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

Barry Wright,[1] David Marshall,[1]* Joy Adamson,[2] Hannah Ainsworth,[2] Shehzad Ali,[3] Victoria Allgar,[3] Danielle Collingridge Moore,[3] Elizabeth Cook,[2] Paul Dempster,[3] Lisa Hackney,[1] Dean McMillan,[3] Dominic Trepél[3] and Chris Williams[1]

[1]Lime Trees Child and Adolescent Mental Health Service Research Team, York, UK
[2]York Clinical Trials Unit, Department of Health Sciences, York, UK
[3]Department of Health Sciences, University of York, York, UK

*Corresponding author

Declared competing interests of authors: Resulting from this work, a manual has been written by Barry Wright and Chris Williams in conjunction with Carol Gray on the writing of Social Stories™ for children in mainstream schools in the UK.

Published January 2016
DOI: 10.3310/hta20060

Scientific summary

Social Stories™ for children with autism spectrum disorder

Health Technology Assessment 2016; Vol. 20: No. 6
DOI: 10.3310/hta20060

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Children with autism spectrum disorders (ASDs) often have difficulties processing social information and may benefit from additional support to understand social situations. Social Stories™ (Carol Gray) are a positive, child-focused intervention designed to fill the gaps of a child's social understanding and promote positive social behaviours.

Objectives

The overall aim was to conduct a feasibility study for a fully powered randomised controlled trial (RCT). Specifically, we aimed:

1. to conduct a systematic review examining the use of social stories for children with ASD, with reference to mainstream school-aged children and challenging behaviour as specified by the Heath Technology Assessment database
2. to conduct a qualitative analysis on information relating to the optimum design and use of Social Stories for children with ASD in a mainstream school setting
3. to form an expert writing panel and develop a manualised toolkit (including a training package) for writing and delivering Social Stories
4. to conduct a feasibility RCT comparing the developed intervention with an attention control on challenging behaviour
5. to establish the acceptability and utility of the developed manualised Social Stories intervention to teachers, parents/carers and children
6. to identify parameters, outcomes and cost-effectiveness from the feasibility study to inform a full-scale RCT.

Methods

Development of the intervention

We conducted a comprehensive systematic review of the evidence for the effectiveness of social stories to use as a basis for developing the intervention. This information combined with qualitative information from parent and professional user groups was fed into an expert panel responsible for fashioning the intervention. A subgroup of the expert panel formed a writing group including two patient and public involvement (PPI) members. This group drafted a training manual through a process of iterative feedback. Once a robust version was achieved, it was piloted with six participant groups before it was finalised for the feasibility study.

Systematic review

A systematic review was conducted to provide a comprehensive description of social story use in practice (overview of practice) and a quantitative estimate of their clinical effectiveness (overview of effectiveness). Searches included free text and controlled terms and were not limited by date range, language or publication status. The following electronic databases were searched: Applied Social Sciences Index and Abstracts, Australian Education Index, British Education Index, British Library integrated catalogue, British Nursing Index, Campbell Library, Cumulative Index to Nursing and Allied Health Literature, ClinicalTrials.gov, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Conference Proceedings Citation Index – Science, Conference Proceedings Citation Index – Social Science &

Humanities, Database of Abstracts of Reviews of Effects, EMBASE, Education Resources Information Center, Health Technology Assessment database, Index to Theses, International Clinical Trials Registry Platform, Library of Congress catalog, MEDLINE, MEDLINE In-Process & Other Non-Indexed Citations, NHS Economic Evaluation Database, OAIster, PsycINFO, Social Care Online, Science Citation Index, Social Science Citation Index, Social Policy & Practice, Social Services Abstracts and Zetoc. Twenty-one additional specialist organisation websites were also searched. Searches were completed in July 2011.

Two researchers independently examined the titles and abstracts of all identified citations using a pre-piloted, standardised screening form. For the overview of practice all study designs were included. This included RCTs, non-RCTs, pre–post designs, quantitative single-case studies, non-quantitative case studies and narrative descriptions of the use of an intervention. For inclusion, the study had to include participants with a diagnosis of ASD between 4 and 15 years and an intervention described as a social story. No specific exclusion criteria were set for the comparator or outcome for the overview of practice. For inclusion into the overview of effectiveness a standardised, numerical measure of behavioural outcomes or non-standardised numerical measure of behaviour was required. Full papers were obtained for any citations passing this first sift and these were independently examined by two researchers. Disagreements were resolved through consensus and referred to a third reviewer when necessary.

For the single-case designs, data on clinical effectiveness were extracted from reported data or graphs of any target behaviour. For between-group designs, information was extracted on the sample size of each group, and post-treatment means and standard deviations (SDs). A quality assessment was carried out for each study using the Single-Case Experimental Design scale for single cases and the Cochrane risk of bias tool for the RCTs.

We used two estimates of effect for single-case designs. The percentage of non-overlapping data (PND) is the percentage of phase B (i.e. intervention phase) data that exceeds the highest point (or lowest depending on the direction of scoring) for the phase A data (i.e. baseline phase). The percentage of data points exceeding the median (PEM) involves calculating the median for the baseline phase and then identifying the proportion of data points in the intervention phase that exceed this value. As in previous studies, both estimates used a value of $\geq 90\%$ to indicate that an intervention was highly effective, ≥ 70 to $< 90\%$ for moderate effectiveness and $< 70\%$ for questionable effectiveness or that the intervention is not effective.

For between-groups designs, the standardised mean difference (Hedges' *g*) was calculated. When insufficient data were reported to make this calculation, the results in the original paper are given. A number of the studies used a design in which the comparator group went on to receive treatment. In these designs, data were analysed up to the point at which the distinction between the treatment and comparator groups was maintained. There were an insufficient number of studies that were comparable in terms of the comparator and the target behaviour to conduct a meta-analysis.

Qualitative methodology

We interviewed users of Social Stories for information for the development of the intervention. This included parents in one group, teachers and other professionals in another and young people in a third. Following this, we interviewed six participant groups who had used the manual and training materials in the pilot phase of the study.

Audio-recordings were fully transcribed and anonymised. The framework method of analysis was used to organise the data. Although the framework approach starts with pre-set aims, it retains some flexibility in order to reflect the experiences and views of the study participants based on a grounded and inductive approach.

Feasibility study

We explored the feasibility of delivering a RCT for children with ASD in mainstream schools comparing the manualised Social Stories intervention with an attention control. We aimed to assess recruitment and retention of participants, the appropriateness of cluster randomisation, the process of delivering the intervention and trial, and to choose a primary outcome measure for a full scale RCT. We also sought to produce estimates of effect sizes to inform the sample size for a future trial.

This was a single-centre cluster randomised feasibility study. The study was given ethical approval by the Leeds East Research Ethics Committee (11/YH/0340). Participant groups were only included if the child had a diagnosis of ASD, were between 5 and 15 years of age, and were attending a mainstream school. They also must have had reported social difficulties or challenging behaviours as identified by the adult participants. They were excluded if they had used a Social Story™ in the preceding 6 months, or were leaving school in the following 4 months.

Baseline measures were taken including the Strengths and Difficulties Questionnaire, the Social Responsiveness Scale-2 (SRS-2), a goal-based outcome measure, a Likert scale of behavioural frequency and a behaviour diary. We also used a bespoke resource utilisation questionnaire, two quality of life measures for use with children [European Quality of Life-5 Dimensions (EQ-5D) and Health Utilities Index 2 (HUI2)], the Parental Stress Index and the Spence Children's Anxiety Scale. They were delivered at baseline, 6 weeks and 16 weeks. Baseline measures were taken before the participant groups were cluster randomised by school, using remote randomisation through the York Clinical Trials Unit. An adverse event monitoring system was followed.

Quality of life was assessed through use of the European Quality of Life-5 Dimensions youth questionnaire (EQ-5DY), EQ-5D (proxy) and the HUI2 (proxy). The ability to assess resource implications was piloted using a bespoke resource-use questionnaire. As social difficulties in ASD may affect provisions for education, supplementary perspectives for an economic analysis were considered in which both parents and teachers were asked to report educational outcomes. A wider societal perspective was also explored.

Results

Development of intervention

Systematic review

The systematic review found that the research into social stories is predominantly based in the USA in under-12-year-olds. Most studies either did not follow Gray's criteria or did not report if they did. Some studies purported to use the criteria but gave example stories that did not meet them.

The review of the single-case studies supported the effectiveness of the intervention. The assessment of effectiveness for single-case designs, using PND, indicated that just under half of the target behaviours met the criteria for highly or moderately effective. The assessment using PEM presents a more positive picture, with most target behaviours meeting these criteria. However, the majority did not meet the criterion for the independence of assessment item.

The review of the between-group studies appears to provide evidence of the clinical effectiveness of social stories. However, there were gaps in the reporting of results and insufficient information to assess a number of important sources of potential bias. The interventions used also did not meet Gray's criteria.

Qualitative work

The qualitative work suggested teaching assistants would be best placed to deliver the intervention and in a quiet place. Focus groups highlighted the importance of parents being involved at all stages. Participants wanted a manual that was short and simple, including checklists and examples of Social Stories. They also wanted training to support story development.

Information about the logistics of the study and intervention led to fine tuning the methodology. This included information about the timing of training, intervention delivery and assessments; the importance of visual elements; and the size of the manual. Challenging behaviour was found to be too limiting as an outcome assessment of effectiveness and was expanded to include social difficulties. Social Stories were considered to be less appropriate in secondary schools and harder to implement.

Feasibility study

We identified 140 pupils with ASD within 134 schools in the local area. Participants were recruited through direct contact, contact with parents and contact with educational and clinical professionals. After exclusion, we recruited 50 pupils (37 male) across 37 schools in a 12-month period. The mean age at study entry was 9.5 years (SD 2.63 years).

School cluster size was on average 1.35 children, questioning the need for cluster randomisation. Delivery of the intervention and trial was feasible. Completion of outcome measures was highest in the teacher group both at baseline (100%) and the follow-up points (6 weeks 90% and 16 weeks 74%). Parent completion rates were comparatively lower (baseline 92%, 6 weeks 74% and 16 weeks 60%) and child rates were very low (baseline 72%, 6 weeks 58% and 16 weeks 34%). Consequently, the teacher group were concluded to be the most appropriate to complete the primary outcome measure. Two outcome measures, the SRS-2 and the custom-made goal-based measure, showed both high levels of completion rates (above 80%) at the primary follow-up point (6 weeks post intervention) and captured social and behaviour skills that professionals and participants reported as important to alter. Power calculations were based on these two outcome measures leading to a total proposed sample size of 180 participant groups.

The feasibility of collecting resource-related information from the educational sector was promising. Teachers gave a good indication of how often school was attended, grades and the level of pupil productivity. It was found that parents gave information on social service use, contact with criminal justice system and issues affecting their own productivity. Most parents (74.3%) who were employed reported their annual income, which can help capture implications of a child's challenging behaviour or social difficulties on their parents' income.

Discussion

The results of the systematic review indicated that the evidence for social stories was predominantly based in the USA and used single-case design methodology. The effectiveness of the intervention was supported by the literature, but there was variation in the extent to which the studies met the Social Stories criteria.

The development of the intervention stage and recruitment of all 50 participant groups were completed within the proposed time scale. Retention and questionnaire return rates were high, although parent return rates dropped across follow-up points. Two outcome measures showed high levels of completion rates, good face validity and a trend in the desired direction, indicating that they would be suitable for use as primary outcome measures in a full-scale trial.

For the economic evaluation, parents gave information on social service use and issues affecting their own productivity. Most employed parents reported their annual income (74.3%), which can help capture implications of a child's condition on their lost income.

This study was particularly effective at including PPI. It involved service users at all stages who provided valuable information to inform a future trial. We used a very wide definition of challenging behaviour to ensure the inclusion of all children who could potentially benefit from the intervention.

Blinding of participants to the intervention was not feasible. However, the feasibility of operationalising behaviours that could be monitored by a blinded independent observer in a full-scale trial was examined. Treatment fidelity with respect to how closely the finalised stories reflected Gray's 10 criteria was not assessed because of low levels of story return rates.

Conclusions

The study suggests that a full-scale trial would be feasible. Power calculations suggest that 180 children would be needed. However, attention should also be paid to treatment fidelity by examining the stories after use.

The developed training package has the potential to help children with ASD overcome some of the social difficulties they experience in a non-expensive, non-intrusive way, subject to its being shown to be effective and cost-effective in a future trial. Were this to be demonstrated, the intervention could be made widely available to educational and community settings across the country.

Some implications for clinical practice have been gleaned from the feasibility trial. It is clear that parents and teachers are well disposed towards the Social Story intervention. Many regard it as effective in practice and are motivated to attend training. Many have also been enthusiastic in requesting accessible training materials for the UK setting.

Study registration

This study is registered as PROSPERO CRD42011001440.

Trial registration

Current Controlled Trials ISRCTN96286707.

Funding

This project was funded by the NIHR Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 20, No. 6. See the NIHR Journals Library website for further project information.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

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

Editorial contact: nihredit@southampton.ac.uk

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

Criteria for inclusion in the *Health Technology Assessment* journal

Reports 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

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

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.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/169/07. The contractual start date was in October 2011. The draft report began editorial review in October 2014 and was accepted for publication in July 2015. 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' report 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 report.

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

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