How effective is TEACCH as a home programme for children with an autism spectrum disorder?

Summary
TEACCH (Treatment and Education of Autistic and related Communication handicapped Children) is a psychoeducational intervention designed for children with autism spectrum disorders. Based on the principles of structured teaching, the approach is popular as both a home and school based programme.

The purpose of this literature review was to establish whether TEACCH is effective as a home-based intervention for children with an autism spectrum disorder. A systematic literature search identified five studies which met the inclusion criteria outlined in Appendix A. These five studies were then analysed using Kratochwill’s (2003) coding protocol, with the Weight of Evidence established according to Gough’s (2007) Weight of Evidence Framework. The review found that there is support for the effectiveness of TEACCH as a home-based intervention, particularly in regards to overall developmental outcomes. However, the significant methodological limitation of current research adds caution to these conclusions. Recommendations for future research are discussed.
Introduction

What is TEACCH?

TEACCH (Treatment and Education of Autistic and related Communication handicapped Children) is a psychoeducational intervention developed in the 1960s by Eric Schopler (Cox & Schopler, 1993). ‘Structured teaching’ underpins TEACCH’s methodological framework, with environmental modification occurring through the use of structured interventions. It provides a practical approach to enhance the learning of children with autism, based on the notion that children with autism have difficulties with skills such as organisation, distractibility, sequencing and generalization (Schopler & Reicher, 1971). Therefore, TEACCH promotes structured teaching to organise the environment and provide clear, concrete and meaningful visual information (Mesibov, 1997).

The principles of TEACCH can be summarised as:

- Understanding the culture of autism.
- Developing an individualized person and family centred plan for each student.
- Structuring the physical environment.
- Using visual supports to make the sequence of daily activities predictable and easy to understand.
- Using visual supports to support individual tasks.

TEACCH suggests that the environment is modified in three key ways to optimize learning (Virues-Ortega, Julio, & Pastor-Barriuso, 2013):
1. Organisation of the physical environment to account for the child’s individual needs (for example, minimising distractions)

2. Predictable arrangement of activities (for example, through the use of visual timetables and schedules)

3. Organisation of materials to promote independence from adult direction (for example, the use of visual prompts rather than verbal prompts).

**Psychological basis**

The TEACCH programme is underpinned by an understanding of autism spectrum disorders. Unlike other educational interventions, the TEACCH programme has been developed specifically for children with autism; through an understanding of the key features of autism it assists learning with the use of structured and continuous intervention, environmental adaptations and alternative communication training (Panerai et al., 2009). The triad of impairments (Wing & Gould, 1979) describes the three main areas of difficulties that all people with autism share (figure 1).
The areas of need highlighted by the triad of impairments (Wing and Gould, 1979) are targeted in the TEACCH approach. It is based on the notion that due to the difficulty with lack of imagination and rigidity of thought, providing a highly structured environment is necessary to accommodate the specific needs of individuals with autistic spectrum disorders.
learning environment will increase the learner’s ability to acquire skills relating to language and communication as well as social awareness and interaction. This further relates to the executive dysfunction theory of autism, with the structure provided by the TEACCH programme aiding the difficulties individuals with autism may have with planning, fluency and flexibility (Hill, 2004).

TEACCH was developed following Schopler’s doctoral dissertation in 1966, which demonstrated that people with autism process visual information more easily than verbal information (Mesibov, Shea and Schopler, 2004). Through this research Schopler successfully showed that autism is not a disorder of emotions but simply an impaired way of experiencing the world and understanding that experience (Mesibov, Shea & Schopler, 2004).

Rather than being based on one distinct psychological model, TEACCH comprises of a series of cognitive, developmental, educational and behavioural strategies that are tailored to the unique learning styles of the individual child (Welterlin, Turner-Brown, Harris, Mesibov, & Delmolino, 2012). TEACCH builds on behavioural paradigms by emphasising structure in teaching new behaviours, targeting specific skills and defining the conditions and consequences of behaviours through shaping (Mesibov, 1997). However, particularly in terms of communication, the TEACCH curriculum has developed past the common criticisms of behaviourism (namely learning occurring in isolation) to aid generalisation by strongly emphasising the meaning of communication (Mesibov, 1997).
Evidence has shown that children with autism learn more successfully in a structured rather than unstructured environment (Panerai, Ferrante, Caputo & Impellizzeri 1998). This is in line with the difficulties with rigidity of thought highlighted by the triad of impairments (Wing & Gould, 1979). A growing body of literature supports the efficacy of TEACCH with research showing the effectiveness in different countries, in a variety of settings and when implemented for varying amounts of time (Welterlin et al., 2012). In terms of the behaviours targeted, improvements have been reported in communication and social skills as well as reductions in inappropriate behaviours (Mesibov, 1997).

**Rationale**

With the prevalence of autism spectrum disorders recently estimated at as high as 2.64% (Kim et al., 2011), the demand for interventions and programmes is ever-increasing. Even a more conservative estimate following a large UK epidemiological study reported an overall prevalence rate of 1 in 100 in school-aged children (Baird et al., 2006).

TEACCH has been a popular approach since its inception, with over 30% of families in the USA reporting that they currently use or had used the programme (Green et al., 2006). According to Humphrey and Parkinson (2006) early intensive behavioural interventions and TEACCH are the most commonly used comprehensive approaches for autism in the UK.

Although a cornerstone of the TEACCH intervention is home-programming, there is currently no review of studies examining the effectiveness of the programme as a
home-based intervention (Ozonoff & Cathcart, 1998). With Educational Psychologists increasingly being viewed as community psychologists, with their work stretching across home, school and communities to provide holistic care for children, the efficacy of home programmes is relevant to Educational Psychology practice (MacKay, 2006).

TEACCH provides a wide-range of services to a large spectrum of people with the primary aim of helping the person with autism live or work more effectively. Due to its breadth, studies examining the efficacy of TEACCH often employ a variety of dependent variables. As such, no single dependent variable was consistent across all studies that examined TEACCH as a home programme. Consequently, it was decided that this review would evaluate the efficacy of TEACCH in regards to the developmental outcomes (in terms of skills and behaviours) for children with autism.

**Review Question**

How effective is TEACCH as a home programme for children with an autism spectrum disorder?

**Critical Review of the Evidence Base**

**Literature search**

A literature search was carried out during January 2014 using the electronic databases PsycINFO, ERIC (Educational Resource Index and Abstracts) and Medline. The search terms ‘TEACCH’ and ‘home’ were used.

The search was first conducted in ‘All Fields’ on PsycINFO, yielding 186 results. The titles and abstracts for the 186 results were screened using the criteria in Table 1. The search in ‘Anywhere’ on ERIC yielded 10 papers and the search on Medline
returned 7 papers (Table 2). Following the exclusion process highlighted in Figure 2, 18 articles were selected for review of the full text. As described in Figure 2, 14 of these studies (13 articles) were excluded.

Table 1

List of Inclusion and Exclusion Criteria

<table>
<thead>
<tr>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Publication Type</strong></td>
<td>The study must be in a peer reviewed journal. This ensures methodological rigour.</td>
</tr>
<tr>
<td><strong>2. Language</strong></td>
<td>The study is written in English, due to lack of resources for translation.</td>
</tr>
<tr>
<td><strong>3. Type of Study</strong></td>
<td>The study must contain primary empirical data, which ensures originality of findings.</td>
</tr>
<tr>
<td><strong>4. Intervention</strong></td>
<td>The study employs TEACCH exclusively as a home programme (not within school settings). TEACCH must be the main component of the intervention. To ensure integrity and consistency.</td>
</tr>
<tr>
<td><strong>5. Population</strong></td>
<td>The study is focused specifically on children (0-16 years) with a diagnosis of an autism spectrum disorder. This is to ensure relevance to Educational Psychology practice and consistency between the studies.</td>
</tr>
<tr>
<td><strong>6. Sample</strong></td>
<td>The study has a sample larger than N=1 to ensure reliability.</td>
</tr>
<tr>
<td><strong>7. Dependent variables</strong></td>
<td>The study measures the efficacy of TEACCH in reference to child outcomes. This separates out confounding variables, ensuring a focus on the development of child outcomes.</td>
</tr>
<tr>
<td><strong>8. Analysis</strong></td>
<td>The study reports quantitative</td>
</tr>
</tbody>
</table>
analysis. This allows confounding variables to be separated out and ensures a focus on the outcomes of intervention.

Table 2

**Search Results**

<table>
<thead>
<tr>
<th></th>
<th>All fields</th>
<th>Abstract</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>PsycINFO</td>
<td>186 results</td>
<td>18 results</td>
<td>1 result</td>
</tr>
<tr>
<td>ERIC</td>
<td>10 results</td>
<td>6 results</td>
<td>0 results</td>
</tr>
<tr>
<td>Medline</td>
<td>7 results</td>
<td>3 results</td>
<td>0 results</td>
</tr>
</tbody>
</table>
Including and excluding studies using criteria

The criteria presented in Table 1 were used to select the included papers. For most articles, inclusion or exclusion was clear through careful screening of the title and abstract. In instances where it was unclear, the full text was reviewed (an examination of the full text was required for 18 articles). Through this process, 13 articles were excluded (see Figure 2 and Appendix A for the specific reasons for...
exclusion). Five papers, listed in Table 3, were selected for critical analysis from the exclusion process.

Table 3

List of Included Studies

<table>
<thead>
<tr>
<th>Included Studies</th>
<th></th>
</tr>
</thead>
</table>

Comparison of selected studies

Prior to comparison, the five studies were summarised (Appendix B). Three of the selected studies employed quasi-experimental designs (Braiden, McDaniel, McCrudden, Janes & Crozier, 2012; McConkey, Truesdale-Kennedy, Crawford, McGreevy, Reavey & Cassidy, 2010; Short, 1984). The remaining two studies used an experimental design (Ozonoff & Cathcart, 1998; Welterlin, Turner-Brown, Harris,
Mesibov, & Delmolino, 2012). All five studies used quantitative data to explore child outcomes.

The methodological quality was assessed using the UCL Educational Psychology Literature Review Coding Protocol, adapted from the APA Task Force on Evidence Based Interventions in School Psychology. The coding protocol was adapted to fit this systematic review with consistent decisions made to fit the review question. Completed coding protocols can be found in Appendix C.

To determine each study’s contribution to answering the review question, the Weight of Evidence Framework by Gough (2007) was used. The framework examines three aspects of a study: Quality of Methodology, Relevance of Methodology and Relevance of Evidence to the Review Question (summarised in Table 4, detailed in Appendix D). An average of these weightings is taken to establish the study’s Overall Weight of Evidence. The Weight of Evidence framework was used to allow objective judgements about the value of each study in answering the review question. Table 5 presents the Weight of Evidence awarded to each of the studies.

Table 4

*Weight of Evidence Framework (Gough, 2007)*

<table>
<thead>
<tr>
<th>Weight of Evidence A</th>
<th>Weight of Evidence B</th>
<th>Weight of Evidence C</th>
<th>Weight of Evidence D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality of Methodology: The accuracy, coherency and transparency of evidence.</td>
<td>Relevance of Methodology: The appropriateness of the methodology for answering the review question.</td>
<td>Relevance of evidence to the review question: The relevance of the focus of the evidence for answering the review question.</td>
<td>Overall weight of evidence: Overall assessment of the extent to which the study provides evidence to answer the review question</td>
</tr>
</tbody>
</table>

12
Table 5 Weight of Evidence Awarded to Each Study

<table>
<thead>
<tr>
<th>Studies</th>
<th>(A) Quality of Methodology</th>
<th>(B) Relevance of Methodology</th>
<th>(C) Relevance of evidence to the review question</th>
<th>(D) Overall Weight of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Braiden, McDaniel, McCrudden, Janes, &amp; Crozier, 2012</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>McConkey, Truesdale-Kennedy, Crawford, McGreevy, Reavey &amp; Cassidy, 2010</td>
<td>Medium</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Ozonoff &amp; Cathcart, 1998</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
</tr>
<tr>
<td>Short, 1984</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Welterlin, Turner-Brown, Harris, Mesibov, &amp; Delmolino, 2012</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
</tr>
</tbody>
</table>

Critical Review

Key features of the five selected studies will be detailed and evaluated. The findings and conclusions will then be discussed.

Participants

The studies were conducted in different countries: three were carried out in the United States (Ozonoff & Cathcart, 1998; Short, 1984; Welterlin et al., 2012) and two in Northern Ireland (Braiden et al., 2012; McConkey et al., 2010). The mean age of
participants ranged from 30.5 months (Welterlin et al., 2012) to 55 months (Short, 1984).

The selection process for participants varied between recruitment from TEACCH centres (Short, 1984; Welterlin et al., 2012) referral from a health trust or professional team (Braiden et al., 2012; McConkey et al., 2010) and volunteers (Ozonoff & Cathcart, 1998). This may have implications for the validity of the findings as there may be confounding variables when participants were recruited on a voluntary basis.

The total number of participants in each study ranged from 15 children (Short, 1984) to 61 children (McConkey et al., 2010). Despite this range, when a power analysis was conducted, all the studies failed to provide an adequate sample size and were considerably under-powered (Appendix C).

Research Design

Three of the studies employed a quasi-experimental design (Braiden et al., 2012; McConkey et al., 2010; Short, 1984) and two studies used an experimental design (Ozonoff & Cathcart, 1998; Welterlin et al., 2012).

The use of control groups varied, with two studies using matched control groups (Ozonoff & Cathcart, 1998; Welterlin et al., 2012), Short (1984) using measures taken during a waiting period as a contrast measure, McConkey et al., (2010) having an un-matched control group and Braiden et al., (2012) simply employing a pre-post design with no control group. For the latter study, the lack of control group means developmental maturation was not accounted for and less of the variance can be attributed to the intervention. The contrast measure employed by Short (1984) was
not sufficient to be counted as a comparison group in the coding protocol and Weight of Evidence framework (Appendix C and D).

In the three studies that employed a control group, assignment to condition was not random for two studies (Ozonoff & Cathcart, 1998; McConkey et al., 2010) with only one study randomly assigning participants to treatment or control (Welterlin et al., 2012) and only two studies matching participants (Ozonoff & Cathcart, 1998; Welterlin et al., 2012). All three studies used ‘no intervention’ as their comparison group. These differences are reflected in the weightings given for ‘Comparison Group’ (Appendix D).

**Intervention**

As specified in the inclusion criteria, TEACCH was a main component of intervention in all the studies. However, there were differences in the emphasis given to the TEACCH approach. These differences affected the ‘Relevance of Evidence to the Review Question (Category C) weightings.

Two studies were awarded high weightings for the relevance of evidence (Ozonoff & Cathcart, 1998; Welterlin et al., 2012) as they used TEACCH exclusively and employed a comparison group. Two studies used TEACCH exclusively but only received a medium weighting due to the lack of comparison group (Short, 1984; Braiden et al., 2012). The remaining study (McConkey et al., 2010) used TEACCH alongside other communication approaches (TEACCH was used in conjunction with Hanen and Picture Exchange Communication System) which affected its weighting for ‘Relevance to the Review Question’ (Appendix D).

The length of intervention varied both between and within the studies. There was a range of 6 to 11 months of intervention in one of the studies, with between 15 and 18 sessions (McConkey et al., 2010). Two studies had a mean of 10 sessions (Braiden
et al., 2012; Ozonoff & Cathcart, 1998), one study ran the intervention for 12 sessions (Welterlin et al., 2012) with Short (1984) running the intervention for only 6 to 8 sessions. This has obvious implications for the findings; the period of intervention may result in a confounding variable. Some measures may also lack the sensitivity to detect and identify subtle progress during shorter intervention periods.

Measures

There was no common dependent variable across the studies, with each study employing different measures to examine their various outcomes. As this review was only interested in child outcomes (as opposed to parental outcomes such as reduction in stress), the measures applicable to that outcome are detailed.

The Psychoeducational Profile-Revised (PEP-R) (Schopler et al., 1990) was used as a measure of the child’s developmental skills for two studies (Braiden et al., 2012; Ozonoff & Cathcart, 1998). This measure relies on a professional’s assessment of the child’s behaviour on seven developmental subscales (PEP-3 has ten subscales). On post-testing, Braiden et al. (2012) only re-administered the measures on the receptive and expressive language domains. As Braiden et al., (2012) failed to report the reliability in text, the study could only be awarded a medium weighting for ‘measure’. However, it did collect data from multiple methods and sources, including parental questionnaires. The study by Ozonoff and Cathcart (1998) was only allocated a low weighting for ‘measure’ as not only did the study omit to report or reference the reliability, it failed to triangulate evidence by collecting data from multiple methods and/or sources (Appendix D).

Although McConkey et al. (2010) also used the PEP-R as a measure of child outcomes, comparable PEP-R data was not available for the contrast group as
researchers felt it would be too intrusive to obtain. Instead, the Vineland Adaptive Behaviour Scale (Sparrow, Cicchetti & Balla, 1989) and Gilliam Autism Rating Scale (GARS) (Gilliam, 1995) were used as a pre and post measure for both treatment and control groups. The Vineland Adaptive Behaviour Scales relies on parental report so may not be as accurate as a measure conducted by a professional. Nevertheless, the paper reported the reliability for the Vineland as .78 to .93. As the reliability was reported and multiple methods were used (the parental questionnaire included perceptions of programme efficacy), this study was awarded a high weighting for ‘measure’.

Short (1984) was the only study to use behavioural observation coding, with a reported reliability of greater than 0.95 for the main variables. These findings were supported by interview ratings and questionnaires which sought parents’ opinion of behaviour change following intervention. As such, this study was also awarded a high weighting for ‘measure’.

The study by Welterlin et al. (2012) assessed child behaviour in the treatment and control group pre and post intervention using the Mullen Scales of Early Learning (Mullen, 1995) and the Scales of Independent Behaviour-revised (Bruininks et al., 1996). The paper states that both scales have demonstrated good reliability and validity but fails to report these figures or cite a reference. The questionnaire administered to parents focused on parental outcomes rather than programme efficacy. Consequently, the study was awarded a medium weighting for ‘measure’.

Analysis

None of the studies received a high weighting for ‘analysis’ as they all failed to use a sufficiently large sample and consequently were considerably underpowered.
All studies were judged to have conducted an appropriate statistical analysis, a key criteria for the ‘Analysis’ weighting (Appendix D). Analysis of variance was selected by four studies; Ozonoff & Cathcart (1998) used a repeated measures multivariate analysis of variance, Short (1984) used a multivariate analysis of variance, Welterlin et al. (2012) used a repeated measures analysis of variance as did McConkey et al., (2010). One study (Braiden et al., 2012) employed a paired-samples \( t \)-test.

Only two studies reported effect sizes (Braiden et al., 2012; Welterlin et al., 2012). McConkey et al. (2010) and Ozonoff and Cathcart (1998) provided sufficient data for the calculation of effect sizes. These studies all received a medium weighting for ‘analysis’ as they were also judged to have used an appropriate statistical analysis (Appendix D). The study by Short (1984) failed to offer sufficient data to compute effect sizes (standard deviations were not provided) so only received a low weighting.

Findings

Four studies reported statistically significant findings (Ozonoff & Cathcart, 1998; Braiden et al., 2012; McConkey et al., 2010; Short, 1984) but this was not necessarily for all child-based outcomes (as seen in Table 4). Welterlin et al. (2012) failed to find a statistically significant group difference across all the measures but did find significant within group differences for the treatment group (Table 4). As the authors report, the insufficient sample size was likely to have affected the significance of statistical testing.

As all of the studies were underpowered, effect sizes were calculated for the four studies which provided sufficient data, regardless of significance (Short (1984) failed
to provide sufficient data for the calculation of effect sizes). However, caution should be applied when interpreting the effect sizes from studies which did not find statistical significance, as the effects may be in the opposite direction.

Effect sizes were calculated for all child-outcomes and grouped by domain, adapted from the Portage Early Education Programme Checklist (NFER-NELSON, 1987) (Table 6). The effect sizes have been interpreted and labelled as small (0.2), medium (0.5) and large (0.8) in line with Cohen’s (1992) recommendations. The significance of statistical testing has been noted next to each effect size following the key in Table 6.

Table 6

Effect Sizes

<table>
<thead>
<tr>
<th>Domain</th>
<th>Study</th>
<th>Weight of Evidence (D)</th>
<th>Pre-test post-test mean difference in treatment group only</th>
<th>Post-test mean difference in treatment and contrast group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive</td>
<td>Ozonoff &amp; Cathcart (1998)</td>
<td>Medium</td>
<td>1.66 <strong>NR</strong> Large</td>
<td>0.32 <strong>S</strong> Small</td>
</tr>
<tr>
<td></td>
<td>(cognitive performance)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ozonoff &amp; Cathcart (1998)</td>
<td>Medium</td>
<td>0.51 <strong>NR</strong> Medium</td>
<td>-0.01 <strong>NS</strong></td>
</tr>
<tr>
<td></td>
<td>(cognitive verbal)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ozonoff &amp; Cathcart (1998)</td>
<td>Medium</td>
<td>1.24 <strong>NR</strong> Large</td>
<td>0.21 <strong>NS</strong> Small</td>
</tr>
<tr>
<td></td>
<td>(perception)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Braiden et al. (2012)</td>
<td>Low</td>
<td>0.30 <strong>S</strong> Small</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>(receptive language)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Braiden et al. (2012)</td>
<td>Low</td>
<td>0.30 <strong>S</strong> Small</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>(expressive language)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>McConkey et al. (2010)</td>
<td>Medium</td>
<td>0.98 <strong>NS</strong> Large</td>
<td>0.60 <strong>S</strong> Medium</td>
</tr>
<tr>
<td></td>
<td>(Vineland - communication)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Welterlin et al. (2012)</td>
<td>Medium</td>
<td>0.33 <strong>NS</strong> Small</td>
<td>0.01 <strong>NS</strong></td>
</tr>
<tr>
<td></td>
<td>(receptive language)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>Measure</td>
<td>Effect Size</td>
<td>Significance</td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-----------------------------------</td>
<td>-------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Welterlin et al. (2012)</td>
<td>(expressive language)</td>
<td>Medium</td>
<td>-0.16</td>
<td></td>
</tr>
<tr>
<td>Welterlin et al. (2012)</td>
<td>(SIB language comprehension)</td>
<td>Medium</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Welterlin et al. (2012)</td>
<td>(SIB language expression)</td>
<td>Medium</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Welterlin et al. (2012)</td>
<td>(SIB language comprehension)</td>
<td>Medium</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Welterlin et al. (2012)</td>
<td>(SIB language expression)</td>
<td>Medium</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Socialisation</td>
<td>Welterlin et al. (2012)</td>
<td>Medium</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>McConkey et al. (2010)</td>
<td>(social interaction)</td>
<td>Medium</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>Self-help/ Behaviour</td>
<td>McConkey et al. (2010)</td>
<td>Medium</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Vineland - daily living)</td>
<td>Medium</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>McConkey et al. (2010)</td>
<td>(Vineland - adaptive bhav)</td>
<td>Medium</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Ozonoff &amp; Cathcart (1998)</td>
<td>(imitation)</td>
<td>Medium</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Motor</td>
<td>McConkey et al. (2010)</td>
<td>Medium</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Vineland - motor skills)</td>
<td>Medium</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Ozonoff &amp; Cathcart (1998)</td>
<td>(fine motor)</td>
<td>Medium</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Ozonoff &amp; Cathcart (1998)</td>
<td>(gross motor)</td>
<td>Medium</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Ozonoff &amp; Cathcart (1998)</td>
<td>(eye-hand integration)</td>
<td>Medium</td>
<td>-0.14</td>
<td></td>
</tr>
<tr>
<td>Overall Measure</td>
<td>McConkey et al. (2010)</td>
<td>Medium</td>
<td>-0.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(GARS autism quotient)</td>
<td>Medium</td>
<td>-0.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(GARS percentile scores)</td>
<td>Medium</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(total PEP-R score)</td>
<td>Medium</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Mullen developmental quotient)</td>
<td>Medium</td>
<td>0.26</td>
<td></td>
</tr>
</tbody>
</table>

*Significant findings from statistical analysis.
NS Non-significant findings from statistical analysis.
NR Statistical analysis not reported.
When examining the effect sizes reported in Table 6, there are large differences between the effect sizes for the pre-test post-test mean difference in treatment group only and the post-test mean difference in the treatment and control group. This reveals how important it is to have a control group as well as the need to establish group equivalence; without group equivalence, the differences between groups may be harder to establish.

Overall, there were moderate effects of TEACCH across the domains, with large outcomes reported within the treatment group for every domain. However, when looking at differences between the treatment and control group, only one large effect size was found - the total PEP-R score in Ozonoff and Cathcart's (1998) study. As this is an overall developmental measure and the study was awarded a ‘medium’ Weight of Evidence, this suggests that TEACCH may be an effective home-based intervention for children with autism, particularly in regards to overall outcomes.

The negative effect sizes found by McConkey et al., (2010) for the GARS autism quotient and percentile scores between the treatment and control group need to be considered. The GARS autism quotient assesses particular features of autism, with a score above 111 highly indicative of autism (McConkey et al., 2010). The authors state that these effects could be due to maturation with features of autism becoming clearer as the child gets older. The negative effect sizes are therefore suggestive of the scores of the children in the contrast group rising more than those in the treatment group, suggestive of the success of TEACCH intervention. Overall, the domain of ‘overall measure’ seems to have produced the most promising outcomes for TEACCH home intervention.
Both Welterlin et al. (2012) and McConkey et al. (2010) found large pre-post treatment group effect sizes for socialisation. Although the effect sizes were small between the treatment and contrast group, these findings show promising effects of TEACCH intervention on a child with autism’s social skills; an important component of autism, as highlighted by the triad of impairments (Wing & Gould, 1979).

There were modest outcomes for the language domain, with the majority of between treatment and contrast group effect sizes rated as small. The only medium effect size between the groups was found by McConkey et al. (2010) who measured communication. Perhaps the lack of clarity of findings results from the different aspects of language explored and the different measurement tools used to measure outcomes. As such, there is only tentative support for the impact of TEACCH on language skills.

The cognitive domain only comprised of Ozonoff and Cathcart’s (1998) study. The small effect size for between group differences for cognitive performance and perception suggests that the intervention may have a small impact on cognitive skills, with no effect on cognitive verbal skills. As group equivalence was established and this study has a medium ‘Weight of Evidence’, TEACCH as a home-based intervention may have small positive effects on cognitive skills.

For self-help skills and behaviour, the effect sizes across treatment and contrast groups are small and medium (McConkey et al., 2010; Ozonoff & Cathcart, 1998).
As both these studies were weighted as ‘medium’, this provides promising support for the efficacy of TEACCH in regards to self-help skills and behaviour.

Although Short (1984) reports a significant increase in appropriate child behaviour compared to the waiting period ($F(1,14, = 29.23, P < 0.001$) there was not sufficient data to calculate effect sizes. The study failed to find a statistical significant decrease in inappropriate child behaviour. In line with the low weightings for ‘Overall Weight of Evidence’, ‘Quality of Methodology’ and ‘Relevance of Methodology’ less weight can be given to these findings. Although promisingly supportive of TEACCH in terms of increasing appropriate child behaviour, additional research would be required with more stringent methodology to support the conclusions.

**Conclusions and Recommendations**

Overall, the studies included in this review provide positive, supportive evidence that TEACCH can be an effective home based intervention for children with autism. However, none of the studies were awarded high weightings for ‘Overall Weight of Evidence’. Specifically, all of the studies failed to employ an active control group, compare the outcomes of TEACCH with an alternative intervention or have a sufficient sample size. As such, the findings need to be interpreted with caution.

Nevertheless, this review does provide support for the efficacy of TEACCH as a home-based intervention, particularly for overall developmental outcomes. As such, TEACCH is recommended as a home-based intervention for children with autism. This needs to be interpreted with caution due to the methodological issues and limitations detailed below.
Through the completed coding protocols and Weight of Evidence Framework (Appendix C and D respectively), it is clear that all the studies included in the review were limited in terms of their methodology. None of the studies employed an active comparison group, with two studies failing to have an appropriate comparison group at all (Braiden et al., 2012; Short, 1984). As such, developmental maturation was not accounted for and the impact of the intervention could not accurately be assessed. Furthermore, Ozonoff and Cathcart (1998) failed to triangulate data by using multiple methods and sources so the behaviour reported could have been impacted by any number of confounding variables. In the future, it is necessary for researchers to triangulate data to reduce bias and the impact of confounding variables. Moreover, the impact of intervention cannot be fully established without active control groups and longer term follow ups; a clear requirement for future research.

Although TEACCH had to be employed exclusively as a home programme as per the inclusion criteria (Appendix A), the studies differed in their use of TEACCH as a sole intervention or part of a wider intervention approach. Only two studies were decided to have used TEACCH exclusively whilst employing a control group (Ozonoff & Cathcart, 1998; Welterlin et al., 2012). Consequently, weightings for the ‘Relevance of Evidence to the Review Question’ were affected for the other studies as it was unclear whether the reported effects could be solely attributed to the TEACCH principles. In the future, it is important for authors to provide clear and transparent information about the type of TEACCH intervention employed to allow a more effective comparison of approaches.

Overall, this review was limited by the lack of consistent dependent variable across the studies. Although each study was interested in child-outcomes as a result of TEACCH intervention, they all focused on slightly different developmental outcomes,
assessed using different measures. Due to this varied focus, the outcomes examined are not necessarily identical. Consequently, in the future when there is more research on TEACCH as a home-programme, it would be interesting to examine the effectiveness of the intervention in regards to one specific child-outcome which is assessed with a consistent standardised measure.

References


## Appendix A: Exclusion Table

<table>
<thead>
<tr>
<th>Study</th>
<th>Rationale for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probst, P &amp; Glen, I. (2011). TEACCH-based interventions for families with children with autism spectrum disorders: Outcomes of a parent group training study and a home-based child-parent training single case study. <em>Life Span and Disability, 14</em>(2), 111-138.</td>
<td>Study 1 was excluded on criteria 7 as it measured parental outcomes. Study 2 was excluded on criteria 6 as the sample was N=1.</td>
</tr>
</tbody>
</table>
| Rickards, A. L, Walstab, J. E, Wright-Rossi, R. | Excluded on criteria 4 as TEACCH

Excluded on criteria 4 as the studies do not employ TEACCH exclusively as a home programme.


Excluded on criteria 4 as the studies do not employ TEACCH exclusively as a home programme.


Excluded on criteria 3 as it does not contain primary empirical data.


Excluded on criteria 1 as could not access full text from a peer-reviewed journal.


Excluded on criteria 4 as TEACCH was implemented in educational settings.


Excluded on criteria 1 as could not access full text from a peer-reviewed journal.

Schopler, E., Mesibov, G., & Baker, A. (1982). Excluded on criteria 6 as the
## Appendix B: Summary of Studies

<table>
<thead>
<tr>
<th>Author and Aim (relevant to review)</th>
<th>Participants</th>
<th>Intervention</th>
<th>Design</th>
<th>Measures</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Braiden, H. J., McDaniel, B., McCrudden, E., Janes, M., &amp; Crozier, B. (2012)</td>
<td>Northern Ireland. 18 children (17 male, one female). Average age 3 years 2 months. All recently diagnosed with autism. Referred to the service from a large health trust. Selected on the basis of age, need and family circumstance (e.g. ability to attend).</td>
<td>TEACCH-based Early Intervention Programme. 10 weeks. Delivered by fully trained and accredited TEACCH facilitators. Children supported through 1:1 sessions and parents are facilitated in understanding their child's autism and in implementing TEACCH methods.</td>
<td>Pre-test/post-test design No control group</td>
<td>Psychoeducational Profile – 3 (PEP-3). PEP subtest of expressive and receptive language used post-test. Non-standardised change questionnaire used to ascertain parent views.</td>
<td>Expressive language: Paired-samples t-test. Significant increase from pre-testing (M=11.88, SD=12) to post-testing (M=15.44, SD=13.23; t(17)=-3.85, p&lt;0.05). Receptive language: Significant increase from pre-testing (M=11.88, SD=13.18) to post-testing (M=15.77, SD=14.40; t(17)=-3.781, p&lt;0.05).</td>
</tr>
<tr>
<td>Mc Conkey, Truesdale-Kennedy, Crawford, Mc Greevy, Reavy &amp; Cassidy (2010)</td>
<td>Northern Ireland: 35 families in intervention group and a contrast group of 26 families to control for time effects. Children aged between 2 and 4 years and were not</td>
<td>Intervention programme known as 'Keyhole', based around TEACCH approaches and principles alongside other methods to enhance communication (Hanen and PECS). Intervention delivered by two different therapists</td>
<td>Quasi-multiple baseline design. In control group, 15 families were offered 5 home visits where they were given general information about autism</td>
<td>Mixture of quantitative and qualitative measures. Child outcome measures: Psychoeducational Profile - Revised (PEP-R)</td>
<td>Vineland Adaptive Behaviour Scores: Repeated measure ANOVA. Communication T: F = 2.89, NS G×T: F = 7.83, p &lt; .01 Socialisation T: F = 13.4, p &lt; .001 G×T: F = 2.10, NS Daily living</td>
</tr>
</tbody>
</table>
**Doctorate in Educational and Child Psychology**

**Jemma Lewis**

| TEACCH approaches and other communication methods (e.g. Hanen and PECS). | attending nursery school. Overall 55 were boys and 6 were girls. Children in contrast group significantly older with mean age 3.4 years compared to 2.8 years in programme group. No random assignment due to ethical considerations Recruitment – offered home-based programme after diagnosis. Offered to all families until all the available places taken. | (qualified speech and language therapists) in two separate geographical areas. Ranged between 18 and 15 home visits over 11 and 6 month period. Same content and schedule of visits were followed but were modified according to the needs of the child and mother. Each visit lasted 90mins on average. and the other 11 families received no additional services or supports. No significant differences between these two groups so participants were combined to form one contrast group. | Gilliam Autism Rating Scale Vineland Adaptive Behaviour Scale T: F = .62, NS G×T: F = 4.54, p < .05 Motor Skills T: F = .08, NS G×T: F = 1.22, NS Adaptive behaviour T: F = 2.30, NS G×T: F = 3.4, p < .08 Gilliam Autism Rating Scale: Autism quotient T: F = 12.05, p < .001 G×T: F = 2.79, p < .100 Percentile scores T: F = 1.00, p < .005 G×T: F = 4.93, p < .05 PEP-R – comparable PEP-R data not available for contrast group as it was felt it was too intrusive. For treatment group, the mean ‘developmental age’ rose significantly from 20.1 months (SD 7.4) to 29.7 months (SD 11.2) (t = 5.57, p < .001). |

| Ozonoff & Cathcart (1998) Evaluated the effectiveness of a USA: 22 children. First 11 subjects to respond to the study announcement were assigned to the Subjects in the treatment group received TEACCH-based home program services from trained graduate students in the Experimental - Matched group design (treatment and control). | Psychoeducational Profile - Revised (PEP-R) Childhood Autism Repeated measures multivariate analysis of variance. Total post-test PEP-R score, F(1, 20) |

**Table**

- **TEACCH approaches and other communication methods**: TEACCH approaches and other communication methods (e.g. Hanen and PECS).
- **Attending nursery school**: Overall 55 were boys and 6 were girls. Children in contrast group significantly older with mean age 3.4 years compared to 2.8 years in programme group. No random assignment due to ethical considerations. Recruitment – offered home-based programme after diagnosis. Offered to all families until all the available places taken.
- **(Qualified speech and language therapists)**: In two separate geographical areas. Ranged between 18 and 15 home visits over 11 and 6 month period. Same content and schedule of visits were followed but were modified according to the needs of the child and mother. Each visit lasted 90mins on average. and the other 11 families received no additional services or supports. No significant differences between these two groups so participants were combined to form one contrast group.
- **Gilliam Autism Rating Scale**: Autism quotient T: F = 12.05, p < .001 G×T: F = 2.79, p < .100 Percentile scores T: F = 1.00, p < .005 G×T: F = 4.93, p < .05
- **PEP-R**: Comparable PEP-R data not available for contrast group as it was felt it was too intrusive. For treatment group, the mean ‘developmental age’ rose significantly from 20.1 months (SD 7.4) to 29.7 months (SD 11.2) (t = 5.57, p < .001).
<table>
<thead>
<tr>
<th><strong>TEACCH-based home program intervention for young children with autism</strong></th>
<th><strong>University of Utah’s Department of Psychology. The mean number of treatment sessions was 10 (range 8-12).</strong></th>
<th><strong>Rating Scale (CARS).</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment group and the latter 11 assigned to the control group. Not randomly assigned.</td>
<td>The children in the control group received no home program but, like the children in the treatment group, regularly attended their day treatment program.</td>
<td>= 4.57, $p &lt; .05$.</td>
</tr>
<tr>
<td>Treatment group: Mean age = 53.3 months. 9 males, 2 females.</td>
<td></td>
<td>Total PEP scores.</td>
</tr>
<tr>
<td>Control group: mean age = 53.5 months. 9 males, 2 females. Children aged 2-6 years. All came from two parent families. All were Caucasian American except one boy in treatment group who was Hispanic American.</td>
<td></td>
<td>Treatment:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pre M = 21.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD = 6.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post M = 28.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD = 11.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Control:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pre M = 24.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD = 7.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Post M = 26.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SD = 10.4</td>
</tr>
<tr>
<td><strong>Short (1984)</strong></td>
<td>USA: 15 children with autism and their families. Mean age = 4.7 years. Families chosen from a series of consecutive referrals to Division TEACCH. Children first followed during a waiting</td>
<td></td>
</tr>
<tr>
<td><strong>An attempt to evaluate an autism treatment model as it functions in an ongoing clinical service.</strong></td>
<td>6-8 treatment sessions, lasting 60-90 min. Treatment the same as is usually used in TEACCH clinics. Waiting period had mean duration of 1 month 29 days, treatment period had mean duration of 4 months 21 days.</td>
<td>Pre-test, post-test design. Repeated measures.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Behavioural observation coding (assessed for inter-rater reliability). 14 behaviour categories were coded as either having occurred or not occurred during 10 second intervals in 1.5 hour observations.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Can’t compute effect size – doesn’t report standard deviation. Multivariate analysis of variance for repeated measures. Appropriate child behaviour</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1st visit = 0.86, 2nd visit = 0.89, post treatment = 1.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$F(1, 14) = 29.23, P &lt; 0.001$</td>
</tr>
</tbody>
</table>
Examine the efficacy of the home TEACCHing program using a dual approach to assessment that combines the advantages of single subject design and those of a group design.

**USA**: Twenty 2-3 year old children. Mean age 30.5 months across groups. 9 males and 1 female per group. Families paired on chronological and mental age (e.g., within 6 months for each) and then randomly assigned to the treatment (Home TEACCHing Program or waitlist group).

Inclusion criteria included a chronological age of less than 42 months and a clinical diagnosis of autism.

Therapists members of TEACCH regular staff.

Employs the home TEACCHing program: an outreach model designed to service the early intervention needs of 2 to 3 year olds with autism and their families.

Runs for 12, 90-minute sessions once a week in which a clinician provides training to the parents on the structured teaching methodology of TEACCH.

Randomised group design with waitlist control group.

Repeated measures analysis of variance for the Mullen Scales of Early Learning (MSEL), Scales of independent behaviour-revised (SIB-R), and Parenting Stress Index 3rd edition.

Treatment group: Mullen developmental quotient baseline to post-test, $t = -2.4, p = 0.0, d = 0.4$. Waitlist group: Mullen developmental quotient baseline to post-test, $t = -2.1, p = 0.1, d = 0.2$.

Results suggest that this study was underpowered to demonstrate significant intervention effects and also suggest that the 12-week timeframe may be too short for these global developmental measures.
Appendix C


Coding Protocol

Name of Coder: Date: 17.01.14

Full Study Reference in proper format:


Intervention Name (description of study): TEACCH informed early intervention

Study ID Number: 01

☐ Type of Publication:
☐ Book/Monograph
☒ Journal Article
☐ Book Chapter
☐ Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)
☐ Completely randomized design
☐ Randomized block design (between participants, e.g., matched classrooms)
☐ Randomized block design (within participants)
☐ Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)
☐ Nonrandomized design
☐ Nonrandomized block design (between participants)
☐ Nonrandomized block design (within participants)
☐ Nonrandomized hierarchical design
☒ Optional coding for Quasi-experimental designs No control group. Pre-post test design.

A3. Overall confidence of judgment on how participants were assigned (select on of the following)
☐ Very low (little basis)
☐ Low (guess)
☐ Moderate (weak inference)
☐ High (strong inference)
☐ Very high (explicitly stated)
☒ N/A
☐ Unknown/unable to code

B Participants

Total size of sample (start of study): 18
Intervention group sample size: 18
Control group sample size: N/A

C. Type of Program

- Universal prevention program
- Selective prevention program
- Targeted prevention program
- Intervention/Treatment
- Unknown

D. Stage of Program

- Model/demonstration programs
  Although TEACCH is an established programme, the format of intervention used in this study is model/new programme.
- Early stage programs
- Established/institutionalized programs
- Unknown

E. Concurrent or Historical Intervention Exposure

- Current exposure
- Prior exposure
- Unknown

Section 2  Key Features for Coding Studies and Rating Level of Evidence/Support

A Measurement (Estimating the quality of the measures used to establish effects)
(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes (see following table for a detailed breakdown on the outcomes)
- Yes
- No
- Unknown/unable to code
  Not reported in paper.

A2 Multi-method (at least two assessment methods used)
- Yes
- No
- N/A
  Used parental evaluation of the programme’s effectiveness through questionnaires.
- Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)
- Yes
- No
- N/A
- Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)
- Yes validated with specific target group
- In part, validated for general population only
- No
- Unknown/unable to code
  Not reported by authors.
Overall Rating of Measurement: □ 3 □ 2 □ 1 □ 0

B Comparison Group

B1 Type of Comparison group
☐ Typical intervention
☐ Attention placebo
☐ Intervention element placebo
☐ Alternative intervention
☐ Pharmacotherapy
☐ No intervention
☐ Wait list/delayed intervention
☐ Minimal contact
☐ Unable to identify type of comparison group

B2 Overall rating of judgment of type of comparison group
☐ Very low
☐ Low
☐ Moderate
☐ High
☐ Very high
☐ Unable to identify comparison group

B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc were counter-balanced across intervention)
☐ By change agent
☐ Statistical (analyse includes a test for intervention)
☐ Other
☐ Not reported/None

B4 Group equivalence established
☐ Random assignment
☐ Posthoc matched set
☐ Statistical matching
☐ Post hoc test for group equivalence

B5 Equivalent mortality
☐ Low attrition (less than 20% for post)
☐ Low attrition (less than 30% for follow-up)
☐ Intent to intervene analysis carried out?
Findings

Overall Level of Evidence = 0
3= Strong Evidence  2=Promising Evidence  1=Weak Evidence  0=No Evidence
C Appropriate Statistical Analysis

Analysis 1

- ☑ Appropriate unit of analysis
- Paired samples t-test.
- Familywise/expermenter wise error rate controlled when applicable
- None reported.
- Sufficiently large N
  - No. Minimum of 64 participants for a medium (0.5) effect size.

Analysis 2

- ☐ Appropriate unit of analysis
- Familywise/expermenter wise error rate controlled when applicable
- Sufficiently large N

Analysis 3

- ☐ Appropriate unit of analysis
- Familywise/expermenter wise error rate controlled when applicable
- Sufficiently large N

Overall Rating of Analysis: ☐ 3 ☐ 2 ☑ 1 ☐ 0 Score of 1 as appropriate unit of analysis used.
Coding Protocol

Name of Coder: Jemma Lewis
Date: 17.01.14

Full Study Reference in proper format:


Intervention Name (description of study): TEACCH based home-intervention

Study ID Number: 02

- [ ] Type of Publication:
  - Book/Monograph
  - Journal Article
  - Book Chapter
  - Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)
- [ ] Completely randomized design
- [ ] Randomized block design (between participants, e.g., matched classrooms)
- [ ] Randomized block design (within participants)
- [ ] Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)
- [ ] Nonrandomized design
- [ ] Nonrandomized block design (between participants)
- [ ] Nonrandomized block design (within participants)
- [ ] Nonrandomized hierarchical design
- [ ] Optional coding for Quasi-experimental designs

Authors stated it was not possible to allocate families randomly due to ethical considerations. Families in contrast groups recruited after intervention programme had ended.

A3. Overall confidence of judgment on how participants were assigned (select on of the following)
- [ ] Very low (little basis)
- [ ] Low (guess)
- [ ] Moderate (weak inference)
- [ ] High (strong inference)
- [ ] Very high (explicitly stated)

- [ ] N/A
- [ ] Unknown/unable to code

B Participants

Total size of sample (start of study): 61
Intervention group sample size: 35

Control group sample size: 26

C. Type of Program

☐ Universal prevention program
☐ Selective prevention program
☐ Targeted prevention program
☐ Intervention/Treatment
☐ Unknown

D. Stage of Program

☐ Model/demonstration programs

*Although TEACCH is an established programme, the format of intervention used in this study is model/new programme.*

☐ Early stage programs
☐ Established/institutionalized programs
☐ Unknown

E. Concurrent or Historical Intervention Exposure

☐ Current exposure
☐ Prior exposure
☐ Unknown

Section 2 Key Features for Coding Studies and Rating Level of Evidence/Support

A Measurement (Estimating the quality of the measures used to establish effects)

(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes (see following table for a detailed breakdown on the outcomes)

☐ Yes

*Authors report reliability for Vineland Adaptive Behaviour Scale as .78 to .92 for each of the domains.*

☐ No

☐ Unknown/unable to code

A2 Multi-method (at least two assessment methods used)

☐ Yes

*Used questionnaire to assess parental perceptions of the programme alongside other measures.*

☐ No

☐ N/A

☐ Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)

☐ Yes

☐ No

☐ N/A

☐ Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

☐ Yes validated with specific target group

☐ In part, validated for general population only

☐ No
Unknown/unable to code

Not cited in paper. Although it was stated that validity was considered in deciding what measures to use.

Overall Rating of Measurement: 3 2 1 0

<table>
<thead>
<tr>
<th>B Comparison Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1 Type of Comparison group</td>
</tr>
<tr>
<td>□ Typical intervention</td>
</tr>
<tr>
<td>□ Attention placebo</td>
</tr>
<tr>
<td>□ Intervention element placebo</td>
</tr>
<tr>
<td>□ Alternative intervention</td>
</tr>
<tr>
<td>□ Pharmacotherapy</td>
</tr>
<tr>
<td>❌ No intervention</td>
</tr>
<tr>
<td>□ Wait list/delayed intervention</td>
</tr>
<tr>
<td>□ Minimal contact</td>
</tr>
<tr>
<td>□ Unable to identify type of comparison</td>
</tr>
<tr>
<td>B2 Overall rating of judgment of type of comparison group</td>
</tr>
<tr>
<td>□ Very low</td>
</tr>
<tr>
<td>□ Low</td>
</tr>
<tr>
<td>□ Moderate</td>
</tr>
<tr>
<td>□ High</td>
</tr>
<tr>
<td>□ Mixture of no intervention and alternative intervention (in the form of home Visits and support) but grouped as one ‘no intervention’ group as statistically analysis showed no difference between the type of comparison group.</td>
</tr>
<tr>
<td>□ Very high</td>
</tr>
<tr>
<td>□ Unable to identify comparison group</td>
</tr>
<tr>
<td>B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc were counter-balanced across intervention)</td>
</tr>
<tr>
<td>□ By change agent</td>
</tr>
<tr>
<td>□ Statistical (analyse includes a test for intervention)</td>
</tr>
<tr>
<td>□ Other</td>
</tr>
<tr>
<td>□ Not reported/None</td>
</tr>
<tr>
<td>B4 Group equivalence established</td>
</tr>
<tr>
<td>□ Random assignment</td>
</tr>
<tr>
<td>□ Posthoc matched set</td>
</tr>
<tr>
<td>□ Statistical matching</td>
</tr>
<tr>
<td>□ Post hoc test for group equivalence</td>
</tr>
<tr>
<td>B5 Equivalent mortality</td>
</tr>
<tr>
<td>□ Low attrition (less than 20 % for post)</td>
</tr>
<tr>
<td>□ Low attrition (less than 30% for follow-up)</td>
</tr>
<tr>
<td>□ Intent to intervene analysis carried out?</td>
</tr>
<tr>
<td>Findings __________</td>
</tr>
</tbody>
</table>

Overall Level of Evidence = 1

3= Strong Evidence  2=Promising Evidence  1=Weak Evidence  0=No Evidence
C Appropriate Statistical Analysis

Analysis 1

☒ Appropriate unit of analysis
☐ Familywise/expermenter wise error rate controlled when applicable
☐ Sufficiently large N

Repeated measures ANOVA

None reported.

No. Minimum of 64 participants for a medium (0.5) effect size.

Analysis 2

☐ Appropriate unit of analysis
☐ Familywise/expermenter wise error rate controlled when applicable
☐ Sufficiently large N

Analysis 3

☐ Appropriate unit of analysis
☐ Familywise/expermenter wise error rate controlled when applicable
☐ Sufficiently large N

Overall Rating of Analysis: ☐  3  ☒  2  ☒  1  ☐  0  Score of 1 as appropriate unit of analysis used.
Coding Protocol

Name of Coder: Date: 17.01.14

Full Study Reference in proper format:


Intervention Name (description of study): TEACCH based home intervention programme.

Study ID Number: 03

- [ ] Type of Publication:
  - [x] Journal Article
  - [ ] Book/Monograph
  - [ ] Book Chapter
  - [ ] Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)
- [ ] Completely randomized design
- [ ] Randomized block design (between participants, e.g., matched classrooms)
- [ ] Randomized block design (within participants)
- [ ] Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)
- [ ] Nonrandomized design
- [x] Nonrandomized block design (between participants)
- [ ] Nonrandomized block design (within participants)
- [ ] Nonrandomized hierarchical design
- [ ] Optional coding for Quasi-experimental designs

A3. Overall confidence of judgment on how participants were assigned (select one of the following)
- [ ] Very low (little basis)
- [ ] Low (guess)
- [ ] Moderate (weak inference)
- [ ] High (strong inference)
- [x] Very high (explicitly stated) *First 11 respondents assigned to treatment group, next 11 to control group.*
- [ ] N/A
- [ ] Unknown/unable to code

B. Participants

Total size of sample (start of study): 22

Intervention group sample size: 11

Control group sample size: 11
C. Type of Program

☐ Universal prevention program
☐ Selective prevention program
☐ Targeted prevention program
☒ Intervention/Treatment
☐ Unknown

D. Stage of Program

☐ Model/demonstration programs
☐ Early stage programs
☒ Established/institutionalized programs
☐ Unknown

E. Concurrent or Historical Intervention Exposure

☒ Current exposure  

Services from local day treatment programmes.

☐ Prior exposure
☐ Unknown

Section 2  Key Features for Coding Studies and Rating Level of Evidence/Support

A Measurement (Estimating the quality of the measures used to establish effects)
(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes (see following table for a detailed breakdown on the outcomes)

☐ Yes
☐ No
☒ Unknown/unable to code  Not reported.

A2 Multi-method (at least two assessment methods used)

☐ Yes
☒ No  Only one measure (Psychoeducational Profile- Revised) was used to examine group changes.

☐ N/A
☐ Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)

☐ Yes
☒ No

☐ N/A
☐ Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

☐ Yes validated with specific target group
☐ In part, validated for general population only
☐ No
☒ Unknown/unable to code  Not reported for Psychoeducational Profile.

Overall Rating of Measurement: 3 2 1 0
B Comparison Group

B1 Type of Comparison group
- Typical intervention
- Attention placebo
- Intervention element placebo
- Alternative intervention
- Pharmacotherapy
- No intervention
- Wait list/delayed intervention
- Minimal contact
- Unable to identify type of comparison

B2 Overall rating of judgment of type of comparison group
- Very low
- Low
- Moderate
- High
- Very high
- Unable to identify comparison group

B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc were counter-balanced across intervention)
- By change agent
- Statistical (analyse includes a test for intervention)
- Other
- Not reported/None

B4 Group equivalence established
- Random assignment
- Posthoc matched set
- Statistical matching
- Post hoc test for group equivalence

B5 Equivalent mortality
- Low attrition (less than 20% for post)
- Low attrition (less than 30% for follow-up)

Intent to intervene analysis carried out?

Findings_____________

Overall Level of Evidence = 2

3= Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence
C Appropriate Statistical Analysis

Analysis 1

☑ Appropriate unit of analysis  Repeated measures ANOVA
☐ Familywise/experimenter wise error rate controlled when applicable  None reported.
☐ Sufficiently large N  No. Minimum of 64 participants for a medium (0.5) effect size.

Analysis 2

☐ Appropriate unit of analysis
☐ Familywise/experimenter wise error rate controlled when applicable
☐ Sufficiently large N

Analysis 3

☐ Appropriate unit of analysis
☐ Familywise/experimenter wise error rate controlled when applicable
☐ Sufficiently large N

Overall Rating of Analysis:  ☐ 3  ☐ 2  ☑ 1  ☐ 0  *Score of 1 as appropriate unit of analysis used.*
Coding Protocol

Name of Coder: Date: 17.01.14

Full Study Reference in proper format:

Intervention Name (description of study): TEACCH as an autism treatment model.

Study ID Number: 04

☐ Type of Publication:  
☐ Book/Monograph  
☒ Journal Article  
☐ Book Chapter  
☐ Other (specify):

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)
☐ Completely randomized design  
☐ Randomized block design (between participants, e.g., matched classrooms)  
☐ Randomized block design (within participants)  
☐ Randomized hierarchical design (nested treatments)

A2. Nonrandomized designs (if non-random assignment design, select one of the following)
☐ Nonrandomized design  
☐ Nonrandomized block design (between participants)  
☐ Nonrandomized block design (within participants)  
☐ Nonrandomized hierarchical design  
☒ Optional coding for Quasi-experimental designs  

Waiting period before treatment provided comparison data

A3. Overall confidence of judgment on how participants were assigned (select on of the following)
☐ Very low (little basis)  
☐ Low (guess)  
☐ Moderate (weak inference)  
☐ High (strong inference)  
☐ Very high (explicitly stated)  
☐ N/A  
☐ Unknown/unable to code

B Participants

Total size of sample (start of study): 15

Intervention group sample size: 15

Control group sample size:
C. Type of Program

- Universal prevention program
- Selective prevention program
- Targeted prevention program
- Intervention/Treatment
- Unknown

D. Stage of Program

- Model/demonstration programs
- Early stage programs
- Established/institutionalized programs
- Unknown

E. Concurrent or Historical Intervention Exposure

- Current exposure
- Prior exposure
- Unknown

Section 2  Key Features for Coding Studies and Rating Level of Evidence/Support

A Measurement (Estimating the quality of the measures used to establish effects)
(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes (see following table for a detailed breakdown on the outcomes)

- Yes
- No
- Unknown/unable to code

A2 Multi-method (at least two assessment methods used)

- Yes
- No
- N/A Questionnaires sought parents’ opinion of the child’s behaviour change.
- Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)

- Yes
- No
- N/A
- Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)

- Yes validated with specific target group
- In part, validated for general population only
- No
- Unknown/unable to code

Overall Rating of Measurement: 3 2 1 0
B Comparison Group

B1 Type of Comparison group
- Typical intervention
- Attention placebo
- Intervention element placebo
- Alternative intervention
- Pharmacotherapy
- No intervention
- Wait list/delayed intervention
- Minimal contact
☑ Unable to identify type of comparison
Contrast data collected from participants in waiting period prior to intervention.

B2 Overall rating of judgment of type of comparison group
- Very low
- Low
- Moderate
- High
- Very high
☑ Unable to identify comparison group

B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc were counter-balanced across intervention)
- By change agent
- Statistical (analyse includes a test for intervention)
- Other
☑ Not reported/None

B4 Group equivalence established
- Random assignment
- Posthoc matched set
- Statistical matching
- Post hoc test for group equivalence

B5 Equivalent mortality
- Low attrition (less than 20 % for post)
- Low attrition (less than 30% for follow-up)
- Intent to intervene analysis carried out?
Findings_____________

Overall Level of Evidence  = 0
3= Strong Evidence  2=Promising Evidence  1=Weak Evidence  0=No Evidence
C Appropriate Statistical Analysis

Analysis 1

☐ Appropriate unit of analysis  multivariate analysis of variance for repeated measures
☐ Familywise/experimenter wise error rate controlled when applicable
☐ Sufficiently large N  No. Minimum of 64 participants for a medium (0.5) effect size.

Analysis 2

☐ Appropriate unit of analysis
☐ Familywise/experimenter wise error rate controlled when applicable
☐ Sufficiently large N

Analysis 3

☐ Appropriate unit of analysis
☐ Familywise/experimenter wise error rate controlled when applicable
☐ Sufficiently large N

Overall Rating of Analysis: ☐ 3 ☐ 2 ☑ 1 ☐ 0 Score of 1 as appropriate unit of analysis used.
Coding Protocol

Name of Coder:  Date: 17.01.14

Full Study Reference in proper format:

Intervention Name (description of study): Home TEACCHing programme.

Study ID Number: 05

<table>
<thead>
<tr>
<th>Type of Publication:</th>
</tr>
</thead>
<tbody>
<tr>
<td>☒ Book/Monograph</td>
</tr>
<tr>
<td>☐ Journal Article</td>
</tr>
<tr>
<td>☐ Book Chapter</td>
</tr>
<tr>
<td>☐ Other (specify):</td>
</tr>
</tbody>
</table>

1. General Characteristics

A. General Design Characteristics

A1. Random assignment designs (if random assignment design, select one of the following)

| Completely randomized design |
| Randomized block design (between participants, e.g., matched classrooms) |
| Randomized block design (within participants) |
| Randomized hierarchical design (nested treatments) |

A2. Nonrandomized designs (if non-random assignment design, select one of the following)

| Nonrandomized design |
| Nonrandomized block design (between participants) |
| Nonrandomized block design (within participants) |
| Nonrandomized hierarchical design |
| Optional coding for Quasi-experimental designs |

A3. Overall confidence of judgment on how participants were assigned (select one of the following)

| Very low (little basis) |
| Low (guess) |
| Moderate (weak inference) |
| High (strong inference) |
| Very high (explicitly stated)  
  
  Randomly assigned to treatment or waitlist.  |
| N/A |
| Unknown/unable to code |

B Participants

Total size of sample (start of study): 20

Intervention group sample size: 10

Control group sample size: 10
C. Type of Program

- Universal prevention program
- Selective prevention program
- Targeted prevention program
- Intervention/Treatment
- Unknown

D. Stage of Program

- Model/demonstration programs
- Early stage programs  Although TEACCH is an established programme, the format of intervention used in this study is model/new programme.
- Established/institutionalized programs
- Unknown

E. Concurrent or Historical Intervention Exposure

- Current exposure  To Speech and Language Therapy and Occupational Therapy.
- Prior exposure
- Unknown

Section 2  Key Features for Coding Studies and Rating Level of Evidence/Support

A Measurement (Estimating the quality of the measures used to establish effects)
(Rating Scale: 3= Strong Evidence, 2=Promising Evidence, 1=Weak Evidence, 0=No Evidence)

A1 The use of the outcome measures produce reliable scores for the majority of the primary outcomes (see following table for a detailed breakdown on the outcomes)
- Yes
- No
- Unknown/unable to code  Paper states that the Mullen Scales of Early Learning and Scales of Independent Behaviour-Revised both have demonstrated good reliability and validity but no scores provided.

A2 Multi-method (at least two assessment methods used)
- Yes  Child outcomes investigated through the Mullen Scales of Early Learning and Scales of Independent Behaviour
- No
- N/A
- Unknown/unable to code

A3 Multi-source (at least two sources used self-reports, teachers etc.)
- Yes
- No  Child outcomes based on professional assessment. Questionnaire for parents did not assess the programme efficacy.
- N/A
- Unknown/unable to code

A4 Validity of measures reported (well-known or standardized or norm-referenced are considered good, consider any cultural considerations)
- Yes validated with specific target group
- In part, validated for general population only
- No
- Unknown/unable to code  Paper states that the Mullen Scales of Early Learning and Scales of Independent Behaviour-Revised both have demonstrated good reliability and validity but no scores provided.

Overall Rating of Measurement: 3  2  1  0
**B Comparison Group**

**B1 Type of Comparison group**
- Typical intervention
- Attention placebo
- Intervention element placebo
- Alternative intervention
- Pharmacotherapy
- No intervention
- Wait list/delayed intervention
- Minimal contact
- Unable to identify type of comparison group

**B2 Overall rating of judgment of type of comparison group**
- Very low
- Low
- Moderate
- High
- Very high
- Unable to identify comparison group

**B3 Counterbalancing of change agent (participants who receive intervention from a single therapist/teacher etc were counter-balanced across intervention)**
- By change agent
- Statistical (analyse includes a test for intervention)
- Other
- Not reported/None

**B4 Group equivalence established**
- Random assignment
- Posthoc matched set
- Statistical matching
- Post hoc test for group equivalence

**B5 Equivalent mortality**
- Low attrition (less than 20 % for post)
- Low attrition (less than 30% for follow-up)
- Intent to intervene analysis carried out?

Findings:__________

**Overall Level of Evidence = 2**

3= Strong Evidence   2=Promising Evidence   1=Weak Evidence   0=No Evidence
C Appropriate Statistical Analysis

Analysis 1

☑ Appropriate unit of analysis  Repeated measures analysis of variance.
☐ Familywise/experimenter wise error rate controlled when applicable  None reported.
☐ Sufficiently large N  No. Minimum of 64 participants for a medium (0.5) effect size.

Analysis 2

☐ Appropriate unit of analysis
☐ Familywise/experimenter wise error rate controlled when applicable
☐ Sufficiently large N

Analysis 3

☐ Appropriate unit of analysis
☐ Familywise/experimenter wise error rate controlled when applicable
☐ Sufficiently large N

Overall Rating of Analysis:  ☐ 3  ☐ 2  ☑ 1  ☐ 0
Appendix D: Weighting of Studies

(A) Methodological Quality

1. Measures

<table>
<thead>
<tr>
<th>Weighting</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High</strong></td>
<td>Used 2 measures which have reliability of .85 or above (reported in text). Collected data from at least 2 sources and methods.</td>
</tr>
<tr>
<td></td>
<td>Used measures that produce reliable scores of .70 for the population under study (do not need to be reported in text).</td>
</tr>
<tr>
<td><strong>Medium</strong></td>
<td>Data collected using either multiple methods AND/OR from multiple sources. A case for validity does not need to be presented.</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>Uses measures that produce reliable scores of 0.50 for the population under study OR that is well known for the target population.</td>
</tr>
</tbody>
</table>

2. Comparison Group

<table>
<thead>
<tr>
<th>Weighting</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High</strong></td>
<td>At least 1 type of ‘active’ comparison group. Initial group equivalency must be established. Counterbalanced change agents. Equal mortality rate between the intervention and comparison group (of under 20% attrition)</td>
</tr>
<tr>
<td></td>
<td>Used at least a ‘no intervention group’ comparison group. Demonstrated 2 of the following: 1. Counterbalancing of change agents, 2. Group equivalency, 3. Equivalent mortality with low attrition (under 20%). If equivalent mortality is not demonstrated, conducted an intent-to-intervene analysis.</td>
</tr>
<tr>
<td><strong>Medium</strong></td>
<td>Uses a comparison group. Demonstrated 1 of the following: 1. Counterbalancing of change agents, 2. Group equivalency, 3. Equivalent mortality with low attrition (under 20%). If equivalent mortality is not demonstrated, conducted an intent-to-intervene analysis.</td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>Demonstrated 1 of the following: 1. Counterbalancing of change agents, 2. Group equivalency, 3. Equivalent mortality with low attrition (under 20%). If equivalent mortality is not demonstrated, conducted an intent-to-intervene analysis.</td>
</tr>
</tbody>
</table>

3. Analysis

<table>
<thead>
<tr>
<th>Weighting</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High</strong></td>
<td>Demonstrated all of the following: 1. Conducted an appropriate statistical analysis 2. Corrected for family wise / experimenter error (if appropriate) 3. Used a sufficiently large sample size 4. Provided enough information for the effect size to be calculated</td>
</tr>
<tr>
<td><strong>Medium</strong></td>
<td>Demonstrated two of the following: 1. Conducted an appropriate statistical analysis 2. Corrected for family wise / experimenter error (if appropriate)</td>
</tr>
</tbody>
</table>
3. Used a sufficiently large sample size
4. Provided enough information for the effect size to be calculated

Demonstrated one of the following:
1. Conducted an appropriate statistical analysis
2. Corrected for family wise / experimenter error (if appropriate)
3. Used a sufficiently large sample size
4. Provided enough information for the effect size to be calculated

Overall Methodological Quality

The following scores were assigned to calculate the overall methodological quality of studies:

**High** weightings = 3
**Medium** weightings = 2
**Low** weightings = 1
Studies which did not meet criteria = 0

<table>
<thead>
<tr>
<th>Studies</th>
<th>Measures</th>
<th>Comparison Group</th>
<th>Analysis</th>
<th>Overall Quality of Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Braiden, McDaniel, McCrudden, Janes, &amp; Crozier, 2012</td>
<td>Medium Score = 2</td>
<td>N/A</td>
<td>Medium Score = 2</td>
<td>1.33</td>
</tr>
<tr>
<td>McConkey, Truesdale-Kennedy, Crawford, McGreevy, Reavey &amp; Cassidy, 2010</td>
<td>High Score = 3</td>
<td>Low Score = 1</td>
<td>Medium Score = 2</td>
<td>2</td>
</tr>
<tr>
<td>Ozonoff &amp; Cathcart, 1998</td>
<td>Low Score = 1</td>
<td>Medium Score = 2</td>
<td>Medium Score = 2</td>
<td>1.67</td>
</tr>
<tr>
<td>Short, 1984</td>
<td>High Score = 3</td>
<td>N/A</td>
<td>Low Score = 1</td>
<td>1.33</td>
</tr>
<tr>
<td>Welterlin, Turner-Brown, Harris, Mesibov, &amp; Delmolino, 2012</td>
<td>Medium Score = 2</td>
<td>Medium Score = 2</td>
<td>Medium Score = 2</td>
<td>2</td>
</tr>
</tbody>
</table>

**B) Relevance of Methodology**

This weighting reviews the appropriateness of the evidence for answering the review question.
<table>
<thead>
<tr>
<th>Weighting</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Uses random assignment of participants. Uses an active comparison group. Takes pre and post measures for both the intervention and control group.</td>
</tr>
<tr>
<td>Medium</td>
<td>Uses a comparison group. Demonstrates group equivalence. Takes pre and post measures for both the intervention and control group.</td>
</tr>
<tr>
<td>Low</td>
<td>Takes pre and post measures.</td>
</tr>
</tbody>
</table>

(C) Relevance of Evidence to Review Question

This weighting is a review-specific judgement about the relevance of the focus of the evidence for the review question.

<table>
<thead>
<tr>
<th>Weighting</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Uses TEACCH exclusively. Employed exclusively as a home programme. Measures child skills/behaviour pre and post intervention alongside a control group.</td>
</tr>
<tr>
<td>Medium</td>
<td>Uses TEACCH alongside other approaches. Employed exclusively as a home programme. Measures child skills/behaviour pre and post intervention.</td>
</tr>
<tr>
<td>Low</td>
<td>Uses elements of TEACCH. Employed as a home programme but not exclusively. Measures child skills/behaviour pre and post intervention.</td>
</tr>
</tbody>
</table>

(D) Overall Weight of Evidence

This is an overall assessment of the extent to which the study contributes evidence to answer the review question. It is calculated by assigning the following scores:

- **High** weightings = 3
- **Medium** weightings = 2
- **Low** weightings = 1
- Studies which did not meet criteria = 0

The scores were then averaged and the following weighting of evidence was awarded for each score:

- **High** weighting = More than 2.5
*Medium* weighting = Between 1.5 and 2.4  
*Low* weighting = Less than 1.4

<table>
<thead>
<tr>
<th>Studies</th>
<th>(A) Quality of Methodology</th>
<th>(B) Relevance of Methodology</th>
<th>(C) Relevance of evidence to the review question</th>
<th>(D) Overall Weight of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Braiden, McDaniel, McCrudden, Janes, &amp; Crozier, 2012</td>
<td>1.33</td>
<td>Low</td>
<td>Medium Score = 1</td>
<td>Low</td>
</tr>
<tr>
<td>McConkey, Truesdale-Kennedy, Crawford, McGreevy, Reavey &amp; Cassidy, 2010</td>
<td>2</td>
<td>Low</td>
<td>Medium Score = 1</td>
<td>Medium</td>
</tr>
<tr>
<td>Ozonoff &amp; Cathcart, 1998</td>
<td>1.67</td>
<td>Medium Score = 2</td>
<td>High Score = 3</td>
<td>Medium</td>
</tr>
<tr>
<td>Short, 1984</td>
<td>1.33</td>
<td>Low</td>
<td>Medium Score = 2</td>
<td>Low</td>
</tr>
<tr>
<td>Welterlin, Turner-Brown, Harris, Mesibov, &amp; Delmolino, 2012</td>
<td>2</td>
<td>Medium Score = 2</td>
<td>High Score = 3</td>
<td>Medium</td>
</tr>
</tbody>
</table>