaluation from the NHS and PSS perspectives) were also conducted. All the results were reported according to the Consolidated Health Economic Evaluation Reporting Standards 2022.⁶⁶

Availability of quality-of-life and cost data

The complete case here refers to the children who completed EQ-5D-Y questionnaires and whose parents and teachers completed resource use questionnaires at both baseline and 6-month follow-up. Details of the QoL and cost data available at each data collection time point are shown in [Table 36](#). As shown, QoL data (EQ-5D-Y) were fully available at any time point for 183 (73.5%), whereas all the cost items from the societal perspective were available for 113 (45.4%). Overall, a total of 112 (45.0%) participants had both EQ-5D and resource use (from the societal perspective) data at the two data collection time points. This sample constitutes the complete-case group, which is one of the two groups used for all the subsequent analyses.

Multiple imputation

To account for the data from the sample other than the complete-case group (55.0%), missing utility scores and costs were further imputed using multiple-imputation via chained equations. The following variables were used in the imputation process to ensure best fit of the imputed results: trial arm, age, sex, stratification factors (SEN status stratum: non-SEN or SEN, Number of participants stratum: ≤ 5 or > 5), parent-completed SRS-2 score, baseline EQ-5D-Y utility score and cost from the societal perspective at

TABLE 36 Availability of QoL and cost data (before imputation)

	Baseline, n (%)	Month 6, n (%)	Complete case, n (%)
Total (n = 249)			
Questionnaire (parent) completion	249 (100.0)	184 (73.9)	184 (73.9)
Questionnaire (teacher) completion	247 (99.2)	212 (85.1)	210 (84.3)
Health economic-related data			
EQ-5D-Y	249 (100.0)	183 (73.5)	183 (73.5)
Costs from the NHS and PSS perspective	245 (98.4)	178 (71.5)	176 (70.7)
Costs from the societal perspective	220 (88.4)	124 (49.8)	113 (45.4)
Economic evaluation			
EQ-5D-Y and costs (NHS and PSS perspective)	245 (98.4)	177 (71.1)	175 (70.3)
EQ-5D-Y and costs (societal perspective)	220 (88.4)	123 (49.4)	112 (45.0)

baseline. The base-case (imputed) sample was 249 (129 allocated to Social Stories and 120 allocated to usual care); this is the sample used for the primary cost-effectiveness analysis.

Baseline characteristics

Descriptive statistics of participating children's characteristics and the additional predictors used in the multiple imputation are presented in [Table 37](#). Around three-quarters of children in the Social Stories and the usual care arms were male. This is in line with the autism population at school age in the UK.⁶⁷ More than 80% in both arms were of primary school age (ranging from 7 to 11 years old). Differences in the parent SRS-2 scores and the EQ-5D-Y utility scores at the baseline were small across arms and samples. Overall, the baseline characteristics are consistent across samples (base case and complete case) and with the main statistical analysis.

Costs

Two types of costs were considered in this study: the cost of intervention (obtained from the interventionists and the study team) and the cost of service use (self-reported by the parents/guardians and the teachers). All the costs were expressed in 2019–20 Great British pounds.

TABLE 37 Key baseline characteristics by trial arm

Baseline characteristics	Base case (n = 249)		Complete case (n = 112)	
	Social Stories (N = 129)	Usual care (N = 120)	Social Stories (N = 58)	Usual care (N = 54)
Gender, n (%)				
Male	95 (73.6)	90 (75.0)	40 (69.0)	36 (66.7)
Age (years), n (%)				
7–11	22 (17.1)	24 (20.0)	11 (19.0)	10 (18.5)
11–15	107 (82.9)	96 (80.0)	47 (81.0)	44 (81.5)
Mean (SD)	8.3 (1.8)	8.5 (1.8)	8.3 (1.7)	8.6 (1.7)
Parent SRS-2 scores				
Mean (SD)	82.0 (8.4)	82.5 (8.1)	80.6 (9.9)	82.6 (8.0)
SEN status, n (%)				
Non-SEN	116 (89.9)	111 (92.5)	53 (91.4)	50 (92.6)
SEN	13 (10.1)	9 (7.5)	5 (8.6)	4 (7.4)
Number of participants, n (%)				
≤ 5	80 (62.0)	70 (58.3)	31 (53.5)	32 (59.3)
> 5	49 (38.0)	50 (41.7)	27 (46.5)	22 (40.7)
Baseline EQ-5D-Y utility				
Mean (SD)	0.54 (0.31)	0.54 (0.31)	0.54 (0.32)	0.51 (0.31)
Number of intervention sessions				
Mean (SD)	4.3 (3.3)	-	5.4 (3.1)	-

Intervention costs

Intervention costs include both training and intervention delivery costs. To reflect the cost of intervention when it rolls out in practice, only the costs that are associated with the intervention were considered. The costs that occur for research purposes (such as researcher time for collecting, recording and analysing data) were not included. The training cost information was obtained from the study team, and the estimated training costs were allocated to each session and each child who received the intervention. In relation to the intervention delivery costs, the information was obtained directly from the interventionists using the self-reported questionnaires. Intervention sessions were costed on the basis of the salary of the professional involved. [Table 38](#) shows the total and average costs of each component that were used in order to deliver the Social Stories intervention. As shown, the main cost driver of training costs was the trainer fee (71.5%), while the main cost drivers of intervention delivery costs were the costs of interventionists' time for preparation and delivery of the intervention (58.2%). On average, the estimated intervention cost per session per child was £15.22 (£12.52 for training and £2.70 for intervention delivery).

Service use and costs

[Table 39](#) shows the summarised service use per child in each trial arm based on the complete cases only. The difference in service use between Social Stories and usual care arms were small at baseline and 6-month follow-up time points. At baseline, children in the Social Stories arm seemed to have used more NHS and school-based health services than those in the usual care arm. After receiving the intervention, resource use in both arms decreased, and such decreases were more prominent in the Social Stories arm than in the usual care arm. Also, children in the Social Stories arm were more likely to use fewer services at follow-up compared to baseline. Caution is needed in interpreting these results, as these summarised service uses include various items with different unit costs.

The total costs broken down by perspective, type of service, trial arm and before and after imputation are included in [Table 40](#). Discounting was not applied due to the short-term nature of the trial. As shown,

TABLE 38 Intervention costs by trial arm

	Total cost (£)	Cost per session per child (£)
Training costs		
Trainer fee (staff time)		
Preparation	922.91	1.65
Training	4088.73	7.30
Trainer's travel costs	1080.96	1.93
Consumable costs	918.05	1.64
Total	7010.65	12.52
Intervention delivery costs		
Intervention (staff time)		
Therapy session (preparation and delivery)	880.33	1.57
Additional work/help from other staff	97.13	0.17
Consumables	313.60	0.56
Supervision ^a	220.50	0.39
Total	1511.56	2.70

^a The calculation method and unit cost were based on Wiles *et al.* (2014)'s study.

TABLE 39 Average service use by the trial arm (complete case, n = 112)

	Unit	Baseline		Month 6	
		Social Stories, N = 58, mean (SD)	Usual care, N = 54, mean (SD)	Social Stories, N = 58, mean (SD)	Usual care, N = 54, mean (SD)
NHS and PSS					
Community-based services					
CAMHS-related	Session	0.55 (2.26)	0.61 (2.02)	0.22 (0.88)	0.44 (1.34)
Non-CAMHS-related					
GP	Appointment	0.98 (2.27)	0.67 (1.17)	0.58 (1.43)	0.46 (0.88)
Allied health professionals	Appointment	1.81 (3.15)	2.26 (3.70)	1.24 (2.79)	1.44 (3.06)
Social care service	Appointment	1.97 (9.13)	0.52 (2.00)	0.45 (1.57)	0.13 (0.67)
Hospital-based services/acute services					
Emergency services	Visit	0.22 (0.80)	0.17 (0.38)	0.16 (0.45)	0.20 (0.59)
Inpatient stay					
Mental health-related	Night	-	-	-	-
Non-mental health-related	Night	-	-	-	0.02 (0.14)
Outpatient visit/day case					
Mental health-related	Visit	0.07 (0.26)	0.07 (0.26)	-	0.09 (0.35)
Non-mental health-related	Visit	0.20 (0.44)	0.20 (0.49)	0.10 (0.31)	0.15 (0.45)
Medication					
Mental health-related	Type	0.31 (0.65)	0.31 (0.72)	0.21 (0.45)	0.26 (0.56)
Non-mental health-related	Type	0.38 (0.81)	0.50 (0.84)	0.22 (0.46)	0.22 (0.50)
Education system-related					
School-based health services	Hour	12.86 (35.45)	5.78 (13.94)	9.58 (23.7)	15.43 (44.24)
Intervention support ^a	Hour	74.17 (156.35)	71.91 (83.56)	48.36 (107.54)	55.24 (95.83)
General support ^a	Hour	1.78 (1.62)	2.28 (2.16)	2.19 (1.74)	2.96 (2.60)
Private expenses – out of pocket					
Privately paid consultation	Session	-	-	0.21 (1.58)	0.07 (0.54)
Child care	Session	5.84 (18.68)	3.02 (11.70)	7.24 (25.24)	6.91 (28.40)
Club	Session	3.76 (10.47)	7.33 (17.62)	1.55 (5.51)	4.72 (12.85)
Productivity					
Parental productivity	Day	0.86 (2.16)	0.81 (1.78)	0.98 (2.90)	1.70 (3.53)

^a Based on the teacher-reported questionnaires.

Notes

CAMHS included child psychiatrist, child psychotherapist, child psychologist, clinical psychologist, mental health nurse, family therapist and primary mental health worker (PMHW).

Allied health professionals for the community-based services included community nurse, community paediatrician, occupational therapist, physiotherapist and speech and language therapist.

Social care services included social care worker, home care worker, family support worker, drug and alcohol support worker and Helpline (e.g. Samaritans).

Club included after-school clubs, religious clubs, sport clubs and special clubs for autism children.

TABLE 40 Average costs of service use in 6 months by trial arm

	Base case		Complete case	
	Social Stories (n = 129), £ (95% CI)	Usual care (n = 120), £ (95% CI)	Social Stories (n = 58), £ (95% CI)	Usual care (n = 54), £ (95% CI)
NHS and PSS	252.5 (177.1 to 328.0)	351.8 (223.1 to 480.5)	250.4 (153.7 to 347.2)	379.0 (157.9 to 600.0)
Community-based services				
CAMHS-related	17.9 (0.1 to 35.6)	53.7 (12.8 to 94.5)	16.0 (-4.3 to 36.3)	43.2 (3.3 to 83.1)
Non-CAMHS-related	151.5 (96.2 to 206.9)	162.6 (86.0 to 239.2)	150.5 (78.8 to 222.2)	123.9 (75.2 to 172.5)
Hospital-based services				
Mental health-related	1.9 (-1.9 to 5.6)	6.4 (-0.9 to 13.7)	-	14.2 (-1.9 to 30.4)
Non-mental health-related	36.5 (13.3 to 59.7)	85.3 (-7.4 to 177.9)	33.4 (3.1 to 63.8)	145.9 (-57.3 to 349.0)
Medications				
Mental health-related	35.0 (14.2 to 55.8)	31.5 (15.0 to 47.9)	41.7 (12.5 to 70.9)	34.7 (12.9 to 56.6)
Non-mental health-related	9.8 (2.4 to 17.1)	12.4 (3.6 to 21.3)	8.8 (-3.7 to 21.4)	17.0 (1.9 to 32.2)
Education system-related	725.3 (392.5 to 1058.1)	819.6 (563.1 to 1076.2)	635.3 (269.3 to 1001.2)	701.2 (387.1 to 1015.3)
School-based health	90.9 (38.4 to 143.4)	253.9 (108.4 to 399.4)	138.8 (46.6 to 231.0)	244.6 (63.7 to 425.5)
Intervention support	497.3 (177.6 to 817.0)	212.6 (93.2 to 331.9)	328.6 (18.3 to 639.0)	230.0 (40.1 to 410.9)
General support	187.1 (104.9 to 269.4)	353.2 (194.5 to 511.9)	167.9 (64.2 to 271.6)	226.7 (87.9 to 365.4)
Private expenses	589.0 (240.0 to 939.0)	425.4 (90.9 to 759.9)	624.6 (114.8 to 1134.3)	575.1 (-3.7 to 1153.9)
Parental productivity loss	80.8 (36.0 to 125.6)	116.8 (45.6 to 188.0)	65.6 (16.7 to 114.4)	116.9 (39.5 to 194.4)
Total costs	1632.4 (1160.3 to 2104.5)	1713.6 (1211.8 to 2215.5)	1575.9 (934.9 to 2207.9)	1772.2 (947.7 to 2596.7)

the costs substantially vary from one child to the other as illustrated by the wide 95% CIs. In detail, the total service costs to the NHS (before imputation) were £250.4 (95% CI £153.7 to £347.2) for the Social Stories arm compared with £379.0 (95% CI £157.9 to £600.0) for the usual care arm. The average total costs for the use of CAMHS-related community-based services, hospital-based services (both mental and non-mental health-related) and non-mental health-related medication were higher for the children in the usual care arm. The average total costs for the use of education services (before imputation) were £635.3 (95% CI 269.3 to £1001.2) for the Social Stories arm compared with £701.2 (95% CI £387.1 to £1015.3) for the usual care arm. It is observed that children in the Social Stories arm incurred less costs in school-based health services (such as educational psychologist and school nurse visits) and in general support from teachers compared to those in the usual care arm. However, higher costs for school-based intervention services (i.e. 1 : 1 mentoring/individual work, Social Communication groups and Social skills work group) were observed in the Social Stories arm. It is worth noting that some of the cost differences were likely to have been driven by the high-cost cases. For instance, the higher average cost of school-based intervention services in the trial arm was driven by two children, the first of which had a total of 540 and the other 624 hours of 1 : 1 mentoring sessions over 6 months. Given that those values were entirely plausible, we have decided to keep them in the data set without any adjustment.

The Social Stories arm also incurred fewer costs in parental productivity losses compared to those in the usual care arm. However, the Social Stories arm incurred slightly higher costs in private expenses. Such difference was likely to have been driven by two high-cost cases in the Social Stories arm. The first of which had 240 and the other 250 sessions (half day = 1 session) of paid child care over the past 6 months. Again, we have decided to keep those high-cost cases in the analysis, as the scenarios are plausible. Overall, the Social Stories arm incurred less costs across all the perspectives. This is observed in both the complete case and the base case. However, owing to the high-cost cases, the cost differences need to be interpreted with caution.

Quality of life

[Table 41](#) shows the mean EQ-5D-Y (3L, proxy version) utility scores between the two arms of the trial at the two time points when scores were not imputed (complete case) and when scores were imputed (base case). As shown, in both arms, a slight increase in EQ-5D-Y scores from baseline to month 6 was observed. Such a small increase was also observed in both the base and the complete cases. Overall, the Social Stories produced similar QALYs compared to the usual care. Further details on the responses in each domain can be found in [Appendix 2, Table 68](#).

Primary analysis

[Table 42](#) shows the incremental ICERs based on the base case for the primary analysis. Before any adjustments, on average, children receiving Social Stories incurred £65.9 less cost from the societal

TABLE 41 Utility scores at baseline and month 6 by trial arm

Time point	Base case		Complete case	
	Social Stories (n = 129), mean (95% CI)	Usual care (n = 120), mean (95% CI)	Social Stories (n = 58), mean (95% CI)	Usual care (n = 54), mean (95% CI)
Utility score				
Baseline	0.54 (0.49 to 0.59)	0.54 (0.48 to 0.59)	0.54 (0.45 to 0.62)	0.51 (0.43 to 0.59)
Month 6	0.59 (0.53 to 0.65)	0.59 (0.53 to 0.66)	0.60 (0.52 to 0.67)	0.59 (0.52 to 0.67)
Total QALYs	0.28 (0.26 to 0.31)	0.28 (0.26 to 0.31)	0.28 (0.28 to 0.32)	0.27 (0.24 to 0.31)

TABLE 42 Base-case cost-effectiveness results (from the societal perspective)

Trial arm	Costs (£), mean (95% CI)	QALYs, mean (95% CI)	Incremental cost (95% CI)	Incremental QALY (95% CI)	ICER
Before bootstrapping					
Social Stories	1698.1 (1225.8 to 2170.3)	0.28 (0.26 to 0.31)	-65.9	0.001	Dominant
Usual care	1713.6 (1211.8 to 2215.6)	0.28 (0.26,0.31)			
After bootstrapping					
Social Stories	1602.9 (1591.8 to 1613.9)	0.28 (0.28 to 0.28)	-191.0 (-767.7 to 337.7)	< 0.001 (-0.013 to 0.015)	Dominant
Usual care	1793.9 (1779.9 to 1807.8)	0.28 (0.28 to 0.28)			

perspective and maintained similar QALYs improvement. To account for the uncertainty and adjust for any imbalanced characteristics at baseline, the estimates of incremental costs and the QALYs from regression were bootstrapped to simulate 5000 pairs of net cost and net outcomes, as recommended by NICE for health technology appraisals.⁵⁴ After bootstrapping, on average, children receiving Social Stories incurred £191.0 (95% CI -£337.7 to £767.7) less costs (not statistically significant) and maintained similar QALYs (mean incremental difference: < 0.001, 95% CI -0.013 to 0.013), compared to those having usual care.

Figure 3 shows the cost-effectiveness plane for Social Stories compared with usual care based on 5000 bootstrapped estimates of incremental costs and incremental QALYs. The red lines represent the willingness to pay thresholds of £20,000 and £30,000 per QALY gained. The simulated estimates were largely below the threshold line, suggesting that if society is willing to pay £20,000 for extra QALY gained, then Social Stories is likely to be the preferred option compared to the usual care, although the

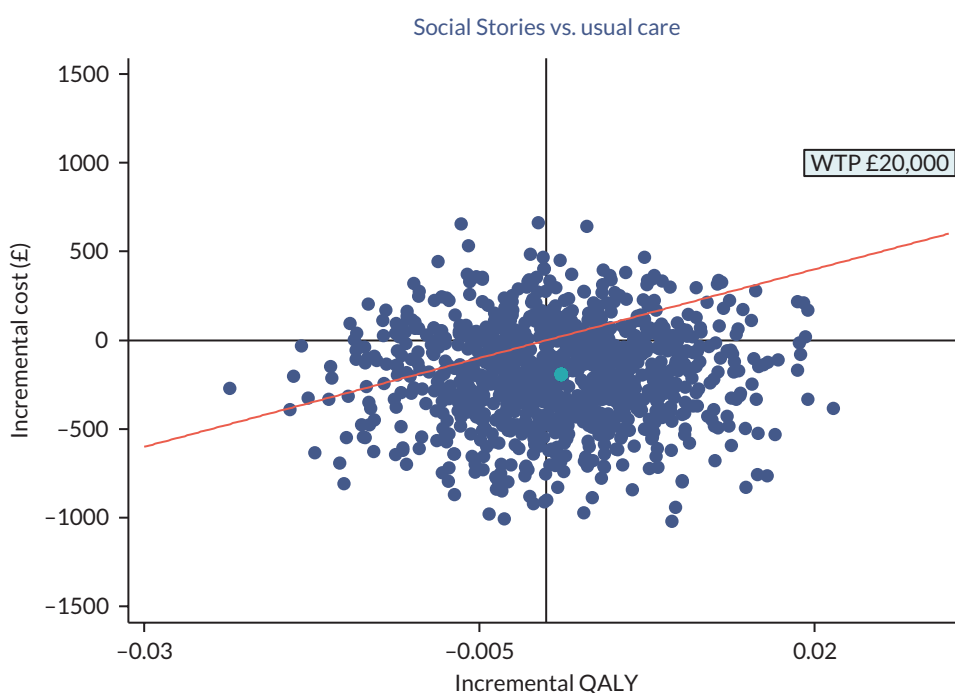


FIGURE 3 Base-case cost-effectiveness plan of Social Stories compared with usual care (from the societal perspective).

incremental cost was small and incremental QALYs was similar. It is worth noting that the dispersion of the results versus the mean is observed, suggesting that the level of uncertainty of the results is slightly high.

The CEAC of Social Stories compared with usual care is presented in [Figure 4](#). The probability of Social Stories being a preferred option is 75% if the society is willing to pay £20,000 for one QALY gained and 74% if the society is willing to pay £30,000 per QALY gained. This is equivalent to stating that, given the data and considering the decision uncertainty, there is a 74–75% chance that the incremental cost–QALY pair of Social Stories compared with usual care is at or below £20,000–30,000 per QALY gained thresholds, and that if the intervention were adopted, there is a 74–75% chance that this is the correct decision. The CEAC also shows that the probability of intervention being cost-effective varies very little by threshold. The main driver of the variance is likely to be the cost difference, as the incremental QALY is very small.

Sensitivity analysis

To evaluate the robustness of the primary analyses, three sensitivity analyses were conducted (see [Table 43](#)). The mean incremental cost and QALY estimates from the complete case were along the line of the base-case scenario, yielding a negative cost per QALY gained. Similar results were also observed from the sensitivity analyses based on the NHS/PSS perspective (scenario 2) and the NHS/PSS and education perspective (scenario 3). All the sensitivity analyses show that Social Stories is dominant, and a good proportion of the bootstrapped estimates lay below the recommended NICE threshold (£20,000 QALY gained) (see [Report Supplementary Material 7, Figure 17](#)). It is worth noting that although a relatively small cost saving was observed in all three scenarios and varies by perspective, like the results of the primary analysis, the incremental costs were not statistically significant, showing the uncertainty of the cost reduction. In contrast, the incremental QALYs in all three scenarios remain the same small (< 0.002) and not statistically significant, suggesting that Social Stories do not appear to improve Children’s QALYs.

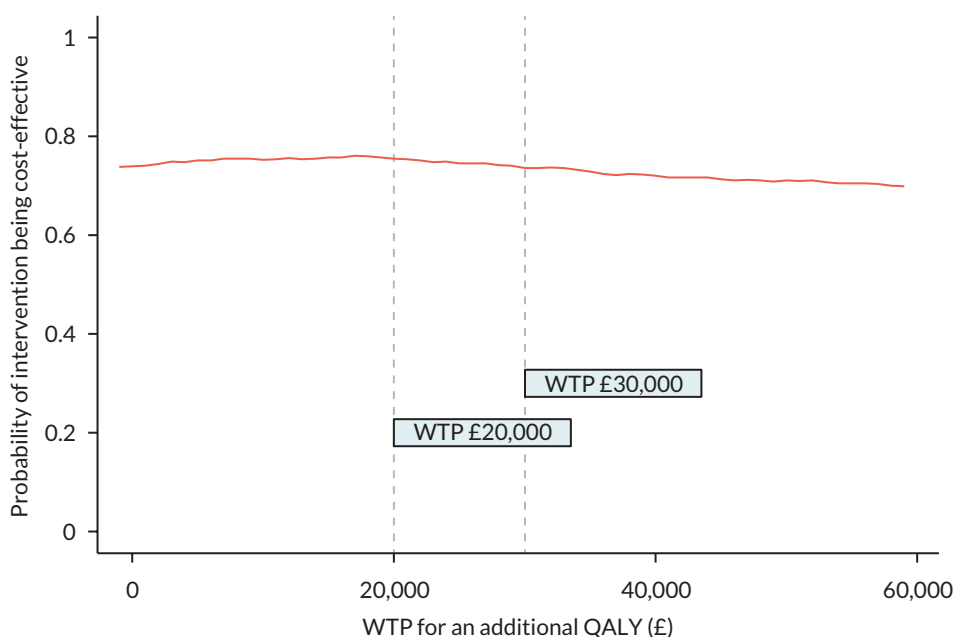


FIGURE 4 Cost-effectiveness acceptability curve of Social Stories compared with usual care.

TABLE 43 Sensitivity analyses

Social Stories vs. usual care	Incremental costs (£) (95% CI)	Incremental QALYs (95% CI)	ICER (£/QALY gained) (95% CI)
Scenario 1: CCA from the societal perspective	-379.7 (-1499.5 to 485.0)	< 0.001 (-0.021 to 0.020)	Dominant
Scenario 2: CEA from the NHS/PSS perspective	-48.4 (-202.3 to 84.6)	0.002 (-0.013 to 0.015)	Dominant
Scenario 3: CEA from the NHS/PSS and education perspective	-240.0 (-578.2 to 79.7)	0.001 (-0.008 to 0.028)	Dominant

CEA, cost-effectiveness analysis.

Summary

Compared to usual care, Social Stories slightly decreased service use costs and maintained similar QALYs. This is evident in both primary and sensitivity analyses, which considered costs derived from various perspectives. It is worth noting that such results are based on the given data, which is limited to the short study time frame, amount of missing data and the disruption of COVID-19.

Chapter 5 Process evaluation

Objectives

What were the main objectives of the process evaluation?

The process evaluation was cross-sectional and longitudinal, encompassing all aspects of the Social Stories intervention. The purpose of the process evaluation was to assess the fidelity of the programme (in terms of delivering the intervention as per the Carol Gray methodology outlined in *Social Stories™ intervention*), consider the views of various stakeholders and identify barriers and facilitators to successful implementation. We aimed to achieve this through a combination of qualitative and quantitative data collection techniques, including interviews, focus groups, questionnaires, surveys and diaries (session logs). A full overview in relation to each of the process evaluation objectives is presented in *Table 2* (see *Chapter 2*). We interviewed five teachers, seven interventionists and nine parents/carers. We conducted three focus groups, one including the trainers and two including the trial RAs. We surveyed teachers, interventionists and parents yielding 32 respondents.

In this chapter, we have applied the components of intervention fidelity as described by Hasson⁴⁴ as a framework for reporting the findings of our qualitative study: participant responsiveness, adherence and exposure/dose, recruitment, quality of delivery, comprehensiveness of the intervention and programme differentiation. Given that the relevance of each component of fidelity varies for a given trial,⁴⁶ our application of the Hasson framework has been broad and flexible. As a result, some of the fidelity components have been interwoven throughout the results section (e.g. context), and some have been presented together (e.g. adherence and exposure/dose). The chapter concludes with an overall discussion of intervention fidelity during the ASSSIST-2 trial and a series of key areas to consider when implementing Social Stories in the future (programme differentiation).

Defining usual practice

What is usual care for autistic children in schools in Yorkshire and the Humber?

Given the high level of variation in the support offered to autistic children, it was important to fully understand what constituted 'usual care' within our sample. Data regarding the usual practices of schools were collected through teacher questionnaires completed at baseline and 6 months post randomisation (see *Appendix 2, Table 59*). As it was a criterion for randomisation, all teachers completed the questionnaire at baseline, providing data on what classroom-based support and specific interventions each child had received within the 6 months prior to the start of the trial.

At baseline, most children within our sample (76.3%) received some form of support within the classroom from someone other than the class teacher. In 96.8% of these cases, support was provided by a TA. Teachers also reported that 16.1% of children were receiving support from someone else during lesson time. Such individuals included behavioural mentors or behavioural leads, learning support, higher-level teaching assistants (HLTAs) and SENCO or SEN support or speech and language therapists. This support was well balanced between the two trial arms. A similar profile of support was also reported in the 6-month questionnaire, which was returned by 212 teachers (85.1% of the total sample), indicating that there was very little change over time.

To capture the interventional support being received by the children, teachers were provided with a pre-specified list of commonly used interventions from which they could select those implemented within the last 6 months. There was also the option for teachers to detail any other interventions delivered that were not in the specified list.

At baseline, 74.7% of children had received at least one intervention within the last 6 months, with 49.4% receiving two or more interventions. The maximum number of interventions received was nine; however, this applied to only one child, and all other children received six or fewer interventions. As per classroom-based support, the number of interventions received within the previous 6 months at baseline was well balanced between the two trial arms.

The interventions received both 6-month pre-randomisation and during the 6-month trial period are presented in [Appendix 2, Table 59](#). The most frequently used intervention at both time points was 1 : 1 mentoring, which was received by 37.8% of children at baseline and 32.5% of children at 6-month follow-up. Other frequently used interventions include LEGO-based therapy, Social Communication Groups or Social Skills Groups and Visual schedules. The proportions of children receiving each of the pre-specified interventions were relatively similar across trial arms. At baseline, 24.9% of children were reported to have received an intervention other than those on the pre-specified list, falling to 15.6% of children at 6 months. The proportion of children receiving these other interventions was again well balanced by the trial arm. Examples of other interventions used include Thrive, Sensory Circuits, Sports Therapy and Games Group.

Description of interventionists

What were the profiles of the interventionists delivering Social Stories?

The full profile of interventionist demographics is presented by the trial arm in [Table 14](#) (see [Chapter 3](#)). Most interventionists who were allocated to deliver Social Stories described themselves as TAs (83.3%). Other educational professionals involved in intervention delivery included a mixture of teachers, higher-level TAs, learning mentors, SENCOs, Emotional Literacy Support Assistants, Specialist autism teachers for the Local Authority (LA) or other individuals such as behaviour mentors, nursery nurses and SEN Specialists. The majority of interventionists were female (94.4%) with a mean age of 42.1 years (minimum 20; maximum 64). On average, interventionists had substantial experience working with CYP, with a mean duration of 12.7 years, though some had as little as 1 year of experience. Very few interventionists considered themselves to have in-depth knowledge and experience of working with children, specifically with ASC (7.4%), with almost a quarter describing themselves as having limited knowledge (22.2%).

Participant responsiveness

How do participants respond to or engage with the Social Stories intervention?

The majority of parents, education professionals and members of the research team in our sample had some prior knowledge of Social Stories before taking part in the ASSSIST-2 trial, including general awareness of Social Stories as a methodology for supporting children on the autism spectrum (e.g. through training events, workshops) and prior experience of using non-personalised (i.e. not Carol Gray) version of Social Stories. For many, this experience was a key motivation for taking part in the trial, with participants keen to expand on their existing knowledge and obtain a more thorough understanding of how Social Stories could be used to support children to cope with a range of social and emotional situations both during the trial and in the future. For instance, parents spoke of wanting to learn how to apply Social Stories to manage challenges encountered at home, while education professionals wanted to further their understanding for their own professional development and to facilitate the more routine use of Carol Gray's Social Stories at their school. Less positively, and speaking to the idea that Social Stories were viewed as one, of a range of tools available to help children on the autism spectrum, Social Stories were seen as a 'last resort' for one parent.

I have worked with children with autism before and had social stories done but it's never been something that I have been particularly part of. It's been something that maybe the well-being team have done or the child's one to one and it's always been quite like, if they were going to the next class there's been a social

story or we're going on a trip and there's been a social story. It's never been for a particular goal. It's always because they've been going to something different or out of routine and I think it was good for us to see that social stories can be used for actual goal setting, not necessarily because there has been a change to routine. Actually, it can be used in a way to develop them and achieve new objectives. So I found it as a teacher, I've been in schools 16 years and have been teaching for 10. I think for me it was good to see it used differently and how actually I can use it as a teacher.

Teacher 3

Irrespective of whether participants had prior knowledge or experience of Social Stories, the intervention exceeded the majority of participants' expectations and was perceived to have positively impacted the child. Across all participant groups, the fact that goals and Social Stories were personalised, delivered one to one and tailored to address the needs of each child was consistently highlighted as a major strength and was deemed pivotal in ensuring children and those delivering the intervention were engaged with and accepting of it. Particularly for school staff, having the opportunity and dedicated time to focus on an individual child and produce tailored stories and goals was a major benefit, while also facilitating the delivery of the intervention.

I think (my child) had an opportunity to be part of a story that is about her and I think that makes it a little bit more special and she has a lot of 1:1 time with various people throughout her week at school but I think it's just given her that little bit more purpose to that time, so she's got something to look forward.

Parent 9

One of the aims of Social Stories was to improve school staff's understanding of potential underlying reasons for the behaviour of the children (e.g. mindblindness, difficulties getting the gist, language/communication difficulties) and to use this knowledge to help them better support the child in the school setting. Reflecting their motivations for taking part in the trial, school staff felt that through training they had a far more detailed and improved knowledge of ASC. Individuals who had previous experience using more generic versions of Social Stories (that do not meet the Carol Gray criteria) and/or had attended previous training about Social Stories praised the ASSSIST-2 trial training for being more detailed and improving their knowledge and understanding of ASC and Social Stories. More specifically, school staff described how, by increasing their knowledge of ASC, they were able to 'think differently' about the child and often spoke of having a greater understanding of the child's behaviour and the reasons for it, which improved how they responded to the child in situations they struggled with. Even those with limited involvement in the delivery of the intervention (e.g. SENCOs or Headteachers) described the clear observable benefits of using Social Stories and taking part in the trial for their own professional development and, in general, they felt that through increasing their knowledge of ASC, their ability to apply and understand Social Stories had been improved.

It's been so beneficial to the child; I think when you take part in things like this you think oh, will it actually impact the child or are we doing it for us and actually it was definitely for the child and it's impacted us as teachers, us as educators because we have developed our own CPD and our knowledge but also it's had a huge impact on the child ... I know with our child it's had a huge impact, positive impact ... I would recommend anyone else to do it because I think it was really, really useful.

Teacher 3

However, support for the intervention was not universal, with some participants describing how they had not observed any impact on their child's behaviour. Criticisms of the intervention were largely related to Social Stories being too repetitive, lengthy or overly simplistic/childish for that specific child. Participants acknowledged that because each child and situation being addressed is different, engagement and responses to the intervention may vary. The importance of not dismissing Social Stories because they were not well received by one child in one specific context was therefore emphasised.

Even for those who did not perceive the intervention to have been beneficial on this occasion, support for Social Stories was strong and is exemplified by school staff and parents describing how they have either already used or plan to continue to use the intervention outside of the trial. For instance, parents explained how they have used Social Stories with siblings of trial participants, and school staff described having used Social Stories to support children who were not in the trial during transitions (e.g. new starters, children returning to school after absences due to COVID-19 school closures and/or sickness). Additionally, some school staff described how they plan to adapt and continue using the stories developed in the trial in the new academic year and/or use the knowledge they have gained to create new Social Stories, highlighting their beliefs that the intervention was of benefit.

I've saved them and put them in a pack to move them up to the next teacher ... I've spoken to mum that it should continue and especially coming back from the break I think he'd really benefit from it. So it is something I'm hoping he will continue but me personally, as a teacher, it is something that I will continue to do with future classes and pupils ... Another child in my class who at the minute really struggles with feeling like she's bad if she gets something wrong ... She'll be really hard on herself and that sort of building resilience, I feel that would really help her as well and I have mentioned it to the next class teacher.

Teacher 4

The questionnaire data completed by interventionists at 6 months post randomisation ($n = 72$) revealed that 44.4% of the interventionists who were trained to write and deliver Social Stories in the trial were still using the intervention. The majority of these interventionists reported using Social Stories weekly. Some interventionists reported only using them 'when needed' and 'for as long as needed'. Other interventionists reported that they had supported colleagues to write Social Stories for other children, either by working on them together or by providing information about the child they work with. The majority of the goals for these additional Social Stories were school-based and focused on managing transitions, classroom behaviour (e.g. asking for help when needed, putting hands up, making mistakes, listening) and managing emotions. When interventionists were asked if they had made any adaptations to the original training they received, 86.11% said 'no'. For the 12.5% of interventionists that responded 'yes', the adaptations included simplifying the text, updating previous stories for a new year group and adding more information/pictures. These adaptations do not deviate from the Carol Gray methodology, which encourages refining the story during implementation if needed as part of her criteria 9. Twenty-five per cent of interventionists reported teaching Social Stories to someone else and this typically involved sharing the training video, PowerPoint slides and materials they had received with other members of staff and guiding other staff to write the stories using the training they received. When asked if they felt the school had supported them in using Social Stories, 88.9% said 'yes'. These results reflect the findings of the qualitative interviews that some staff had planned to continue using the intervention within the school setting. However, as 55.6% of interventionists said they were no longer using Social Stories 6 months post randomisation, it is important to consider the reasons for this, which are explored under 'quality of delivery'.

It is also important to reflect on how experiences of being involved in the ASSSIST-2 trial may have influenced an individual's willingness to engage with and deliver the intervention. Mirroring the largely positive views of stakeholders towards the intervention, participants enjoyed partaking in the trial and the opportunity to gain detailed knowledge of Social Stories and ASC. Negatives of trial participation related mainly to the additional time associated with completing questionnaires and difficulties understanding how to answer some of the questions that were posed within them. For instance, one parent raised an objection to rating themselves as a parent. However, for many, these concerns, particularly those relating to time, were outweighed by the perceived benefits of the intervention and/or mitigated through the support, communication and availability of the research team to support participants throughout the trial period.

I've enjoyed taking part in the study. I think it has been interesting meeting people who are knowledgeable about these things because sometimes I think in school you kind of get oh we're doing social stories and you're like 'and that is?' and you're expected to do something with it and you're like I don't really know how it works. So I feel that has been really nice. I know it is a commitment on our time and things from our point of view but I feel like we've had the training and feel more confident and have it as another strategy in the toolkit to kind of help support children who might be facing some challenging behaviours and don't really know why. So yeah, no I've enjoyed taking part. I think my TA has as well and I think it has really benefited the child.

Teacher 5

Fidelity: adherence, exposure and dose

Did the children on the autism spectrum receive the intervention as intended?

As described in [Social Stories™ intervention](#), there were very few specific criteria to which delivery needed to adhere to. Stories needed to be generated according to the 10 Carol Gray criteria; however, because the trial team worked collaboratively with schools to generate stories, we were able to ensure that all stories had maximum fidelity to the criteria. During the ASSSIST-2 trial, participants were asked to deliver the intervention a minimum of six times across the 4-week intervention period. Further highlighting the importance of considering the unique needs and preferences of each child when creating and delivering social stories, the pre-specified frequency for delivering social stories was met with a mixed response from participants. For example, while some interventionists reported that children enjoyed reading the same story, for others, the novelty of stories was reported to have 'worn off' after the first time reading them. Repeating the same Social Story was deemed to have made it difficult for school staff and parents to make the intervention 'exciting' and to have had a detrimental impact on some of the children's engagement and enjoyment. As a result, some children were reported to have not seen the value in social stories and questioned why they were being asked to engage with them. Other children really enjoyed the repetitive nature of the intervention. The divided opinion surrounding the frequency of the intervention meant that whether participants delivered Social Stories to the prescribed frequency varied during the intervention period; indeed, over one-tenth of the sample is known to have received less than the prescribed dosage, with a further quarter missing the compliance data. Some altered the number of sessions according to the perceived engagement and enjoyment of each child and, in rare cases, stopped using stories altogether.

It was quite interesting because the young man concerned when we started it I was reading to him and by the end of it because he knew it so well because we'd done it so many times, he was reading it to me quite happily and he'd say to me, again? Again you want to do it again? I was like aha, I want to do it again.

Interventionist 3

He didn't really respond other than the first time, because it just keeps repeating itself, to do it six times, he just doesn't want to read the same book six times. Whether it was about him or not, he just wasn't interested in reading it that many times It's very difficult to make it exciting to say, we're going to go and read this book again, do you know what I mean, which is what I said at the beginning, you know, maybe we need to get around it, you, if he can keep reading it.

Interventionist 2

Social Stories are written in a specific way using a defined formula and sentence types. However, the appropriateness of this approach was perceived to vary depending on the preferences, age and receptive and expressive skills of each child. For example, the length and amount of language used within stories could mean they are pitched beyond a child's level of understanding, and some children were perceived to respond and engage better with shorter, concise stories. Equally, others criticised stories for 'over-explaining' key messages, which was considered either 'patronising' or to mean key messages were lost

and not understood depending on the individual child. The importance of tailoring Social Stories to the preferences and needs of each child was emphasised in line with the Carol Gray recommendations.

There was so much of that [preamble] and kind of qualifying things and explaining things that when I read it the message was totally lost.

Parent 2

Despite these challenges, participants were keen to ensure that Social Stories were not dismissed as a method due to an individual child not engaging in the intervention. Instead, participants suggested raising awareness of the fact that there are certain children for whom the intervention is less likely to be effective, for example, children who struggle with repetition or older children who may recognise the intent of the story or find it overly simplistic. Similarly, some participants suggested speaking to teachers and parents before developing the goals and Social Stories to ensure that they are not being developed for children who are unlikely to engage with them. Adapting the Social Stories training to include role-play scenarios of delivering stories to a child who is not engaged in the intervention was also suggested.

To have identified children that would be sort of receptive to it because, you know, I can think of a couple of cases where a child refused to read the story more than once and the whole point of it is you read the thing over and over again. But this child didn't like repetition, so in that case a social story is going to be completely useless to them because the whole idea is you read this thing so many times that the information goes in. But this child didn't like repetition, so the second time that they tried to read through the story the child was like, nope. I can think of another child who received the social story who thought it was patronising because they understood it, you know, they had a higher level of understanding and said it was really patronising.

Trainer focus group

The COVID-19 pandemic also had an impact on the perceived benefits of the intervention and the frequency with which it was delivered. Periods of school closures naturally meant that, in some cases, the intervention could not be delivered or was stopped. Some parents/carers perceived there to be a reduced need for the intervention during this time, as children were considered to have reduced stress levels when being home-schooled. In fact, this may have indicated there is even more of a need for the intervention if removing the school environment is reducing the stress levels for the child. Additionally, the usefulness of goals that were set during the COVID-19 lockdown was considered to have been reduced post lockdown, at which point, if the story had not been written, the goal needed to be discussed again with the teacher and a new one was set. However, parents acknowledged that their ability to assess the impact of Social Stories on their child's behaviour due to the significant changes to their lifestyle that happened at this time made it difficult to attribute and isolate any perceived benefits or dis-benefits to Social Stories.

To be honest we've not used them at home because he's not been interested in it and with lockdown he's not actually needed it because all the stresses of school have not been there and we've not had any issues at all ...

Parent 3

Given the importance that was placed on considering the context (e.g. impact of COVID-19) and varying needs, preferences and abilities of children on the autism spectrum, participants considered it difficult to prescribe a universal frequency and duration for delivering Social Stories in the future – suggestions varied between daily, weekly and more infrequent use of Social Stories. Indeed, some participants provided examples of how they either are or plan to continue to use Social Stories, albeit less frequently, according to the needs of each child and the resources they have available.

it is on a different basis but I would say at least 2 or 3 times a week on average. So they see it every day because obviously it's behaviour that you would want them to see and they're young children and they forget.

Teacher 2

Suitability of Social Stories for children on the autism spectrum

Which goals are most appropriate for Social Stories?

Participants felt that Social Stories had the potential to support children with a wide range of social and emotional developmental needs and that their use should not be restricted to children on the autism spectrum. Participants were specifically asked which behaviours and support needs they felt Social Stories could be used for and, in response, they listed a wide range of appropriate behaviours relating to areas such as reassurance, emotional well-being, social skills, understanding others' perspectives, personal care information (puberty, toileting), transitions (e.g. new school year, returning following periods of absence), support seeking, behavioural regulation and understanding sensory issues. One group, which participants felt may experience particular benefits from Social Stories, was children with anxiety. More specifically to the school environment was the potential use of Social Stories in supporting children around transitioning to new schools and/or classes/year groups or returning to school after a period of prolonged absence (e.g. due to sickness or COVID-19 school closures), topics that some teachers commented they had seen Social Stories used for the most prior to the study. Additionally, participants stressed that Social Stories should not be viewed as a 'one-stop-shop' and that it was important to consider creating new stories with the same child to address different and/or new challenges that they may face.

I think I would say nearly every child, I think it would be quite easy to set a goal for them of some sort because I think there's always something in schools that children need a little bit of help with.

Interventionist 1

There were, however, some situations or types of behaviours for which Social Stories were considered less appropriate by participants. For example, Social Stories were not deemed appropriate by educational professionals for use in 'crisis situations' and/or for resolving complex issues such as those pertaining to child protection or a child's sexuality where it was felt that strategies and/or interventions designed specifically for these situations should be used. Carol Gray's Social Stories are designed to be a flexible intervention, and although some educational professionals felt more complex topics could not be addressed within the stories, this may not be the case if the individual writing them has taken time to gather sufficient information about the situation and the child so that the story is personalised and tailored to the child's receptive understanding.

I suppose personal to them as in things that would need to be discussed further. Your child protection type of things wouldn't be suitable. Anything sexually related definitely not ... I think a story is okay if it's ... the simple one step problem, I suspect that any more complex wouldn't work.

Interventionist 3

Despite the focus of ASSSIST-2 being on the use of Social Stories in a school setting, parents were invited to the training in order to help shape the development of the story and to gain the skills necessary to develop Social Stories at home in the future. To maintain consistency, educational professionals and teachers were asked to keep the Social Story developed for the trial specific to the school environment. By the time parents were invited to the interview, some had tried using Social Stories in the home setting too. As a result, during interviews, parents and school staff described their views on the use of Social Stories at home, both during the trial and in the future, giving us some insight into their perceived applicability in this setting. In some cases, the uniqueness of the home environment precluded the use of the intervention at home, and this was largely due to concerns of 'singling out'

a child from their siblings by delivering the Social Stories intervention with them and finding the time to write and deliver the stories at home. Other parents described making adaptations to facilitate the use of Social Stories at home and so deviated from delivering them strictly according to the Carol Gray method. Many of these changes involved making pragmatic changes so that stories were more informal and less intensive, and this was considered to be more realistic as to how they would be used if developed for the home environment in future, particularly if different stories were being used at home and school. For example, participants described reducing the frequency with which stories were delivered, reducing the content or creating stories using only pictures, and, in some cases, involving older siblings in creating and delivering stories.

When I did write one I was careful to write it so that it would work for both kids because I'm quite aware of not trying to single [child] out as having problems. He knows that he has autism but ... I don't want him to feel embarrassed about it or that there's anything wrong with that ... How to introduce them in a way that doesn't make a child feel like you've done something wrong and I'm going to read this to correct you. So at school they know that they have lots of different situations where they go with the teacher to do something ... so it's not word if a teacher just bring it out ... we're going to read through this together. Whereas at home it is a bit weird. When and how do you use this in your day if you're not having a session with a professional?

Parent 2

Quality of delivery

Which factors of delivering Social Stories were particularly important?

To support the delivery of the intervention, many school staff described how they had developed a routine, which ensured that the Social Story was delivered at the same time each week. Reasons for adopting this routine-based approach to delivery were largely pragmatic and were devised as a way to overcome difficulties in finding the staff and time needed to deliver Social Stories regularly within a busy school environment. Having regular, dedicated times for delivering the intervention therefore allowed staff to be allocated and time to deliver Social Stories to be protected.

As a teacher obviously we are directing the teaching assistant, so we can say, look 10 minutes on a Monday morning needs to be this social story, every Monday morning. Normally we'd do a timetable as well, so that can literally go on a timetable so it's a regular thing. That just makes it a bit easier.

Teacher 4

Despite advocating for protected time slots for delivering Social Stories, school staff also highlighted the importance of ensuring that the intervention is delivered at a time that is right for the child. Interventionists in particular emphasised the importance of ensuring that any planned times for delivering the intervention do not coincide with overly busy, stimulating and/or pressurised periods in the classroom. To ensure children are willing to engage with the story and maximise the potential benefits of the intervention, participants also described the need to adjust when the intervention is delivered in response to a child's specific needs, preferences and behaviour. For example, one participant described how they delivered the intervention while a child was doing a craft-based activity, as this increased the likelihood that that child would sit and listen to the story for its duration.

The only problem was sometimes if I'd timed it and there was something going on in the classroom he was more wanting to be engaged in that was tricky because he'd want to rush through the story to get back out to whatever was going on. So I just adjusted the timings when I went it to make sure that wasn't happening.

Interventionist 6

Interventionists also highlighted the importance of ensuring that Social Stories are delivered in a quiet, private, neutral, comfortable and familiar environment to the child without interruptions to promote discussion and for children to feel as though they are not in a formal lesson. This helped children's concentration and engagement with the story while having the added benefit of ensuring that children were provided with dedicated 1 : 1 time which was protected from distractions and interruptions from other children who may want to be involved and avoided others overhearing personal and, in some cases, sensitive story content. While such an environment was considered ideal, the challenges of finding such a space regularly within a busy school environment were acknowledged.

I think a more ideal space would be somewhere you can kind of be focussed on and feel kind of comfortable and able to ask questions because really you want to promote that discussion and get to the bottom of what is causing some of these because often it is something underlying isn't it that you're struggling to get to the bottom of and I just don't know if they'd be willing to kind of come out with that or have that discussion when there are people toing and froing all the time.

Teacher 5

When reflecting on the way the intervention was delivered, it is also important to consider the role of 'who' is responsible for delivery. While some participants stressed the importance of ensuring that those involved in writing and delivering social stories are experienced in dealing with children with varying emotional and social support needs and had experience and training knowledge of ASC and Carol Gray Social Stories, others felt that this knowledge was secondary to knowledge of the child. Indeed, across participant groups, knowledge of the child was described as the most crucial prerequisite for maximising the benefits of the intervention. This was largely as close relationships with the child were seen to facilitate exploration of the underlying reasons behind a chosen behaviour and mean that Social Stories are focused on an area that is of benefit to that specific child. More specifically, the process of creating stories and identifying and narrowing down goals was considered easier when undertaken with someone who understands the child, their behaviour and what motivates them. Indeed, trainers described how during training those without knowledge of the child struggled to 'put pen to paper' and made it difficult for trainers to ascertain whether the story and goal that had been created was right for that child and, in these cases, participants were advised to gather further information from other teaching staff and the child's parents. Ensuring that the Social Stories were created to address the specific needs of each child was also considered paramount to the intervention's success, with children described as more accepting and willing to engage with an intervention that was tailored to their specific needs and delivered by someone they know and trust. Perhaps exemplifying this point, a couple of interventionists described how the goal-setting and story development was more challenging at the start of the academic year when school staff had less familiarity with the child and is something to consider when planning to use the intervention in future.

I always found it much easier when the person setting the goal really knew the child. Because typically we'd have the teacher setting the goal and then their teaching assistant would be the one reading the social story and it could be different ways around. Some schools were a bit more flexible with that but especially like September, that was a really difficult time to set goals because even though we tried to ask the teacher from the year before, it wasn't always possible to facilitate kind of getting them out of the class to do. But when a person really knew that kid and really understood what motivated them, what they enjoyed and also like maybe taking a bit of time to think through what they were experiencing, was a much easier process to set that goal. So I definitely feel like it needs to be someone that knows that child and maybe September is not a good time for a teacher to be setting the goal.

Focus group with research assistants

In a school environment, teachers, TAs and SENCOs (if trained appropriately) were considered to have the potential to deliver the intervention and develop goals and Social Stories. However, TAs were considered best placed to create and deliver a personalised intervention that was tailored to the needs

of that specific child. By working with the child directly on a daily basis, TAs were deemed to have the closest relationship with the child and may have experience delivering other interventions with the child, allowing a sense of normality for the child. RAs and trainers highlighted the importance of ensuring that those involved in setting the goals and creating the Social Stories also deliver them, not only for continuity and engagement but also to ensure that those delivering the intervention understand the purpose of the intervention is to support the child.

I think it was really worthwhile for the people who actually write the social stories who work with the children ... They [TA] are the right person because obviously they're the people who work with the children day in and day out and they are the one that has the relationship.

Teacher 2

For school-based interventions, some participants also felt that there needed to be greater involvement of teachers to work alongside TAs when identifying goals and creating the Social StoriesSM. The generally higher levels of computer literacy amongst teachers were considered particularly important, as this was described as a barrier to some individuals being able to add pictures and personalise social stories. Additionally, given that teachers have overriding responsibility and are the main point of contact for a child, they often receive key information (e.g. child focus plans) that they pass onto TAs, their involvement is important.

The teachers really enjoyed the sessions. They were really competent in writing the sentences and they were more competent with technology. But then you often need the sort of personalised input that the TA might have because they work so closely with the child. So it sort of needs to be a collaborative process. The teachers really can't do it without the TA but equally most TAs can't do it without the teacher. So it does need to be a sort of combined effort to get these stories together and that's often what was lacking.

Focus group with trainers

There was some uncertainty amongst participants as to whether, as a school-based intervention, parents should be involved in creating the Social Stories. Some participants felt that parents should be involved irrespective of whether the behaviour and/or situation being addressed applied to school, home or both environments. This was largely due to beliefs that there was a need for transparency and to acknowledge that most parents would want to be aware of any interventions/strategies that were being used with their child, even if they were not involved in the process directly. For instance, even in situations where a goal was only applicable at school, some felt it was important to keep an open dialogue with parents about the purpose and delivery of the intervention and to offer them training and support in designing their own stories at home. A number of parents reinforced this and felt that to help their child's social and emotional development, they needed to be aware of any support their child was receiving at school. Some teachers and interventionists echoed this and felt they would not feel comfortable delivering interventions if the child's parent was not aware of it and in support. Parent involvement was also considered beneficial for facilitating the writing of the Social Story, with the unique knowledge that a parent brings of their child considered crucial to providing insight and context to a child's behaviour. Similarly, involving parents in the development of the Social Stories was considered beneficial for improving relationships and communication between parents and school staff.

So I spoke to the parent about the goal and then gone to the school and set the goal which was to do with hiding under the school table during lesson time and then when I went back to the teacher. So the story would have ended up going down the road of, try not to hide under the table da, da, da, and then when I went back to the parent she was saying, actually that's a behaviour that the child is probably doing because she's feeling anxious about something. So it was very similar. It was like explaining that behaviour and it would have, again, gone down the wrong route had the parent not been involved.

Research assistant focus group

While many advocated the inclusion of parents, others felt they had a limited role in creating and delivering a Social Story in a school environment. For some, involving parents during the goal-setting process had led to some discussions and disagreements about the focus of the goal for the child. In these situations, the fact that the trial was focused on addressing school-based behaviours was perceived to have potentially exacerbated tension between parents and school staff, particularly if a child's needs differed across the two environments and parents viewed the research project as their 'one chance' to get help with a behaviour. Participants also acknowledged that there may be less relevance of including parents for children on the autism spectrum for which it is important to separate school and home life, with home for these children considered a 'safe zone' where the pressures of school are not addressed. While researchers and trainers found that pre-empting these disagreements and explaining to parents in advance of the goal-setting meeting, the school-based nature of the research helps in relieving this tension. Researchers/trainers also described themselves as having a mediation role when disagreements occurred, and it is important to note that it may be beneficial to have a third party, such as a teacher or TA not involved in the intervention to provide this role, although in practice this may be difficult.

I think it probably depends on the child with what you're wanting to tackle really. I think for school-based issues, I think it's just best that school do it because they know what he's like in that situation but then I wouldn't expect school to write one for a home based situation because they wouldn't know really. I think whatever situation you're trying to tackle you want to choose the best person to write the story.

Parent 7

Mirroring the need to tailor the intervention and consider the needs of each individual child, participants had varying views on the appropriateness of involving the child in the creation of goals and Social Stories and recommended that this should be considered on a case-by-case basis. Interventionists (typically TAs) and the research team who advocated involving the child felt adopting this approach may be of particular benefit for children where it is possible to anticipate issues with engaging them (i.e. older children, those who feel 'babied' by the intervention, and/or those who do not like repetition) as it may give them a sense of ownership and, where appropriate, an understanding of the purpose of Social Stories. Some participants felt it was difficult for a child to understand the 'intricate' process of goal-setting and the need to create Social Stories within the parameters specified by Carol Gray. There were also certain discussions around goals that were considered to be inappropriate with the child present (e.g. the impact of the behaviours on parental stress, the educational professionals or the wider classroom).

I don't know whether it's an age thing but certainly it's nice isn't it to brainstorm with them and get them ... It's their goals, it's not my goal ... It should be something that he wants to do and I think if you've got more buy-in at that stage rather than just something arriving, oh we've written a story for you because we think you need help ... It's like a performance review without the person who's actually doing the work.

Parent 6

Ultimately, when considering who should be involved in goal-setting and Social Story development and delivery in the future, it is important to ensure that the person/people with the closest relationship to the child are involved. Despite the various pros and cons of involving different stakeholder groups, the majority of participants advocated adopting a collaborative approach (e.g. TAs, classroom teachers, parents) during goal-setting and story development to enhance the overall quality of the story and to encourage reflection and different experiences of the child and their behaviour to be discussed and incorporated into Social Stories.

Comprehensiveness of the intervention

Which factors were particularly important for developing and delivering the overall intervention?

To support the use of Social Stories in the future, it is important to learn from participants' experiences during the trial and identify what helped or hindered them to deliver the intervention as well as to identify areas for improvement. These are each discussed in turn below.

Unique features of Carol Gray Social Stories as facilitators to intervention delivery

The ability to personalise Social Stories to the needs of each child was consistently identified as a main benefit of Social Stories. For example, incorporating pictures and photographs was perceived to facilitate learning and enhance children's engagement with and enjoyment of the intervention. This ability to tailor the intervention was also perceived to accommodate the individual learning needs of each child. For instance, one parent described how using pictures was particularly helpful for their child, who they considered to be more visual in their learning and who it was felt would not have engaged with or understood an intervention that was solely word-based. Equally, the ability to make adaptations or adjustments to the content of Social Stories to reflect changes to the environment or an individual child's needs was considered a huge benefit, particularly for children on the autism spectrum, who tend to prefer information that is literal and accurate. For example, changing the names of teachers throughout stories if substitutes were in place.

The fact it was personalised he really enjoyed that. He saw his pictures and his friends and she made it have a train aspect to it as well and he loves trains, so he was wanting to flick to the next page and find out more about himself.

Parent 8

The structure of the Social Stories also facilitated the development and delivery of the intervention. Breaking stories down into 'manageable chunks' (e.g. clear title, positive introduction, different sentence types to include) gave school staff an understanding of the importance of including each story component. Participants felt that the structure of the story and how it was tailored to the individual needs of the child were beneficial for helping the child to understand complex social situations. Many stories focused on gently guiding children to communicate their emotions to an adult. Participants reported that they had observed clear changes to a child's behaviour following the story, and they felt that this was due to the information provided in the Social Story which had helped the child to better understand why it is important to tell an adult how they feel and potential methods for doing so.

Navigating the goal-setting process

Despite the goal-setting process being considered pivotal for understanding what underlies a behaviour or situation and praised for providing individuals with the chance to talk about and understand a child and their behaviour, specific challenges with goal-setting were discussed. One of the main difficulties participants faced during the goal-setting process was identifying a goal that was specific, beneficial to the child and realistic to achieve within the 6-week intervention period in a school setting. Goals based around routines within the school day and playtime, as well as those around understanding and communicating emotions and making mistakes, were all considered to work well and be child-oriented goals. The 'goal-setting sheet' and SMART framework (used to ensure the goals were Specific, Measurable, Achievable, Realistic and Time-specific) were considered crucial in overcoming many of the challenges identified with the goal-setting process, particularly in terms of 'solidifying' goals and for understanding how to break goals down into achievable steps from umbrella goals. The goal sheet was also deemed useful for ensuring goals were suitable, of high quality and designed in line with Carol Gray methodology. Additionally, while considered time-consuming to use, as many school staff were already familiar with the 'SMART' framework, the research team felt that this helped participants to identify and refine goals and ensure they were positively framed. Conflicting priorities between parents and school

staff and difficulties redirecting goals away from something negative and/or vague were commonly reported issues. The goal sheet and SMART framework helped the research team to navigate difficult conversations and differences of opinion and ensure goals were focused on addressing the individual needs of each child and were not too vague. For example, if a goal around 'managing playtime better' was proposed, these resources were used to help determine if this is something that is in the child's interest, and if so, the goal was converted into something more specific, for example, 'to try to talk to some of the other children during playtime'.

It [the SMART framework] was good and it makes you think sometimes a little bit outside the box to make sure that it was achievable with the sort of parameters in mind I think. What we maybe wanted to go with perhaps sometimes didn't exactly fit right ... I think that was good.

Teacher 2

However, there were differences within and across stakeholder groups in terms of their confidence in creating goals, reflecting the different degrees of familiarity with autism interventions and, more specifically, Social Stories, goal-setting and resources such as the SMART framework. As such, the extent that the SMART framework and goal sheet were considered useful varied, with some trainers describing how they felt participants would not have used these resources unless encouraged to do so. While some found these resources ideal prompts to facilitate goal-setting as per the Carol Gray method, others found them time-consuming. Participants made a number of suggestions for the future to facilitate the creation of achievable, child-oriented, positively framed goals. For instance, including a section on 'general information about the child' on the goal sheet to ensure that, at the start of the goal-setting process, individuals are given the opportunity to discuss the behaviours and needs of that child, increasing the likelihood of creating a tailored and child-orientated goal. Additionally, providing examples of goals that have been broken down and emphasising the need for positively framed goals in resources was proposed. In the current trial, the goal for each individual child was set before randomisation and in cases where there had been a significant gap between baseline measures and receiving Social Stories training on occasion educational professionals asked for the goal to be updated when they met the trainers. Outside of a trial, it would be possible to set the goal as part of receiving Social Stories training, and therefore this issue could be avoided.

I think in terms of making it a bit easier, I think for the sheets it might be good to have a section that is just like, generally like about information about the child. So, when we do the kind of training sessions to write the stories, we often say okay can you tell me a bit about the child and then we'll start writing the story. I think that's just quite a nice opportunity for the teacher to just, or the intention is to just like chat about the child because in the training session that's where we get their kind of likes and dislikes and all that from. But I think for the goal setting it would be nice to do that as well and then you can just ... because you end up scribbling loads of stuff down on the sheet and then you're like trying to set the goal and you look at it and it's like all over the place. So, I think if there was a section that was just like kind of just general information and then moving onto, okay what is the problem. Because sometimes you'll write it all down after you've said, do you know, what's the problem at the moment and they'll say loads of stuff and then they'll say, oh actually so what we'd like to focus on is this or you have to like pick something from what they've said. So, I guess that's more of a admin thing for us that might make it a bit easier but I think it would also make it easier for the teacher as well just to give them that time to talk about the child a bit more.

RA focus group

Training model

How were the research team trained to deliver Social Stories training and their experiences of this?

Profile of trainers

Four academic staff were trained to deliver Social Stories training to school staff (all female, mean age 31). All trainers had an undergraduate or master's degree in psychology and three had completed a PhD in the field. Through their qualifications and work experience, the trainers had prior knowledge of autism and had some experience delivering workshops/training/presentations to staff and higher education students.

The training model

A cascade model of training (train-the-trainers approach) was implemented. The former Chief Investigator (Professor Barry Wright) who is a consultant child and adolescent psychiatrist and has extensive experience of autism and Social Stories trained the first trainer by delivering mock Social Stories training sessions, role-play scenarios, answering any queries and then observing the trainer deliver the session themselves. The trainer was only 'signed off' to deliver training independently when the Chief Investigator was satisfied with the performance. Due to staff changes, two further trainers were trained by the first trainer. This involved shadowing Social Stories training sessions in schools, becoming familiar with all training materials and resources and then being observed delivering the training themselves before being 'signed off' to deliver training alone. The second set of trainers were then able to train a fourth trainer following the method described above. The quotes below describe some of the trainers' reflections on the training process.

My training will probably have been a little bit different from everybody else's because it was [Chief Investigator] who delivered to me ... [Chief Investigator] delivered the training with sort of his slides that have since been slightly amended but none of the content really has been amended ... So it was like imagine a child that you know and, you know, what might you put in these different sections [worksheets] to get us a little bit more familiar with what it was that we were looking for when it came to like what do we mean by mind blindness, how would this, so on and so on. It could be a real child that we knew that had a diagnosis or it could be one that we just kind of made up and try and fill in to understand the sections. I didn't have any training 'in practice', if that makes sense. So after that training it was just go out to schools and deliver it and then for a long time it was just me delivering it until [name] came on the scene and sort of did training.

Trainer focus group

Just before I started as a Trial Coordinator on the study ... I shadowed three times and I think [name] you came long to maybe a couple of those as well. So we saw the presentation and the training getting delivered in practice, which was really useful, and it was good to be able to see it at a larger school because I think we went to one and it was very well attended by additional people who didn't know anything about the trial. So it was good to see that it changes in every school that you go to. But, yeah it was basically shadowing three times and then having a go yourself. I think, I also did some mock training with [name] and I sat with [name] who was a research assistant on the study at the time and we've had a look through the slides and just made sure that I understood the content before going out to deliver it for the first time.

Trainer focus group

All trainers reported feeling slightly anxious or nervous before delivering the training to school staff, but this did not stem from lack of understanding of autism or Social Stories but more from the uncertainty about how the session would go and the needs of the attendees. Worries discussed included the unpredictable nature of whether the goals for the children may need to be changed during the session, whether the stories may address difficult topics that are more challenging to write and the level of support required by the attendees within the short time frame of the training session.

As to how I felt about the training. I think for me what I felt uncertain about was less to do with the autism specific parts because the team that we sat within COMIC, like [Chief Investigator] is very familiar with autism, it's one of his sort of competencies. So, from that point of view I felt like I already knew quite

a lot from how we were talking about that sort of stuff in the team, you know, we had other trials running around autism and adapting assessment instruments for autism and so on. From that point of view I can't say that I felt unprepared. I did feel unprepared in the sense that I didn't know what to expect, so it was kind of like straight out into school alone and there was no checks and balances there. So like, am I doing it right, am I not doing it right and so on. Then after a while I wouldn't even say it took that long to sort of feel like, actually I feel okay with this, you know, in the sense that kind of similar to when I first started teaching, you know, when you first start going into it. Am I going to be able to answer any questions? Wonder about your own competency and so on and then you realise that the sort of questions that they tend to ask and the sort of guidance that they tended to need, I realised I was quite able to give that. But, yeah for sure, I was a bit nervous in the beginning that was a bit into the deep end but then again that's also how I like to learn.

Trainer focus group

Trainers felt that it is important that anyone who is trained to become a Social Stories trainer in the future has a background in autism and/or psychology in order to be able to deliver high-quality psychoeducation about autism that can help school staff to understand the possible underlying reasons behind a child's behaviour and challenges.

I personally don't think someone that didn't have a psychology background could just pick this up and run with it, you know. I think it would need to be like an educational psychologist or something like that that was delivering the training that would have that background level of autism because the training in autism and going through what is different about a child with autism and getting the school staff to really sort of understand it from a psychological perspective is almost more beneficial than the sort of story-writing process because you saw so many moments where we were explaining, you know, when a child behaves like this it's not because they're naughty, it's not because they're selfish and you could sort of see the cogs turning for them thinking ah. One of the sort of indirect ways that I feel like this trial has impacted is their learning about autism. So if you get someone who didn't have that background about autism that knew how to write a social story and be able to tick the Carol Gray criteria then the school staff wouldn't be getting the same sort of knowledge that's really going to benefit the children.

Trainer focus group

The trainers felt that the package of training for school staff needed strengthening so that school staff could confidently continue to use the intervention and train others following the trial. In order to do this, trainers proposed simplifying some of the resources provided to schools (e.g. worksheets and the terminology used within them) and ensuring that the school staff attending the training had sufficient time to process the psychoeducation about autism and think about how it relates to the children they are working with.

You've got large numbers of children with autism, so it wouldn't be practical for a practitioner to be coming in and writing social stories for all of them. So we need to move away from this really high level of handholding that we've had in this trial, so really intensely training, you know, one or two people within a school to really roll with this and not need that post support that we've given. So I think we really need to strengthen the training package for the practitioners, so that they can go forward and train the school staff.

Trainer focus group

I really liked the training on ASSSIST-2 because it emphasises not just positivity around the story and so on but also genuine understanding of the child, the child's needs and how we have to adapt and respond to the child's needs as well. I think that's quite important and not always recognised particularly as [name] says, if they don't have that much experience with autism and they kind of get seen as a bit difficult and like, you know, they're being naughty and like they're just really rigid and it's like, well for a reason.

Trainer focus group

Participants' views on training

As previously discussed in the section about participant responsiveness, irrespective of whether a participant had prior knowledge of social stories, training was regarded positively and considered to have increased participants' confidence in writing and delivering Social Stories and setting goals. Increased knowledge of ASC and the underlying philosophies of Social Stories was considered a key benefit of taking part in ASSSIST-2 training. Despite this, a small number of participants criticised the in-person training for being rushed and including too much content, particularly the volume of written text in the PowerPoint presentation about autism and the Carol Gray criteria for writing Social Stories.

The training was really thorough and because of the thorough training it just meant that the writing of the social stories was much easier to do.

Teacher 1

Participants were particularly supportive of the practical and interactive components of training, which were thought to greatly facilitate their ability to write and deliver the intervention with confidence. A number of participants emphasised that the reality of producing Social Stories and identifying goals was very different to hearing about them, even if you understood the process, and that it was only through the interactive elements of training that they realised how challenging the process was and identified gaps in their knowledge. In particular, having the opportunity to create a Social Story during training was considered crucial to intervention delivery and meant participants felt less daunted about delivering the intervention. Creating the Social Stories during training also provided a personal, supportive space to learn and for those that received training in person, doing so in small groups, through exposure to combined expertise (e.g. trainers, TAs, teachers and parents) and different perspectives gave individuals the confidence and knowledge to create a personalised Social Story that was tailored to the child as well as an understanding of the language and structure that is used within the Carol Gray methodology.

I think what was also great it was just three of use literally, so we could ask the stupid questions and we could kind of, you know, you could get quite individual about what it was about and kind of what was on your mind really which was nice as well. It didn't really feel like training. It's just sort of felt somebody guiding you through a process really which was good.

Parent 4

Of those who were critical of training, the majority were parents. This may reflect the fact that parents were not directly involved in delivering the intervention and that training was focused on preparing educational professionals to deliver the intervention in an educational information. The perceived complexity of creating a Carol Gray Social Story also caused participants to feel daunted at the prospect of writing their own story and stemmed, particularly for parents, from a lack of familiarity and prior experience of using Social Stories or similar interventions in comparison to school staff. For example, participants were nervous at the idea of remembering how to create and deliver a Social Story as per the Carol Gray method (i.e. the sentence types to use and style) and felt that while they may be able to 'give it a go' that 'it might not pass the hundred percent authenticity check'. Access to example Social Stories that were created both during training and in the manual mitigated these concerns to some extent. However, some parents still felt that to use Social Stories at home, they would need more examples and/or additional support and training before feeling confident to write a social story using the Carol Gray methodology.

I suppose it's given me the opportunity to build a social story in the future about other things. I haven't actually done that but I think it's knowledge and training and I've kept all the information that would enable me to do that ... If I was to create a social story at home ... I'd kind of feel a bit alone doing it because I was doing it in partnership with someone who knew how my son is at school. I'd kind of feel like I'd want that again I suppose, a couple of hours of someone's time just to work on it and I don't know if it would be feasible ... yes I would be interested, now and in the future, in being involved in some kind of, you know, once a year or whenever a meeting with school to try and keep that training and knowledge up and maybe work together.

Parent 6

During the ASSSIST-2 trial, training was originally designed to be delivered in person in a 3-hour session which took place either in the morning or afternoon. Due to restrictions put in place in response to the COVID-19 pandemic, from March 2020 the Social Stories training was delivered online. Changing the mode of delivery naturally had implications for how training was structured and delivered. During focus groups, trainers described how the adaptations that were made to training when transitioning to the online model appeared to result in less difficulty supporting individuals to write their own Social Stories. For instance, in comparison to the in-person model where trainers were supporting multiple groups of participants to develop Social Stories, the online model provided each interventionist with a dedicated 1-hour 30-minute time slot to create a Social Story for the child they were working with alongside a trainer. In advance of these sessions, trainers pre-populated standardised templates (e.g. a positive introduction), ensuring that these individual sessions were spent expanding and tailoring the Social Story with specific content relevant to each child. Educational professionals and parents attending the training were sent a 30-minute presentation to watch in advance of the 1 : 1 story-writing session. This presentation covered everything that was delivered in the presentation, which took place in person prior to the COVID-19 pandemic (e.g. psychoeducation about autism and the Carol Gray methodology). Additionally, at the start of the 'live' online session, participants were provided with a summary presentation, which included a brief explanation to recap Social Stories, an example Social Story. While it was acknowledged that the online model meant trainers had to allocate sufficient time to pre-populate a story template, this was considered preferable to the amount of post-training support trainers provided during the in-person model. Those who delivered training in person described regularly having to provide up to an additional half a day of support to ensure that Social Stories were completed and/or reported having to rewrite some of the Social Stories that participants had produced. In the future, regardless of whether training is delivered online, in person or using a hybrid approach, trainers emphasised that significant support is required to ensure that people are creating and delivering Social Stories as per the Carol Gray methodology and to ensure that the amount of pre- and post-training support required to produce social stories is not underrated.

Once we moved to the on-line where we've tended to sort of have one story at a time, so you know, we maybe [during the in-person model] had three interventionists in the room writing three stories and it was quite tricky to sort of manage that. Whereas, when we moved to the on-line model, we'd block out one hour slots with each of the interventionists which made things a lot easier, and because we sort of put together this template, which had a lot of the standard sentences. So you know, all the introductory pages were the same for each child so we sort of pre-populated that so we could just type in the details on the day. So I think once we moved to that model we had less sort of going over time, if you know what I mean, you know the story writing fit more into the session.

Focus group with trainers

Despite participants generally being accepting of online delivery models and acknowledging that due to COVID-19, training packages are likely to be delivered remotely in the future, for some their preference would always be for face-to-face delivery. Equally, concerns were raised about the potential for 'zoom fatigue' with such a comprehensive training package, and a very small number of participants experienced technical difficulties, which meant that due to poor audio quality, they may have missed some of the content during training sessions. Additionally, while the online model included interactive components, in-person events were considered more conducive to engaging and interactive discussions.

It's going to be the norm anyway isn't it from now on. Yeah, absolutely fine. I personally prefer face to face training but, yeah I thought it was absolutely fine. I was not left in anyway thinking, oh I didn't get that or I didn't understand that.

Interventionist 2

Irrespective of the model chosen to deliver training in the future, the importance of ensuring participants engage and most crucially that the interactive elements of training remain and are not removed for the sake of time and resource was highlighted. Participants stressed the importance of

ensuring participants are able to ask questions and allocating sufficient time to facilitate discussions. Keeping the ratio of participants to trainers as small as possible was also suggested as important for ensuring a positive and supportive learning environment and for facilitating discussion. Trainers highlighted the importance of not underestimating the amount of time that is needed for them to prepare to deliver Social Stories training (online model) or to provide post-training support (in-person model).

In addition to training, education professionals were provided with a number of additional resources (a Social Stories manual, example Social Stories and worksheets focused on understanding autism, setting goals for the stories, sentence types and intervention delivery) which were described as crucial in supporting the creation and delivery of Social Stories.

Making a series of example Social Stories available, in addition to those created during training, was considered particularly beneficial not only for consolidating knowledge, but also for providing a reference point and reassurance to participants that they were creating Social Stories as per the Carol Gray method. For those who described themselves as less confident in creating an entirely new social story or for those with limited time, example stories provided a source of inspiration and/or useful templates that could be tailored to the needs of an individual child. As such, educational professionals and parents advocated developing a booklet, with a large number of Social Stories to support the continued and wider use of the intervention in the future but stressed the importance of ensuring this does not prevent future stories from being tailored to the needs of each child. Exemplars and other resources (e.g. goal-setting worksheets) were therefore not only seen as a resource for supporting individuals who had taken part in the trial to continue to use social stories but also to act as a mechanism for encouraging others to be involved in the development and delivery of Social Stories.

Although individuals were provided with training, support and resources (e.g. the Social Stories manual, exemplar stories), there was some uncertainty within the group of trainers as to whether individuals were likely to be sufficiently prepared to write and deliver Social Stories outside of the trial. This uncertainty stemmed from the fact that each goal and story is unique to the needs of a given child, and some interventionists did not have a strong knowledge base in autism and lacked confidence during writing sessions. RAs and trainers received a slightly different model of training from participants, which involved Cascade, or a train the trainer model, involving a combination of formal training, that is, shadowing Social Stories training sessions and spending time with an experienced trainer to review all materials in detail and gain feedback on their performance. This experience of shadowing and delivering training was perceived to have given them an ability to describe, understand and apply how to set a goal and create a Social Story. Observing and shadowing the goal-setting process in particular was perceived to have exposed RAs/trainers to the practical aspects of goal-setting, such as using information about a child provided by an individual to tailor goals. Additionally, shadowing more experienced individuals was thought to have given them exposure to different stories and insight into how to challenge teachers and parents about their choice of goals and have confidence in discussing the underlying reasons behind a child's behaviour. Ultimately, they viewed the process as unpredictable due to the uniqueness of each goal and story. As a result, they placed significant importance on 'learning on the job' to accumulate the experience, but most importantly confidence needed to be able to translate skills and experiences of previous goals and stories. To mimic this experience in the future, participants highlighted the importance of including role plays, particularly those of challenging situations (e.g. narrowing down broad goals, dealing with conflicting priorities), to help trainers feel more confident.

I think getting that decision, you know, getting that end goal, the find, it's like a skill that you have to keep ... so the shadowing definitely helped and then just going in and actually doing the sessions helped because every teacher comes in with different things and different behaviours and when you have all these guidelines on how it has to be, like a smart guide it has to be specific, it has to be positively framed. Sometimes it's really easy and sometimes it's quite difficult. Yeah a lot of practice was definitely helpful.

Focus group with research assistants

Difficulties associated with delivering a school-based intervention

What were the challenges faced in delivering Social Stories?

A number of barriers to delivering the intervention within a school environment were identified. Even for those who had allocated dedicated timeslots for delivering the stories, ensuring that there was sufficient time within the school day and a TA available to deliver the intervention was challenging. Particular difficulties surrounded trying to balance the need to ensure the intervention was delivered 1 : 1 in a quiet, private environment against not removing the child from an important lesson and ensuring Social Stories were delivered at the right time for the child. Some teachers described having engaged with wider professional groups during the trial, such as autism teams, to highlight the benefits of Social Stories. School staff suggested that to deliver Social Stories outside of the trial, they would need to allocate one member of staff per child to facilitate the delivery, and that dedicated time would need to be given from the headteacher to school staff (TA, teacher) for them to write an effective and personalised story together for each child and for staff to attend training.

From my point of view as a class teacher, I'll probably have a bit more confidence in saying, well is there any chance we could have a bit of time in the day to do [the intervention] because actually if we give it a bit of time to it, it is more meaningful and effective and actually more likely to get you the outcome that you're looking to achieve. So I think sometimes that is the issue and maybe doing the social story trial has made me and the SENCO think that a little bit.

Teacher 5

Given the amount of additional support (e.g. through the research team) and time that had been made available to support the delivery of the intervention during the trial, interventionists raised concerns about the amount of resources required to support the continued and wider use of Carol Gray Social Stories outside the trial, particularly in a school setting. As such, participants were keen to avoid being reliant on a small number of school staff to create and deliver stories at each school, with this approach deemed particularly problematic in larger schools where it is likely that multiple children would need support despite limited resources. Additional and/or refresher training was proposed to ensure that the knowledge and skills participants had obtained during the trial were maintained and passed on to other school staff. However, a popular suggestion, for facilitating the wider use of Social Stories within schools, was through 'Social Stories Champions'. While there was some speculation as to who could take on these roles (e.g. SENCOs), having someone who is embedded within the school to co-ordinate the facilitation of Social Stories was considered important. This was largely because the amount of time required to support staff create stories was considered too great for an 'outsider' (e.g. local authority personnel). Participants also felt that champions would require specific, intensive training to ensure that after training, champions felt confident and were able to support staff to create and deliver Carol Gray Social Stories without requiring additional support. Participants also saw a key element of the champion role to be to encourage best practice and a supportive environment, underpinned by peer support. As such, signposting staff to resources and training and encouraging sharing of experiences, within and between schools, for instance, through the creation of Social Stories networks, was proposed.

... social stories champions within schools because I don't think it would be feasible that educational psychologists or local authority practitioners or whoever be dealing by case-by-case basis, especially like [name] says when you've got these big schools. I mean in that school it was eleven participants, that wasn't eleven children with autism, you know You've got large numbers of children with autism, so it wouldn't be practical for a practitioner to be coming in and writing social stories for all of them. So we need to move away from this really high level of handholding that we've had in this trial, so really intensely training, you know, one or two people within a school to really roll with this and not need that post support that we've given. So I think we really need to strengthen the training package for the practitioners, so that they can go forward and train the school staff.

Additionally, school staff described the challenges of delivering such a resource-intensive intervention amidst a high workload and staff shortages and likened doing so to a ‘juggling programme’. Pressure around staffing was particularly heightened during the COVID-19 pandemic and resulted in some staff describing how they had had to rush the delivery of the intervention, or in some cases, to stop or miss sessions entirely due to child and/or staffing absences or school closures.

I think mainly it is the workload as in there is a lot of with COVID there’s been a lot of staff absences so there’s been, yes this is your plan and it fits in, but then on a daily basis a lot of things change. Either the child is absent or there is a staff absence so you’re filling in a slot. Same with most things and you’ve just got to prioritise what can be dropped and what can stay. So I think it became a bit of a juggling programme.

Interventionist 5

Poor communication between staff surrounding the logistics of delivering the intervention, Wi-Fi issues in schools and low levels of computer literacy were also described as barriers to implementing and delivering the intervention. Low levels of computer literacy among TAs and an unfamiliarity with PowerPoint created unanticipated difficulties in personalising Social Stories, particularly in regard to adding pictures. The implications of this for future roll-outs on resourcing warrant consideration. These barriers were not unique to school staff, as parents also found time and capacity, particularly during the COVID-19 pandemic, and access to materials and equipment (printers, laminators) a barrier during the trial.

You need to be relatively IT savvy in order to be able to use them. I think particularly for some children. The one that I did was just word but I know the other two we did in communicate and print as well because that is what was needed for our children to access it so that requires a certain level of IT skills to be able to communicate and print and then be able to personalise it and put pictures in as well.

Interventionist 6

As part of the interventionist 6-month follow-up questionnaire, interventionists were asked to report whether they had encountered any problems/barriers to using Social Stories. Of the 73 respondents, 68.1% said ‘no’, and of the 29.2% that said ‘yes’ the barriers generally fell within the following categories: the time to write and deliver Social Stories, COVID-19 disruption and school closures, staff changes (e.g. changing classrooms or the children they are working with), school holidays disrupting the intervention period and the child refusing to engage (e.g. not interested after the first read, inability to focus on a story). When asked if they had stopped using Social Stories completely, 56.9% said ‘yes’, and the reasons were very similar to the barriers described above, for example, COVID-19 disruption, stories no longer needed as the child had moved classroom/school and time constraints within the school.

For the use of Carol Gray Social Stories to become more widespread, some RAs/trainers felt that some schools may need convincing as to why they should invest their time and resources into delivering the Carol Gray method, particularly when considering that many of their staff will already have some familiarity and experience of delivering more generic versions of social stories that do not meet the Carol Gray criteria. Some participants described that since the trial rather than using PowerPoint and the structured format of Carol Gray’s Social Stories, they had condensed stories to one page, asked children to create their own stories or used non-personalised stories with a group highlighting that not all participants have continued following the Carol Gray criteria. To avoid adoption of the Carol Gray methodology being ‘hit and miss’, trainers/RAs proposed clearly communicating the benefits of Carol Gray social stories over alternatives. Accredited training and/or certification, delivered through established networks and/or professional groups, for example (SENCO, educational psychologists, academies, educational mental health practitioners), was considered one way to achieve this, alongside limits on the time that training is valid for to ensure that quality of social stories remains high and as per the Carol Gray methodology.

The other thing I think that is really confusing and I found that as a teacher, and I find that in my role now, is that actually most educational professionals don't know that there's like this Carol Gray social stories and there's a lot of information on the internet on things like Twinkle. That isn't a proper social story as we would say now that we've been trained in it but it's still called a social story and it's almost like too common a word now. So if you say, oh a social story, everyone is like, yeah I know what that is, we've used one of them. People haven't and I know it's got like the TM on but no one notices that. It almost needs a bit of rebranding or be called 'a something social story', so people know that it is something that you've got to be specifically trained in and there's one way of doing it and you have to do it that way. You can just Google social stories and download them off the internet and not personalise them at all and most staff still think they're using a proper social story.

Research assistant focus group

Some individuals reported that the main benefit of the intervention was a greater understanding of autism and how to improve their discussions about social situations with their child/the child they were working with. For example, some participants described how they are applying what they have learnt, specifically around the use of simple, declarative sentence types and are using this to break down social information in a number of situations in daily life, not necessarily in the form of a Social Story.

I think the process and the construction of social story could be broken down in verbal ways, in more formal ways I suppose that I've tended to do, that's what I've taken from it in daily life ... sometimes with children you just sort of tell them what to do, you don't really say why you're doing it, what the benefits are, what might happen if you do this. You sort of skip all of that presuming they've got the knowledge and then just go in with what you need them to do. The message was to try and ask for help in class, have the confidence to do that and there was sort of so many levels of why that would be good and what would happen ... I think it just increased awareness around why that was a good thing to do and it helped me with other things really. Kind of communicating other messages to my son rather than just skipping to the purpose of the request and kind of actually taking the time to explain and maybe verbalise a social story sometimes when doing things.

Parent 6

The 6-month follow-up questionnaire asked interventionists to rate how useful they felt Social Stories have been ('0' not useful at all, '5' somewhat useful and '10' extremely useful). The most frequent ratings fell between 5 and 8, with 23.6% of interventionists responding '8'. No interventionists rated the usefulness of Social Stories between 0 and 3, reflecting the mostly positive views of the intervention held by the interventionists.

Programme differentiation

What are the considerations to make when implementing Social Stories in the future?

Lastly, we use our qualitative and quantitative findings to propose a series of areas to consider when informing how Social Stories could be used and implemented outside of the ASSSIST-2 trial. [Table 44](#) outlines areas for consideration when using and implementing Social Stories in the future and is based on the qualitative study results and informed by our public involvement representatives, who met to review the results on 27 April 2022.

Summary

Core to the underpinning philosophy of the Carol Gray method is the importance of personalising stories to the needs of each child. While participants were overwhelmingly positive about the intervention, the degree of benefit observed with each child was variable, demonstrating the uniqueness of each child and importance of considering and designing Social Stories to reflect this. As such, the series of success

TABLE 44 Areas for consideration when using and implementing social stories in the future

Area for consideration	Suggestion(s)
Social Stories training	<p><i>Training the trainers</i> A cascade training model for training the Social Stories trainers appears to be effective. Social Stories trainers should have a background in psychology/similar subject and/or have an in-depth understanding of autism.</p> <p><i>Training model</i> A hybrid approach where attendees can choose between online or in-person training to suit the needs and preferences of the school/parents may be beneficial. The interactive elements of the training are important and beneficial, e.g. during the psychoeducation section about autism. Provision of resources/reference materials was helpful for educational professionals and parents (e.g. the Social Stories manual, worksheets and example Social Stories).</p> <p><i>Post-training support</i> It is important to consider what level of follow-up support can be offered following training for parents/school staff and who is best placed to deliver this.</p>
Who should be involving in setting the goals and creating and delivering Social Stories in a school setting?	<p>Individuals who have knowledge of the individual child and have insight into the potential underlying reasons for the child's difficulties or behaviours. It is important that individuals writing and delivering Social Stories have received formal training and have access to resources (e.g. a Social Stories manual) to refer to. For goal-setting and writing the Social Stories, a collaborative approach between educational professionals and parents/carers appears to be beneficial, as different perspectives about the child and the challenge they are facing can be taken into account.</p>
Frequency and duration	The frequency that a Social Story is delivered should be tailored to the individual child and should take into account their preferences and like/dislike for repetition.
Setting	The individual preferences of the child should be taken into account when choosing a setting. Quiet, private, informal and comfortable settings to facilitate 1 : 1 conversations appear to be most suitable for the majority of children on the autism spectrum.
When	Scheduling a dedicated time within the school day/week to read the Social Story may help to avoid disruption and can become part of the child's timetable. However, it is important to be flexible where possible and respond to the preferences of the child so that the story sessions are always a positive experience.
Structure, format and content of Social Stories	As per the Carol Gray Social Stories criteria, stories should be highly personalised and tailored to the individual child. Structured format of the stories was generally considered to facilitate the creation and delivery of Social Stories.
Potential barriers to delivery within a school setting	Limited time to write and deliver Social Stories within a school setting. Dedicated time is required for staff to set the goal, write the story and deliver the sessions. Staffing issues may affect the writing and delivery of the stories, and each child may not be able to be supported by someone who knows them well. IT literacy and access to equipment can greatly vary amongst school staff and should be considered when arranging training with schools. Consider having dedicated staff and staff time allocated to support the creation and delivery of Social Stories (e.g. Social Stories Champions) outside of the trial, where there will be no further support available from the research team.

factors that we propose above should not be considered as a prescriptive list from which individuals cannot deviate, but rather areas of consideration to inform future implementations. Indeed, for each area proposed, there will be situations, children and contexts for which these suggestions may not apply or would need to be adapted. There is unlikely to be a one-size-fits-all approach to Social Stories in terms of how they 'should' be delivered or indeed how children on the autism spectrum respond to them.

Chapter 6 Discussion

Summary of key findings

Trial population

This trial aimed to evaluate whether Social Stories is a clinically effective and cost-effective method of improving social responsiveness in children on the autistic spectrum. A total of 87 schools were randomised, representing 249 autistic children. This was lower than the initial target sample size of 278 pupils, with the target sample size being revised following the impact of the COVID-19 pandemic. Participants were recruited via multiple pathways, although direct recruitment of schools was the most successful. Despite all children being under the care of the NHS, this pathway yielded only three pupils from three schools, as most schools were already in discussion with the trial team through the direct school recruitment pathway by the time NHS referrals were received.

The average age of the participating children was 8.6 years, and three-quarters of children were male, as is typical for the population of individuals with ASC.⁶⁷ Most children were described by their parents as having a white ethnic background (88.3%). People delivering the intervention were mostly female (93.5%) and TAs (83.7%), with an average age of 42.6 years. Most considered themselves to have moderate or greater levels of knowledge and experience of working with autistic children (83.8%). Exhibiting behavioural challenges in school was criteria of participation in the trial. The challenges faced by autistic children that were supported through the Social Stories broadly fell into five categories: social interactions; emotional development; routines and change; classroom behaviours; and more general life skills.

Primary outcome

At baseline, children scored 72 points on average on the SRS-2, indicating a moderate level of difficulty. The average SRS-2 scores decreased in both groups at 6 weeks and 6 months. At 6 weeks, children allocated to Social Stories scored on average 1.14 points lower than those who received usual care (95% CI -3.35 to 1.06; $p = 0.310$). This difference increased slightly at 6 months to 1.61 points (95% CI -4.18 to 0.96; $p = 0.220$). However, at both time points, the observed difference was much lower than the 3 points postulated as clinically important, and neither difference was considered statistically significant.

After accounting for compliance with the intervention, defined as children receiving the intended six or more story sessions, a CACE analysis found that Social Stories led to a -3.37-point reduction in social responsiveness at 6 months for participants in the complier principal stratum (95% CI -6.65 to -0.10). However, some caution is required in interpreting these estimates because a substantial number of intervention group participants were excluded from the estimation due to having missing compliance data. Two post hoc sensitivity analyses to assess the possible influence of the missing compliance data suggested that the magnitudes of the CACE estimates were partly explained by selection bias arising from exclusion of intervention participants with missing compliance data. However, the results suggest the data were most compatible with hypotheses positing moderate effects of the intervention when Social Stories is received as per protocol.

Given that much of the compliance issues were thought to be due to the impact of COVID-19, rather than feasibility or systemic issues delivering the intervention, two sensitivity analyses were undertaken to specifically explore the magnitude of this impact. In the first, we investigated the extent to which disruption of follow-up due to the pandemic was associated with variation in treatment effects. The estimates generated through the statistical model provided some evidence that disruption to follow-up due to school closures was associated with variation in treatment effects at both 6 weeks and 6 months. This finding was somewhat replicated in the second analysis which explored the extent to which disruption of follow-up and the shifting of school training to an online model were associated with

variations in treatment effects. Again, estimates generated from this model provided some evidence that the impact of Social Stories was greater at the 6-month time point amongst those who were randomised and followed up prior to the beginning of lockdown.

Taken together, the CACE and COVID-19 sensitivity analyses do indicate that poor compliance may have reduced the effectiveness of the intervention. This is further supported through the process evaluation, where teachers and interventionists frequently reported COVID-19, staff shortages and time issues to be key barriers for successful intervention implementation. Of the interventionists that returned 6-month questionnaires, a large number (29.2%) reported encountering problems during story sessions, with many reporting difficulties keeping the children engaged and issues with the repetitive nature of the intervention.

Secondary outcomes

There were no differences in any of the measures assessing children's mental health and well-being or parental stress. However, in terms of the specific goals against which behavioural outcomes were measured, children who received Social Stories typically met their behavioural goal more frequently after 6 months than children who received care as usual ($p = 0.012$).

Cost-effectiveness

Compared with usual care, Social Stories slightly decreased the mean service use costs while maintaining similar QALY improvement. This was shown in both the primary and sensitivity analyses but subject to the given data with the uncertainty related to short study time frame, amount of missing data and the disruption of COVID-19. The results show a small reduction in overall costs across different perspectives, particularly through the cost reduction from the education sector, such as the costs of school-based health services and the costs of general support from teachers and/or TAs at school. Although school-based intervention costs were found to be higher in the Social Stories arm, it appeared to be more related to a small number of children receiving high levels of high-tariff 1 : 1 mentoring sessions in the Social Stories arm than a general cost increase across the whole group (see [Chapter 4](#)). This is a relatively positive finding in that Social Stories seems to be able to reduce the care needed in schools without affecting children on the autistic spectrum receiving school-based interventions. Overall, within a 6-month time period, compared with receiving usual care, there is around £200 saving per child (from the societal perspective) for receiving Social Stories. The cost saving is relatively small, but this cost saving may be relevant for commissioners and education sectors when considering resource utilisation if multiplied by the number of children on the autism spectrum in the UK. Owing to the fact that there are currently 166,041 autistic CYPs in schools in England and more than 70% of autistic CYPs are educated in mainstream schools,⁶⁸ the potential cost savings to the NHS and education system could reach around £28 million for the CYP population as a whole. It is worth noting that although the cost reduction, especially in the education sector, is a rather positive finding, it needs to be interpreted with caution. This is because the cost reductions were relatively small no matter which perspective was taken, and the cost differences were not statistically significant, showing the uncertainty of the cost results.

On the other hand, the difference in QALY improvements between the two trial arms was very small (see [Table 41](#)). After taking uncertainty into consideration and adjusting for the imbalanced utility, cost and other characteristics at baseline, the difference in QALYs remains small (< 0.001 QALYs), suggesting that the Social Stories do not appear to improve the QoL of children on the autism spectrum based on the dimensions measured by the EQ-5D-Y.

Process evaluation

The intervention was facilitated by way of a cascade or 'train-the-trainers' training model whereby a child psychiatrist (Professor Barry Wright) trained one member of the trial team, who in turn trained other members of the trial team. These members of the team then became the 'trainers' responsible for training the educational professionals within schools. Training and support appeared to be well received

at each level. Trainers felt sufficiently prepared to deliver training to educational professionals, who, in turn, felt well-prepared to deliver the intervention to pupils. Schools generally felt supported by trainers. Owing to the COVID-19 pandemic, the training model had to be adapted part-way through the trial to move the training online. There were mixed feelings around this, with some educational professionals and parents valuing the shorter sessions and improvements in access; for example, some parents were able to join videoconferences while at work or while caring for other children at home. Others valued the face-to-face interaction.

The process evaluation provided data surrounding the implementation of the Social Stories intervention. The majority of educational professionals appeared to grasp fully the core principles of Social Stories and, in most cases, were able to deliver them with fidelity. Around 30% of interventionists reported that they encountered some barriers in delivering Social Stories. These barriers were largely reported during the latter half of the trial, with professionals citing the COVID-19 pandemic and associated difficulties families, schools and the research team encountered. Other logistical issues, such as staff changes and school holidays, were also cited.

Comparison with previous studies

The results of the process evaluation broadly align with the findings of the ASSSIST feasibility study which demonstrated a high degree of acceptability with young people, families and schools.¹³ The study was similar in design to the full-scale RCT;¹ however, there were some differences to the trial procedures and intervention delivery. A key difference is the training model. In the feasibility study, a Social Stories in-person training day was provided for teachers and parents so that the Social Stories could be created for the participating children. Although 26 schools in the current trial were trained via this model, due to the COVID-19 pandemic, 18 schools randomised after 23 March 2020 (because of restrictions imposed by the COVID-19 pandemic) were trained via a combination of an online training presentation and a virtual story-writing session. This meant that important components of the face-to-face training procedure were lost, and it is unclear how much impact this had on educational professionals' and parents' understanding of the intervention.

The timing and intensity of the intervention delivery were also modified for the present trial, lengthening the intervention delivery period from 2 to 4 weeks and the number of reading sessions from 3 to 6. Although potentially more practicable for schools to deliver the intervention in this way from a logistical perspective, extending the duration of the intervention over a longer period does increase the risk of disruption which may have led to increased problems in the present trial. A reduced dose of the intervention may also have impacted its potency.

In the feasibility study, outcome measures were collected over a shorter period, with the final follow-up data being collected at 16 weeks as opposed to 6 months in the present trial. In the feasibility study, the SRS-2 and goal-based measures showed change scores in the desired direction¹³ which aligns with the findings of the present trial.

Comparison with current evidence base

Several systematic reviews have been conducted to synthesise the existing evidence base around Social Stories. The first review published in this area concluded that the effects of Social Stories in general (i.e. not specifically Carol Gray's Social Stories) are highly variable. Twelve of the 16 studies examined in the review used a single-subject design, and the authors highlighted that it is difficult to interpret the results due to inadequate participant descriptions and because Social Stories were often used in combination with other interventions.²³ Another systematic review concluded that Social Stories in general (i.e. not specifically Carol Gray Social Stories) may be substantially more effective when used to target behaviour reduction than to teach appropriate social skills and appeared more effective when the story focused on a single behaviour rather than more complex social situations. The authors suggested that this may

be because social behaviours are more abstract and complex, and thus more difficult for children on the autism spectrum to understand.²²

In the most recent systematic review conducted by Wright *et al.*,²⁴ any study of Social Stories that used a standardised, numerical measure of outcomes or a non-standardised, numerical outcome measure was included. Single-case designs reporting repeated measurement of the target behaviour were included alongside seven between-group studies, four of which were RCTs.²⁵⁻²⁸ Outcomes included social abilities and awareness, communication, restricted behaviours, life skills, emotional development and sustained attention. This review highlighted that although randomised studies appear to provide broadly supportive evidence of an effect of the Social Stories interventions, the findings from the studies are compromised as there were gaps in the reporting of results, which made it difficult to quantify the level of effect of the intervention and to assess the risk of bias in the studies. The Social Stories interventions examined in the RCTs included in the review were not individually tailored to the child or young person and therefore did not meet the Carol Gray criteria and were delivered under experimental conditions. It is therefore difficult to compare the current findings with those of the previous RCTs, with the exception of our feasibility RCT,²⁵⁻²⁸ which demonstrated a high degree of acceptability and some positive findings that were not subjected to statistical analysis because of the feasibility nature of the study.

Taken together, these reviews demonstrate that the effect of Social Stories is highly variable. These studies are likely to have included stories of varying quality, as none specifically evaluated Carol Gray's Social Stories. This study examined Social Stories that met the Carol Gray criteria and were delivered within a school setting that was familiar to the child, with the intervention delivered six times over 4 weeks, whereas in the previous RCTs, they were delivered over a much shorter time frame (e.g. exposure to the story five times over the course of 1 day).²⁸ Another important factor to consider is that the intervention is complex. The Social Story is tailored to the child and may target a range of different goals that are difficult to combine in measuring outcomes. Our goal-based outcomes showed more promising results than our validated measure of social responsiveness. Not all of our goals were directed at social responsiveness, as many teachers or parents may not have chosen this as important for the child at this time. Goal-based outcome measures are more flexible in this regard but are less respected in some areas of the scientific community.^{69,70} The intervention also involves other potential sources of benefit for the child, including the fact that the intervention usually requires the teacher and parent to converse about the child's needs and to involve the child where possible, all of which may carry benefits that are difficult to disentangle from the reading of the story itself. The quality of these latter processes was greatly impacted by the COVID-19 pandemic. This study was a pragmatic trial in the real world involving educational professionals, parents/carers and children from a diverse set of mainstream and SEN schools across Yorkshire and the Humber and therefore the results may be more representative of the effectiveness of the Social Stories intervention for supporting children on the autism spectrum compared to previous results. Overall, this full-scale RCT has added to the current literature which includes mainly single-case studies and a small number of RCTs (with methodological limitations) conducted primarily in the USA.

Strengths and limitations

To our knowledge, this is the first pragmatic trial evaluating the effectiveness and cost effectiveness of Carol Gray's Social Stories for improving social responsiveness in children on the autistic spectrum. Response rates to the primary outcome were high. In total, 211 (84.7%) teachers provided primary outcome data at the 6-month follow-up point, with similar proportions in each group. Equally, despite being a typically harder-to-reach group, we achieved a response rate of 184 (73.9%) parents, again with similar proportions in each group. The trial was reported in line with CONSORT, CHEERS and other relevant guidelines. An independent TSC as well as a DMEC provided oversight to ensure that the trial was conducted as planned and that participant safety issues were considered.

A key strength of the methodology of the trial was in our approach to minimising the risk of bias. All data were collected by blinded RAs to reduce the risk of bias in the data-gathering process. There were no cases where data were collected by someone who was not blind to allocation. The process evaluation was overseen by an independent researcher who joined the trial part way through with no prior knowledge and thus no pre-conceptions of Social Stories. Furthermore, the trial statistician remained blind to allocation until the final analysis.

The multiperspective approach of the economic evaluation was a particular strength. The evaluation accounted for the costs from a range of perspectives (NHS/PSS, NHS/PSS/education and societal perspectives) making the evaluation results useful to health policy-makers, healthcare providers, education sectors and the parents/guardians of children with ASC. We also explored the impact of missing data by analysing the complete case. This approach not only ensures the robustness of our findings but can also help policy-makers from different sectors to make informed decisions. There were, however, some limitations in the economic evaluation. Firstly, and most importantly, cases that could not receive Social Stories or where Social Stories was interrupted as a result of the COVID-19 pandemic were a concern. This could potentially introduce bias to our results and underestimate the potential cost savings and QALY improvement. However, since the disruption occurred to both arms, it is expected that the impact on our results is limited. Secondly, there were a good number of missing data. Overall, complete EQ-5D-Y and resource use data were fully available for around 45.0% of children. This number of missing data may introduce bias and limit the accuracy of the conclusions. However, our results are unlikely to be affected, as the results of the complete case were consistent with the base case (imputed cases). Thirdly, a few high-cost cases were observed, and they may affect the interpretation of certain cost comparison outcomes (see [Cost-effectiveness](#)). However, these high-cost cases are unlikely to affect the direction of the economic results about Social Stories. This is because these high-cost cases are in the Social Stories arm, and the cost savings of Social Stories would have been larger if those high-cost cases were removed. Finally, this economic evaluation measured the short-term (6-month) cost effectiveness of Social Stories with the long-term cost-effectiveness of Social Stories remaining unknown. Although it is outside the scope of the current study, a model-based economic evaluation would be desirable in future research to allow lifetime cost effectiveness and children's lost productivity during adulthood to be measured.

There are several limitations that should be considered in the interpretation of our findings. First and foremost, a substantial portion of the study occurred while a worldwide pandemic was occurring. This led to large disruptions to schools with closures, staff and pupil absences and changes to the intervention construction and delivery, as well as care as usual. The results of this trial should therefore be taken with caution in both potential directions. Owing to COVID-19, we were not able to quality-assure the teacher-reported outcome through independent observation. Given that the intervention, due to its nature, was delivered unblinded, there is some potential to introduce reporting bias into teacher ratings in favour of those receiving the intervention. By having the interventionist as a separate member of staff from the teacher completing the main outcome measure, we attempted to minimise this, but some potential for bias remains.

Because of the restrictions imposed as a consequence of the pandemic, the training model was updated to rely on some level of independent learning by way of following an online presentation. Although trainers asked schools whether they watched the presentation at the start of the story-writing session, we are unable to accurately assess whether interventionists did watch the presentation and whether they effectively processed this information. This was much clearer during the face-to-face training, where sessions were interactive and activities were used to check understanding. Given the importance of the psychoeducation training, failing to access and engage with this independent learning could have had an impact on the effectiveness of the intervention.

As we were initially planning to follow the same training model as the feasibility trial, no related process evaluation activities were built into the design and funding plan, hence we were unable to explore the

cascade model of training through observations. However, we did gain some insights into this from the trainer focus group and through interventionist interviews.

To strengthen the process evaluation in light of all the changes to the trial design associated with COVID-19, we developed a series of surveys to complement the qualitative interviews and reach a larger number of participants. As these surveys did not form part of the core data collection process, they were distributed by e-mail rather than on paper, and there was limited capacity within the research team to follow up non-response. The response rates for these surveys were thus much lower than anticipated, yielding only 32 responses from a mixture of educational professionals and parents. Nevertheless, we were able to integrate the data into the process evaluation, and all themes overlapped with data collected through other methods, such as interviews or questionnaires, thus further strengthening the security of the findings.

The analyses of outcome measures showed some impact in terms of social skills and specific behavioural outcomes. There is some possibility that these benefits may not have been due to the Social Stories intervention itself, but rather an increase in one-to-one time spent with a TA. This could be further explored in the future.

The results are somewhat limited in terms of generalisability as the majority of the schools included in our sample were mainstream primary schools, unlikely to serve children with more profound autistic traits, hence more research around the suitability of the intervention for these children may be useful.

Finally, our preference would have been to use the goal-based outcome measure as our primary outcome measure; however, a validated instrument was favoured by the funder. The challenge with a single validated outcome measure is that Social Stories are a complex intervention that can target a number of goals related to an autistic child, and no single measure is likely to capture outcomes well. Furthermore, using a generic child QoL measure has advantages related to comparability but may not be a well-suited measure to a population of autistic children. Improved ways of considering outcome measures in pragmatic trials of complex interventions are needed.

Equality, diversity and inclusion

All primary schools within Yorkshire and the Humber region were invited to participate in the ASSIST-2 trial, and there were very few exclusion criteria applied to the children who could participate. Although there may be multiple levels of volunteer bias in that both the schools and families who agreed to participate may be different from those who declined, we ultimately included a diverse range of schools across the Yorkshire and the Humber region, including both rural and urban schools, mainstream and specialist settings and single and multiform intake schools. The results are thus largely generalisable to other primary-aged children on the autism spectrum.

We would have liked to have recruited more families from diverse backgrounds and we made considerable efforts to do this; however, only 10.9% of families identified as non-white. It was raised during a PPI event that Asian families are particularly hard to reach, often due to cultural beliefs surrounding fear of stigma, which may explain some of the challenges recruiting from this population. Given the flexibility of the Social Stories intervention, the results are still likely to be generalisable to other ethnic groups as the stories could be tailored accordingly.

The training of educational professionals was conducted by the research team. This was not initially planned. In the original trial design, it was planned that local authority educational psychologists would conduct the training as per the feasibility trial; however, due to the strict time constraints applied with regards to randomisation, this was not practicable. The members of the research team responsible for training were able to afford greater time to follow-up support after training sessions took place and

were able to support and guide schools through revisions to the stories and more technical issues, such as adding pictures and other content to stories. Given the limited resources within the Local Authority service provision, educational psychologists are unlikely to be able to mirror such high levels of support. However, we feel that the need for ongoing support could be lessened by strengthening the initial training model.

Implications for practice

Social Stories are already frequently used as a means of supporting autistic children within primary schools and are recommended by the National Autistic Society. Within the UK, there are several Carol Gray-certified trainers, known as satellite trainers, who offer training packages to schools. The trial did not show meaningful effects on the SRS-2, the primary outcome, and thus we cannot recommend Social Stories as a means of improving social skills in autistic children. However, positive benefits were seen in tailored individual goals that were broader than social skills. Furthermore, some cost savings of approximately £200 per child were evident, which, given the large population of autistic children in schools, would be a saving that may be attractive to local authorities if multiplied up. No negative impacts were found, and benefits were also reported in terms of facilitating dialogue between child and interventionist. Given these findings and the fact that both parents and educational professionals look favourably upon Social Stories and that delivery incurs only minimal costs, this may provide sufficient justification for schools to continue to use the intervention.

If schools were to choose to implement Social Stories, there are several lessons learnt during the course of the trial that should be applied. In the trial, the training was conducted by members of the research team. Under usual circumstances, we anticipate that school staff would be trained by a member of the Local Authority such as an autism specialist teacher or educational psychologist. This is the approach utilised during the feasibility study and was feasible using cascade training. However, conditions have changed with reduced resources in local authorities leading to cuts in services (e.g. less educational psychologists available for intervention and training work). This suggests that the offer of a more flexible approach to training, online or in person or potentially a combination, is desirable and leads to necessary flexibility (e.g. online training plus in-person Q&A and story writing in person). We could also work towards offering more guidance to schools around the child characteristics best suited to Social Stories (e.g. those not adverse to repetition). Goal-setting and training were undertaken by separate people in the RCT, whereas in the real-world context, both would be undertaken by the same person, and we would need to ensure they are adequately trained to deliver/facilitate both components. Carol Gray has a team of approved trainers who could potentially be utilised to facilitate wider uptake.

Recommendations for research

Given the uncertainty of our findings in light of the COVID-19 pandemic, which substantially impacted the delivery of the trial post March 2020, we cannot definitively rule out the possibility that Social Stories are beneficial for developing social skills in autistic children, especially when delivered at an adequate dose and with maximum fidelity. Because of this, there may be value in further research around Social Stories, particularly with regards to the impact on specific goals or subgoals. Indeed, a systematic review exploring the impact of Social Stories concluded that they are most beneficial when used to target specific behaviours rather than teach general social skills, although this moves away from the intention of Carol Gray's original design and purpose.⁷¹ Improved research designs should attempt to capture wider outcomes, such as the importance of teacher/parent interaction in understanding and planning support for the child.

Future research could explore the value of psychoeducation training for educational professionals and parents. The school-level training model included an aspect of psychoeducation, and this was

highlighted as particularly valuable. During the PPI/stakeholder event, it was suggested that this alone would likely be a beneficial intervention to support parents whose child/children have been recently diagnosed with autism or educational professionals with limited experience by raising their awareness of autistic traits and empowering them with information and tools about how to best support children, for example, using or avoiding specific types of language.

Our sample comprised only 10.9% non-white families, and it was highlighted during the PPI/stakeholder event by educational psychologists and charity representatives that accessing ethnic minorities, particularly Asian communities, can be exceptionally challenging, with children often missing out on support. Future research could explore ways to establish trust with child mental health services within communities and provide psychoeducation around autism with the view to increase the uptake of healthcare services.

Finally, one of the key outcomes of the process evaluation was that Social Stories may serve as a useful tool to facilitate conversations with children around difficult topics. However, it is unclear how this approach compares to other approaches already used in schools. This could be explored further.

Conclusion

We found no impact of Social Stories on autistic children's overall social responsiveness skills. There was some evidence that Social Stories are well supported by parents and teachers, may be effective at addressing a wider set of individual goals and could reduce costs. Based on the evidence generated through this trial, we cannot recommend Social Stories for the purposes of improving social skills, anxiety and/or depression, parental stress and general health in autistic children. However, we did not find any negative effects, and Social Stories are already frequently used in schools to support autistic children and represent a low-cost and potentially cost-saving intervention. Despite limited evident impact on global social skills, Social Stories may serve as a useful tool for facilitating dialogue between children and school staff to better understand the needs of autistic children, and usage should be at the school's discretion. Findings within this report should be treated with caution given large effects on schools during the COVID-19 pandemic when this trial took place.

Additional information

Contributions of authors

Barry Wright (<https://orcid.org/0000-0002-8692-6001>) was a chief investigator.

Kerry Jane Bell (<https://orcid.org/0000-0001-5124-138X>) was responsible for trial management and co-drafted the report.

Jane E Blackwell (<https://orcid.org/0000-0002-5878-8959>) was responsible for trial management and co-drafted the report.

Catarina Teige (<https://orcid.org/0000-0002-8958-6487>) was responsible for trial management.

Laura Mandefield (<https://orcid.org/0000-0002-4219-5673>) was a trial statistician responsible for conducting the statistical analysis.

Han I Wang (<https://orcid.org/0000-0002-3521-993X>) was a health economist responsible for the economic evaluation.

Charlie Welch (<https://orcid.org/0000-0002-2421-5538>) was a trial statistician responsible for conducting the statistical analysis.

Arabella Scantlebury (<https://orcid.org/0000-0003-3518-2740>) was the qualitative lead.

Jude Watson (<https://orcid.org/0000-0003-0694-3854>) was a co-applicant and contributed to the trial design and trial management group.

Dean McMillan (<https://orcid.org/0000-0002-2901-8410>) was a co-applicant and contributed to the trial design and trial management group.

Emma Standley (<https://orcid.org/0000-0001-8702-4379>) was a research assistant responsible for data collection.

Leah Attwell (<https://orcid.org/0000-0002-8308-6465>) was a research assistant responsible for data collection.

Hayley Carrick (<https://orcid.org/0000-0001-6700-0783>) was a research assistant responsible for data collection.

Amelia Taylor (<https://orcid.org/0000-0002-7314-2654>) was a research assistant responsible for data collection.

Olivia Taylor (<https://orcid.org/0000-0001-7043-1795>) was a research assistant responsible for data collection.

Rachel Hodkinson (<https://orcid.org/0000-0002-3972-2890>) was a research assistant responsible for data collection.

Hannah Edwards (<https://orcid.org/0000-0002-0500-2123>) was a research assistant responsible for data collection.

Hannah Pearson (<https://orcid.org/0000-0003-0698-9457>) was a research assistant responsible for data collection.

Steve Parrott (<https://orcid.org/0000-0002-0165-1150>) was a health economist responsible for the economic evaluation.

David Marshall (<https://orcid.org/0000-0001-5969-9539>) was a co-applicant and contributed to the trial design and trial management group.

Danielle Varley (<https://orcid.org/0000-0001-7208-8716>) was a co-applicant and contributed to the trial design and trial management group.

Rebecca Hargate (<https://orcid.org/0000-0001-6532-2375>) was the COMIC programme manager and contributed to the trial design.

Ann McLaren was a PPI representative and contributed to the trial design.

Catherine Elizabeth Hewitt (<https://orcid.org/0000-0002-0415-3536>) was a chief investigator.

All authors within the writing group reviewed and agreed the contents of the report.

Acknowledgements

The authors wish to thank all the participating schools and families for taking part in the trial.

We would specifically like to thank Shirley Paul, Natasha Mitchell, Matthew Bailey, Tanya Pawson, Sarah Gardner, James Craggs, Julie Ross, Sharon Bird and Natalie Day for their contribution to data management and assistance with the day-to-day running of the study. We would also like to thank Saleema Rex and Professor Victoria Allgar for their respective input into the design of the statistical analyses and Professor Simon Gilbody for his clinical expertise.

We would also like to thank the following independent members of the Trial Steering Committee: Michael Morton (Chair), James Shearer (Health Economist), Cliona McDowell (Statistician), Holly Judge (Public Involvement Representative), Nikki Totton (Medical Statistician), Sarah Cooper (Sponsor Representative), Alison Thompson (Sponsorship Representative) and Sinead Audsley (Sponsor Representative) and the Data Monitoring and Ethics Committee: David Sims (Chair), Lee Shepstone (Medical Statistician) and Linda Davies (Health Economist) for their support and guidance throughout the trial.

Our thanks are also given to members of our public involvement panel who attended our results meeting on 27 April 2022, for their invaluable feedback and guidance. We would also like to thank Carolyn Roberts, Ann McKelvey, Jo Whitehead and Louise Meehan for their helpful contributions to our patient and public involvement work throughout the study.

We would like to thank the following NHS Trusts for acting as Participant Identification Centres (PIC sites): Bradford District Care NHS Foundation Trust, Harrogate and District NHS Foundation Trust, Humber Teaching NHS Foundation Trust, Sheffield Children's NHS Foundation Trust and Tees, Esk and Wear Valleys NHS Foundation Trust.

Finally, we would like to thank the NIHR HTA programme for funding the trial.

Initially, the trial opened as a non-NHS trial, recruiting schools and families directly with no NHS involvement. Ethics approval for the trial was obtained from the Health Sciences Research Governance Committee embedded within the Department of Health Sciences at the University of York on 2 July 2018. At a later stage of the research, ethical approval was sought and awarded from the Health Research Authority (HRA) on 24 July 2019 and the North East – York Research Ethics Committee (REC) on 23 July 2019 (REC reference number 19/NE/0237) to allow opening of an additional recruitment stream via the NHS. Substantial and non-substantial amendments to approve changes to the protocol and study documentation were submitted to the REC, HRA, the Department of Health Sciences Research Governance Committee and each NHS site's research and development office as required during this trial. The trial sponsor was Leeds and York Partnership NHS Foundation Trust.

Patient data statement

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it is important that there are safeguards to make sure that they are stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

Data-sharing statement

Requests to access the ASSSIST-2 data should be made to the corresponding author and will be considered on a case-by-case basis by the Chief Investigator. All data requests will be managed in accordance with York Trials Unit, University of York, processes and procedures.

Ethics statement

Ethics approval for the trial was obtained from the Health Sciences Research Governance Committee embedded within the Department of Health Sciences at the University of York on 2 July 2018. At a later stage of the research, ethical approval was sought and awarded from the Health Research Authority (HRA) on 24 July 2019 and the North East – York Research Ethics Committee (REC) on 23 July 2019 (REC reference number 19/NE/0237) to allow opening of an additional recruitment stream via the NHS.

Information Governance statement

Under the Data Protection legislation, LYPFT and the University of York are joint Data Controllers, and you can find out more about how we handle personal data, including how to exercise your individual rights and the contact details for our Data Protection Officer here (<https://www.leedsandyorkpft.nhs.uk/advice-support/how-we-use-information-about-you/> and <https://www.york.ac.uk/records-management/dp/your-info/generalprivacynotice/>).

Disclosure of interests

Full disclosure of interests: Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/JBTM8017>.

Primary conflicts of interest: Catherine Hewitt is a member of: NIHR clinical trials unit Standing Advisory Committee, HTA Post-funding Committee teleconference (POC members to attend), HTA Funding Committee Policy Group (formerly CSG), HTA Commissioning Committee, HTA – Fast Track-Funding Committee, HTA Fast Track Committee – June 2021. Barry Wright is a Trustee of Biomation CEC and co-author of *How to Write Social Stories* (Jessica Kingsley Publishers), the reference manual distributed to schools as part of the intervention. All royalties are donated to Martin House Children's Hospice. Charlie Welch works on other NIHR-funded trials: NIHR HTA 16/111/91, NIHR HTA 15/102/04, NIHR HTA 15/130/84, NIHR HTA 16/73/03, NIHR PHR 131745, NIHR PGfAR RP-PG-0109-10020, NIHR RfPB PB-PG-0416-20035, NIHR HTA NIHR131784, NIHR PGfAR RP-PG-0615-20003, NIHR HTA NIHR128625. Charlie Welch sits on oversight committees for the following NIHR-funded trials: NIHR HTA 17/23/02, NIHR HTA 17/18/02, NIHR RfPB PG-0817-20045. Jude Watson has received funding for the following: NIHR EME – NIHR132622, NIHR HTA Programme, NIHR131483, NIHR RfPB – NIHR201559, NIHR RfPB – NIHR201176, NIHR HTA Programme: NIHR128625, NIHR TCC (Clinical Lectureship) – ICA-CL-2018-04-ST2-015, NIHR Programme Grant – RP-PG-0618-20001, NIHR Programme Grant: NIHR200607, NIHR Programme Grant – RP-PG-0216-20002, NIHR PGfAR Programme-RP-PG-0615-20003, NIHR PGfAR Programme – RP-PG-1016-20003, NIHR Post-Doctoral Fellowship – NIHR301807, Cancer Research UK – C11232/A23434, NIHR PGfAR Programme – RP-PG-0109-10020, NIHR RfPB – PB-PG-0816-20029, NIHR RfPB – PB-PG-0317-20047. Steve Parrott declares work on the following projects: Yorkshire Cancer Research ESCAPE: Electronic cigarettes as smoking cessation aid for smokers with Mental Illness – Pilot + full RCT; Medical Research Council GCRF: Tobacco control capacity programme and Medical Research Council; Children Learning About Second-hand Smoke: Cluster randomised-controlled trial.

References

1. Wright B, Teige C, Watson J, Hodkinson R, Marshall D, Varley D, *et al.* Autism Spectrum Social Stories In Schools Trial 2 (ASSIST2): study protocol for a randomised controlled trial analysing clinical and cost-effectiveness of Social Stories™ in primary schools. *BMC Psychol* 2020;**8**(1):1–10.
2. Roman-Urrestarazu A, van Kessel R, Allison C, Matthews FE, Brayne C, Baron-Cohen S. Association of race/ethnicity and social disadvantage with autism prevalence in 7 million school children in England. *JAMA Pediatr* 2021;**175**(6):e210054.
3. Lai M, Lombardo M, Baron-Cohen S. Autism. *Lancet* 2014;**383**(9920):896–910.
4. Baron-Cohen S. Editorial perspective: neurodiversity a revolutionary concept for autism and psychiatry. *J Child Psychol Psychiatry* 2017;**58**:744–7.
5. National Autism Society. *How to Talk and Write about Autism*. 2022. URL: www.autism.org.uk/what-we-do/help-and-support/how-to-talk-about-autism (accessed 13 June 2022).
6. Kim JA, Szatmari P, Bryson SE, Streiner DL, Wilson FJ. The prevalence of anxiety and mood problems among children with autism and Asperger syndrome. *Autism* 2000;**4**(2):117–32.
7. Koegel LK, Koegel RL, Hurley C, Frea WD. Improving social skills and disruptive behavior in children with autism through self-management. *J Appl Behav Anal* 1992;**25**(2):341–53.
8. Vivanti G, Kasari C, Green J, Mandell D, Maye M, Hudry K. Implementing and evaluating early intervention for children with autism: where are the gaps and what should we do? *Autism Res* 2018;**11**(1):16–23.
9. Williams C, Wright B. *A Guide to Writing Social Stories™: Step-by-step Guidelines for Parents and Professionals*. London: Jessica Kingsley Publishers; 2016.
10. Simpson RL, de Boer-Ott SR, Smith-Myles B. Inclusion of learners with autism spectrum disorders in general education settings. *Top Lang Disord* 2003;**23**(2):116–33.
11. Dean M, Chang Y-C. A systematic review of school-based social skills interventions and observed social outcomes for students with autism spectrum disorder in inclusive settings. *Autism* 2021;**25**(7):1828–43.
12. Lord C, Wagner A, Rogers S, Szatmari P, Aman M, Charman T, *et al.* Challenges in evaluating psychosocial interventions for autistic spectrum disorders. *J Autism Dev Disord* 2005;**35**(6):695–708; discussion 709.
13. Marshall D, Wright B, Allgar V, Adamson J, Williams C, Ainsworth H, *et al.* Social Stories in mainstream schools for children with autism spectrum disorder: a feasibility randomised controlled trial. *BMJ Open* 2016;**6**(8):e011748.
14. Gray C. *Comparison of Social Stories™ 10.0–10.2 Criteria*. 2014. URL: <https://carolgraysocial-stories.com/wp-content/uploads/2015/09/Social-Stories-10.0-10.2-Comparison-Chart.pdf> (accessed October 2018).
15. Bledsoe R, Smith B, Simpson RL. Use of a social story intervention to improve mealtime skills of an adolescent with Asperger syndrome. *Autism* 2003;**7**(3):289–95.
16. Barry LM, Burlew SB. Using social stories to teach choice and play skills to children with autism. *Focus Autism Other Dev Disabil* 2004;**19**(1):45–51.
17. O'Connor E. The use of social story DVDs to reduce anxiety levels: a case study of a child with autism and learning disabilities. *Support Learn* 2009;**24**(3):133–6.

18. Adams L, Gouvousis A, VanLue M, Waldron C. Social story intervention: improving communication skills in a child with an autism spectrum disorder. *Focus Autism Other Dev Disabil* 2004;**19**(2):87–94.
19. Lorimer PA, Simpson RL, Smith Myles B, Ganz JB. The use of social stories as a preventative behavioral intervention in a home setting with a child with autism. *J Posit Behav Interv* 2002;**4**(1):53–60.
20. Kuttler S, Myles BS, Carlson JK. The use of social stories to reduce precursors to tantrum behavior in a student with autism. *Focus Autism Other Dev Disabil* 1998;**13**(3):176–82.
21. Ozdemir S. The effectiveness of social stories on decreasing disruptive behaviors of children with autism: three case studies. *J Autism Dev Disord* 2008;**38**(9):1689–96.
22. Kokina A, Kern L. Social Story™ interventions for students with autism spectrum disorders: a meta-analysis. *J Autism Dev Disord* 2010;**40**(7):812–26.
23. Reynhout G, Carter M. Social Stories™ for children with disabilities. *J Autism Dev Disord* 2006;**36**(4):445–69.
24. Wright B, Marshall D, Adamson J, Ainsworth H, Ali S, Allgar V, *et al.* Social Stories™ to alleviate challenging behaviour and social difficulties exhibited by children with autism spectrum disorder in mainstream schools: design of a manualised training toolkit and feasibility study for a cluster randomised controlled trial with nested qualitative and cost-effectiveness components. *Health Technol Assess (Winchester, England)* 2016;**20**(6):1–258.
25. Quirnbach LM, Lincoln AJ, Feinberg-Gizzo MJ, Ingersoll BR, Andrews SM. Social stories: mechanisms of effectiveness in increasing game play skills in children diagnosed with autism spectrum disorder using a pretest posttest repeated measures randomized control group design. *J Autism Dev Disord* 2009;**39**(2):299–321.
26. Andrews SM. *Increasing Game Playing Skills and Social Comprehension in School-aged Children with Autism using Social Stories*. San Diego, CA: Alliant International University; 2004.
27. Bader R. *Using Social Stories to Increase Emotion Recognition and Labeling in School-age Children with Autism*. San Diego, CA: Alliant International University; 2006.
28. Feinberg MJ. *Using Social Stories to Teach Specific Social Skills to Individuals Diagnosed with Autism*. San Diego, CA: California School of Professional Psychology; 2001.
29. Moudry Quilty K. Teaching paraprofessionals how to write and implement social stories for students with autism spectrum disorders. *Remedial Spec Educ* 2007;**28**(3):182–9.
30. Knapp M, Romeo R, Beecham J. Economic cost of autism in the UK. *Autism* 2009;**13**(3):317–36.
31. Lindsay S, Proulx M, Thomson N, Scott H. Educators' challenges of including children with autism spectrum disorder in mainstream classrooms. *Int J Disabil Dev Educ* 2013;**60**(4):347–62.
32. Hedges LV, Hedberg EC. Intraclass correlation values for planning group-randomized trials in education. *Educ Eval Policy Anal* 2007;**29**(1):60–87.
33. Constantino JN, Gruber CP. *Social Responsiveness Scale: SRS-2*. Torrance, CA: Western Psychological Services; 2012.
34. Abidin R. *Parenting Stress Index – Fourth Edition (PSI-4)*. Lutz, FL: Psychological Assessment Resources; 2012.
35. 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**(6):875–86.
36. Chorpita BF, Yim L, Moffitt C, Umemoto LA, Francis SE. Assessment of symptoms of DSM-IV anxiety and depression in children: a revised child anxiety and depression scale. *Behav Res Ther* 2000;**38**(8):835–55.

37. Law D, Jacob J. *Goals and Goal Based Outcomes (GBOs)*. London: CAMHS Press; 2013.
38. StataCorp. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC2021.
39. Moher D, Schulz KF, Altman DG, CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Clin Oral Investig* 2003;**7**(1):2–7.
40. Rubin DB. Inference and missing data. *Biometrika* 1976;**63**(3):581–92.
41. White IR, Carpenter J, Horton NJ. A mean score method for sensitivity analysis to departures from the missing at random assumption in randomised trials. *Stat Sin* 2018;**28**(4):1985–2003.
42. White I. *RCTMISS: Stata Module to Analyse a Randomised Controlled Trial (RCT) Allowing for Informatively Missing Outcome Data, Statistical Software Components S458304*. Boston College Department of Economics; 2018.
43. Wang R, Lagakos SW, Ware JH, Hunter DJ, Drazen JM. Statistics in medicine – reporting of subgroup analyses in clinical trials. *N Engl J Med* 2007;**357**(21):2189–94.
44. Hasson H. Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci* 2010;**5**(1):67.
45. Mars T, Ellard D, Carnes D, Homer K, Underwood M, Taylor SJ. Fidelity in complex behaviour change interventions: a standardised approach to evaluate intervention integrity. *BMJ Open* 2013;**3**(11):e003555.
46. Scantlebury A, Cockayne S, Fairhurst C, Rodgers S, Torgerson D, Hewitt C, *et al*. Qualitative research to inform hypothesis testing for fidelity-based sub-group analysis in clinical trials: lessons learnt from the process evaluation of a multifaceted podiatry intervention for falls prevention. *Trials* 2020;**21**(1):1–10.
47. Dusenbury L, Brannigan R, Falco M, Hansen WB. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ Res* 2003;**18**(2):237–56.
48. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci* 2007;**2**(1):1–9.
49. Cockayne S, Rodgers S, Green L, Fairhurst C, Adamson J, Scantlebury A, *et al*. Clinical effectiveness and cost-effectiveness of a multifaceted podiatry intervention for falls prevention in older people: a multicentre cohort randomised controlled trial (the REducing Falls with ORthoses and a Multifaceted podiatry intervention trial). *Health Technol Assess* 2017;**21**:1–198.
50. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qual Res Sport Exerc Health* 2021;**13**(2):201–16.
51. Ravens-Sieberer U, Wille N, Badia X, Bonsel G, Burström K, Cavrini G, *et al*. Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Qual Life Res* 2010;**19**(6):887–97.
52. Willems DC, Joore MA, Nieman FH, Severens JL, Wouters EF, Hendriks JJ. Using EQ-5D in children with asthma, rheumatic disorders, diabetes, and speech/language and/or hearing disorders. *Int J Technol Assess Health Care* 2009;**25**(3):391–9.
53. Burström K, Bartonek A, Broström E, Sun S, Egmar AC. EQ-5 D-Y as a health-related quality of life measure in children and adolescents with functional disability in Sweden: testing feasibility and validity. *Acta Paediatr* 2014;**103**(4):426–35.

54. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal*. 2018. URL: www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/technology-appraisal-processes-guide-apr-2018.pdf (accessed November 2022).
55. Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic Evaluation in Clinical Trials*. Oxford: OUP; 2014.
56. Wang H-I, Wright BD, Bursnall M, Cooper C, Kingsley E, Le Couteur A, et al. Cost-utility analysis of LEGO based therapy for school children and young people with autism spectrum disorder: results from a randomised controlled trial. *BMJ Open* 2022;**12**(1):e056347.
57. NHS Improvement. *National Cost Collection 2019/20*. NHS Improvement. 2020. URL: www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/ (accessed November 2022).
58. Curtis L, Burns A. *Unit Costs of Health and Social Care 2016*. PSSRU. 2016. URL: www.pssru.ac.uk/project-pages/unit-costs/2012 (accessed November 2022).
59. Prescribing and Medicines Team ND. *Prescription Cost Analysis, England 2019*. London: Health and Social Care Information Centre; 2020.
60. Department for Education. *DfE Consolidated Annual Report and Accounts 2019 to 2020*. London: Department for Education; 2020.
61. Rubin DB. Statistical matching using file concatenation with adjusted weights and multiple imputations. *J Bus Econ Stat* 1986;**4**(1):87-94.
62. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Econ* 1997;**6**(4):327-40.
63. Tyrer P, Salkovskis P, Tyrer H, Wang D, Crawford MJ, Dupont S, et al. Cognitive-behaviour therapy for health anxiety in medical patients (CHAMP): a randomised controlled trial with outcomes to 5 years. *Health Technol Assess* 2017;**21**(50):1-58.
64. Wright B, Tindall L, Scott A, Lee E, Biggs K, Cooper C. One-session treatment compared with multisession CBT in children aged 7-16 years with specific phobias: the ASPECT non-inferiority RCT. *Health Technol Assess* 2022;**26**:1-74.
65. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;**10**(8):779-87.
66. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *Int J Technol Assess Health Care* 2022;**25**(1):3-9.
67. Loomes R, Hull L, Mandy WPL. What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry* 2017;**56**(6):466-74.
68. Department for Education. *Academic Year 2020/21: Special Educational Needs in England*. London: Department for Education; 2021.
69. Jacob J, Edbrooke-Childs J, Costa da Silva L, Law D. Notes from the youth mental health field: using movement towards goals as a potential indicator of service change and quality improvement. *J Clin Psychol* 2021;**79**:697-710.
70. Flannery H, Jacob J. Measuring psychological outcomes in paediatric settings: making outcomes meaningful using client-defined perspectives. *Clin Child Psychol Psychiatry* 2020;**25**(3):594-603.
71. Gray C. *Social Stories™ 10.1 Definition, Criteria, & Sample Stories*. 2010. URL: <http://blogs.4j.lane.edu/communityaccess/files/2013/05/Social-Story-Criteria.pdf> (accessed November 2022).
72. Curtis L, Burns A. *Unit Costs of Health and Social Care 2018*. Canterbury: Personal Social Services Research Unit, University of Kent. URL: <https://doi.org/10.22024/UniKent/01.02.70995>

Appendix 1 Additional tables

TABLE 45 Progression criteria (Stop/Go)

Criterion	Evaluation
Recruitment	110 children expected to be recruited by the end of 10 months following the start of recruitment. Recruitment rate by school and students will be reported.
Retention	44 children expected to have completed final follow-up. Follow-up and retention rate will be reported.
Safety	All AEs during the internal pilot period will be reported.

TABLE 46 Components of intervention fidelity as described by Hasson^a

Component of intervention fidelity	Explanation
Adherence	Was the Social Stories intervention delivered as intended?
Exposure or dose	Was the amount of the intervention received by participants (frequency, duration) as intended?
Quality of delivery	How did those responsible for delivering the intervention deliver it?
Participant responsiveness	To what extent were participants engaged by or responsive to the intervention?
Program differentiation	Which elements of the intervention and its delivery are important to consider incorporating it into routine practice?
Intervention complexity	How complex is the intervention – this can be a barrier to adoption and use
Facilitation strategies	When evaluating implementation fidelity, what are the specific strategies to support implementation. For example, training. How did those involved in delivery perceive these interventions?
Recruitment	What recruitment procedures were used? What were the challenges of delivering the intervention to this population?
Context	What factors at political, economic, organisational and work group levels affected implementation?

a Hasson⁴⁴

TABLE 47 Interviews conducted with parents, teachers and interventionists

Role	Time point interviewed (either 3 months post randomisation or 6 months post randomisation)	Training model attended (online vs. in person)	School type (SEN school vs. mainstream)	Number of schools represented	School size: small < 200 pupils and large > 200 pupils	School location (rural or urban)	Gender	Age (range and mean)	Years working with young people (range and mean)	Self-reported level of experience with children with ASC
Teacher (n = 5)	Three were interviewed 3 months post randomisation; two were interviewed 6 months post randomisation	Two attended in-person training; three attended online training	All non-SEN	Four schools	Two teachers from small schools (< 200 pupils) and three teachers from large schools (> 200 pupils)	One teacher from a school in a rural location; four from schools in an urban location	Female (n = 5)	Range: 30–44 years old M = 38 years old	Range: 9–19 years M = 14 years	Three 'sound' level of experience, one 'moderate' level of experience and one 'in-depth' level of experience
Interventionist (n = 7)	Five were interviewed 3 months post randomisation; two were interviewed 6 months post randomisation	Seven attended online training	Six non-SEN and one SEN	Six schools	Four interventionists from small schools and three interventionists from large schools	Four interventionists from schools in a rural location and three from schools in an urban location	Female (n = 7)	Range: 21–53 years old M = 42 years old	Range: 8–30 years M = 14 years	Five 'sound' level of experience and two 'limited' experience
Parent (n = 9)	Three were interviewed 3 months post randomisation; six were interviewed 6 months post randomisation	Six attended in-person training; three attended online training	All non-SEN	Eight schools	Three parents of children that attend small schools and six parents of children that attend large schools	Five parents of children that attend schools in a rural location and four parents of children who attend schools in an urban location	Female (n = 9)	Range: 30–47 years old M = 39 years old	N/A	N/A

TABLE 48 Missing primary outcome data patterns by allocation (0 = Missing, 1 = Observed)

Allocation	Baseline	Week 6	Month 6	N (%)	Including in primary analysis
Intervention (N = 129)	0	0	0	0 (0.0)	No
	1	0	0	9 (7.0)	No
	0	1	0	0 (0.0)	No
	0	0	1	1 (0.8)	No
	1	1	0	17 (13.2)	Yes
	1	0	1	15 (11.6)	Yes
	0	1	1	2 (1.6)	No
	1	1	1	85 (65.9)	Yes
Control (N = 120)	0	0	0	0 (0.0)	No
	1	0	0	2 (1.7)	No
	0	1	0	0 (0.0)	No
	0	0	1	0 (0.0)	No
	1	1	0	10 (8.3)	Yes
	1	0	1	8 (6.7)	Yes
	0	1	1	0 (0.0)	No
	1	1	1	100 (83.3)	Yes

TABLE 49 Key baseline characteristics by inclusion in the primary analysis

	Excluded (N = 14)	Included (N = 235)	Total (N = 249)
Age at randomisation (years)			
N	14	235	249
Mean (SD)	8.3 (2.0)	8.5 (1.7)	8.5 (1.7)
Median (Q1, Q3)	7.4 (6.9, 10.5)	8.8 (7.2, 9.9)	8.7 (7.2, 9.9)
Minimum, maximum	5.3, 11.0	4.5, 11.9	4.5, 11.9
Sex, n (%)			
Male	13 (92.9)	172 (73.2)	185 (74.3)
Female	1 (7.1)	63 (26.8)	64 (25.7)
SRS-2 total raw score (teacher-reported)			
N	11	235	246
Mean (SD)	84.7 (35.5)	92.8 (30.1)	92.5 (30.3)
Median (Q1, Q3)	90.0 (57.0, 104.0)	94.0 (69.0, 116.0)	94.0 (69.0, 115.0)
Minimum, maximum	26.0, 140.0	22.0, 165.0	22.0, 165.0

continued

TABLE 49 Key baseline characteristics by inclusion in the primary analysis (*continued*)

	Excluded (N = 14)	Included (N = 235)	Total (N = 249)
SRS-2 total T-score (teacher-reported)			
N	11	235	246
Mean (SD)	67.5 (11.7)	72.3 (10.9)	72.1 (10.9)
Median (Q1, Q3)	69.0 (58.0, 74.0)	73.0 (63.0, 80.0)	73.0 (63.0, 80.0)
Minimum, maximum	50.0, 86.0	47.0, 90.0	47.0, 90.0
SRS-2 total raw score (parent-reported)			
N	14	235	249
Mean (SD)	120.9 (24.1)	115.1 (23.8)	115.4 (23.8)
Median (Q1, Q3)	118.0 (104.0, 142.0)	118.0 (101.0, 130.0)	118.0 (102.0, 130.0)
Minimum, maximum	78.0, 156.0	34.0, 170.0	34.0, 170.0
SRS-2 total T-score (parent-reported)			
N	14	235	249
Mean (SD)	82.8 (7.2)	82.2 (8.3)	82.3 (8.2)
Median (Q1, Q3)	83.5 (78.0, 90.0)	84.0 (78.0, 90.0)	84.0 (78.0, 90.0)
Minimum, maximum	68.0, 90.0	50.0, 90.0	50.0, 90.0
PSI total stress score (parent-reported)			
N	13	231	244
Mean (SD)	103.5 (24.7)	104.1 (19.9)	104.1 (20.1)
Median (Q1, Q3)	101.0 (94.0, 117.0)	104.0 (91.0, 118.0)	104.0 (91.0, 118.0)
Minimum, maximum	56.0, 141.0	47.0, 147.0	47.0, 147.0
RCADS total score (parent-reported)			
N	13	235	248
Mean (SD)	47.0 (23.8)	45.1 (22.4)	45.2 (22.4)
Median (Q1, Q3)	34.0 (34.0, 59.0)	44.0 (27.0, 59.0)	43.5 (28.5, 59.0)
Minimum, maximum	13.0, 86.0	2.0, 102.0	2.0, 102.0
EQ-5D-Y VAS (parent-reported)			
N	14	235	249
Mean (SD)	75.3 (23.0)	80.1 (18.6)	79.8 (18.9)
Median (Q1, Q3)	82.5 (50.0, 99.0)	85.0 (70.0, 95.0)	85.0 (70.0, 95.0)
Minimum, maximum	40.0, 100.0	25.0, 100.0	25.0, 100.0

TABLE 50 Teacher-reported SRS-2 total raw scores by allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Baseline ^a			
N	120	126	246
Mean (SD)	90.4 (28.7)	94.5 (31.8)	92.5 (30.3)
Median (Q1, Q3)	89.5 (72.0, 113.0)	98.0 (69.0, 120.0)	94.0 (69.0, 115.0)
Minimum, maximum	22.0, 162.0	26.0, 165.0	22.0, 165.0
Week 6 ^a			
N	110	104	214
Mean (SD)	88.5 (32.0)	89.6 (31.7)	89.1 (31.8)
Median (Q1, Q3)	91.0 (62.0, 112.0)	93.0 (72.0, 110.5)	92.5 (65.0, 112.0)
Minimum, maximum	17.0, 167.0	23.0, 150.0	17.0, 167.0
Month 6 ^a			
N	108	103	211
Mean (SD)	87.5 (28.1)	86.5 (31.6)	87.0 (29.8)
Median (Q1, Q3)	89.0 (69.0, 107.5)	86.0 (61.0, 109.0)	89.0 (68.0, 108.0)
Minimum, maximum	19.0, 142.0	20.0, 161.0	19.0, 161.0

^a Higher score indicates 'worse' outcome.

TABLE 51 Teacher-reported scores for the goal-based outcome measure by time point and allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Baseline ^a			
N	114	126	240
Mean (SD)	2.5 (2.0)	2.6 (2.2)	2.6 (2.1)
Median (Q1, Q3)	2.0 (1.0, 4.0)	2.0 (1.0, 4.0)	2.0 (1.0, 4.0)
Minimum, maximum	0.0, 9.0	0.0, 10.0	0.0, 10.0
Baseline, ^a n (%)			
0	19 (15.8)	19 (14.7)	38 (15.3)
1	20 (16.7)	24 (18.6)	44 (17.7)
2	26 (21.7)	30 (23.3)	56 (22.5)
3	17 (14.2)	18 (14.0)	35 (14.1)
4	12 (10.0)	12 (9.3)	24 (9.6)
5	12 (10.0)	9 (7.0)	21 (8.4)
6	3 (2.5)	6 (4.7)	9 (3.6)
7	3 (2.5)	3 (2.3)	6 (2.4)
8	1 (0.8)	3 (2.3)	4 (1.6)
9	1 (0.8)	1 (0.8)	2 (0.8)
10	0 (0.0)	1 (0.8)	1 (0.4)
Missing	6 (5.0)	3 (2.3)	9 (3.6)

continued

TABLE 51 Teacher-reported scores for the goal-based outcome measure by time point and allocation (*continued*)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Week 6^a			
N	109	104	213
Mean (SD)	4.4 (2.2)	5.2 (2.5)	4.8 (2.4)
Median (Q1, Q3)	4.0 (3.0, 6.0)	5.0 (3.0, 7.5)	5.0 (3.0, 7.0)
Minimum, maximum	0.0, 10.0	0.0, 9.0	0.0, 10.0
Week 6,^a n (%)			
0	3 (2.5)	2 (1.6)	5 (2.0)
1	6 (5.0)	2 (1.6)	8 (3.2)
2	13 (10.8)	18 (14.0)	31 (12.4)
3	20 (16.7)	10 (7.8)	30 (12.0)
4	16 (13.3)	7 (5.4)	23 (9.2)
5	13 (10.8)	17 (13.2)	30 (12.0)
6	19 (15.8)	12 (9.3)	31 (12.4)
7	9 (7.5)	10 (7.8)	19 (7.6)
8	5 (4.2)	18 (14.0)	23 (9.2)
9	4 (3.3)	8 (6.2)	12 (4.8)
10	1 (0.8)	0 (0.0)	1 (0.4)
Missing	11 (9.2)	25 (19.4)	36 (14.5)
Month 6^a			
N	105	102	207
Mean (SD)	5.2 (2.8)	6.2 (2.2)	5.7 (2.5)
Median (Q1, Q3)	5.0 (3.0, 7.0)	6.0 (5.0, 8.0)	6.0 (4.0, 8.0)
Minimum, maximum	0.0, 10.0	0.0, 10.0	0.0, 10.0
Month 6,^a n (%)			
0	5 (4.2)	1 (0.8)	6 (2.4)
1	7 (5.8)	1 (0.8)	8 (3.2)
2	12 (10.0)	5 (3.9)	17 (6.8)
3	6 (5.0)	7 (5.4)	13 (5.2)
4	13 (10.8)	4 (3.1)	17 (6.8)
5	10 (8.3)	17 (13.2)	27 (10.8)
6	14 (11.7)	18 (14.0)	32 (12.9)
7	12 (10.0)	15 (11.6)	27 (10.8)
8	16 (13.3)	22 (17.1)	38 (15.3)
9	5 (4.2)	8 (6.2)	13 (5.2)
10	5 (4.2)	4 (3.1)	9 (3.6)
Missing	15 (12.5)	27 (20.9)	42 (16.9)

^a Higher score indicates 'better' outcome (i.e. goal being met more of the time).

TABLE 52 Treatment effect estimates from post hoc semiparametric analysis of the week 6 teacher-reported goal-based outcome scores

Covariates (X)	RD [Pr (Y > 5 X)] ^a (95% CI) ^b	RR [Pr(Y > 5 X)] ^c (95% CI) ^c
Baseline score = 5, sex = male	0.17 (0.01 to 0.32)	1.37 (1.01 to 1.87)
Baseline score = 1, sex = male	0.15 (0.01 to 0.28)	1.58 (1.03 to 2.44)
Baseline score = 5, sex = female	0.13 (-0.09 to 0.34)	1.28 (0.84 to 1.93)
Baseline score = 1, sex = female	0.11 (-0.08 to 0.29)	1.42 (0.79 to 2.56)

a Pr (goal score at w6 > 5|intervention, X) – Pr (goal score at w6 > 5|control, X).

b Based on delta method standard errors.

c Pr (goal score at w6 > 5|intervention, X)/Pr (goal score at w6 > 5|control, X).

Note

All reported contrasts are calculated assuming attendance at a non-SEN school with fewer than six participants and age equal to the mean of the trial sample in addition to the covariate patterns given in the leftmost column.

TABLE 53 Treatment effect estimates from post hoc semiparametric analysis of the month 6 teacher-reported goal-based outcome scores

Covariates (X)	RD [Pr (Y > 5 X)] ^a (95% CI) ^b	RR [Pr (Y > 5 X)] ^c (95% CI) ^b
Baseline score = 5, sex = male	0.17 (0.03 to 0.32)	1.33 (1.02 to 1.72)
Baseline score = 1, sex = male	0.18 (0.03 to 0.33)	1.44 (1.04 to 1.98)
Baseline score = 5, sex = female	0.30 (0.12 to 0.48)	1.56 (1.14 to 2.14)
Baseline score = 1, sex = female	0.33 (0.15 to 0.52)	1.81 (1.24 to 2.62)

a Pr (goal score at w6 > 5|intervention, X) – Pr (goal score at w6 > 5|control, X).

b Based on delta method standard errors.

c Pr (goal score at w6 > 5|intervention, X)/Pr (goal score at w6 > 5|control, X).

Note

All reported contrasts are calculated assuming attendance at a non-SEN school with fewer than six participants and age equal to the mean of the trial sample in addition to the covariate patterns given in the leftmost column.

TABLE 54 Parent/carer-reported SRS-2 total T-scores and raw scores by time point and allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Baseline T-score ^a			
N	120	129	249
Mean (SD)	82.5 (8.1)	82.0 (8.4)	82.3 (8.2)
Median (Q1, Q3)	84.0 (78.0, 90.0)	84.0 (78.0, 89.0)	84.0 (78.0, 90.0)
Minimum, maximum	57.0, 90.0	50.0, 90.0	50.0, 90.0
Baseline raw score ^a			
N	120	129	249
Mean (SD)	116.5 (24.0)	114.4 (23.6)	115.4 (23.8)
Median (Q1, Q3)	118.0 (102.0, 132.0)	119.0 (101.0, 129.0)	118.0 (102.0, 130.0)
Minimum, maximum	50.0, 170.0	34.0, 161.0	34.0, 170.0

continued

TABLE 54 Parent/carer reported SRS-2 total T-scores and raw scores by time point and allocation (*continued*)

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Week 6 T-score ^a			
N	97	97	194
Mean (SD)	110.6 (24.3)	112.2 (25.1)	111.4 (24.7)
Median (Q1, Q3)	115.0 (98.0, 127.0)	113.0 (98.0, 131.0)	114.0 (98.0, 129.0)
Minimum, maximum	31.0, 153.0	48.0, 175.0	31.0, 175.0
Week 6 raw score ^a			
N	97	97	194
Mean (SD)	110.6 (24.3)	112.2 (25.1)	111.4 (24.7)
Median (Q1, Q3)	115.0 (98.0, 127.0)	113.0 (98.0, 131.0)	114.0 (98.0, 129.0)
Minimum, maximum	31.0, 153.0	48.0, 175.0	31.0, 175.0
Month 6 T-score ^a			
N	89	95	184
Mean (SD)	112.9 (26.3)	110.3 (24.4)	111.6 (25.3)
Median (Q1, Q3)	114.0 (98.0, 133.0)	115.0 (95.0, 127.0)	114.5 (97.0, 130.0)
Minimum, maximum	33.0, 174.0	41.0, 167.0	33.0, 174.0
Month 6 ^a raw score			
N	89	95	184
Mean (SD)	112.9 (26.3)	110.3 (24.4)	111.6 (25.3)
Median (Q1, Q3)	114.0 (98.0, 133.0)	115.0 (95.0, 127.0)	114.5 (97.0, 130.0)
Minimum, maximum	33.0, 174.0	41.0, 167.0	33.0, 174.0

^a Higher score indicates 'worse' outcome.

TABLE 55 Parent-reported RCADS total scores by time point and allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Baseline^a			
N	120	128	248
Mean (SD)	45.7 (23.3)	44.7 (21.6)	45.2 (22.4)
Median (Q1, Q3)	45.0 (27.6, 59.5)	41.0 (28.5, 59.0)	43.5 (28.5, 59.0)
Minimum, maximum	2.0, 102.0	5.0, 100.0	2.0, 102.0
Week 6^a			
N	97	97	194
Mean (SD)	44.2 (23.6)	43.9 (22.4)	44.0 (23.0)
Median (Q1, Q3)	42.0 (25.0, 61.0)	44.0 (28.0, 57.0)	44.0 (26.0, 61.0)
Minimum, maximum	3.0, 99.0	5.0, 106.0	3.0, 106.0
Month 6^a			
N	88	94	182
Mean (SD)	42.1 (23.6)	43.9 (21.2)	43.1 (22.4)
Median (Q1, Q3)	39.0 (22.5, 58.0)	43.9 (29.7, 58.2)	42.0 (27.0, 58.2)
Minimum, maximum	1.0, 94.0	1.0, 93.2	1.0, 94.0

^a Higher score indicates 'worse' outcome.

TABLE 56 Parental stress index total stress scores by time point and allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Baseline^a			
N	117	127	244
Mean (SD)	103.2 (20.9)	104.9 (19.5)	104.1 (20.1)
Median (Q1, Q3)	104.0 (90.0, 117.0)	104.0 (94.0, 119.0)	104.0 (91.0, 118.0)
Minimum, maximum	47.0, 147.0	52.0, 147.0	47.0, 147.0
Week 6^a			
N	95	89	184
Mean (SD)	103.9 (23.5)	102.9 (20.5)	103.5 (22.0)
Median (Q1, Q3)	104.0 (90.0, 120.0)	103.0 (88.0, 118.0)	103.0 (88.0, 118.0)
Minimum, maximum	38.0, 158.0	53.0, 164.0	38.0, 164.0
Month 6^a			
N	84	90	174
Mean (SD)	103.3 (23.1)	103.2 (18.4)	103.3 (20.8)
Median (Q1, Q3)	104.5 (86.5, 120.0)	104.0 (93.0, 115.0)	104.0 (89.0, 118.0)
Minimum, maximum	48.0, 155.0	55.0, 152.0	48.0, 155.0

^a Higher score indicates 'worse' outcome.

TABLE 57 European Quality of Life-5 Dimension Youth Questionnaire general health VAS scores (parent-reported) by time point and allocation

	Control (N = 120)	Intervention (N = 129)	Total (N = 249)
Baseline ^a			
N	120	129	249
Mean (SD)	79.0 (19.0)	80.6 (18.8)	79.8 (18.9)
Median (Q1, Q3)	80.0 (70.0, 95.0)	90.0 (70.0, 95.0)	85.0 (70.0, 95.0)
Minimum, maximum	30.0, 100.0	25.0, 100.0	25.0, 100.0
Week 6 ^a			
N	97	97	194
Mean (SD)	79.9 (16.3)	79.9 (17.8)	79.9 (17.0)
Median (Q1, Q3)	80.0 (70.0, 91.0)	85.0 (70.0, 95.0)	80.0 (70.0, 93.0)
Minimum, maximum	10.0, 100.0	30.0, 100.0	10.0, 100.0
Month 6 ^a			
N	89	93	182
Mean (SD)	82.1 (18.4)	80.7 (18.2)	81.4 (18.2)
Median (Q1, Q3)	85.0 (75.0, 95.0)	85.0 (75.0, 95.0)	85.0 (75.0, 95.0)
Minimum, maximum	0.0, 100.0	10.0, 100.0	0.0, 100.0

a Higher score indicates 'better' outcome.

TABLE 58 Treatment effect estimates at 6 weeks and 6 months for the post hoc semiparametric analysis of the EQ-5D-Y general health VAS score (parent-reported)

	Estimated difference (intervention – control) in expected value (95% CI) ^a
Week 6	0.46 (–2.97 to 3.73)
Month 6	–1.19 (–5.39 to 2.70)

a Based on bootstrap standard errors (1000 replicates). Bias corrected intervals reported.

Appendix 2 Unit cost tables

TABLE 59 The interventions received by time period

Usual care	Baseline			6 months		
	Total (n = 249)	Social Stories (n = 129)	Usual care (n = 120)	Total (n = 212)	Social Stories (n = 104)	Usual care (n = 108)
1 : 1 mentoring/individual work	94 (37.8)	46 (35.7)	48 (40.0)	69 (32.5)	32 (30.8)	37 (34.3)
Emotional Literacy Support Assistant (or equivalent)	15 (6.0)	6 (4.7)	9 (7.5)	8 (3.8)	5 (4.8)	3 (2.8)
Functional communication training	1 (0.4)	0 (0)	1 (0.8)	1 (0.5)	0 (0)	1 (0.9)
Intensive interaction	14 (5.6)	8 (6.2)	6 (5.0)	8 (3.8)	4 (3.8)	4 (3.7)
LEGO	36 (14.5)	16 (12.4)	20 (16.7)	21 (9.9)	6 (5.8)	15 (13.9)
Music, art or drama therapy	16 (6.4)	10 (7.8)	6 (5.0)	5 (2.4)	2 (1.9)	3 (2.8)
PECS	13 (5.2)	4 (3.1)	9 (7.5)	10 (4.7)	3 (2.9)	7 (6.5)
Responsive teaching	5 (2.0)	3 (2.3)	2 (1.7)	3 (1.4)	1 (1.0)	2 (1.9)
Social communication groups	53 (21.3)	25 (19.4)	28 (23.3)	50 (23.6)	25 (24.0)	25 (23.1)
Social skills work group	50 (20.1)	23 (17.8)	27 (22.5)	36 (17.0)	12 (11.5)	24 (22.2)
Social stories	15 (6.0)	6 (4.7)	9 (7.5)	55 (25.9)	45 (43.3)	10 (9.3)
SULP	3 (1.2)	2 (1.6)	1 (0.8)	1 (0.5)	1 (1.0)	0 (0)
SCERTS	12 (4.8)	10 (7.8)	2 (1.7)	3 (1.4)	1 (1.0)	2 (1.9)
TEACCH	5 (2.0)	3 (2.3)	2 (1.7)	2 (0.9)	1 (1.0)	1 (0.9)
Verbal behaviour approach	2 (0.8)	1 (0.8)	1 (0.8)	0 (0)	0 (0)	0 (0)
Video modelling	1 (0.4)	0 (0)	1 (0.8)	1 (0.5)	0 (0)	1 (0.9)
Visual schedules	54 (21.7)	26 (20.2)	28 (23.3)	36 (17.0)	17 (16.3)	19 (17.6)
Other	62 (24.9)	38 (29.5)	24 (20.0)	33 (15.6)	19 (18.3)	14 (13.0)

PECS, picture exchange communication systems; SCERTS, social-communication, emotional regulation and transactional support; SULP, social use of language programming; TEACCH, treatment and education of autistic and communication-handicapped children.

TABLE 60 General health community services

Item	Unit cost (£)			Source
	At home ^a	At clinic/surgery	Via phone/e-mail	
GP	45.9	39.2	37.6	PSSRU 2019 (chapter 10.3, 10.4)
Community nurse ^b	55.7	49.0	49.0	PSSRU 2019 (chapter 6.1)
Community paediatrics	161.7	155.0	155.0	PSSRU 2019 (chapter 6.1)
Social care worker ^b	28.7	22.0	22.0	PSSRU 2019 (chapter 11.2)
Home care worker ^b	20.7	14.0	14.0	PSSRU 2019 (chapter 11.5)
Family support worker ^c	23.2	16.5	16.5	PSSRU 2018 (chapter 6.9)
Helpline ^d	–	–	6.0	PSSRU 2019 (chapter 10.5)
Occupational therapist	151.7	145.0	145.0	PSSRU 2019 (chapter 6.1)
Physiotherapist	87.7	81.0	81.0	PSSRU 2019 (chapter 6.1)
Speech and language therapist	89.7	83.0	83.0	PSSRU 2019 (chapter 6.1)
Dentist ^b	73.2	66.5	66.5	PSSRU 2019 (chapter 10.6)
Dietician	92.7	86.0	86.0	PSSRU 2019 (chapter 7.1)

a Including the average travel time of 12 minutes per visit (PSSRU 2015).

b Assume the duration of visit is 30 minutes.

c Unit cost was inflated to 2019 prices based on originally reported values from PSSRU 2018/19 (Curtis and Burns 2018)⁷².

d Assume the cost is the same as telephone triage led by nurse.

TABLE 61 Mental health-related, community-based services

Item	Unit cost (£)	Source
Child psychiatrist	133.0	PSSRU 2019 (chapter 9)
Child psychotherapist	94.0	PSSRU 2019 (chapter 6.10)
Child psychologist or clinical psychologist ^a	109.0	PSSRU 2019 (chapter 9)
Mental health nurse or CAMHS therapist ^a	60.0	PSSRU 2019 (chapter 10.1)

a Assume 1 hour per session.

TABLE 62 Hospital-based services

Item	Unit cost (£)	Source
Accident and emergency	142.4	National Cost Collection 2019/20 (service code: T01NA)
NHS walk-in centre	34.7	National Cost Collection 2019/20 (service code: T04NA)
Urgent care centre	34.7	National Cost Collection 2019/20 (service code: T04NA)
Outpatient visit – paediatrics	232.0	National Cost Collection 2019/20 (service code: 420)
Outpatient visit – ear, nose and throat	124.0	National Cost Collection 2019/20 (service code: 215)
Outpatient visit – allergy	247.0	National Cost Collection 2019/20 (service code: 255)
Outpatient visit – epilepsy	222.0	National Cost Collection 2019/20 (service code: 223)
Outpatient visit – paediatrics	232.0	National Cost Collection 2019/20 (service code: 420)

TABLE 62 Hospital-based services (continued)

Item	Unit cost (£)	Source
Outpatient visit – dental medicine	171.0	National Cost Collection 2019/20 (service code: 450)
Outpatient visit – dermatology	170.0	National Cost Collection 2019/20 (service code: 257)
Outpatient visit – cardiology	142.0	National Cost Collection 2019/20 (service code: 320)
Outpatient visit – occupational therapy	73.0	National Cost Collection 2019/20 (service code: 651)
Outpatient visit – podiatry	62.0	National Cost Collection 2019/20 (service code: 653)
Outpatient visit – ophthalmology	103.0	National Cost Collection 2019/20 (service code: 216)
Outpatient visit – psychiatry	361.0	National Cost Collection 2019/20 (service code: 711)
Day case – bladder condition	1031.0	National Cost Collection 2019/20 (HRG* code: PX08C)
Day case – non-surgical ophthalmology	558.0	National Cost Collection 2019/20 (HRG code: PP64B)
Day case – ear condition	1021.0	National Cost Collection 2019/20 (HRG code: PP64B)
Day case – bone condition	771.0	National Cost Collection 2019/20 (HRG code: PX05C)
Day case – dental procedure	683.0	National Cost Collection 2019/20 (HRG code: CD03B)
Inpatient stay – arm fracture	5281.0	National Cost Collection 2019/20 (HRG code: HE51C)

HRG = Health Resource Group

TABLE 63 Medications

Item	Chemical name	Dosage	Unit cost (£)	Source
Melatonin_Tab 2 mg	Melatonin	2 mg tablet (once daily)	1.33 per quantity	PCA 2019 (0401010ADAABKBK)
Melatonin_ 2 mg/5 ml oral solution	Melatonin	2 mg/5 ml (once daily)	0.17 per quantity	PCA2019 (0401010ADAABYBY)
Clenil Modulite_Inha 50 mcg ^a	Beclometasone Dipropionate	50 mcg (2 puffs daily)	3.70 per quantity	PCA 2019 (0302000COBPAABE)
Salbutamol_Inha 100 mcg ^a	Salbutamol	Two puffs (as required)	1.70 per quantity	PCA 2019 (0301011R0AAAAAA)
Movicol_Paed Pdr Sach	Macrogol 3350	6.9 g sachet (2 sachets a day)	0.30 per quantity	PCA 2018 (0106040M0BBAIAB)
Equasym XL_tab_ 30 mg	Methylphenidate hydrochloride	30 mg tablet (once daily)	1.17 per quantity	PCA 2019 (0404000M0BCAFV)
Methylphenidate_tab 10 mg	Methylphenidate hydrochloride	10 mg tablet (once daily)	0.12 per quantity	PCA 2019 (0404000M0AAAAAA)
Dulcolax Pico_liquid 5 mg/5 ml	Sodium picosulphate	5 mg/5 ml (once daily)	0.02 per quantity	PCA 2019 (0106020POBDACAB)
Cetirizine_tab 10 mg	Cetirizine hydrochloride	10 mg tablet (once daily)	0.03 per quantity	PCA 2019 (0304010I0AAAAAA)
Lactulose_oral solution 10 g/15 ml	Lactulose	10 g/15 ml (once daily)	0.25 per quantity	PCA 2019 (0106040G0AAAGAG)
Montelukast_tab 5 mg	Montelukast	5 mg tablet (once daily)	0.05 per quantity	PCA 2019 (0303020G0AAAAAA)

a Assume each inhaler contains 200 puffs.

Note

The medication costs were based on the Prescription Cost Analysis (PCA) 2019. Due to the significant amount of information, further details are available upon request.

TABLE 64 School-based health services

Item	Unit cost (£)	Source ^a
Educational psychologist	23.6/hour	NJC Green Book 2019 (SCP 43)
Education welfare officer	12.4/hour	NJC Green Book 2019 (SCP 17)
School or college nurse	12.4/hour	NJC Green Book 2019 (SCP 17)

a Based on the average SCP from the NJC for Local Government Services – Green Book 2019 (<https://neu.org.uk/support-staff>). Hourly rate was calculated by dividing annual salary by 52.143 weeks (which is 365 days divided by 7) and then divided by 37 hours (the standard working week in the National Agreement 'Green Book').

TABLE 65 Education services/support

Item	Unit cost (£)	Source ^a
Teacher	16.6/hour	NJC Green Book 2019 (SCP 29)
TA	9.9/hour	NJC Green Book 2019 (SCP 6)
HLTA	11.2/hour	NJC Green Book 2019 (SCP 12)
SENCO	19.6/hour	NJC Green Book 2019 (SCP 36)
SEN	11.9/hour	NJC Green Book 2019 (SCP 15)
LM	11.6/hour	NJC Green Book 2019 (SCP 14)
Emotional Literacy Support Assistant head	10.1/hour	NJC Green Book 2019 (SCP 7)
Literacy support assistant	11.0/hour	NJC Green Book 2019 (SCP 11)
STA	14.5/hour	NJC Green Book 2019 (SCP 24)
Other	9.9/hour	Assume the same as TA

LM, learning mentor; NJC, National Joint Council; SCP, spinal column point.

a Based on the average SCP from the NJC for Local Government Services – Green Book 2019 (<https://neu.org.uk/support-staff>). Hourly rate was calculated by dividing annual salary by 52.143 weeks (which is 365 days divided by 7) and then divided by 37 hours (the standard working week in the National Agreement 'Green Book').

TABLE 66 Private expenses

Item	Unit cost (£)	Source
Privately paid mental health services	360.0/session	www.psychiatry-uk.com/fees/
After-school club	57.4/week	Childcare survey 2019
Holiday club	133.4/week	Childcare survey 2019
Day care	158.6/day	PSSRU 2019 (chapter 6.11)
Child care – home support	25.8/hour	PSSRU 2019 (chapter 6.11)
PA	8.2/hour	National minimal wage 2019
Swimming club	4.58/session	2019 State of the UK swimming industry report www.leisuredb.com/publications/

PA, personal assistant.

TABLE 67 Productivity cost

Item	Unit cost (£)	Source
AWE of employed people in the UK	585/week	Office for National Statistics 2019
AWE, average weekly earnings.		

TABLE 68 European Quality of Life-5 Dimension Youth Questionnaire responses by trial arms by data collection time points

	Baseline ^a			Month 6 ^a		
	Level 1, n (%)	Level 2, n (%)	Level 3, n (%)	Level 1, n (%)	Level 2, n (%)	Level 3, n (%)
Social Stories (n = 58)						
Mobility	39 (67.2)	15 (25.9)	4 (6.9)	43 (74.1)	13 (22.4)	2 (3.5)
Self-care	14 (24.1)	27 (46.6)	17 (29.3)	16 (27.6)	29 (50.0)	13 (22.4)
Usual activity	25 (43.1)	26 (44.8)	7 (12.1)	26 (44.8)	28 (48.3)	4 (6.9)
Pain/discomfort	41 (70.7)	15 (25.9)	12 (3.4)	42 (72.4)	15 (25.9)	1 (1.7)
Anxiety/depression	20 (34.5)	33 (56.9)	5 (8.6)	22 (37.9)	32 (55.2)	4 (6.9)
	Baseline ^a			Month 6 ^a		
Usual care (n = 54)	Level 1 n (%)	Level 2 n (%)	Level 3 n (%)	Level 1 n (%)	Level 2 n (%)	Level 3 n (%)
Mobility	38 (70.4)	14 (25.9)	2 (3.7)	39 (72.2)	15 (27.8)	-
Self-care	12 (22.2)	29 (53.7)	13 (24.1)	14 (25.9)	29 (53.7)	11 (20.4)
Usual activity	19 (35.2)	31 (57.4)	4 (7.4)	19 (35.2)	29 (53.7)	6 (11.1)
Pain/discomfort	39 (72.2)	13 (24.1)	2 (3.7)	38 (70.4)	15 (27.8)	1 (1.8)
Anxiety/depression	22 (40.7)	19 (35.2)	13 (24.1)	24 (44.4)	25 (46.3)	5 (9.3)

a Level 1: none; Level 2: some; Level 3: extreme.

EME
HSDR
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 and Care 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