Imagining Autism: Feasibility of a Drama-Based Intervention on the Social, Communicative and Imaginative behaviour of Children with Autism.

Julie Beadle-Brown[1], David Wilkinson[2], Lisa Richardson[1], Nicola Shaughnessy[3] Melissa Trimingham[3], Jennifer Leigh[1], Beckie Whelton[1] and Julian Himmerich[1]

[1]The Tizard Centre, University of Kent, Canterbury, Kent, UK, CT2 7NZ

[2]School of Psychology, University of Kent, Canterbury, Kent, UK, CT2 7NP

[3]School of Arts, University of Kent, Canterbury, Kent, UK, CT2 7UG.

Corresponding author: Julie Beadle-Brown, The Tizard Centre, University of Kent, Canterbury, Kent, UK, CT2 7NZ. Tel: (0)1227 827763. Email: j.d.beadle-brown@kent.ac.uk.

Accepted for publication in *Autism: International Journal of Research and Practice* 28th April 201

Abstract

We report the feasibility of a novel, school-based intervention, coined 'Imagining Autism', in which children with autism engage with drama practitioners though participatory play and improvisation in a themed multi-sensory "pod" resembling a portable, tent-like structure. 22 children, aged 7–12 years, from three UK schools engaged in the 10week programme. Measures of social interaction, communication, emotion recognition, along with parent and teacher ratings, were collected before and up to 12 months after the intervention. Feasibility was evaluated through 4 domains: (1) process (recruitment, retention, blinding, inter-rater reliability, willingness of children to engage), (2) resources (space, logistics), (3) management (dealing with unexpected changes, ease of assessment), and (4) scientific (data outcomes, statistical analyses). Overall, the children, parents and teachers showed high satisfaction with the intervention, the amount of missing data was relatively low, key assessments were implemented as planned, and evidence of potential effect was demonstrated on several key outcome measures. Some difficulties were encountered with recruitment, test administration, parental response, and the logistics of setting up the pod. Following several protocol revisions and the inclusion of a control group, future investigation would be justified to more thoroughly examine treatment effects.

Background

There is no cure for autism, no single effective intervention and a lack of formal trials validation for many of the interventions that claim to be effective (Lord, 2000; Schopler, 2001; McConnell, 2002; Matson, Matson and Rivet., 2007; Ospina et al., 2008; Eldevik et al., 2009; Eikeseth, 2009; Seida et al., 2009; Virués-Ortega, 2010; Beadle-Brown, Mills and Marchant, 2011; Warren et al., 2011). These treatment failings are best conveyed by the fact that children with autistic spectrum disorder (ASD) are typically enrolled on between 4 and 6 different treatments at any one time, and have tried between seven and nine treatments of different types (Goin-Kochel, Myers and Mackintosh., 2007; Green et al., 2006). The aim of the current study is to make a preliminary assessment of both the feasibility and likely treatment effect of a novel drama-based intervention that we call '[name removed for Blinding]'.

The most common treatment approaches tend to be psycho-educational and behaviour based, including Early Intensive Behaviour Intervention (e.g. Lovaas, 1987; McEachin et al., 1993; Eikeseth et al., 2012) the Early Start Denver Model (Rogers and Dawson, 2009 a; b), social skills teaching, relational development training (Gutstein et al., 2007) and communication based approaches such as PECS (Frost and Bondy, 2002). Interventions with a slightly stronger evidence-base such as those incorporating applied behaviour analysis also tend to be intensive, require substantial resources to implement, and attract criticism from the autism community for trying to "cure" the autism, control or change the child or for focusing on only some domains such as educational achievement and not overall well-being (The Guardian, 2015).

Recently there has been increased focus on interventions that improve social skills. These have included the NETT program (Soorya et al, 2015) and parent- or caregiver- mediated approaches such as the PACT (e.g. Green et al., 2010; Pickles et al., 2016) and the PEERS programmes (Laugeson et al., 2015). Peer-mediated, peer-modelling, video-modelling and video-instruction techniques have also been utilised (Corbett, 2003; Corbett and Abdullah, 2005; Shukla-Mehta et al., 2010; Wang et al., 2011; Kasari et al., 2012). These interventions have primarily focused on improving communication, social competence and social interaction, and have produced some encouraging evidence of potential effectiveness. However, most have only been tested on children and young people with autism who are more cognitively and verbally able and some have entailed at least some methodological limitations, such as inadequate participant allocation, lack of blinding procedures or not using validated outcome measures.

Given the current need to develop more effective treatment approaches available to a wider population of people with autism, we suggest that there is value in further exploring the contribution of drama techniques. Developmental psychologists and educationalists have long held the view that dramatic play can help people learn how to read others' beliefs and intentions within a safe, structured and reinforcing environment (Guss, 2005; Vygotsky, 1987; Gupta, 2009).

The current evidence-base for drama techniques in autism is mostly a mixture of anecdote and qualitative feedback from parents and teachers who have participated in single-case or small-group studies. (Laugeson, Frankel, Mogil and Dillon, 2009; Lerner, Mikami and Levine, 2011; Godfrey and Haythorne, 2013; O'Sullivan, 2015; Trowdale

and Hayhow, 2015). Nevertheless, preliminary results give reason to continue empirical enquiry, and recent research is beginning to elucidate how efficacious drama interventions can be when coupled with a robust research design (Guli, Semrud-Clikeman, Lerner and Britton 2013; Lerner and Mikami, 2012; Corbett et al., 2011; 2014; 2016). For example, Corbett et al. (2011 and 2014) evaluated the Social Emotional NeuroScience Endocrinology (SENSE) Theatre, a programme designed to improve socioemotional functioning and reduce stress in children with ASD, using live and video peer modelling that took place through the medium of a musical theatrical performance. They reported improvements in face identification, theory of mind, social awareness, and interaction with familiar peers as well as a decrease in cortisol levels indicating reduced stress. However, there were no changes in parent or occupational therapist ratings. Corbett et al. (2016) added to the evidence base for SENSE with a randomised control trial which found treatment effects for cognitively and verbally able children with ASD on social ability, communication symptoms, group play with toys in the company of peers, memory of faces and theory of mind. The treatment effects were maintained for communication symptoms at 2 months follow-up.

Guli et al. (2013) also focused on social competence and found significant improvements in observed social behaviour in a natural setting following participation in the Social Competence Intervention Program (SCIP). SCIP blends drama-based techniques with more traditional behavioural techniques in a 16 session programme for children with autism aged 8 to 14. It includes substantial discussion and reflective exercises as well the requirement for students to keep a diary. As such, it is primarily accessed by those with higher functioning autism.

Finally, Lerner and Mikami (2012) in a small scale, randomised controlled trial, compared a traditional social knowledge training intervention (Skillstreaming) and a social performance training intervention (Sociodramatic Affective Relational Intervention, SDARI) delivered over 1 day for 4 weeks. While both groups increased in reciprocated friendship nominations and staff-reported social skills, individuals in the SDARI group were faster to both like and interact with one another. However by the end of the four weeks, group differences had disappeared and parents continued to report no change in social functioning at home.

The distinction made by Lerner and Mikami between social knowledge training and social performance training is an important one in the current context. All of the drama-based intervention studies used drama or theatre as the medium by which the skills were taught by combining a skills teaching approach (modelling, instruction, drawing, discussion) with role play and other performance related activities. Although the Corbett et al. (2011, 2014, 2016) intervention was carried out in the theatre with the end aim of producing a musical performance, the intervention was nonetheless set up to teach particular skills, primarily through peer- and video-modelling and reinforcement schedules. Two of the recognised limitations of traditional social skills- or knowledge-based training approaches are difficulties with generalisation to other contexts and the fact that the behaviour or skill does not necessarily become part of the person's natural repertoire but rather remains rule-governed and context dependent, or at least becomes natural much more slowly (Rao, Beidel & Murray, 2008; Lerner and Mikami, 2012). By contrast, the aim of performance-based approaches such as SDARI is to give the person the opportunity to experience 'doing the behaviour' in as natural a way as possible in a

controlled and safe environment; an approach that appears to support the development of behaviours in a more naturalistic way (Lerner and Mikami, 2012).

The present intervention, termed '[name removed for blind review], seeks to build on the growing evidence-base for performance-based approaches. The aim of [the intervention] is not to teach the children skills per se but to draw out relevant behaviours and support their development in a play-based environment, allowing the child to initiate and lead the action as much as possible. The targets for improvement extend beyond social skills and interaction to communication and imagination. Peter (2003) outlines, from a practitioner perspective, some of the theoretical basis for drama as an intervention in autism, a basis that is also demonstrated by Shaughnessy (2012) in the parallels between the triad of concepts in drama - *Imagination* through the construction of a fictional or alternative reality, *Communication* through the dialogue between performer(s) and audience; and *Interaction* in the physical engagement between performer(s) and audience - and the triad of impairments traditionally described in autism (*imagination, communication, social interaction,* Wing and Gould *1979*). As such, Imagining Autism is designed to elicit and support the child's communication, social interaction and imagination, in particular though techniques of interactive performance practices which are experiential, physical and immersive in nature.

Imagining Autism is school-based and takes place within an enclosed, indoor tent (or pod) that provides an multi-sensory, themed environment. The 'environments' (forest, arctic, outer space, under water, under the city), were designed to facilitate communication (verbal and physical), social interaction (with practitioners and peers), imagination (participating in fictional frameworks) and creativity (through

improvisation). Working in conjunction with performers, autistic participants (in groups of 4) encounter a range of stimuli, triggers and responsive technologies to include physical action, puppetry, lighting, sound, costume and masks, digital media (e.g. live feed) and responsive technologies. Although practitioners work to a rough script during each session, they very much follow the lead of the children. These interactive sessions are intended to transport the child into an 'alternative' physical reality which engages their imagination and facilitates communication by providing a stimulating and unique environment in which they can share and direct a narrative, and safely explore the social consequences of their actions.

The present evaluation centred on a pre- vs. post-intervention comparison in which all children were entered into a single-tier design. That is, there was no 'dose' manipulation or control group – at this stage we just wanted to determine the feasibility and best design of a larger, effectiveness study and measure if any behavioural change could be observed over the duration of study that might signal therapeutic potential. Prior to study, it was difficult to predict how easily and accurately the intervention could be both administered and assessed. Following the checklist for assessing feasibility proposed by Thabane et al. (2010) we therefore assessed the case for, and nature of, future study along four dimensions: (1) process, which addressed recruitment and retention, missing data, assessor blinding, inter-rater reliability, willingness of children to engage, (2) resources, which addressed the time and human/physical infrastructure needed to implement and sustain the study, (3) management, which addressed dealing with unexpected changes in the intervention and ease of assessment and (4) scientific, which addressed the data outcomes and potential effect.

Methods

Participant selection. To be included, children had to be aged between 7 and 12 years old at the start of the intervention, with a diagnosis of autism recorded on their school record and checked via a background questionnaire completed by parents. Children also had to fall within the cut-off for an ASD diagnosis during administration of the Autism Diagnostic Observational Schedule (ADOS) at study baseline. Children were excluded if they were reported by the school to have a severe physical or sensory disability or if either they or a parent expressed unwillingness to participate and/or the latter failed to provide written informed consent. See Table 1 for participant characteristics.

The target enrolment figure of 22 was determined by both resource-limitations and to enable properly-powered estimates of potential effect on the primary outcome measures employed, the ADOS and Vineland Adaptive Behaviour Scale (VABS). Based on the data reported by Gutstein et al. (2007), a sample size of at least 18 was deemed sufficient to detect clinically significant changes (with power set at 0.8 $p<0.01$) in both the communication and reciprocal social interaction domains of the ADOS. Additional sample size calculations based on pilot work by Beadle-Brown, Murphy and Dorey (2004) indicated that a sample size of 20 would be sufficient to pick up change over time on the VABS, with a medium to large effect size with power set at 0.8 (i.e., the recommended level for preliminary studies of this nature; Lee, Whitehead, Jacques and Julious, 2014).

The intervention. After an initial play based session to meet the practitioners and some of the 'characters' to be encountered in the sessions, the children participated in a 45 minute session, each week, for 10 weeks as part of a group of 3-4 children[1]. Across the ten weeks, each child experienced each of the five environments (space, under the sea, under the city, in the forest, and arctic) twice. The order in which they experienced each environment was determined by the practitioners. However, all children experienced all five environments once (before the half term break) and then experienced all five environments a second time. The immersive environment or 'pod' was brought to the school each week and erected in the school's hall. The practitioners worked to a rough narrative structure, a loose scenario, often based on a journey (e.g. to the moon, boat trip) with opportunities for children to lead the action as it developed.

There were usually 4-5 practitioners in the pod for each session, in character, as the technologist or operating video camera for documentation, so there was a practitioner child ratio of at least 1:1. The practitioners were recruited specifically for the project and were already experienced in the techniques to be used due to their previous professional experience or their current participation in a postgraduate theatre and performance programme at their university. All practitioners were trained by the 4th and 5th authors in the importance of play,

[1] Most children knew at least some of the other children in their group and the membership of the groups stayed the same throughout the sessions.

turn-taking, liveness (being 'in the moment'), physicality, improvisation, shared attention, imitation, reading non-verbal cues and working as an ensemble. They also had training from the first author on autism and providing autism friendly approaches. Practitioners were *also* trained to be responsive to the children's cues, scaffolding their play to facilitate language, peer interaction and imaginative engagement, as well as mimetics (Trowsdale and Hayhow, 2013).

Children could enter and leave the pod as they wish during the 45minute sessions – if they left, a practitioner would go with them. Most of the time they would come back into the pod of their own accord, however, for a small number of children, some of the interaction with practitioners occurred outside of the pod, usually at the beginning of the sessions.

Outcome Measures.

- Social-communicative, interactive and imaginative skills were measured using the ADOS (Lord et al., 2000). This assessment tool has been used in a number of other autism intervention studies (Gutstein et al., 2007; Howlin et al., 2007; Green et al., 2010). Comparisons were made using the algorithm sub-scale scores i.e. Communication, reciprocal social interaction, creativity and restricted and repetitive behaviour. ADOS raw scores were also transferred to ADOS 2 severity scores using the process described in Gotham et al. (2007, 2009).
- Adaptive behaviour and cognitive functioning were measured using the VABS 2nd Edition, another commonly employed test in autistic populations

(Sparrow et al, 2005). The sub-domain standard scores were the primary point of comparison. The age equivalents for subdomains were also used.

- Emotional expression recognition was measured using a modified version of the Ekman faces task (1993). Faces depicted either happy, sad, angry, disgusted, frightened, surprised or neutral expressions (see Table 6 figure legend for further details). Comparisons were made on the basis of the number of emotions correctly identified across the four subsets of emotions presented.
- Parent and teacher views were measured via interviews at post intervention and by questionnaires[2]. Parents were asked to rate on a five point scale (1 = does not apply to my child to 5 = applies all the time) the impact of the child's autism in terms of their social communication, interaction, imagination and repetitive and stereotypical behaviours. This rating scale was adapted from existing screening measures for autism with the addition of some specific items related to play, imagination and social interaction. The rating scale was piloted and then completed by the parents at three time points: baseline, post-intervention and follow up[3]. Cronbach alpha for the parent questionnaire was found to be above 0.8 or above at all three time

[3] Copies of both the parent and teaching rating scales can be requested from the first author.

points. Teachers and teaching assistants completed a similar measure to the one that parents completed, rating the impact of the child's autism in terms of their social communication, interaction, imagination and repetitive and stereotypical behaviours on a five point rating scale (1 = a major issue to 5 = no problem). The questionnaire was piloted with teachers from different schools and then ratings by teachers were completed at baseline, once a month during the intervention, post intervention and at one or more follow-up assessments[2]. Cronbach's alpha was above 0.8 at all time-points for both teachers and teaching assistants.

Procedure. Following ethical and research governance approval from the university research ethics committee and local education authority, staff contacts at each school were asked to send out project information to parents of children within the target group. Prior to participation, written informed consent was requested from parents for both the intervention and the evaluation. Children's willingness to take part in both the intervention and assessment was carefully monitored and they were able to withdraw from the environment or testing sessions if needed.

Baseline data were scheduled for collection at least once, although the timing of terms in schools 2 and 3 and their early commencement of study meant that 2 or 3 baselines (each separated by 2-3 months) could be captured to give a more stable measure of natural performance. Re-assessments were scheduled for between 4-8 weeks post-intervention and then, if possible at 3, 6 and 12 months.

The evaluation activities for each time point were generally timetabled across three sessions spaced over two weeks. As far as possible, no child had more than one

assessment session per day. Depending on the child's level of focus and interest, each session lasted up to 1 hour. All direct assessments with the children were conducted outside of the classroom in a separate room.

Table 1 here

Study Evaluation

Process – Recruitment: In school 1, 8 children were recruited after parents were contacted, sometimes several times, by the school to follow-up the information letters. In school 2, recruitment was relatively straightforward and eight children were identified and consent gained from parents within a month from having sent out the letters. In school 3, recruitment was much harder – at the scheduled time of their first baseline, only 3 children had been recruited. By second baseline, a further 3 children had been recruited. However to achieve this we had to increase the upper age limit from 11 to 12 years. It was not possible to recruit a further two children from this small school within the project timeframe. Two key problems seemed to underlie this recruitment problem: (1) the school had recently hosted a number of research studies which seemed to reduce the interest of some parents and teachers for further research involvement and, (2) many of the children were residential with the result that some of them had misplaced the forms by the time they were able to take them home at the weekend.

Process - Adherence to test schedule: As shown in Table 2, data from all schools were successfully captured at the key time-points; that is, at least one baseline session, at the post-intervention session, and then at least one longer-term follow-up.

Table 2 here

Process - Missing data: Table 3 summarises the missing data. As can be seen, assessments that were conducted by the researchers were least likely to have incomplete data. Missing data was most apparent at final follow-up, especially when parental involvement was needed. This was better in school 3 where the follow-up period was shorter. Although there was at least one teacher questionnaire returned for each child at each time point, it was difficult to get questionnaires back from teachers and teaching assistants. At some baseline, post intervention and follow-up time points we had asked for 3 questionnaires (one each week) and then once a month during the intervention phase – we rarely received all of the questionnaires back. In retrospect, this may have been quite a heavy demand on teachers even though the questionnaire only took 10-15 minutes to complete. For any subsequent study it will be important to reduce the number of forms given to parents and teachers.

We initially attempted to measure IQ using the *Wechsler Intelligence Scale for Children 4th Edition (WISC-IV, Wechsler 2003)* but its language requirement proved too complex and as such the Weschler Preschool and Primary School Scale of Intelligence 3rd Edition was attempted (WPSSI III, Wechsler, 2002). However, for most children it was not possible to complete all subtests or, where children did complete them, it was not possible to calculate scaled scores and composite scores because the children were chronologically older than the test allowed for calculating the scaled scores. As such, an estimate of IQ was calculated using their test age equivalents on the six domains most commonly completed (Block design, Information, Matrix reasoning, Receptive

vocabulary, Object assembly and Picture naming). Language was assessed using the Clinical Evaluation of Language Fundamentals (CELF) (Semel, Wiig and Secord 2003) but, again, it was often not possible to complete it fully due to the complexity of the items.

Table 3 here

Process - Assessor blinding and inter-rater reliability: The assessors did not see any footage of the children in the pod until after the data had been collected for that school. However, they were aware that children were undergoing the intervention and had a basic understanding of the methods. In addition, staff at school did sometimes talk to the researchers about perceived changes in the children, despite being asked to desist. To strengthen the findings, for all but one child, one baseline ADOS video and the post-intervention ADOS video was coded by someone outside the research team who was very experienced with the ADOS and who was blind to whether the session was a baseline or post intervention session.

Intra-class correlations (absolute agreement) were calculated and were generally found to be acceptable, with an average Interclass correlation co-efficient of 0.75 across the three modules (range 0.56 to 0.915). The average Intraclass Correlation Coefficient across each of the four algorithm subscales (communication, reciprocal social interactions, creativity and sensory and repetitive behaviour) was 0.73 for absolute agreement (range 0.44 – 0.92) and 0.77 for consistency (range 0.70 – 0.92).

Willingness of children to engage: All but one child, who was discovered to have a much more severe visual impairment than had previously been recognised, took part in all intervention sessions. A small number of children initially interacted with practitioners outside of the pod but gradually chose to spend more and more time in the pod as the time went on.

Thematic analysis of post-intervention interviews with teachers and parents identified the willingness of children to engage and enjoyment of the session as a key theme. Teachers described how keen the children were to go to the sessions and how they continued to request the sessions once they had stopped. For example, one teacher commented *' he always came out smiling, always wanted to go, was happy to go.. so he must have totally enjoyed it and had fun.'*

Parents too reported that the children had enjoyed the sessions, e.g.

> '*she loved it, she really did…just the fact that she told me about it, ... she would talk in the morning about "the ladies are gonna come today and we are gonna have fun and we are gonna do this or that or the other"'.*

Analysis of the reports completed by practitioners at the end of each session also indicated, for the most part, that children appeared to willingly engage in the sessions and that for many children engagement within the sessions increased over time.

Resources – space and time: Space within most schools was limited and finding a space big enough that was not used for other purposes (e.g. as a dining hall) was difficult in these small special schools. This meant that intervention sessions had to

fit in around other activities and lunchtimes. There had to be enough space for the pod, and for the pod to be large enough for practitioners to work with 3-4 children at a time; working with a group was felt to be a defining ingredient of the intervention. It was also felt important that having the pod in a familiar environment was helpful for children to understand that they were "pretending" – i.e. that they knew that in reality they were still in the school dining hall but for the 45 minutes of the session they were pretending to be on the moon.

The pod itself and all the equipment took a long time to assemble and take down which restricted the length of time children could be in the pod and the number of groups that could be run. Also, it required substantial storage space and required transportation to each setting as in most schools it was not possible to store it from one week to the next. A much more portable version of the pod, with more easily erected lighting and fixed cameras, would greatly improve the ease of conducting the sessions. A greater use of cameras would also increase the consistency of data capture.

Management: By necessity the intervention required a holistic and flexible approach. To fit in with seasons of the year or events that were happening at the time of the intervention, practitioners had to adapt the order of the different scenarios. They had to respond flexibly to the lead of the child rather than working to a script set in stone at the beginning. The environments also needed to be physically changeable – whilst the environments were multi-sensory, they had to be able to adapt to different levels of tolerance and able to respond to both hypersensitivities and hypo-sensitivities in the children. As such it was important that sound could be turned up or down and that lighting could be varied depending

on the child that was in the environment. Being responsive to individuals and their sensory difficulties and preferences was essential. Some children responded much more to visual elements of the environments while others engaged more actively with the auditory elements. Materials that could be used both imaginatively and functionally were needed and there had to be the possibility of primarily non-verbal as well as verbal interactions.

Ease of child assessment: To deal with limitations in the space available and to keep some children engaged it was necessary to present subtests in different orders and change the toys or books available during the ADOS. Sometimes the assessments had to be conducted over shorter sessions or conducted on the floor rather than a table.

Scientific evaluation (data outcomes and statistical analyses)

Trends at the group level were interrogated using Wilcoxon signed-rank and Friedman tests.

To estimate significant change for individual children, exploratory single case analyses were also conducted. We adopted the recommendation made by a number of trial statisticians to use lower statistical estimation thresholds for pilot studies of this nature (Kianifard and Islam, 2011; Lee, Whitehead, Jacques and Julious 2014; Schoenfeld, 1980; Stallard, 2012). Specifically, Lee and colleagues propose combining confidence intervals with a statistical threshold of 75%-80% on the basis that, by virtue of likely being under-powered, pilot studies should be more about learning than confirming. They suggest that the emphasis should be on

showing effects in the predicted direction rather than definitively testing hypotheses. To this end, we adopted the procedure described in Wilkinson et al (2013) in which confidence intervals were first calculated from an allied population where available[4] or from the whole sample of IA participants and then placed around the relevant baseline score for each individual. A statistically significant change from baseline was inferred if the post-intervention score fell outside the 80% confidence interval.

Tables 4, 5, and 6 illustrate the descriptive analysis on the ADOS, VABS and emotion expression task.

On the Communication subdomain of the ADOS, there was no statistically significant difference between the baseline and post-intervention scores ($z=1.007$, $p=.314$, $d = 0.16$). However, confidence interval analysis indicated that four children showed significantly decreased scores (i.e. improvement) from baseline on the communication sub-domain.

---

[4] Standard errors for each ADOS module were obtained from the Gutstein et al. (2007) evaluation of the Relational Development Intervention. This was because the standard errors for that sample fell between the higher values found in Zachor et al. (2007) study and the lower SEs found in the Howlin et al. (2007) study and thus were felt to represent a middle ground in terms of published comparisons. They were also higher than the [name of intervention] SEs and thus represented a more conservative measure. The sample size for each ADOS module was also greater than the current sample, with 12 children completing each module. Finally, the Gutstein intervention was also the one that was slightly closer in nature to the current intervention.

Although there was a trend towards decreasing scores on the reciprocal social interaction domain, this did not reach significance (z = .664, p=.057, d=0.26). However, there was a statistically significant reduction for those children completing Module 3 (z = 1.973, p=.049, Cohen's d = 1.96). Confidence interval analysis illustrated that 7 individual children (2 from school 1, 4 from school 2 and 1 from School 3) showed significant reductions between baseline and post-intervention.

There were no significant group changes in the creativity/play domain (z=.863 p=.405, d=0.08).

The ADOS 2 severity score was calculated using Gotham et al. (2007, 2009). The resulting severity scores are presented in Table 4. Although the trend towards a reduction over time across the whole sample did not meet statistical significance (Friedman non-parametric related measures ANOVA X^2 = 5.833, df 2, p=0.054), for those doing Module 3 the reduction in severity over time was significant (X^2 = 6.643, df 2, p=0.036). Post-hoc analysis using Wilcoxon signed ranks test found that the difference between Follow-up and Baseline was significant (z=2.20, p=0.028, n=8) but with the biggest change between baseline and post-intervention (Z=1.70, n=8, p=0.089 compared to z=0.679, p=0.497 for post-intervention to follow-up).

Finally, it is worth noting that although the majority of children changed within the Autism severity band (severity scores of 6 to 10), two children changed severity category at post-intervention from Autistic to not on the spectrum. For one of these children this change was maintained. For the other the score increased from 3 to 4 at follow-up but did not return to the original severity score of 7. For two other children, one of whom had stayed the same and one of which decreased slightly (but stayed

within the autism band) between baseline and post-intervention, their follow-up scores had decreased to within the non-autistic band.

On the VABS (See Table 5) there was a statistically significant increase in scores on the Communication domain (Wilcoxon signed ranks test $z = 3.825$, $p<0.001$, n=19, Cohen's d=6.07) and on the Socialisation domain scores at post-intervention compared to baseline ($z = 3.180$ $p<0.001$, n=19, Cohen's d=3.42).

On the Emotion recognition task (see Table 6) the increase in number of emotions correctly identified between baseline and final follow-up was statistically significant (Friedman $X^2 = 8.645$ df = 2 p = .013, d=2.12). Although a similar trend was apparent between baseline and post-intervention, it was not statistically significant (z=1.355, p=.175, d=0.82).

Tables 4, 5 and 6 here

Finally, ratings from both parents and teachers of autistic severity showed favourable changes relative to those recorded at baseline – see Tables 7 and 8.

Tables 7 and 8 here

Discussion

In this study we set out to establish the feasibility of implementing and evaluating "Imagining Autism", which aimed to develop the social interaction, communication and imagination of young people with autism through drama-based methods. We also sought to establish whether there was any preliminary evidence that children benefited

from Imagining Autism in the form of changes in social, communicative or imaginative skills – evidence that would warrant further investigation. The key finding was that, despite some practical issues around space, storage and portability of the pod, it was possible to successfully provide the intervention in three different schools, involving a range of children with different needs and abilities. Not only was it possible to do, but teachers and parents reported that the children expressed enjoyment in the sessions, looked forward to the sessions and expressed disappointment when they had finished. Even children who did not normally like to transition to new environments or sessions over time went willingly to the pod. Participant drop-out post-enrolment was low, as was the frequency of missing data with the exception of the IQ tests administered at baseline and the parent questionnaires administered at final follow-up.

Good practice would normally suggest that children with autism benefit most from a low arousal environment, one that takes account of their sensory sensitivities. For this reason, their tolerance of our highly arousing, multi-sensory environment of the pod was far from guaranteed. However, despite being exposed to this multi-sensory environment, the children with autism responded favourably to it, evidenced by an eagerness to enter the pod, active participation in the sessions and limited overt expressions of anxiety. It is likely that the child-led nature of the sessions contributed to this success; children were able to move in and out of the pod if they needed and also often had some control over some of the levels of noise/light, and the nature of the interactions with practitioners and characters. Their knowledge that the session was limited to 45minutes may also have made them more tolerable of the heightened sensory stimulation.

One current study limitation that will not affect future studies is the relatively frequent and repeated administration of the ADOS which may have induced practice/learning effects. Although the ADOS has been used in many intervention evaluations, it has usually been applied with a longer delay between fewer administrations (e.g. Gutstein et al., 2007; Howlin et al., 2007; Green et al., 2010). In line with this more typical procedure, the use of a separate control group in the next study will reduce the number of times that any individual performs the test, and follow-up will occur later to better assess clinical relevance.

Although preliminary in nature, the study did show statistical evidence of cognitive benefit. Evidence of change at the group level was observed on one ADOS domain, ADOS2 severity scale, the Vineland, and the Emotion Recognition test. This was accompanied by favourable changes on the parent and teacher observations. For some children, improvement was only evident from the parental accounts so capturing home-based behaviours will be important in any future study.

As is commonly the case for proof-of-principle and feasibility studies at the present stage, we believe that the next study should adopt a randomised, controlled design in which the potential effects (and allied confidence intervals) of the intervention are compared to a group of age-, gender- and severity-matched children who are not exposed to the pod. There are many aspects to the pod that may contribute to positive effect (e.g. multi-media, narrative structure, interaction with the practitioners) which makes it difficult to reduce the effects to a core set of elements. Rather than administering multiple versions of the same intervention, a more efficient means of inference may be to compare the intervention to those others currently under development and which have shown particular promise. One possibility currently being

explored is to compare Imagining Autism to the Skillstreaming and SDARI interventions evaluated by Lerner and Mikami (2012). We note that efficacy pilot studies of this nature, in which there is little prior evidence with which to accurately determine the sample size needed to show a potential between group treatment effect with corresponding 80% confidence intervals should follow the 'rule of 12' in which at least 12 participants are allocated to each group (Julious, 2005; Lakens and Evers, 2014; Moore, Carter, Nietart and Stewart, 2005). Downstream study can then more accurately estimate optimum sample size and determine, for example, whether meaningful effects will require the 60 participants-per-arm that were recruited to the recent Pickles et al. (2016) study.

Before moving to the next stage of development, several other methodological changes will be needed: (1) at study debrief, some parents and teachers told us that the questionnaires would be easier to complete on-line rather than on paper; (2) teachers also told us that participant sign-up would be quicker if we held a recruitment event in each school so that parents and children could come to 'try out' the pod for themselves; (3) given the need to maximise the consistency with which practitioners made their observations, it will be important to retain the same practitioners throughout the study (several staff replacements had to be made) and install remote-controlled, rather than hand held or static, video-cameras within the pod to properly monitor activity; (4) although data collection was adequate for the main study outcomes, too few children were able to complete the WISC and WPSSI. It may be more sensible to use the KBIT2 which is briefer and partly non-verbal in nature; and (5) We noted that the children who seemed to show greatest ADOS change were those who were higher functioning. It is possible that those who are more severely affected by their autism need more frequent

and longer exposure to reap the same benefits. Further research will need to consider the cost/benefit ratio of having more weekly sessions and/or a longer intervention phase.

In sum, despite the methodological and inferential challenges identified above, we believe that Imagining Autism shows sufficient proof of feasibility and enough preliminary evidence of impact to progress to larger-scale validation. To gain momentum and widespread acceptance, drama-based interventions for autism must hold appeal beyond those within the field of drama-practice to teachers, educationalists, psychologists, clinicians and caregivers. By employing an inter-disciplinary team that draws from psychology and the performing arts, and which has continually sought to engage with stakeholders throughout the local community, we hope that Imagining Autism will fulfil this key criterion.

References

Ayres A (1972) *Sensory integration and learning disabilities.* Western Psychological Services: Los Angeles.

Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, Meldrum D and Charman T (2006) Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet* 368 (9531):210-215.

Beadle-Brown J, Murphy G, and Dorey H (2004) Evaluation of early intervention in autism: A pilot study. *Journal of Intellectual Disability Research* 48 (4-5): 328.

Mills R and Marchant S (2011) Intervention in autism: a brief review of the literature. *Tizard Learning Disability Review* 16 (4): 20-35.

Bowers K Lin P and Erickson C (2014) Pharmacogenomic Medicine in Autism: Challenges and Opportunities. *Pediatric Drugs*: 1-10

Case-Smith J, Weaver LL and Fristad MA (2015) A systematic review of sensory processing interventions for children with autism spectrum disorders *Autism: the international journal of research and practice* 19 (2): 133-148.

Corbett, B. A. (2003). Video modeling: A window into the world of autism. *The Behavior Analyst Today*, 4(3), 367–377.

Corbett, B. A., & Abdullah, M. (2005). Video Modeling: Why does it work for children with autism? *Journal of Early and Intensive Behavior Intervention*, 2(1), 2–8.

Corbett BA, Gunther JR, Comins D, Price J, Ryan N, Simon D, Schupp CW and Rios T (2011) Brief report: theatre as therapy for children with autism spectrum disorder.*Journal of Autism and Developmental Disorders* 41 (4): 505-511.

Corbett BA, Swain DM, Coke C, Simon D, Newsom C, Houchins-Juarez N, Jenson A, Wang L and Song Y (2014) Improvement in Social Deficits in Autism Spectrum Disorders Using a Theatre-Based Peer-Mediated Intervention. *Autism Research* 7 (1): 4-16.

Corbett BA, Key AP, Qualls L, Fecteau S, Newsom C, Coke C & Yoder P. (2016) Improvement in Social Competence Using a Randomized Trial of a Theatre Intervention for Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 46(2):658-72.

Dow AW, Leong D, Anderson A and Wenzel RP (2007) Using theater to teach clinical empathy: a pilot study. *Journal of general internal medicine* 22 (8): 1114-1118.

Eikeseth S (2009) Outcome of comprehensive psycho-educational interventions for young children with autism. *Research in developmental disabilities* 30 (1): 158-178.

Eikeseth S. et al. (2012) Outcome for children with autism receiving early and intensive behavioral intervention in mainstream preschool and kindergarten settings. Research in Autism Spectrum Disorders. 6(2), pp. 829-835

Ekman P (1993) Facial expression and emotion. *American psychologist* 48 (4): 384.

Eldevik S, Hastings RP, Hughes JC, Jahr E, Eikeseth S and Cross S (2009) Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child and Adolescent Psychology* 38 (3): 439-450.

Fletcher-Watson S, McConnell F, Manola E and McConachie H (2014) Interventions based on the Theory of Mind cognitive model for autism spectrum disorder (ASD). *The Cochrane Library.*

Frost, L., & Bondy, A.(2002) The Picture Exchange Communication System training manual, 2nd ed. Pyramid Educational Consultants, Inc.

Godfrey E and Haythorne D (2013) Benefits of dramatherapy for Autism Spectrum Disorder: a qualitative analysis of feedback from parents and teachers of clients attending Roundabout dramatherapy sessions in schools. *Dramatherapy* 35 (1): 20-28.

Goin-Kochel RP, Myers BJ and Mackintosh VH (2007) Parental reports on the use of treatments and therapies for children with autism spectrum disorders. *Research in Autism Spectrum Disorders* 1 (3): 195-209.

Goldstein TR and Winner E (2012) Sympathy for a Character's Plight: Sex Differences in Response to Theatre. *Empirical Studies of the Arts* 30 (2): 129-141.

Goldstein TR, Lerner, M and Winner E (in press) The Arts as a Venue for Developmental Science: Realizing a Latent Opportunity. *Child Development*

Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS Scores for a Measure of Severity in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 39(5), 693-705.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised Algorithms for Improved Diagnostic Validity. *Journal of Autism & Developmental Disorders*, 37(4), 613-627.

Green VA, Pituch KA, Itchon J, Choi A, O'Reilly M and Sigafoos J (2006) Internet survey of treatments used by parents of children with autism. *Research in developmental disabilities,* 27 (1): 70-84.

Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., … the PACT Consortium. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. Lancet, 375(9732), 2152–2160. http://doi.org/10.1016/S0140-6736(10)60587-9

Guli LA, Semrud-Clikeman M, Lerner MD and Britton N (2013) Social Competence Intervention Program (SCIP): A pilot study of a creative drama program for youth with social difficulties. *The Arts in Psychotherapy* 40 (1): 37-44.

Gupta A (2009) Vygotskian perspectives on using dramatic play to enhance children's development and balance creativity with structure in the early childhood classroom. *Early Child Development and Care* 179 (8): 1041-1054.

Guss F (2005) Reconceptualizing play: Aesthetic self-definitions. *Contemporary Issues in Early Childhood* 6 (3): 233-243.

Gutstein SE, Burgess AF and Montfort K (2007) Evaluation of the relationship development intervention program. *Autism: the international journal of research and practice* 11 (5): 397-411.

Howlin P, Gordon R K, Pasco G, Wade A and Charman T (2007) The effectiveness of Picture Exchange Communication System (PECS) training for teachers of children with autism: a pragmatic group randomised controlled trial. *Journal of Child Psychology and Psychiatry* 48(5): 473-48.

Julious, S. A. (2005). Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics*, 4(4), 287-291.

Kasari C, Gulsrud AC, Wong C, Kwon S, Locke J. (2012) Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. *Journal of Autism and Developmental Disorders*, 40(9):1045-56.

Kianifard, F., & Islam, M. Z. (2011). A guide to the design and analysis of small clinical studies. *Pharmaceutical Statistics*, 10(4), 363-368.

Lakens, D., & Evers, E. R. K. (2014). Sailing From the Seas of Chaos Into the Corridor of Stability: Practical Recommendations to Increase the Informational Value of Studies. *Perspectives on Psychological Science*, 9(3), 278-292.

Laugeson EA, Frankel F, Mogil C and Dillon AR (2009) Parent-assisted social skills training to improve friendships in teens with autism spectrum disorders. *Journal of Autism and Developmental Disorders,* 39 (4): 596-606.

Laugeson, E. A., Gantman, A., Kapp, S. K., Orenski, K., & Ellingsen, R. (2015). A Randomized Controlled Trial to Improve Social Skills in Young Adults with Autism Spectrum Disorder: The UCLA PEERS® Program. *Journal of Autism and Developmental Disorders*, 45(12), 3978-3989. doi:10.1007/s10803-015-2504-8

Lee EC Whitehead AL Jacques RM and Julious SA (2014) The statistical interpretation of pilot trials: should significance thresholds be reconsidered? *BMC medical research methodology* 14 (1): 41.

Lerner MD and Mikami AY (2012) A preliminary randomized controlled trial of two social skills interventions for youth with high-functioning autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities* 27: 145-155.

Lerner MD, Mikami AY and Levine K (2011) Socio-dramatic affective-relational intervention for adolescents with asperger syndrome and high functioning autism: pilot study. *Autism : the international journal of research and practice* 15 (1): 21-42.

Lord C, Risi S, Lambrecht L, Cook Jr EH, Leventhal BL, DiLavore PC, Pickles A and Rutter M 2000 The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders* 30 (3): 205-223.

Lovaas O. I. (1987) Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*. 55 pp. 3-9

Lubar K and Halpern BL (2004) *Leadership presence* Penguin

Matson JL, Matson ML and Rivet TT (2007) Social-skills treatments for children with autism spectrum disorders: an overview. *Behavior modification* 31 (5): 682-707.

McEachin J. J. et al. (1993) Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal of Mental Retardation.* 97(4), 359-372

McConnell SR (2002) Interventions to facilitate social interaction for young children with autism: Review of available research and recommendations for educational intervention and future research. *Journal of Autism and Developmental Disorders* 32 (5): 351-372.

Mohiuddin S and Ghaziuddin M (2013) Psychopharmacology of autism spectrum disorders: a selective review. *Autism : the international journal of research and practice* 17 (6): 645-654.

Moore, C. G., Carter, R. E., Nietert, P. J., & Stewart, P. W. (2011). Recommendations for planning pilot studies in clinical and translational research. *Clinical Translational Science*, 4(5), 332-337.

Ospina MB, Seida JK, Clark B, Karkhaneh M, Hartling L, Tjosvold L, Vandermeer B and Smith V (2008) Behavioural and developmental interventions for autism spectrum disorder: a clinical systematic review. *PloS one* 3 (11): e3755

O'Sullivan C (2015) Drama and Autism. In: R F Volkmar (Ed) *Encyclopedia of Autism Spectrum Disorders* New York NY: Springer New York pp.1-13.

Peter M (2003) Drama narrative and early learning. *British Journal of Special Education* 30 (1): 21-27

Pickles, A., Le Couteur, A., Leadbitter, K., Salomone, E., Cole-Fletcher, R., Tobin, H., . . . Green, J. Parent-mediated social communication therapy for young children with autism (PACT): long-term follow-up of a randomised controlled trial. The Lancet. doi:10.1016/S0140-6736(16)31229-6

Rao PA, Beidel DC and Murray MJ (2008) Social skills interventions for children with Asperger's syndrome or high-functioning autism: A review and recommendations. *Journal of Autism and Developmental Disorders* 38 (2): 353-361.

Reichow B, Steiner A and Volkmar F (2012) Social skills groups for people aged 6 to 21 years with autism spectrum disorders (ASD). *Cochrane Database Syst Rev* 7 doi:10.1002/14651858.CD008511.pub2

Rogers SJ and Dawson G. (2009a) *Play and Engagement in Early Autism: The Early Start Denver Model. Volume I: The Treatment*. New York: Guilford Press. Also available in Japanese, Italian, Dutch, French, Spanish, and Arabic.

Rogers SJ and Dawson G. (2009b) *Play and Engagement in Early Autism: The Early Start Denver Model. Volume II: The Curriculum*. New York: Guilford Press. Translated into Japanese, Italian, Dutch, French, Arabic, and Spanish.

Schopler E (2001) *Treatment for autism.* Springer

Seida JK, Ospina MB, Karkhaneh M, Hartling L, Smith V and Clark B (2009) Systematic reviews of psychosocial interventions for autism: an umbrella review. *Developmental Medicine and Child Neurology* 51 (2): 95-104.

Semel E, Wiig E and Secord W (2003) *Clinical evaluation of language fundamentals-IV* Marickville: Harcourt Assessment. Shaughnessy N (2012) Knowing Me Knowing You: Cognition, Kinesthetic Empathy and Applied Performance. In: Reynolds D and Reason M (eds) *Kinesthetic Empathy in Creative and Cultural Practices* Bristol: Intellect pp. 33-50.

Shukla-Mehta, S., Miller, T., & Callahan, K. J. (2010). Evaluating the Effectiveness of Video Instruction on Social and Communication Skills Training for Children with Autism Spectrum Disorders: A Review of the Literature. *Focus on Autism and Other Developmental Disabilities*, 25(1), 23-36.

Soorya LV, Siper PM, Beck T, Soffes S, Halpern D, Gorenstein M, Kolevzon A, Buxbaum J, Wang AT (2015) Randomized comparative trial of a social cognitive skills group for children with autism spectrum disorder. *Journal of American Academy of Child and Adolescent Psychiatry*. 2015, 54(3):208-216

Sparrow SS, Balla DA, Cicchetti DV And Doll EA (2005) *Vineland-II adaptive behavior scales*. Circle Pines, Minn: AGS Publishing.

Stallard, N. (2012). Optimal sample sizes for phase II clinical trials and pilot studies. Statistical Medicine, 31(11-12), 1031-1042.

Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, Robson R, Thabane M, Giangregorio L and Goldsmith CH (2010) A tutorial on pilot studies: the what why and how. *BMC Medical Research Methodology* 10 (1): 1

The Guardian (2015) Autism: does ABA therapy open society's doors to children, or impose conformity? Article written by Sydney Parker, 20th March 2015. https://www.theguardian.com/society/2015/mar/20/autism-does-aba-therapy-open-societys-doors-to-children-or-impose-conformity

Trowsdale, J. and Hayhow, R. (2013), Can mimetics, a theatre-based practice, open possibilities for young people with learning disabilities? A capability approach. British Journal of Special Education, 40: 72–79.

Trowsdale J and Hayhow R (2015) Psycho-physical theatre practice as embodied learning for young people with learning disabilities. *International Journal of Inclusive Education* 19 (10) Published online 22nd April 2015 DOI:101080/1360311620151031832.

Virués-Ortega J (2010) Applied behaviour analytic intervention for autism in early childhood: Meta-analysis meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical psychology review* 30 4: 387-399.

Vygotsky, L. S. (1978). Interaction between learning and development (M. Lopez-Morillas, Trans.). In: M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological processes* Cambridge, MA: Harvard University pp. 79-91.

Warren Z, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A and Veenstra-Vanderweele J (2011) A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics* 127 (5): e1303-11.

Wang, S.-Y., Cui, Y., & Parrila, R. (2011). Examining the effectiveness of peer-mediated and video-modeling social skills interventions for children with autism spectrum disorders: A meta-analysis in single-case research using HLM. *Research in Autism Spectrum Disorders*, 5(1), 562-569.

Wechsler D (2002) WPPSI-III technical and interpretive manual. *The Psychological Corporation.*

Wechsler D (2003) Wechsler intelligence scale for children–Fourth Edition (WISC-IV). *San Antonio TX: The Psychological Corporation*

Wilkinson D, Morris R, Milberg W and Sakel M. (2013). Caloric vestibular stimulation in aphasic syndrome. *Frontiers in Integrative Neuroscience* 7 (Article 19). http://dx.doi.org/10.3389/fnint.2013.00099

Wing L and Gould J (1979) Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders* 9 (1): 11-29.

Zachor DA, Ben-Itzchak E, Rabinovich A and Lahat E (2007) Change in autism core symptoms with intervention *Research in autism spectrum disorders* 1 (4): 304-317.

	School 1 (n=8)	School 2 (n=8)	School 3 (n=6)	Overall (N=22)
Gender (number male)	6	6	6	18
Age (years)	9	8.9	9.7	9.16 (7.33-12.92)
Age equivalent on Vineland (means and range for sample overall)				
Receptive communication	1.32	2.48	1.89	1.90 (0.05-3.11)
Expressive communication	1.67	2.8	1.88	2.14 (0.01-4.04)
Socialisation-interpersonal	1.55	2.70	2.22	2.15 (0.03-7.09)
Socialisation play and leisure	1.04	2.56	1.55	1.73 (0.01-6.08)
Socialisation- coping	2.09	3.18	2.39	2.57 (0.03-5.10)
Wechsler Intelligence scale for children (mean index scores and range)				
Verbal Comprehension		78 (50-91)		
Perceptual reasoning		82 (67-94)		
Working Memory		75 (65-99)		
Processing Speed		75 (62-97)		
Full-scale		74 (63-87)		
Estimated IQ from WPSSI test age equivalents	43 (31-58)		39 (29-50)	
Clinical Evaluation of Language Fundamentals (CELF), Semel et al., 2003: Range (SD) and range (max 120 with high score = more problems)				
Score on parent observational measure at baseline	54 (32) 17 – 94 n=6	72 (15) 52-97	78 (27) 34-106	68 (26) 17-106 n=20
Score on teachers observational measure at baseline	67 (28) 25-11	60 (16) 35 – 7	59 (29) 15 – 91	62 (24) 15 – 111

Table 1. Characteristics of the children at baseline. School 1 is a special school for children with mixed disabilities, School 2 is a special school for higher functioning

children with autism, and School 3 is a special school for children across the autism spectrum.

	B1	B2	B3	PI	FU1	FU2	Final FU
School 1 (Intervention Spring 2012) N=8	✓	-	-	✓	✓, at 3 months	✓, at 6 months	✓, at 12 months

School 2 (Intervention Summer 2012) N=8	✓	✓	-	✓	✓, at 3 months	✓, at 6 months *	✓, at 9 months
School 3 (Intervention Autumn 2012) N=6	✓**	✓***	✓***	✓	-	-	✓, at 5 months

*Not ADOS **for 3 children ***for all 6 children

Table 2. Summary of evaluation time points by school

Abbreviations: = evaluation occurred at this time point. B1 = first baseline session (month 0), B2 = second baseline session (month 2-3), B3 = third baseline session (month 4-5), PI = post-intervention (4-8 weeks after intervention), FU1 = first follow-up (3 months after intervention), FU2 = second follow-up (6 months after intervention), Final FU = final follow-up (6 months after intervention). Note that the delay between the last baseline session and start of the 10 week intervention programme was always the same regardless of the number of baseline sessions administered.

Evaluation element	**Missing data**
Post intervention data (4-8 weeks post intervention)	VABS – 2 children Parent questionnaire – 3 children
Final follow-up data (at least 5 months post intervention)	VABS – 15 children Parent questionnaire – 14 children Teacher questionnaire – 6 children Emotion recognition task – 6 children
WISC at baseline	13 children
WPSSI (used when WISC not possible)	9 children attempted but did not complete 2 children completed it but they were older than the top

	age in the conversion tables). 2 children did not attempt any IQ test.
Summary of missed observations	Baseline 1 – 29/1215 (2%) Post-intervention – 39/1314 (3%) Final follow-up – 44/1290 (3%) Across all time points (3 baselines, post-intervention and three follow ups) – 180/6414 (3%)

Table 3. Cases of missing data

ADOS module and domain	Final baseline Average (SD)	Post intervention Average (SD)	Follow-up Average (SD)
Communication			
Module 1 (n=4)	5.25 (0.5)	6.25 (1.26)	7.33 (0.58) (n=3)
Module 2 (n=9)	5.75 (1.12)	6.55 (1.01)	5.89 (1.27)
Module 3 (n=8)	3.62 (2.20)	2.87 (2.70)	2 (1.93)
Reciprocal social interaction			
Module 1 (n=4)	9 (3.74)	10 (4.55)	8 (3.60) (n=3)
Module 2 (n=9)	9.44 (2.24)	10.56 (3.13)	9.44 (2.24)
Module 3 (n=8)	9.13 (2.53)	6.13 (4.32)	4 (4.57)
Creativity/Play			
Module 1(n=4)	3.50 (1.0)	3.25 (1.50)	3.67 (0.58) (n=3)
Module 2(n=9)	0.78 (0.84)	1.33 (0.71)	0.89 (0.60)
Module 3 (n=8)	0.29 (0.49)	0.13 (0.35)	0.75 (1.75)
ADOS 2 Severity scores – Mean (SD), min-max			
Module 1(n=4)	6.5 (2.52), 3-9	7.5 (2.38), 4-9	6.00 (1.73), 4-7 (n=3)[5]
Module 2(n=9)	6.56 (0.82), 6-8	7.33 (1.5), 6–10	6.33 (1.32), 4-9
Module 3 (n=8)	7.63 (2.64), 3-10	5.75 (3.37), 1–10	4.63 (3.07), 1-10
Overall sample	6.95 (1.83), 3-10[6]	6.76 (2.53), 1-10	5.60 (2.28), 1-10

[5] N.B. the child who was not tested at Follow-up had not changed between baseline and post-intervention (severity score of 9). Leaving this child out of the sample at baseline reduces the range to 3-7 (mean 5.67). The range does not change at Post-intervention

Table 4. ADOS scores by module and ADOS 2 severity scores by module and overall sample.

[6] Two children did not quite meet the ASD criteria on the ADOS2 algorithm although they had met the criteria on ADOS and had a clinical diagnosis.

Time point		**School 1 (n=5)**	**School 2 (n=8)**	**School 3 (n=6)**	**Total (n=19)**
Socialisation domain score (Interpersonal relationships, play and Coping skills combined)					
Baseline	Mean (SD) Range	44 (10.4) 27-54	58(6.7) 51-67	49(13.9) 31-72	51(11.7) 27-72
Post-intervention	Mean (SD) Range	60 (10.5) 48-75	65(6.9) 53-75	61(20.5) 40-98	63(12.9) 40-98
Communication domain score (receptive, expressive and written combined)					
Baseline	Mean (SD) Range	39 (9.6) 23-48	51 (7.8) 42-64	40 (13.8) 26-64	44 (11.3) 23-64
Post-Intervention	Mean (SD) Range	61 (11.5) 43-74	71 (10.1) 56-88	61 (15.1) 45-82	65 (12.6) 43-88

Table 5. Scores on the Vineland Adaptive Behaviour Scales Socialisation and Communication domains at baseline and at follow up.

	Baseline Mean (SD)	Post Intervention Mean (SD)	Follow-up Mean (SD)
School 1 (n=4)	7 (7.02)	11.25 (9.14)	15.5(3.79)
School 2 (n=8)	18.5 (4.04)	17.5 (2.39)	20.9 (1.89)
School 3 (n=5)	9.4 (6.23)	14.6 (3.21)	15.25 (2.75) (n=4)
All (n=17)	13.12 (7.36)	15.18 (5.23)	18.13 (3.76) (n=16)

Table 6. Mean number of facial expressions correctly identified on the Ekman photos.

In phase one of the task, seven cards with cartoon/emoticon faces were individually presented to the child and the child was asked "how is this face feeling?" Each child was given 15 seconds to respond. If no answer or a wrong answer was given then the child was given the correct answer. At the end of the first presentation, the incorrectly identified emotions were presented again. If the children correctly identified 4 of the 7 emotions (happy, sad, angry, disgusted, frightened, surprised and neutral) then the real face task was administered. In the real face task (data presented above in Table 6), 2 sets of male faces and 2 sets of female faces from the Ekman Pictures of Facial Affect Series were presented to each child, with a total of 28 trials administered at baseline, post-intervention and follow-up. Which sets of photos each child saw at which time point was determined randomly and the order in which the emotions were presented was determined randomly for each of the four sets of seven photos. Performance was scored as the total number of emotions correctly identified.

	Parental ratings (decrease in score is favourable)	
	Baseline Mean (SD)	Post Intervention Mean (SD)
School 1 (n=4)	49.75 (18.57)	43.5 (23.95)
School 2 (n=8)	57.87 (15.80)	58.87 (14.82)
School 3 (n=6)	52.5 (6.16)	42.67 (5.05)
All (n=18)	54.28 (13.59)	50.06 (17.11)

Table 7. Parental ratings of autism impact/severity at baseline and post-intervention (NB: a *reduction* in score is a *positive* outcome).

	Baseline Mean (SD)	Intervention Mean (SD)	Post Intervention Mean (SD)		Follow-up Mean (SD)	
			Whole sample	Those with baseline data	Whole sample	Those with baseline data
School 1	2.62 (0.73) N=6	3.19 (0.78) (n=7)	3.22 (0.85)	3.21 (0.93)	3.06 (0.85)	3.08 (0.93)
School 2	3.13 (0.50) N=4	3.36 (0.50) (n=8)	3.53 (0.58)	3.08 (0.17)	3.79 (0.68)	3.29 (0.56)
School 3	2.53 (0.81) N=3	2.94 (0.54) N=6	2.70 (0.70)	2.63 (0.86)	2.64 (0.62)	2.48 (0.57)
All	2.76 (0.73) N=13	3.16 (0.62) N=21	3.19 (0.76)	3.04 (0.74)	3.06 (0.85)	3.01 (0.77)

Table 8. Summary of severity of autism ratings by teaching assistants in each school. (NB: an *increase* in score is a *positive outcome*).