APPENDIX 4: EVIDENCE TABLES

Evidence tables for included studies assessing the effectiveness of applied behavioural therapy for people with autism spectrum disorder.

1. Level I studies:

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Parr 2007) BMJ Clinical evidence review UK SR of SRs and RCTs (lower order studies only included if no RCTs)	**Evidence level:** I **Aim:** To assess the effects of interventions in children with autism. **Search period**: May 2006: different periods for different databases. **Databases:** Medline: 1986 to May 2006 Embase: 1986 to April 2006 The Cochrane Library, Issue 2, 2006 NHS Centre for Reviews and Dissemination (CRD): May 2006 DARE: May 2006 HTA: May 2006 TRIP: May 2006 NICE website: May 2006. (search terms not published but always used by BMJ Clinical	**Inclusion criteria:** ▪ systematic reviews, RCTs, quasi-randomised trials, controlled clinical trials and prospective and retrospective cohort studies (only for interventions for which no RCTs were identified) ▪ studies including at least 20 participants ▪ participants either children or adolescents with autism or studies including children with other autism spectrum disorders that provided a subgroup analysis of at least 20 individuals with autism. **Exclusion criteria:**	Applied Behavioural Analysis was defined as early intensive behavioural intervention (EIBI) or Lovaas therapy. Autism Preschool Programme was defined as a programme offering parents and caregivers support in behavioural and language development methods that are then carried out at home or during day care. PECS was defined as a behaviourally based program designed to help young children learn to initiate requests and communicate their needs. ABA: One systematic review (Diggle et al. 2002) – *not included* and one study described as an RCT (Eikeseth et al. 2002) but categorised as non-randomised experimental study in this report, were identified. The systematic review did not identify any RCTs meeting the inclusion criteria of the review. The Eikeseth et al (Eikeseth et al. 2002) trial (25 children with autism aged 4 to 7 years) compared ABA to eclectic treatment. Benefits were found for ABA vs eclectic treatment for: IQ, language comprehension, expressive language, communication and overall adaptive behaviour but not for daily living and socialization. Recommendation: *Compared with eclectic treatment, ABA may improve IQ and language skills (low quality evidence – from GRADE grading [RCT with 1 point deducted for sparse data, 1 point deducted for conflicting results from different outcomes)* Autism Preschool Programme: One systematic review (Diggle et al. 2002) was identified. The	▪ diagnosis of autism spectrum disorders (ASD) includes children with AS and PDD NOS. This review refers only to children and adolescents with autism. ▪ findings of the review may not be generalisable to children with other ASDs ▪ strict inclusion criteria meant that the authors excluded many studies with small numbers of participants, a range of ASD diagnoses and a range of abilities as they considered that combining data from these studies was unlikely to be scientifically valid or clinically useful ▪ clearly defined methodology and comprehensive searches ▪ a number of studies were identified in this synthesis that meet the inclusion criteria of this review (Eikeseth et al. 2002) (Howlin et al. 2007) (Yoder and Stone 2006b) (Yoder and Stone 2006a) ▪ suggestions made for future research. **Quality score:** very good (++)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
	Evidence searching team) Hand-searching of all issues of the Journal of Child and Adolescent Psychopharmacology was performed and reference lists of papers were cross-checked.	▪ children/adolescents with AS or PDD NOS. Appraisal of studies not formally undertaken but inclusion criteria limits studies to RCTs and SRs where available and summary statements reflect informal evaluation of the quality of the studies.	Diggle systematic review did not identify any RCTs meeting the inclusion criteria of the review. Recommendation: *We found no clinically important results about the effects of Autism Preschool Programme compared with no active treatment or with other treatments in children with autism.* Picture Exchange Communication System: No studies were identified that met the inclusion criteria. However, after the publication of this issue, (Howlin et al. 2007) and (Yoder and Stone 2006a) (Yoder and Stone 2006b) published data supporting the use of PECs. Recommendation: *We found no clinically important results about the effects of the PECS on symptoms of autism in children.* **Authors' conclusions:** "Despite the lack of robust RCT evidence, there is consensus based on clinical experience that applied behaviour analysis. . . and PECs [are] likely to be beneficial in children with autism. Further large well designed RCTs with comparable control groups and long term follow up are required to assess the effectiveness of ABA . . PECs . . . .and the Autism preschool programme". The authors categorised ABA, PECs and Autism Preschool Programme as "likely to be beneficial". The authors have added a rider to their conclusion: "In the absence of robust RCT evidence in children with autism, categorisation is based on observational evidence and strong consensus belief that these interventions are likely to be beneficial".	

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Scottish Intercollegiate Guidelines Network 2007) UK SR	**Evidence level:** I **Aim:** To review the assessment, diagnosis and clinical interventions for children and young people up to the age of 16 years and make evidence based recommendations to inform clinical service provision. **Search period**: 1996 – 2006. **Databases:** Medline, Embase, Cinahl, PsychINFO, Caredata, Social Work Abstracts and the Cochrane Library. (no search terms provided but detailed clinical questions and terms listed for each question which presumably were used in the searching) Also, internet searching of various websites including NZGG, NeLH Guidelines Finder, GIN, NICE and the US National Guidelines Clearinghouse. The main searches were supplemented by material identified by individual members of the guideline group.	**Inclusion criteria:** ▪ based on relevance to pre-specified clinical questions ▪ children and young people up to the age of 16 years with a diagnosis of ASD ▪ only educational interventions that may influence clinical outcomes ▪ other criteria not reported. **Exclusion criteria:** ▪ epidemiology of ASD. Recommendations graded according to study design and quality.	1 systematic review of RCTs (Bassett et al. 2000) was identified that assessed whether intensive behavioural intervention for preschool children with ASD could achieve normalisation (capacity to follow a normal academic curriculum in a mainstream school). The other included studies had considerable methodological flaws. The review concluded that: *A causal relationship could not be established between a particular program of intensive behavioural intervention and the achievement of 'normal functioning'.* A comprehensive literature search did not find any good quality evidence for other intensive behavioural interventions. 1 systematic review of observational studies (Matson 1996) was identified that assessed focal treatments for children and young people with ASD. The included studies varied in their quality. The authors concluded that: *Focal behavioural interventions consistently resulted in positive behavioural outcomes across a wide range of target areas*. **Guideline recommendations:** A: The Lovaas programme should not be presented as an intervention that will lead to normal functioning. B: Behavioural interventions should be considered to address a wide range of specific behaviours in children and young people with ASD, both to reduce symptom frequency and severity and to increase the development of adaptive skills. ✓: Healthcare professionals should be aware that some aberrant behaviours may be due to an underlying lack of skills and also may represent a child's strategy for coping with their individual difficulties and circumstances.	▪ inclusion criteria not clear but guideline developed according to explicit SIGN methodology which includes defined inclusion and exclusion criteria so this it likely to be because of insufficient reporting ▪ no indication that primary studies also considered in the drafting of recommendations but SIGN methodology manual requires a comprehensive literature search with recommendations drafted according to consistent methodological standards, so likely to be a lack of sufficient reporting ▪ SIGN grading system used to grade recommendations ▪ insufficient detail provided on the type of behavioural interventions in the recommendation that were considered useful or the particular outcomes that were improved ▪ one study was identified which met the inclusion criteria of this review (Bassett et al. 2000). **Quality score:** very good (++)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Roberts and Prior 2006) Australia SR	**Evidence level:** I **Aim:** To review the research literature relating to the management and treatment of young children with autism in order to identify the most effective models of best practice. **Search period**: 2003 - 2006 **Databases**: CINAHL ISI Web of Science Wiley Interscience Medline PsychINFO Embase PUBMED ERIC (search terms included for all). Also, a number of reviews and guidelines were accessed: (Bassett et al. 2000) Crewther 2003 Evans 2003 Jordan 1998 Librera 2004 (Ludwig and Harstall 2001) MADSEC 2000 Medical Research Council 2001 (McGahan 2001)	**Inclusion criteria:** not described **Exclusion criteria:** not described No formal appraisal of included studies but shortcomings of studies discussed in narrative format.	15 primary studies and 3 reviews were identified that assessed the effectiveness of programs using behavioural principles in children with ASD. The majority of the primary studies were case series; there were 4 quasi experimental studies and 2 RCTs. 13 primary studies assessed the effects of ABA and 2 primary studies assessed the effects of PECS. The authors outlined the results from the two seminal studies (Lovaas 1987 and McEachin 1993) and other examples of "partial replications" of the Lovaas method (Anderson 1987; Birnbrauer 1993); (Smith et al. 2000) (Eikeseth et al. 2002) (Sallows and Graupner 2005). There were major methodological flaws in the earlier studies and 2 out of 3 of the later studies had comparisons between 2 types of intensive treatment based on Lovaas. The lack of a non treatment control group or comparison group receiving a different type of intervention meant that conclusions could not be drawn in relation to outcomes compared to no treatment or different treatment programs. 4 case series studies assessed the role of behavioural interventions in classroom settings (Fenske 1985; Harris 1991; Meyer 2001 and Romanczyk 2001). There were positive results from classroom based behavioural intervention programs but lack of a control group meant that there was no conclusive evidence to support the efficacy of the approach. One case series (Bibby 2002) found a benefit for children with ASD with parent managed home based intensive early intervention after a period of almost 3 years but the results did not support the findings of the original Lovaas study. It is not clear what factors led to the improvement and there was no control group. Intensity: The authors concluded that there was no conclusive evidence for the optimal intensity of EIBI. In the (Sallows and Graupner 2005) RCT, there was little difference between children who received more intensive clinic directed behavioural intervention and a group who received slightly less intensive parent-directed therapy. Recovery The authors noted that claims of "recovery" have led to the most controversy. The authors of (Bassett et al. 2000) suggest that	▪ comprehensive narrative discussion of research ▪ inclusion and exclusion criteria for studies not described ▪ very short search period for literature and some of the included studies were outside the described search period ▪ good discussion of the methodological limitations of the included studies ▪ some of the included studies meet the criteria for the current review (Bassett et al. 2000) (Ludwig and Harstall 2001) (McGahan 2001) (Smith et al. 2000) (Sallows and Graupner 2005) (Eikeseth et al. 2002) **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
	Newfoundland 2003 Osbourn 2004 Perry 2003 Special Programs Branch 2000		Lovaas et al have not limited their effectiveness claims to achieving developmental gains but have claimed that "normal functioning" is achievable with ABA. Specific skills MADSEC 2000 and DeMeyer 1981 note that research evidence suggests that behavioural interventions have been the most effective of all interventions used with ASD children. Most educational programs with ASD children incorporate at least some behavioural strategies. PECS 2 case studies found an effect of PECS on children's verbal skills but there was no control group to put these findings into perspective. **Authors' conclusions:** EIBI programmes produce positive outcomes for children with autism but "there continues to be a great deal of controversy about particular behavioural interventions and programs and differences in the interpretation of research findings. . . The controversy revolves around outcome claims, exclusivity, extensive use and personnel". "Blanket statements about the effectiveness of behavioural interventions may be misleading given the variation in children, families, therapists, contexts and methods". Conclusions based on those of (Bassett et al. 2000): ▪ The original Lovaas and McEachin studies are inadequate to determine whether DTT achieves "normal functioning" ▪ There is insufficient evidence of effectiveness to establish a relationship between the amount of any form of EIBI and overall outcome ▪ RCTs of alternative EIBI are ethical and feasible ▪ There is insufficient evidence of effectiveness to conduct a cost benefit analysis of EIBI in terms of normalisation of children with autism	

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Burrows 2004) Canada Society evidence based report	**Evidence level:** I Canadian Paediatric Society (CPS) Position Statement, which includes a review process. **Aim:** To briefly describe the main educational interventions (programmes) that are intended to result in global improvement in autism and to review the status of the evidence regarding their effectiveness. **Search period**: 1995 – September 2003 **Databases**: Medline, PsycINFO (basic search terms provided).	**Inclusion criteria:** ▪ not specified apart from scope of review aim. **Exclusion criteria:** ▪ behavioural techniques that limit their aim to changing specific areas of functioning in autism.	The review identified and described 9 primary studies, three of which are include in the current review: (Sheinkopf and Siegel 1998); (Smith et al. 2000); (Jocelyn et al. 1998), and 5 reviews, one of which was included in this report: (Bassett et al. 2000). Key points: ▪ critically describes Lovaas's 1987 trial, outlining deficiencies including non-random group assignment, poor description of treatment implementation, variety of assessment tests, including people with a unusually favourable prognosis, and outcome measures limited to IQ and class placement. Author concludes, "the Lovaas study does not meet criteria for an empirically supported treatment." ▪ summarises other studies of Lovaas model, concluding, "as a group, these studies show that the children who received intensive therapy for a minimum of 15 to 20 hours per week for one to two years had greater functional gains than controls who received no extra treatment." ▪ identifies the RCT by Smith et al (Smith et al. 2000) as methodologically the strongest, but notes its small sample size, and lack of a standardised diagnostic instrument and follow-up diagnostic assessment ▪ notes that most published studies that demonstrated treatment efficacy involved children younger than 48 months ▪ the effectiveness of parent-mediated treatment is unknown. ▪ reviews various examples of "normalised teaching" interventions including TEACCH, LEAP, and Floor Time, all described as being "in a preliminary phase of investigation." Makes following practice recommendations: ▪ set a target of a minimum of 15 hours per week of structured, individualised teaching ▪ involve the family in service provision ▪ including ongoing programme evaluation and adjustment to meet the child's needs. ▪ resources must be sufficient and accessible ▪ funded programmes should include a research arm **Authors' conclusions:** "The quality of the studies on educational treatment programs for children with autism are	▪ position statement including a narrative review (with some evidence of systematic searching) on early intervention for children with ASD ▪ only two databases searched and no other searching methods used ▪ no checklists or appraisal details given ▪ findings critically summarised in text, noting strengths and weaknesses of individual studies ▪ no tables and description of study methods very brief ▪ recommendations for future research and implications for practice briefly given ▪ conclusions based on various primary and secondary studies referred to in text ▪ included 4 studies eligible for inclusion in current review: (Bassett et al. 2000); (Jocelyn et al. 1998); (Sheinkopf and Siegel 1998); (Smith et al. 2000). **Quality score:** fair (-)

			suboptimal. However the studies do show a trend toward showing a positive outcome from intervention. There is no evidence to support adopting a single autism treatment programme as the gold standard".	

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Doughty 2004) New Zealand SR, including RCTs	**Evidence level:** I **Aim:** Not explicitly stated, but focus was to consider the "most recent and best evidence" for the effectiveness of behavioural and skill-based interventions that are used to manage ASD in young children. **Search period**: 2000 - 2003 **Databases**: Medline, Embase, CINAHL, PsycINFO, Amed, ERIC, Current Contents, Web of Science, Index New Zealand, Cochrane Database of Systematic Reviews, Evidence-based medicine reviews, DARE, NHS Economic Evaluation Database, Health Technology Assessment Database, and various evidence-based, guideline, government and related health professional websites (search terms provided).	**Inclusion criteria:** ▪ English language publications ▪ at least 75% of sample diagnosed with ASD, or results reported separately ▪ sample mean age < 8 years, or age range 1-7 years ▪ sample size 5 or more people in each study arm ▪ standardised and/or quantitative outcome measure relating to core symptoms of ASD or associated features for people with ASD or carers/family ▪ NHMRC evidence level III.2 or above. **Exclusion criteria:** ▪ pharmacological, surgical, physical therapy or dietary interventions, auditory and sensory integration, studies reporting only on specific learning difficulties ▪ absence of behavioural techniques or outcomes ▪ SCED studies.	Retrieved 108 publications, 10 of which met selection criteria: 5 primary studies and 5 systematic reviews. The 5 primary studies included 3 RCTs, one non-randomised experimental studies (with concurrent controls), and one "cohort study" with no concurrent comparator. All considered ABA interventions except one study that evaluated a social-developmental intervention. Reports that preliminary evidence suggests that early intervention may lead to gains in specific domains, however further research is required to overcome methodological limitations. These include: providing detailed manual of intervention, using random assignment or matching, blind assessment of outcome, variety of standardised assessment tools, measures of treatment intensity, larger samples and multi-site collaborations. Reviewer noted need to evaluate active ingredients of treatment as comparators can vary in critical ways (e.g., whether one-to-one or group, or whether home, school or clinic based), which may relate to treatment effectiveness. All five reviews appraised concluded that to date there was insufficient evidence to allow conclusions to be drawn about best practice, or one programme's effectiveness over another. Researchers also have not been able to establish a relationship between intensity of programme delivery (amount of treatment hours per day) and outcome (Bassett et al. 2000). **Authors' conclusions:** "The majority of recent primary studies reviewed here document some improvement associated with intervention," "however it remains to be determined if any one early/and/or intensive intervention programme is more effective than another". "The primary studies included in this Tech Brief cover a range of interventions and comparisons and it was not clear that the definition of intensive behavioural treatment, parent training or parent-managed behavioural therapy were uniform across individual studies evaluating similar approaches to early intervention. The intensity and duration of the intervention offered was only documented in some of the studies and sample sizes were mostly small. "	▪ focus on early behavioural and skill-based interventions for children with autism ▪ narrow publication period considered and excluded SCED studies and SR's of SCED studies ▪ extensive range of databases searched, but no other search methods employed ▪ no appraisal checklists mentioned ▪ studies described critically in the text, and detailed evidence tables with full description of interventions, methods and study results ▪ textual summary of appraised studies, their strengths and limitations ▪ list of ongoing, unpublished trials ▪ brief summary of methodological issues and areas for future research ▪ conclusions based on 5 primary (4 ABA) and 5 secondary appraised studies ▪ for ABA studies, 3 RCTs, one concurrent controlled, non-randomised experimental study, and 1 "cohort study", however no concurrent comparator and results compared to historical data, therefore a prospective case series. ▪ 3 primary studies (Drew et al. 2002; Eikeseth et al. 2002; Smith et al. 2000) and 4 secondary studies (Bassett et al. 2000; Finch and Raffaele 2003; Ludwig and Harstall 2001; McGahan 2001) met criteria for inclusion in current review. **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Finch and Raffaele 2003) SR Canada	**Evidence level:** I **Aim:** To critically appraise the evidence on intensive behavioural intervention (IBI) for children with autism. **Search period**: not stated **Databases**: AMED, CINAHL, Cochrane Reviews, Medline, and "occupational therapy specific databases" (search terms provided). In addition, searched "relevant books". Consulted clinicians for other sources.	**Inclusion criteria:** ▪ published in Journals ▪ experimental trials of IBI ▪ children under age of 8 years ▪ children with a diagnosis of autism or PDD. **Exclusion criteria:** ▪ none reported Studies appraised using the Critical Review Form for Quantitative Studies (Law et al, 1998).	7 publications included in the review: including an RCT, 4 matched-pair studies, a cohort study, and a retrospective case review study. All studies reported positive gains in IQ and class placement for the majority of children, but gains varied considerably between individuals. There were inconsistent findings in social and adaptive functioning, varying from 47% of children with EIBI being indistinguishable from average children (McEachin et al, 1993) to no significant improvement from an RCT (Smith et al, 2000). Found fewer positive gains for those in community or home settings compared with those under the controlled setting of a university setting. Children receiving less intensive treatment (under 25 vs 40 hours) also attained less significant results. Children receiving EIBI from less experienced instructors made smaller gains. One study investigated long-term results of EIBI and found gains in IQ, class placement and adaptive behaviour persisted after 7 years (McEachin et al, 1993). However, the study suffered from methodological several limitations and excluded children who were profoundly mentally retarded. Smith et al (2000) reported that children with PPD-NOS may especially benefit from EIBI compared to those with autism. Methodological limitations included that not all children had received a clinical diagnosis of autism, diagnostic criteria for autism had changed, measures and outcomes varied greatly between studies, comparators were not similar to EIBI in intensity and duration, and only one study reported long term follow-up. **Authors' conclusions:** "Research indicates some positive gains in IQ, class placement and adaptive behaviour for most children receiving IBI. However it is difficult to conclude strongly that IBI is effective for all children with autism based on limitations in the evidence."	▪ focus on IBI in early childhood (or EIBI) ▪ broad range of databases searched (but not PsycINFO) and supplemented by opportunistic sources. ▪ appraisal checklists used ▪ studies described critically in the text, no tables, relatively brief report ▪ conclusions based on 7 studies ▪ described range of quasi experimental and experimental designs including one RCT (Smith et al, 2000) ▪ 2 studies met criteria for inclusion in current review - (Smith et al. 2000); (Sheinkopf and Siegel 1998) **Quality score:** very good (++)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Chorpita et al. 2002) Hawaii SR	**Evidence level:** I **Aim:** To examine the efficacy and effectiveness of child treatments for anxiety disorders, depression, attention deficit hyperactivity disorder, conduct and oppositional disorders, and autistic and related disorders. **Search period**: 1980 onwards for database search **Databases**: PsycINFO (search terms not given). Studies identified in major literature reviews (including those from reference lists of retrieved studies), personal communication with national scholars, and nominations from the Lonigan and Elbert Taskforce on Empirically Supported Psychosocial Interventions for Children.	**Inclusion criteria:** ▪ appropriate treatment descriptions ▪ post-treatment data. **Exclusion criteria:** ▪ psycho-pharmacological treatments ▪ follow-up investigations, prevention studies, uncontrolled efficacy trials. Results were reported for each condition domain under efficacy and effectiveness, and rated based on explicit study design characteristics as "well-established" (level 1), "probably efficacious" (level 2), "possibly efficacious" (level 3), "unsupported" (level 4), or "possibly harmful" (level 5). Variables for coding effectiveness were clearly defined.	115 studies provided results reported in review. Under "autistic and related disorders" section, review considered treatments relating to auditory integration training, discrete trial training, FCT/ABA, Playschool Programme, caregiver based intervention, and TEAACH. Interventions were considered as being either "comprehensive" (addressing overall functioning, multiple symptoms over the long term) or "focal" (aimed to eliminate problematic or undesired behaviours). Efficacy No comprehensive treatments were found to have support in terms of efficacy. Studies finding clinical improvements couldn't rule out alternative explanations including group selection procedures, maturation, misdiagnosis, or non-therapy factors. As a focal treatment, FCT/ABA was supported at Level 3 as possibly efficacious based on 15 controlled single-subject experimental designs. Under caregiver-based interventions, Jocelyn et al (1998) evaluated a programme providing weekly training over 12 weeks to parents and demonstrated positive impacts (median effect size=0.81), offering level 3 support as a focal treatment. Effectiveness FCT/ABA was demonstrated as appropriate for school-based implementation (with teachers managing the programme under supervision). FCT/ABA was often associated with clinically important changes in behaviour (such as termination of self injury). **Authors' conclusions:** Found results "suggesting that behaviour analytic techniques appear to provide the greatest chance of focal improvement in children with autism." "Most services were provided by Bachelor's-level practitioners, graduate students, or teachers, under the supervision of the investigator".	▪ broad review of several child disorders rather than for ASD specifically ▪ very limited range of databases searched, supplemented by reference checking and more opportunistic sources ▪ explicit attempt to rank studies by level of evidence based on study design ▪ studies described critically in the text for each disorder category and tables presented of main features ▪ only identified SCED studies relating to ABA. The RCT by Jocelyn et al (1998) was considered a caregiver-based intervention rather than ABA. ▪ one study was identified which met selection criteria for the current review - (Jocelyn et al. 1998). **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Ludwig and Harstall 2001) Canada SR of 'reviews'	**Evidence level:** I **Aim:** To review the published research on the effectiveness of intensive behavioural intervention (EIBI) programs for children with ASD **Search period**: See below – different for different databases **Databases:** Medline 1985 – 2000 Premedline: 2000 Embase: 1990 – 1999 Best Evidence 2000 HTA 2000 EED 2000 DARE 2000 CDSR 2000 ISTAHC 2000 HealthSTAR 1985 – 2000 PsychInfo 1985 – 2000 Cinahl: 2000 ERIC 1985 – 2000 Dissertation Abstracts 2000 CMA practice guideline (infobase) 2000 US National Guideline Clearinghouse 2000 ECRI 2000 Globe and Mail 2000 Numerous HTA databases 2000 (search terms included)	**Inclusion criteria:** ▪ 'critical' reviews (i.e. reviews that appraised the scientific validity of primary research studies according to a set of criteria (not given). **Exclusion criteria:** ▪ none stated. No formal quality appraisal of included studies undertaken but shortcomings discussed in narrative format.	3 critical reviews were identified (Bassett et al. 2000), ECRI 2000 (*not included in this report),* (Smith 1999). All reviews evaluated a number of comprehensive treatment programs for young children with ASD. ECRI analysed studies on Lovaas therapy, TEACCH, the Rutgers program, the Denver program, LEAP and the Autism Preschool Program. Bassett focused primarily on Lovaas therapy but also critically reviewed one study on the effectiveness of TEACCH. Smith critically reviewed 9 studies on behavioural therapy, one on TEACCH and 2 on the Denver model. ECRI 2000 concluded that Lovaas therapy appears to increase scores on IQ tests and behavioural adaptation, at least in some children with autism; however, given the studies' designs and methodological flaws, they could not determine if the changes in IQ and functional parameters could be attributed to the Lovaas therapy. (Bassett et al. 2000) concluded that the study conducted by Lovaas (1987) and the follow up study done by McEachin and colleagues (1993) were methodologically stronger than other published studies; however, they were still inadequate to establish the degree to which this form of therapy resulted in "normal" children. (Smith 1999) stated that methodological flaws in the research hinder the ability to draw conclusions; however, the studies by Lovaas (1987) and McEachin (1993) had the strongest study design. Children with autism in these studies made "long lasting improvements as a result of the treatment that they underwent". Smith also stated that 2 of the 3 studies which attempted to replicate the Lovaas study produced favourable results. **Authors' conclusions:** "It appears that children improve in functioning with intensive intervention programs, but it remains to be determined if any one program is more effective than another. There is insufficient evidence to establish a relationship between amount (intensity and duration) of any intensive intervention treatment program and outcomes measures (intelligence tests, language development and adaptive behaviour tests)".	▪ this HTA assessment is an evaluation of other reviews of interventions for ASD (a review of reviews). ▪ the review does not state the criteria used to select studies from their comprehensive search. ▪ in all reviews, most of the intensive interventions were shown to be effective in producing developmental gains, increases in IQ and less restrictive school placement. However, because of methodological limitations, the evidence remains limited on the efficacy and effectiveness of one intervention in comparison to another. It does appear that children improve in functioning with behavioural intervention programs but remains to be determined if any one program is more effective than another program. ▪ the outcome measure instruments used in the studies assessed in the critical reviews were very similar: standardized measures of IQ tests, adaptive functioning and language development. So results from these reviews are applicable only to these outcome measures. ▪ the following studies were identified that meet the inclusion criteria of this review: (Bassett et al. 2000) (Smith 1999) **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(McGahan 2001) Canada SR of consensus statements and other SRs	**Evidence level:** I **Aim:** To present a summary of evidence and expert opinions regarding behavioural therapy for preschool children with autism or pervasive developmental disorders. **Search period**: 1995 - 2000 **Databases:** DIALOG MEDLINE EMBASE HealthSTAR ERIC PsycINFO CINAHL CURRENT CONTENTS The Cochrane Library (search terms included) Database searching was supplemented with searches of websites of health technology assessment and related agencies, specialized databases such as the University of York NHS Centre for Reviews and Dissemination, GOOGLE and handsearching of selected journals and documents in the CCOHTA library collection	**Inclusion criteria:** ▪ subjects with an identifiable diagnosis of autism or a related PDD or the presence of 'autistic-like disorder' ▪ intervention accepted as a valid behavioural procedure by professionals in the field of behaviour modification or ABA ▪ publication in a refereed professional journal, professionally reviewed book or a paper presented at a conference following peer review – limited to secondary reviews. **Exclusion criteria:** ▪ none stated. No formal appraisal of included studies but narrative discussion of shortcomings of	3 summaries of recommendations by working groups were identified: (New York 1999; California 1997; MADSEC 2000 [*not included in this report*). These were considered to be level III studies (based on clinical experience, descriptive studies or reports of expert committees without explicit critical appraisal). ▪ New York 1999 recommends (based on graded evidence) that principles of ABA and behavioural strategies be included in any intervention program for young children with autism. They recommended that intensive behaviour programs include a minimum of approx 20 hours per week and that parents be trained to incorporate behavioural techniques in daily routines. ▪ California 1997 recommends (based on consensus) that a program be founded on current research and effective practices, to include a variety of methods and approaches consistently applied in a variety of settings as determined by a multidisciplinary team. They recommend that ABA be used to assist a child in gaining skills and reducing negative behaviour but that an individualised approach be used for each child where several intervention methods are used. ▪ MADSEC 2000 concluded that competently delivered early intensive behavioural intervention can be cost effective but did not provide details or specifics of programs. 5 systematic reviews were also identified (ANAES 1994; ECRI 1999 [*not included in this report]*; (Bassett et al. 2000)*;* (Smith 1999), (Ludwig and Harstall 2001). ▪ ANAES 1994 suggested that EIBI ameliorates the symptoms associated with autism but further research was necessary to determine whether the effects of therapy remain long term. ▪ ECRI 1999 identified 4 studies on Lovaas therapy, and one study each on Rutgers, LEAP and Autism Preschool Programs. ECRI emphasised that the evidence on Lovaas therapy is only suggestive of treatment effectiveness. ECRI concluded that the results of one study on the Rutgers program ((Weiss 1999) [*not included]* could not be interpreted because of significant flaws in the study. ECRI noted that one RCT of the Autism Preschool program (HTAIS 1999) [*not included]* found that only language development in the treatment group showed significant improvement compared to the control group. ▪ (Bassett et al. 2000) reviewed the same studies of ECRI (with the exception of 1 study) and had differing conclusions. ECRI found that (Sheinkopf and Siegel 1998) provided some support for the Lovaas	▪ a review of both consensus statements of varying quality and other secondary systematic reviews ▪ no formal assessment of the included studies in each review (except for the New York guideline which provided graded recommendations based on quality). However, a lot of narrative comment included on the quality of the studies ▪ the methodological flaws of the included studies were generally acknowledged in all included studies and conclusions in general were suitably cautious ▪ it was not possible to synthesize the results of the studies but conclusions of the full report were based on the body of the evidence provided by all the included studies ▪ the following studies were included that met the inclusion criteria of this review: (Ludwig and Harstall 2001) (Smith 1999) (Bassett et al. 2000). **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
	and the bibliographies of selected papers.	studies.	seminal study in 1987 (less intensity but was also effective) but Bassett found it was too unreliable to draw support for Lovaas in terms of IQ benefits. Bassett concluded that while many forms of behavioural therapy clearly benefit children with autism, there is insufficient scientifically valid effectiveness evidence to establish a causal relationship between a particular regimen of EIBI and the achievement of normal functioning. ▪ (Smith 1999) concluded that behaviour analysis treatment programs based on Lovaas yielded short term benefits in intellectual and adaptive functioning and less restrictive school placements although long term benefits were not evaluated. ▪ (Ludwig and Harstall 2001) concluded that the methodological limitations and weaknesses of existing primary research meant that evidence regarding the effectiveness and efficacy of any single treatment program is not available. While it appears that children who receive behavioural intervention exhibit functional improvement it is not clear that any single program is more effective than another. **Authors' conclusions:** "Though limited, findings of existing studies suggest that preschool children with autism exhibit cognitive and functional improvement when receiving behavioural intervention with ABA for approx 20 hours per week or more. It is not clear, however, which subset of children with autism derive the most benefit, which components of therapy are integral to positive outcomes, whether similar results would be observed in older children with autism, whether there are definable long term functional benefits or whether reported gains in IQ translate into happier people with greater functioning in the community." "The heterogeneity associated with autism may explain the variability in the effectiveness of therapies. While early identification and early behavioural intervention are important factors in effective treatment, specific components of treatment may be more suitable for children with specific characteristics. Certain communication deficits may be more or less responsive to certain behavioural procedures. Thus, individualized case management and programming are ideal for these children (but implementation can prove challenging)."	

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Bassett et al. 2000) Canada SR of non randomised studies and systematic reviews	**Evidence level:** I **Aim:** To assess the effectiveness evidence of the Lovaas treatment program (not other types of ABA) for children with autism. **Search period**: varies according to database searched (see below). **Databases:** Medline 1966 – 1999 Healthstar: 1975 – 1999 Embase: 1988 – 1999 CINAHL: 1982 – 1999 Current Contents: 1996 – 1999 Science and Social Sciences Citation Index: 1989 – 1999 Detailed search terms were provided. . References of retrieved articles also searched for additional citations. Other databases (commercial): Cochrane Library HSTAT HSRProj (NLM) Dissertation Abstracts Article1st (OCLC) Papers1st (OCLC) – conferences and paper	**Inclusion criteria:** ▪ human paediatric (preschool) populations with autism (no exclusion due to presence of comorbidity) ▪ early applied behaviour analysis, behavioural therapy or intensive home-based program ('early' defined as initiation of therapy when the diagnosis is made and prior to age 5 years) ▪ measurement of overall function (including intellectual functioning, language, social interaction and play, adaptive or self-care skills or maladaptive behaviour ▪ study design included a treatment and control group. **Exclusion criteria:** ▪ studies limited to 'training trials'	4 primary studies identified which were not randomised (Lovaas 1987/McEachin 1993, Birnbauer and Leach 1993, Ozonoff and Cathcart 1998 – *all not included in this report* and (Sheinkopf and Siegel 1998)). 5 systematic reviews were identified that included critical appraisal of the primary data (Rogers 1998, Tregear et al 2000, Howlin 1997, Green 1996 and Smith 1998 – *all not included in this report).* Other secondary data were included that critiqued the Lovaas and McEachin studies. The authors critically appraised the studies and came to conclusions, based on other critiques. Lovaas 1987/McEachin 1993: The authors concluded that these studies had major methodological shortcomings including non random assignment of children to treatment and control, an unrepresentative control group in terms of sex ratio, inadequate documentation of treatment integrity, different assessment tools administered under non standard conditions and concerns about external validity. Birnbauer and Leach (1990) : The authors considered this study too small, too short and too methodologically weak to provide evidence either for or against Lovaas therapy. It did not validate the claims of Lovaas and McEachin. Ozonoff and Cathcart (1998): The authors concluded that this study was worth considering. The intervention period was only 10 to 12 weeks and it suffered from many methodological flaws but it was a prospective controlled trial using overall outcomes that studied an intensive home based treatment alternative to Lovaas therapy. The results suggested that auxillary home interventions increase developmental functioning in young autistic children above and beyond the gains due to school based services. Sheinkopf 1998: Study used a weak, partially retrospective observational study design. It provided some support of benefit from the Lovaas method but no children achieved normal functioning. **Authors' conclusions:** "While many forms of intensive behavioural therapy clearly benefit children with autism, there is insufficient, scientifically-valid effectiveness evidence to establish a causal relationship between a particular program of intensive, behavioural treatment and the achievement of 'normal functioning'". The authors also noted that	▪ report compiled in response to legal proceedings brought against the government of British Columbia on behalf of several children seeking an intensive behavioural program ▪ thorough and comprehensive search strategy, clear inclusion criteria, independent selection of studies, data extraction and rigorous critical appraisal by two researchers ▪ use of secondary research to support author's conclusions ▪ the title of the report is 'Autism and Lovaas treatment' but the inclusion criteria include ABA, behavioural therapy and intensive home based programs (which may not be behavioural). For example, the Ozonoff study compares TEACCH delivered in a home based program with a school based program using DTT and finds evidence of benefit with TEACCH (which this review does not regard as ABA). The authors seem to be equating 'early intensive behavioural treatment programs' with Lovaas therapy because the clinical question for the report is: 'What is the effectiveness evidence that early, intensive behavioural treatment programs for preschool children with autism result in improved overall outcome versus alternative management strategies?'. ▪ good discussion of previous research (prior to this report) **Quality score:** very good (++)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
	abstracts TRIP (evidence-based medicine) Ebsco Academic Search Ebsco Canadian MAS Best Evidence CPG Infobase CRISP National Guideline Clearing House HTA database Additional searches were made of in-house databases, web library catalogues, internet peer-reviewed sites, internet search engines, directories and a large number of organisations were contacted.	▪ individual case series or reports.	the benefits of overall functioning found by Lovaas and McEachin could have been achieved by assembling a high functioning group of autistic children and the results have not been corroborated by independent researchers. They recommend further research with RCTs before effectiveness claims can form the basis of public funding decisions. They also state that there is insufficient evidence to conduct a cost benefit analysis.	

Level II studies

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Howlin et al. 2007) UK RCT	**Evidence level**: II "pragmatic, group RCT" **Intervention:** 2 treatment groups: 1. Immediate Treatment Group (ITG): Expert training and consultation in the use of PECS offered to teachers and parents and received immediately after baseline (Time 1) assessment. 2. Delayed Treatment Group (DTG) – watchful waiting condition, received PECS treatment post Time 2 **Comparator:** No Treatment Group (NTG) – standard care with classes **Study setting**: Specialist (autism specific) classes or schools.	**Participants:** Of 38 schools contacted, 18 classes (15 schools) were eligible. N=84/88 (95%) children with autistic disorder participated to Time 3. Each class required to have minimum of 3 children meeting selection criteria: diagnosed with autism, non-verbal (not exceeding single words), have no sensory impairment, aged between 4 and 11 years, not be using PECS beyond Phase 1 (able to exchange symbols only if prompted) ▪ After randomisation, one class (4 children) dropped out of ITG, 1 girl did not meet ASD criteria at baseline: data excluded ▪ 7 children did not have treatment in DTG group (but were included in intention-to-treat analysis) ▪ 1 girl joined DTG at Time 2.	18 classes stratified by number of eligible children and randomised by online randomisation programme (6 classes each group). **Follow up**: Three assessments: ▪ Time 1 (baseline) ▪ Time 2 (straight after intervention for ITG) ▪ Time 3 (straight after intervention for DTG). Actual time varied across classes and groups. Time 1 – Time 3 follow-up: ▪ ITG: M=17.9 mths (SD=0.4) ▪ DTG: M=14.6 mths (SD=1.9) ▪ NTG: M=15.3 mths (SD=0.7) For ITG: ▪ Time 2 was Mean=7.6 mths post baseline; ▪ Time 3, M=10.4 mths post Time 2 (maintenance phase). **Intervention intensity:**	Children assessed and video-taped during snack times (which were coded) to record rates of: ▪ communicative initiations ▪ use of PECS symbols ▪ speech in the classroom (including non-word vocalisations) ▪ Autism Diagnostic Observation Schedule-Generic (ADOS-G) domain scores for Communication and Reciprocal Social Interaction ▪ scores on formal tests (EOWPVT and BVS) to assess expressive and receptive language. Snack-time sessions video-taped varied in length, ranging 1.2-15 minutes, M=11, SD=3.4). Rates observed per minute recoded in ordinal categories as ordinal data were significantly skewed to low rates.	**Pre-treatment** No difference at baseline on ADOS between groups. However, DTG group had higher language impairment than ITG ($p<.05$) and NTG groups ($p<.05$), whilst ITG group had higher non verbal developmental quotient (NVDQ) than DTG ($p<.05$). **Post-treatment** Controlling for baseline age, developmental quotient (NVDQ) and language, significantly increased rates immediately post-treatment observed directly (from video) for: ▪ frequency of initiations: OR (of being in a higher ordinal rate category)=2.73 (95%CI 1.22–6.08, $p<.05$) ▪ frequency of PECS usage: OR=3.90 (95%CI 1.75–8.68), $p<.001$). Increases were not maintained over time; the ITG group was no more likely to have increased rates of initiations or	▪ classes randomly allocated to 3 groups ▪ "PECS-naïve" classes not possible but teachers hadn't received direct, in-class PECS training ▪ assessors, treatment facilitators and parents not blinded to group ▪ power calculations determined sample size ▪ treatment groups varied in language impairment and DQ at baseline; but adjusted in analyses. ▪ treatment fidelity not measured, but PECS manual followed ▪ no implementation data on use of PECS ▪ time intervals between assessment not matched across groups and was greatest for treatment groups, introducing potential bias due to maturation effects ▪ only one assessment at each time point. Snack times are when children motivated to request food and behaviour may have changed during videotaped sessions. Gains may not

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		Sample characteristics (of children with ASD): Sex: 87% male ▪ ITG: N=26/30: 21 boys, 5 girls ▪ DTG: N=30 (29 at baseline as extra girl joined at time 2): 27 boys, 3 girls ▪ NTG: N=28/29: 25 boys, 3 girls. **Diagnosis**: All newly diagnosed with autism (n=75) or ASD (n=9) using Autism Diagnosis Observation Schedule – Generic Module (ADOS-G)	Intervention: A 2-day (13 hours) PECS workshop for teachers and parents, followed by 6 half-day, school-based training, monitoring and feedback sessions with expert consultants (one per month over 5 months), delivered over 2 school terms. Attendance at workshops varied for teachers (range=4-6, mean=5.1, SD=0.6) and parents (range=0-7, mean=3.2, SD=2.4). Class training followed manual, but teachers' practice varied widely. **Statistics** Multilevel regression model allowed within child and within class correlations to be taken into account.		PECS usage at T3 compared with the untreated NTG group. No increases in observed frequency of speech, improvements in ADOS-G ratings, or language test scores between groups. ITG had decreased severity of symptoms on ADOS-G RSI at T3 (OR=.28, 95%CI .09-.89, p<.05), but not at Time 2. **Authors' conclusions:** **"**The results indicate modest effectiveness of PECS teacher training/consultancy. Rates of pupils' initiations and use of symbols in the classroom increased, although there was no evidence of improvement in other areas of communication. Treatment effects were not maintained once active intervention ceased." "There was no evidence that 7 months' experience with PECS resulted in increases in spoken language".	generalise to other situations or settings. ▪ snack sessions video-taped were brief and varied in length. Ordinal data may reduce sensitivity to detect change. ▪ treatment effect had clinical meaning, post-PECS median rates of initiations increased from 15 to 26 per hour, and of PECS use increased from 12 to 40 per hour ▪ intention-to-treat analysis used for 7 children in DTG who didn't have treatment ▪ odd finding of delayed decrease in severity of symptoms for ITG at T3 may be artifact of longer follow-up than other groups ▪ no longer term follow-up **Quality score**: good (+)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Kasari et al. 2006) USA RCT	**Evidence level**: II **Interventions:** Children in experimental groups received usual care Early Intervention Programme (EIP) *as well as* daily intervention sessions using principles of behaviour analysis. Sessions aimed to shape treatment goal through prompting, modelling, and reinforcement. Goals determined from baseline of a skill that the child was not independently conducting, relevant to focus of intervention. Once mastery achieved, more advanced skill targeted. Two parts. Table activity: adult-directed instruction targeting joint attention goal while engaging with child in social interaction. Floor activity: naturalistic milieu child-driven instruction on the floor to enhance generalisation of same skill. <u>Joint Attention intervention (JA)</u>: taught	**Participants:** N=58/65 (89%) children with autism participated, recruited from an existing EIP. Children approached if aged between 3 and 4 years, had received clinical diagnosis of autism. Children excluded if they had had seizures, had additional medical diagnoses, were geographically inaccessible for follow-up visits or did not plan to stay in EIP for at least 4 weeks. After randomisation, 7 children were excluded from further analyses (2 in JA, 1 in SP and 4 in UC) because they did not meet diagnostic criteria for autism, refused final assessments, or "left the programme unexpectedly". **Sample characteristics (of children with ASD)**: Sex: 79% male JA: N=20: 15 boys, 5 girls SP: N=21: 16 boys, 5 girls	Eligible children randomised to one of three groups (before diagnosis assessed by researchers). **Follow up**: 5-6 weeks from baseline (immediately post intervention) **Intervention intensity:** Attendance for 6 hours per day, 5-6 weeks, for EIP for all participants. For intervention groups only, in addition to EIP, daily intervention sessions (5 per week) involving table activity and floor activity lasting about 30 minutes, including 5-8 minute discrete trial training to 'prime' the particular treatment goal. Intensity in number of sessions varied between intervention groups (mean=28.6, SD=15.6 for JA group; mean=34.7, SD=16.4, for SP group), but not significantly. Treatment fidelity assessed for random 106 sessions across five interventionists: mean of	Following measures administered: ▪ Language: Reynell Development Language Scales: expressive language and comprehension. ▪ Nonverbal communication: Early Social Communication Scales (ESCS) – video-taped observational instrument of rates of: "Initiating joint attention", & "responding to joint attention". ▪ Structured Play Assessment: video-taped observational instrument of rates of different novel functional play acts, and symbolic play types, and rating of mastered play level (ranging from physical and conventional combinations to socio-dramatic and thematic/fantasy play). In addition, video-tape of care-giver interacting with child and standard toys for 15 minutes, coded for types or	**Pre-treatment** No differences at baseline on age, mental age, development quotient, and expressive language and comprehension. **Post-treatment** ESCS: Joint attention initiation and responding skills: JA and SP groups showed greater improvement in initiating shows than UC (but were not different from each other) ($p<.05$). However, no significant group differences in initiating points, initiating gives, or coordinated joint looks. JA had more improvement in responding to joint attention over time than UC or SP groups ($p<.01$). For structured play assessment, SP group improved more than UC group for "mastered level of play". No group differences for functional and symbolic levels of play. In mother-child interactions, joint attention: JA group made more gains on gives and shows than SP group (both $p<.05$). JA and SP made gains in coordinated	▪ method of randomisation not reported. ▪ no breakdown of reasons for dropping out by group, but as control group had lowest sample size, probable that more people dropped out in that group than others. ▪ all groups included some ABA of similar intensity; the interventions targeted specific skills. ▪ according to parent reports, no other *behavioural* interventions were offered during study. ▪ assessment blinded ▪ EIP staff blinded ▪ Intervention facilitators randomised to group. ▪ Inter-rater reliability acceptable. ▪ many comparisons with no adjustment to p value. ▪ power calculations not mentioned ▪ treatment groups did not vary at baseline. ▪ intervention brief ▪ treatment fidelity high. Manual

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
	joint attention skills, incorporating imitation of the child and engineered play routines. Symbolic Play intervention (SP): taught symbolic play skills, focusing on object combinations that were increasingly symbolic. **Comparator:** Usual Care (UC) control: Early Intervention Programme (EIBI) with staff ratios nearly 1:1, 6 hours per day, based on applied behavioural analysis principles following typical preschool curriculum. Joint attention and symbolic play skills were not assessed or taught in the EIBI. **Study setting**: Early Intervention Programme pre-school/community setting	UC: N=17: 15 boys, 2 girls. **Diagnosis**: All newly diagnosed with autism using ADOS and ADI-R.	92% fidelity for floor and 95% for table exercise. **Statistics** Mixed effect regression models testing for intervention and time effects.	functional and symbolic play acts and mastered play level. Inter-rater reliability calculated with Cohen's Kappa of 0.79 for 20% of ESCS, and intra-class correlation coefficients ranged 0.94-1.00 for structured play, and 0.65-0.95 for caregiver-child assessments.	joint looks than UC group ($p<.05$) (but were not different from each other). JA group engaged in more child-initiated joint engagement than UC group ($p<.001$). No group differences on pointing, or mother initiated joint engagement. In mother-child interactions, symbolic play: SP groups showed more types of symbolic play over time ($p<.001$), and had improved levels of play than JA and UC groups ($p<.001$). **Authors' conclusions:** "In this study, play and joint attention groups showed some specificity of treatment, but they also yielded some similarities in outcomes. Most notably, both groups improved in some aspects of joint engagement and functional play skills." "Compared to the control group, both experimental groups engaged in higher-level engagement states with their mothers".	followed ▪ maintenance not measured with follow-up limited to immediately post intervention. **Quality score**: good (+)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Yoder and Stone 2006b) (Yoder and Stone 2006a) USA RCT	**Evidence level**: II **Intervention:** Picture Exchange Communication System (PECS) **Comparator:** Responsive Education and Prelinguistic Milieu Teaching (RPMT) **Study setting**: University clinic in the Special Education Department of Vanderbilt University, Nashville, Tennessee.	**Participants:** N=36 children (out of 120 screened) (PECS, n=19, RPMT, n=17). No dropouts. **Inclusion criteria:** ▪ diagnosis of autism or PDD NOS ▪ aged between 18 and 60 months ▪ evidence of being non verbal or low verbal. **Exclusion:** ▪ severe sensory or motor deficits ▪ English not primary language. **Sample characteristics**: Age (mean (SD)): 33.6 (8.4) months Non verbal age: 18.6 (3.7) months Verbal mental age: 11.9 (2.8) months IQ: 54-55 (6-7) Groups balanced in baseline characteristics. **Diagnosis**: 33 with clinical diagnosis of autism and 3 with clinical diagnosis of PDD NOS (diagnoses verified by the ADOS)	Randomised according to computer program schedule. Allocation concealed. At entry, the children were administered: ▪ ADOS ▪ Mullen Scales of Early Learning (MSEL) ▪ abridged version of the Early Social Communication Scales (ESCS) ▪ an examiner-child free play session (SFPE) ▪ the Developmental Play Assessment ▪ a turn-taking measure ▪ parent child free play During the treatment phase (6 months) children attended 3 20-minute play sessions per week (total of 24 hours over 6 months). Parents were offered up to 15 hours of training to complement what the children learned in the sessions. Parents filled out a questionnaire describing the children's participation in non project treatments every month during the Rx phase. The free play	Outcomes assessed by an examiner that was independent of staff in the intervention sessions: • ESCS (measured at time 1 and time 2) • SFPE (measured at entry, time 2 and time 3) measured spoken communication. No prompts were allowed. • The Developmental Play Assessment measured initial object exploration (measured at time 1 only) • Turn taking assessed the frequency of object-exchange returns (measured at time 1 and time 2) • Parent child free play (measured at time 1 and time 2)	**Pre-treatment:** 2 variables significantly different between groups: MSEL expressive language score significantly higher for RPMT; frequency of object exchange turns significantly higher for PECS (controlled for in analyses). **Post-treatment:** Chn experienced an average of 20 hours of staff implemented therapy. Main effects: There was a strong growth on both measures of spoken communication from time 1 to time 3 (no overall difference between Rxs at time 3). At time 2, PECS>RMPT for frequency of non imitative spoken acts: t(34)=2.30, p=0.03 PECS>RMPT for number of different non imitative words. T(34)=2.10, p=0.04 RMPT>PECS for object exchange turns: t(34)=2.46, p=0.19.	▪ sample fully randomised and allocation concealed ▪ no dropouts ▪ treatment groups equivalent on 55 pre treatment variables and on both non project treatment attendance variables ▪ ITT analysis regardless of number of sessions children attended ▪ high level of treatment fidelity ▪ outcomes "ecologically valid" – i.e. measures of spoken communication were assessed within a context that required generalisation across locations, persons etc ▪ no environmental arrangement was permitted for elicit communication ▪ maintenance of the treatment effects was tested and detected ▪ examiners and coders were not blinded to treatment assignment (although reliability was assessed with a blinded independent coder) ▪ no control group to allow comparison of children who received treatment to those only receiving non project treatment ▪ suggestions made for

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
			measure was repeated at the end of the Rx phase (time 2) and 6 months after the end of Rx (time 3) together with the non project Rx questionnaire. PECS: 6 phases with ratio of 2 adults to 1 child, instructing children to make a request by teaching them to hand a picture of a desired object or food to a message recipient. Prompts were faded as progress toward independent picture exchange was made RPMT: 2 components: one for parents (responsive education) and one for children (prelinguistic milieu teaching). PMT is a child-led play based incidental teaching method designed to teach gestural, nonword vocal, gaze use and later word use as forms of clear intentional communication for turn taking, requesting and commenting pragmatic functions. Once per month, each		Exploratory analysis to test whether initial object exploration interacted with Rx group: Fixed effect statistically significant after controlling for the effects of initial number of different non imitative words and main effects for initial object exploration and treatment group: t(29)=-2.7, p=0.01. RMPT>PECS for initiating joint attention (for chn who used at least 7 initiating joint attention acts across time 1 procedures) PECS>RMPT for initiating joint attention (for chn who used at most 1 initiating joint attention act across both procedures before Rx), t(33)=3.25, p=0.003. (ie. there is an interaction). **Authors' conclusions:** 1st study: "PECS had a more rapid effect on spoken communication than RPMT. . Relative treatment efficacy varied by initial object exploration level. If the	future research specifically in this context, acquiring spoken communication. **Quality score**: very good (++)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
			clinician-child session was coded for fidelity of treatment. Interobserver reliability was also estimated for at least 20% of coded data (where coder was blind to Rx assignment). **Intensity of Rx:** Total of 24 hours over 6 months for both Rxs. **Timing of assessment:** Time 1: baseline Time 2: post treatment at 6 months from study entry Time 3: 6 months after treatment finished. **Statistics:** ANCOVA, mixed level modelling. Multiple regression procedures.		children began treatment with high object exploration, the initial advantage of PECS maintained 6 months after treatment ended. If children began treatment with low object exploration levels, RPMT facilitated a number of different non imitative words faster than did PECS and these effects maintained 6 months after the end of treatment." 2nd study: "RPMT facilitated object exchange turns and initiating joint attention more than did the PECS. The latter results occurred only for children who began treatment with at least some initiating joint attention. In contrast, the PECS facilitated requests more than the RPMT in children with very little initiating joint attention prior to treatment".	

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Sallows and Graupner 2005) USA RCT	**Evidence level**: II **Interventions:** Clinic directed group of early intensive behavioural treatment developed at UCLA (EIBI) **Comparator:** Parent directed group that received intensive hours but less supervision by equally well-trained supervisors (PT). **Study setting**: The Autism and Pervasive Developmental Disorder Clinic, part of a teaching hospital associated with the University of Toronto.	**Participants:** N=23/24 (96%) autistic children aged 24 to 42 months and neurologically normal completed the study (EIBI, n=13; PT, n=10). **Sample characteristics**: Sex (male/female): EIBI: 11/2 PT: 8/2 Mean age (SD): EIBI: 33.2 (3.9) months PT: 34.2 (5.1) months **Diagnosis**: 100% with autism by independent child psychiatrists known for their experience with autism and meeting the criteria for autism based on DSM-IV and ADI-R both administered by trained examiner.	Children matched on pre-treatment IQ and randomly assigned to groups, 13 in 1996 and 11 in 1997 (14 in 1998-1999 not included in this publication). One child from IBG dropped out (not clear which group). All children received treatment based on the UCLA model (except that no aversives were used) and parents were involved in extending the treatment during the day. Treatment began at 35 to 37 months. Children in EIBI group had a mean of 39 hours in year 1 and 37 hours in year 2 of direct treatment, 6 to 10 hours/wk of in home supervision from a senior therapist and weekly consultation by the senior author or clinic supervisor. Children in the PT group had a mean of 32 hours/wk during year 1 and 31 hours/wk during year 2 (except for one child, 14 hours/wk each). Children in this grp had less supervision: 6 hours/month of in home supervision from a	**Pre-treatment:** ▪ Bayley Scales of Infant Development (BSID) (pre-treatment IQ) ▪ Merrill-Palmer Scale of Mental Tests (non verbal IQ) (MPSMT ▪ Reynell Developmental Language Scales (language ability) (RDLS) ▪ Vineland Adaptive Behaviour Scales (adaptive functioning) (VABS) ▪ Early Learning Measure (acquisition of skills during the first several months of treatment) (ELM) Information regarding developmental history, use of supplemental Rxs and pre-treatment presence of functional speech gathered from parents, professionals and direct observation As children grew older the following age appropriate measures were used: ▪ Wechsler Preschool and Primary Scale of Intelligence – Revised (WPPSI), Wechsler Intelligence Scale for	**Pre-treatment:** No significant differences between groups. **Post-treatment:** The average IQ for all children increased from 51 to 76, a 25 point increase. There were no significant differences between groups at pre and post test. Combining children in both groups, there were significant differences btwn pre and post test for: ▪ Full scale IQ ▪ Verbal IQ ▪ Performance IQ ▪ Receptive language ▪ Vineland Communication ▪ Vineland Socialization ▪ ADI-R Social Skills ▪ ADI-R Communication The authors divided the sample into rapid (48% of sample) and moderate learners. For rapid learners, there was an increase in mean IQ from 55 pre treatment to 104 post treatment. At age 7, these children were succeeding in regular classrooms. The increase	▪ comparison of 2 different modes of delivery of early intensive behavioural treatment (EIBT) ▪ the authors used the BSID for pre-treatment IQ and the WPPSI and WISC for post treatment IQ. The observed increase in IQ may have reflected the use of different tests instead of treatment effects. However, comparisons of test results at different follow up times indicated that the effects were not likely to be attributable to using different tests ▪ most pre and post testing of moderate learners was done by the second author and this could be a source of bias ▪ sample was randomised and groups were matched on age and IQ but samples in each group were unequal ▪ the small number of children in the study limited the power of statistical tests to detect differences ▪ the division of the

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
			senior therapist and consultation every 2 months by the senior author or clinic supervisor. All children had fewer hours of Rx as they entered school. Therapists were trained according to UCLA principles. Treatment fidelity was ensured by requiring therapists to meet quality control criteria set at UCLA. **Timing of assessment**: Follow up testing administered annually for 4 years. **Statistics** ANOVA with a least squares solution for unequal group size.	Children (WISC-III) and Bayley II (for cognitive functioning) (BSID-II) ▪ Leiter-R and Merrill Palmer (non verbal cognitive functioning) (MPSMT ▪ Clinical Evaluation of Language Fundamentals (CELF-III) and Reynell (language) (RDLS) ▪ Vineland (adaptive functioning) (VABS) **Post treatment:** ▪ ADI-R and Personality Inventory for Children (social functioning) (PIC) (by parents) ▪ Achenbach Child Behaviour Checklist (ACBC) and Vineland (VABS) (by parents and teachers) ▪ Classroom placement (from teachers) ▪ Woodcock-Johnson III Tests of Achievement (WJTA III) (academic skills of children in regular classes at age 7)	in IQ for the moderate learners was not significant. Treatment outcome was best predicted by pre treatment imitation, language, daily living skills and socialization. **Authors' conclusions:** "We demonstrated that the UCLA EIBT could be implemented in a clinical setting outside a university . . . without aversives". "Parent directed children . . . did about as well as clinic-directed children, although they received much less supervision." "These results are consistent with those reported by Lovaas and colleagues".	sample into 'rapid' and 'moderate' learners was done post hoc (i.e. according to scores achieved) and so this needs to be tested in further research ▪ method of randomisation not explained, allocation concealment unlikely and no blinding of outcome assessment ▪ the results reported in this publication only from the analysis of part of the fully randomised group ▪ conclusion of the authors that UCLA EIBI is feasible is not pertinent to our report as there is no proper control. The only outcome of interest is that parent directed children did as well as clinic directed children. **Quality score**: good (+)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Drew et al. 2002) UK RCT	**Evidence level**: II "Pilot RCT" **Intervention:** Parent training (PT) group: psycho-linguistic and social-pragmatic approach to language development emphasising the development of joint attention skills and joint action routines, and advice about behavioural management in promoting compliance. Included principles of reinforcement, interrupting unwanted behaviours and teaching alternative behaviours. **Comparator:** Local services of eclectic care (ECI) only: children received speech and language therapy sessions, occupational therapy, physiotherapy, home worker input and Lovaas/ABA therapy. **Study setting**: An outpatient hospital in conjunction with local services group.	**Participants:** N=24/31 (77% participation rate) children with autistic disorder participated in the study. **Sample characteristics (of children with ASD)**: Sex: 80% male Age (mean): 23 months PT: N=12: 11 boys, 1 girl ECI: N=12: 8 boys, 4 girls **Diagnosis**: 100% with autistic disorder based on consensus clinical judgement of 2 clinicians, using the ICD-10 and the ADI-R.	Groups were matched for age at the pre-intervention assessment. Allocation to intervention was via random number table. Intention-to-treat approach. **Follow up**: 12 months (at completion of intervention) **Intervention intensity:** PT group: Speech and language therapists visited parents at home every 6 weeks for 3 hours, and were available for telephone support. Both interventions extended over 12 months.	▪ Non-verbal IQ from subscales of Griffiths Scale of Mental Development (NVIQ) ▪ MacArthur Communicative Development Inventory (CDI) ▪ ADI-R domains including reciprocal social interaction, nonverbal communication, repetitive and stereotyped behaviour ▪ Parent Stress Inventory (PSI). Note: The ADI-R, CDI and PSI were completed by parents. An activity checklist was also completed every 3 months to measure input from other health and education services (eg, therapy, pre-school).	**Pre-treatment:** PT group had higher non-verbal IQ than ECI group: F[1,23]=14.8, p<0.001. **Post-treatment:** No differences in age, NVIQ, initial words or gestures, symptom severity or parent stress. A trend towards a difference was observed on CDI: F[1,16]=3.1, p=0.09. More children in PT group moved from being nonverbal to single word/phrase speech (n=7) than ECI group (n=2), & 1 child went from single words to <5 words (Fisher exact test, p<0.05). However there was variation between children; increase in words at follow-up in PT group ranged from 0-365. Investigating potential confounders/intensity effects: The groups didn't differ in time per week in playgroup or nursery, time in speech and language therapy, or time parents spent in	▪ sample randomised ▪ groups varied in non-verbal IQ at baseline ▪ not possible to maintain blindness to intervention ▪ most outcomes relied on parental self-report, susceptible to biases ▪ significant missing data on activity checklists ▪ parents of 3 children in ECI group opted for home-based 1-to-1 behavioural, discrete-trial programmes (mean=33 hours p/wk). Note, parental choice reflects real-life clinical situation. ▪ no implementation data on PT programme or systematic assessment of parent-child interaction, a critical treatment variable. ▪ maintenance not measured ▪ small sample reduced statistical power, particularly to investigate individual characteristics associated with benefit ▪ parent-as-therapist model is relatively low-cost, but many parents found difficulty in maintaining activities, and frequency of therapist visits were

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
					one-to-one structured activities. Trend for children in ECI group to receive more other activities than Parent Training group; F[1,23]=3.6, p=0.07. **Author's conclusions: "**Not possible to rule out that the marginally significant finding of greater language gains in the Parent Training group was due to difference in initial characteristics."	considered sub-optimal. **Quality score**: fair (-)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Moore and Calvert 2000) USA RCT	**Evidence level**: II **Intervention:** Behavioral programme (BP): Vocabulary acquisition began with teaching skills that were gradually expanded by chaining mastered skills. Began with object labeling drills, which were rewarded if correct by brief praise, or brief play with desired object (M=7.4 seconds). Verbal prompts given if incorrect. An object name was mastered when children correctly responded three times in a row without prompting. **Comparator:** Educational software programme (ESP) that builds upon behavioral learning principles. The computer software program paralleled the BP drill, but added sensory reinforcement. After an object was mastered, the computer delivered an 8-second reinforcement of visual stimulation (including colour, animation) and interesting sounds.	**Participants:** N=14 children with autism participated in the study. **Sample characteristics (of children with ASD)**: Sex: 86% male Age range: 3-6 years BP: N=7 ESP: N=7 **Diagnosis**: "With autism" – diagnostic criteria used (author contacted).	Children at a school were grouped in classes according to their verbal skill levels. Within skill level and gender groups, children were randomly assigned to one of the two treatment conditions. All children began treatment as soon as they could sit for at least 10 minutes and attend on command. The two children in the ESP who were unfamiliar with using computers were given an initial practice to familiarise them with the 'mouse'. **Follow up**: 1 week after completion of intervention **Intervention intensity:** Not described.	Learning measure: before the intervention, children were shown two flash cards of nouns simultaneously (of group of 18) and asked to identify the asked for object, three times in a row. Six unknown nouns were targeted in the intervention and tested using same flashcard procedure at post-test. Attention measure: treatment sessions video-taped and total percentage time child attended ("on") or looked away ("off) from the teachers or materials (in BP condition) or computer (in ESP condition) calculated. Reliability calculated for 3 randomly selected children from each group (inter-observer agreement was 96%). Motivation measure: after last session, children asked whether they would like to keep working on the drill activity or go play. Those choosing (verbally, by pointing, or moving) to continue	**Pre-treatment:** There were no baseline measures reported apart from using the learning measure to identify 6 unknown nouns targeted in the intervention phase. **Post-treatment:** Learning: children recalled more nouns (unknown at baseline) after exposure to the computer presentation (ESP group) (M = 4.43 or 74%) than to the teacher (BP group) (M = 2.43 or 41%), $F(1, 13) = 10.89$, $p < .01$. Attention: Children were more attentive in the computer (ESP) than to the teacher (BP) (M = 97 vs. 62%, respectively), F(1, 13) = 13.28, $p < .01$. From a regression analysis, the more children attended in either condition, the more they learned, $F(1, 13) = 38.45$, $p < .001$. Motivation: more children were interested in continuing treatment in the computer/ESP group (57%) than in the	▪ sample randomised but method not described ▪ participation rate not reported ▪ study setting not described ▪ little information on sample characteristics ▪ groups not compared at baseline ▪ all children in the BP had previously experienced some behavioral training, and 5 of 7 in the ESB were familiar with computers ▪ not possible to maintain blindness to intervention for child or tester ▪ outcomes not tested at baseline and differences may relate to sample differences ▪ no implementation data on programme or detail of duration or intensity ▪ maintenance not measured ▪ outcomes narrowly defined relevant to task-specific learning ▪ very small sample (7 per arm) reduces statistical power **Quality score**: fair (-)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
	Study setting: School (not described)			with drill were scored as being motivated, those choosing to play were scored as not being motivated.	teacher-led/BP group (0%), X^2= 3.818, p < .05. **Authors' conclusions: "**Children with autism were more attentive, more motivated, and learned more vocabulary in the computer than in the behavioral programme."	

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Smith et al. 2000) USA RCT	**Evidence level**: II **Interventions:** High intensity clinic-directed EIBI (HI) **Comparator:** Parent training (PT) group where parents had received training in EIBI approach. **Study setting**: The Autism and Pervasive Developmental Disorder Clinic, part of a teaching hospital associated with the University of Toronto.	**Participants:** All referrals to the UCLA Young Autism Project between 1989 and 1992 who met the ff criteria: **Inclusion criteria:** ▪ between 18 and 42 months of age ▪ residence within a 1 hour drive of the research/treatment site ▪ IQ between 35 and 75 ▪ diagnosis of autism or PDD NOS ▪ absence of major medical problems other than autism or MR. HI: n=15 PT: n=13 **Sample characteristics**: Gender (male/female): HI: 12/3 PT: 11/2 Age (mean (SD)): HI: 36.1 (6.0) months PT: 35.8 (5.4) months Diagnosis (autism/PDD NOS) HI: 7/8 PT: 7/6 **Diagnosis**: 50% diagnosed with autism; 50% diagnosed with PDD NOS (made independently of the	Assignment based on matched pairs by random number tables in 2 cohorts depending on diagnosis. Children in both groups received intervention based on Lovaas 1981 manual. HI: Intensive treatment was defined as 30 hours per week for each child for 2 to 3 years from 4 to 6 student therapists working under supervision (child's caregiver also involved for 5 hours per week during first 3 months of treatment. PT: The goal was to teach parents to use Rx approaches as described in the 1981 Lovaas manual. The families received 5 hours per week of training for 3 to 9 months and were supervised. Between training sessions parents worked an additional 5 hours per week implementing the programs with their child. For the rest of the training time, the child was enrolled in special education classes in public schools for 10 to	Standardised tests were administered by doctoral students in clinical psychology at UCLA (trained by author(s) and blind to children's assignment). Parents also completed a questionnaire and progress in Rx was evaluated by senior staff members of the UCLA Young Autism Project (not blind to chn assignment but reliability measured by independent raters). The following measures were assessed: • Intellectual functioning (SBIS, BSID or MPSMT) • Language functioning (RDLS) • Adaptive functioning (VABS) • Socioemotional functioning (ACBC, ATRF) • Academic achievement (WIAT) • Class placement (Children's report cards and/or	**Pre-treatment:** No statistically between group differences – all children had major developmental delays. **Post-treatment:** No of hours/wk of treatment: HI: m=24.5, gradually reducing over the next 1 to 2 yrs. PT: not reported. At follow up, the HI group had a statistically significant advantage over the PT group in: ▪ IQ: 66.5 vs 49.7, p<0.05 ▪ Visual spatial skills: 64.3 vs 49.2, p<0.05 ▪ Language development: 87.4 vs 61.3, p<0.05 There were no significant differences between groups on: Adaptive functioning The HI group had significantly less restrictive school placements than the PT group (4 regular ed and 24 regular ed with support vs no regular ed and 3 regular ed with support) (p value not given).	Limitations: ▪ allocation concealment not reported ▪ small sample size and heavy tailed skewed distributions of scores precluded conducting some statistical procedures such as factorial analyses of variance to examine interactions that might have helped in interpretation of results ▪ low power to detect predictors of Rx response ▪ assessment instruments included only one measure of social skills (VABS socialization domain), a parent satisfaction questionnaire with untested psychometric properties, no measure of childrens' or parents' quality of life and no measure of parents' participation in Rx ▪ assessment measures were mostly instruments developed for both typically and atypically developing children rather than ones specifically designed for children with developmental disabilities (eg. ACBC)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		study by licensed psychologists at the California State regional centres)	15 hours per week. Treatment fidelity was measured only for the EIBI group **Intensity** Number of hours of intervention was recorded. **Timing of assessment**: Pre-treatment evaluations occurred in the 3 months prior to Rx. Follow up evaluations occurred at a CA of 7 to 8 years. **Statistics** 1 and 2 tailed t tests (with Dunn Bonferroni corrections)	individualised education plans (IEP) to determine whether the child was suitable for a regular classroom, regular classroom with support or self contained classroom) • Progress in treatment (ELM) (Smith et al 1995) • Parent evaluation (Family Satisfaction Questionnaire)	The HI group had higher WIAT scores than the PT group (75.7 vs 58.4) (p value not given). There was little difference between groups in behaviour problems (as measured by the Child Behaviour Checklist). **Authors' conclusions** "Intensively treated children outperformed children in a parent training group at follow up on measures of intelligence, visual-spatial ability, language, and academic achievement. Also as a group they had less restrictive school placements . . . Parents in both groups held highly positive views about the services their children received. Children with PDD NOS benefited at least as much from intensive treatment as did children with autism . . . Intensively treated children did not differ from children in the parent training group on standardised tests of behaviour problems and adaptive functioning in everyday settings at follow up."	▪ study lacked a standardised diagnostic instrument and had no follow up for diagnostic assessment at all ▪ no measurement of Rx fidelity in PT group to determine what type of Rx given ▪ large variations in response within each group makes it difficult to determine whether Rx useful for all children. Strengths: ▪ manualised treatments supervised by experienced personnel ▪ blind assessors ▪ long-term follow up. ▪ study attempted to address some of the shortcomings of the original Lovaas studies but Rx intensity less and study did not find the same gains as Lovaas. **Quality score**: good (+)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Jocelyn et al. 1998) Canada RCT	**Evidence level**: II **Intervention:** Autism Preschool Program (APP): Integrated treatment model directed primarily at the parents and day-care staff, including educational seminars, reading, on-site consultations at day-care centres, and psycho-educational and supportive work with the family in their homes by social worker, including watching videotapes of child at day-care. **Comparator:** Eclectic Care (ECI). A usual, standard care control consisting of attendance at community day care centre. Families received input from community Family Services Workers. Offered APP after study period. **Study setting**: Community-based day-care centre. Children were recruited from referrals to the Child Development Clinic of the children's hospital.	**Participants:** N=35/36 (97% participation rate) children with autistic disorder participated in the study. One child dropped out of intervention group. Parents and child care workers (CCW) funded to work with the child at the day-care centre received the intervention. Children excluded if already attending day-care/school, aged >24 or <72 months, outside region, or have a serious physical disability. **Sample characteristics (of children with ASD)**: Sex: 97% male Childhood Autism Rating Scale: 14 severe and 21 mild/moderate symptoms APP: N=16: 15 boys, 1 girl; mean age: 43 months ECI: N=19: all boys; mean age: 44 months **Diagnosis**:	The Childhood Autism Rating Scale (CARS) stratified children into mild/moderate and severe symptom groups. Children were randomly assigned to treatment group from the two categories by random number table. **Follow up**: 12 weeks (at completion of intervention) **Intervention intensity:** Treatment period over 12 weeks. Seminars over 5 weekly 3-hour classes. Autism Behavior Specialists visited each day-care centre for 3 hours per week for 10 weeks. Three case conferences.	Assessment conducted by psychologist blind to group assignment of child. ▪ Knowledge of autism: TRE-ADD Autism Quiz (TAQ) ▪ Measure of autism symptomatology: The Autism Behavior Checklist (ABC) ▪ Developmental measures: Early Intervention Developmental Profile (EIDP); and the Preschool Developmental Profile (PSDP). ▪ Family (self-report) measures: Stress-Arousal Checklist; the Family Assessment Measure; and the Client Satisfaction Questionnaire.	**Pre-treatment:** No difference at baseline on autism symptoms (CARS) or IQ (Leiter International Performance Scale). **Post-treatment:** Autistic symptoms (ABC) rated by psychologist did not improve over time and did not differ between groups, whilst parents generally reported improvement over time ($p=.0001$). Mothers ($p<.02$) and CCWs ($p<.008$) in the intervention group (but not in the control ECI group, or fathers) reported a significant increase in understanding of autism. Development increased over time, and was greater in the language area for the experimental group (mean change=5.3 +/- 5.0 months) compared with the ECI control (mean change=1.1 +/- 4.6 months). Stress, and family functioning did not vary	▪ sample randomised ▪ groups did not vary at baseline on IQ or autism symptom measures ▪ Bonferroni correction not used to adjust for multiple number of statistical tests performed (ie; some could be chance effects) ▪ blind assessment at baseline and follow-up ▪ not possible to maintain blindness to group for treatment facilitators and parents ▪ one drop-out from intervention group but no intention-to-treat analysis used ▪ maintenance over time not measured ▪ no longer term follow-up to see whether changes result in socially valid outcomes ▪ no power calculations, limited sample size ▪ parents' assessment of autistic behaviour tend to overestimate symptoms at baseline compared to psychologist, and rated an improvement over time not found by the psychologist. **Quality score**: good (+)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		100% with PDD or autism newly diagnosed by developmental paediatrician by DSM III-R			by intervention group. Experimental group parents were more satisfied. **Authors' conclusions:** **"**We conclude that the research design demonstrated that the intervention was significantly superior to day care alone."	

3. Level III-1 studies

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Bellini and Akullian 2007) USA SR/MA of single case experimental design (SCED) studies	**Evidence level:** III-1 **Aim:** To examine the effectiveness of video modelling and video self-modelling (VSM) interventions for children and adolescents with ASD. **Search period**: 1985 – 2005. **Databases**: ERIC, PsycINFO (search terms provided). In addition, hand searching of three relevant Journals, and ancestral searching of reference lists of retrieved studies and the Ayres and Langone (2005) review.	**Inclusion criteria:** ▪ data reported on people aged 3-21 years with ASD ▪ outcome measures targeted behavioural functioning, social-communication skills or functional skills. ▪ assessed video modelling or VSM ▪ SCED studies only ▪ individual data points reported ▪ published in the English language in peer reviewed Journals. **Exclusion criteria:** ▪ studies involving video priming, visual cueing or in-vivo modelling ▪ fewer than three probes/questions per data point. Coding system and inter-rater agreement described. Percentage of non-overlapping data points (PND) computed for each participant and study, per outcome.	Included 23 SCED studies of 73 children aged 3-20 years; 15 investigated video modelling, 8 investigated video self-modelling. Key findings: ▪ Videos are relatively brief interventions: Median video treatment length=9.5 sessions, duration median=3 minutes. ▪ Results indicate a moderate intervention effect (n=22 studies, *M* PND=80%, range 29-100). ▪ There was moderate effect for maintenance (n=18 studies, *M* PND=83%, range 35-100). ▪ There was moderate effect for generalisation (n=7 studies, *M* PND=74%, range 22-100). ▪ Highest intervention effects were for functional skills however subsamples were small, and differences may be an artifact of measurement issues related to different skill domains. ▪ No significant difference was found in intervention effects across age groups, or between video modelling and VSM, though again the sub-samples sizes were small. ▪ Discussion of study results highlighted that there were sometimes no effects for individuals, and posited possible reasons, including disruptive behaviour or lack of interest causing lack of attendance to video, and/or lack of visual learning capabilities of participant with ASD. ▪ Charlop-Christy et al (2000) compared in-vivo with video modelling and found video led to faster acquisition of skills and larger generalisation of effects, and was less costly and time consuming, than live modelling. Results may be due to a child's over-selectivity to irrelevant features in live models, and/or anxiety about interacting with a live person. ▪ Sherer et al (2001) compared self versus other video modelling with no difference found in rate of task acquisition. **Authors' conclusions:** "Results suggest that video modelling and VSM are effective intervention strategies for addressing social-communication skills, behavioural functioning and functional skills in children and adolescents with ASD." "Further research is needed to elucidate the participant, setting and procedural features that lead to the most effective intervention	▪ focus on video modelling and VSM in children and adolescents with autism ▪ only two databases searched but some hand searching and reference checking conducted ▪ details of coding given ▪ inter-rater agreement was 100% for study features, and after review of discrepancies for first 10 studies, was 100% for remaining 23 studies, and 100% when assessed by independent third reviewer. ▪ description of individual study methods and results given in Tables for video modelling and VSM separately. ▪ findings critically summarised in detail in text, noting researchers' hypotheses for lack of positive results. ▪ there were variations between interventions, including whether combined with other therapeutic strategies (eg; instruction, visual cuing, consequent strategies, self-monitoring techniques), whether video was from first person viewpoint, whether edited to present successful performance of task (eg; prompts or off task behaviour removed), or whether scripted or naturalistic behaviour viewed. ▪ future research directions and implications for practice discussed ▪ conclusions based on 23 SCED primary studies ▪ no appraised studies eligible for inclusion in current review (all SCED). **Quality score:** very good (++)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
			outcomes for children with ASD" and "to elucidate features that improve generalisability effects".	

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Bellini et al. 2007) USA SR of SCED studies	**Evidence level:** III-1 **Aim:** To examine the effectiveness of school based social skills (SS) interventions for children and adolescents with ASD by: 1. providing a quantitative synthesis of existing single subject research studies on school based social skills interventions for children with ASD 2. examining the aggregated outcomes of these studies and identifying the participant, setting and procedural features that lead to the most effective intervention outcomes, and 3. comparing the intervention, maintenance and generalisation effects of the studies to the outcomes of similar studies involving social skills interventions with other populations of children. **Search period**: 1980 - 2005. **Databases:**	**Inclusion criteria:** ▪ participants with ASD ▪ outcome measures that targeted social functioning ▪ assessment of the efficacy of social skills interventions ▪ social skills interventions implemented in a school setting ▪ single subject research design ▪ inclusion of dichotomous dependent variables (e.g. yes-no, correct-incorrect) with at least 3 probes or questions per data point or 3 data points per intervention phase ▪ presentation of data in graphical displays that depicted individual data points rather than aggregated data (such as means) ▪ studies only published in peer reviewed journals ▪ published in English. **Exclusion criteria:** Not reported. **Methods:** Independent data extraction and	55 studies (with a total of 157 participants) published between 1986 and 2005 were included in the meta-analysis. 42 studies used a variation of a multiple baseline or probe design, 6 studies used a reversal design, 3 studies used an A-B design, 2 studies used a changing conditions design, 1 study used an alternating treatment design and 1 study used an alternating treatment and reversal design. Interobserver reliability was reported in all 55 studies. Social validity was measured in 12 studies, intervention fidelity in 14 studies and experimental control was demonstrated in 49 studies. Interventions ranged in length from 8 to 73 sessions, hours of intervention ranged from 2.5 to 28 hours and length of intervention ranged from 10 to 210 days. Only 1 study systematically matched the type of intervention strategy with the type of skill deficit of the participants. Low to questionable treatment effects: n=52, PND M=70%, range=17 to 100% Low to questionable generalisation effects: n=15, PND M=53%, range=17 to 100% Moderate maintenance effects: n=25, PND M=80%, range=17 to 100%. There were no significant differences in the intervention, maintenance and generalisation effects across different types of SS interventions and also for individual vs group interventions, age group (preschool, elementary and secondary) and features of the interventions (length, hours and number of sessions). There were significant differences in effects according to location of the interventions, classroom or pullout – maintenance and generalisation effects were significantly lower for interventions implemented in pullout settings. No significant differences were reported between PND scores for the 2 categories of dependent variables, collateral skills and specific social behaviours. **Authors' conclusions:** "The results . . . suggest that school-based social skills interventions are minimally effective for children with ASD.	▪ relatively small number of studies included in the MA precluded a thorough analysis of covariation between participant characteristics (e.g. age specific diagnosis, cognitive level, language level), setting characteristics (classroom vs pullout), intervention features (length and type, group vs individual) and the outcomes associated with SS interventions ▪ the small sample size precluded examining interaction effects among intervention features, participant features and outcomes ▪ the small number of studies precluded a comparison of different types of SS strategies, e.g. social stories, prompting, video modelling etc. Thus, difficult to interpret how ABA principles have contributed to the results reported ▪ possibility of publication bias which might have inflated the effects of the SS interventions ▪ only 15/55 studies measured the generalisation effects of SS interventions ▪ unclear how the results from this study compare to SS interventions implemented in different settings such as the home. ▪ attempted to use some kind of quantitative analysis, in contrast to the other reviews in the report evaluating social skills ▪ future research should elucidate features that can improve generalisation effects, such as programming for generalisation, teaching self-monitoring, or providing instruction in the natural setting **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
	PsycInfo ERIC (search terms given) Also manual searches of the journals, Focus on Autism and Other Developmental Disabilities and Journal of Autism and Developmental Disorders. The authors also examined the reference sections of each study located via the database search.	measurement of inter rater agreement. The effectiveness of the interventions was determined by computing the percentage of non overlapping data points (PND) (rather than a conventional effect size) for each study (calculating the percentage of data points that do not overlap with the highest baseline data point). They were calculated for intervention effects, maintenance effects and generalisation effects and across 2 categories of dependent variables, collateral and social interaction skills. The Kruskal Wallis procedure was used to test for significant differences in PND across dependent variables and Pearson product-moment correlations were conducted to examine relationships.	Specifically, social skills interventions produced low treatment effects and low generalisation effects across participants, settings and play stimuli. Moderate maintenance effects were observed, suggesting that gains made via social skills interventions are maintained after the intervention is withdrawn. . Similar intervention, maintenance and generalisation effects were observed between interventions targeting collateral skills (eg, play skills, joint attention and language skills) and interventions targeting specific social behaviours (eg, social initiations, social responses and duration of interaction)".	

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Delano 2007) USA SR of single case experimental design (SCED) studies	**Evidence level:** III-1 **Aim:** To summarise empirical studies that evaluated the use of video modelling interventions with children with ASD. **Search period**: 1985 – March 2005. **Databases**: ERIC, PsycINFO (search terms provided). In addition, ancestral searching of reference lists of retrieved studies.	**Inclusion criteria:** ▪ published in a peer-reviewed journal ▪ described experimental research in which an independent variable was manipulated ▪ included quantitative measures of a dependent variable ▪ participants in the studies were identified as having an ASD ▪ the primary independent variable was a video modelling intervention ▪ the videotapes were individualised and created specifically for the research participants. **Exclusion criteria:** ▪ studies without a carefully defined experimental design or studies without quantitative data ▪ studies that evaluated the use of commercial Videotapes ▪ studies in which a video modelling intervention was part of a treatment package or part of a computer-based instructional program.	19 included studies 15 of which had fewer than 4 participants. All were SCED studies, with a total of 55 children aged 3-20 years; 12 video interventions presented other (adult or peer) as model, 5 of self as model and 2 involved comparisons. Key findings: ▪ all studies reported acceptable inter-observer agreement ▪ no study provided a measure of treatment fidelity, and 5 studies reported social validity. ▪ video modelling interventions were related to positive gains in social–communicative skills, functional skills, perspective taking skills, and problem behavior. 50 of 55 participants experienced positive gains in one or more target skill. ▪ 5 studies reported mixed results, 3 found video modelling was not associated with an increase in social initiations or novel responses. ▪ maintenance was assessed and positive in 14 studies, with half measuring it immediately after (not really maintenance), half 1-3 months after intervention, and one 15 months after intervention phase. ▪ generalisation assessed and generally positive in 10 studies ▪ {Charlop-Christy, 2000 #338) compared in-vivo with video modelling and found video led to faster acquisition of skills and larger generalisation of effects than live modelling. Results may be due to a child focusing on relevant cues in videos, watching videos may be reinforcing, and the intervention makes no social demands on children. ▪ {Sherer, 2001 #342} compared self versus other video modelling with no difference found in rate of task acquisition ▪ review suggests that video modelling tapes are relatively easy to create and intervention may take only minutes per day. **Authors' conclusions:** "The findings suggest that video modelling interventions are effective in teaching a variety of skills to children with autism". "It is unclear at this time whether video modelling is more or less effective than other models of instruction for learners with autism, and too soon to make detailed recommendations for practice". "Including a larger	▪ focus on video modelling in children with ASD ▪ only two databases searched but also reference checking conducted ▪ no checklists or appraisal details given ▪ description of skill areas targeted and dependent measures summarised ▪ summary of individual study sample, setting, skill, and whether or not maintenance, generalisation, social validity, treatment fidelity, and inter-observer agreement assessed and presented in a table ▪ findings critically summarised in text, noting hypotheses for mixed results. ▪ recommendations for future research and implications for practice discussed at length ▪ conclusions based on 19 SCED primary studies ▪ no appraised studies eligible for inclusion in current review (all SCED). **Quality score:** good (+)

			number of participants, assessing treatment fidelity, and addressing social validity issues are important components of future evaluations of video modelling interventions."	

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Kroeger et al. 2007) USA Quasi-randomised study	**Evidence level**: III-1 Quasi-randomised study **Intervention:** Direct teaching (DT) group: used a video-modelling format to teach play and social skills with facilitators prompting children's practice of modeled play skills, and offering primary (edible) reinforcement. Video models were same aged male peers. Included a built-in animated video-clip to promote attention. **Comparator:** Play activities (PA) group: engaged in unstructured play during the sessions without direct instruction. Both groups included beginning and ending circle times, visual schedules during each session to transition activities, 2:1 student to facilitator ratio, secondary (social) reinforcement for pro-social behaviors,	**Participants:** N=25/27 (92%) children with autistic disorder participated, recruited by local autism society with newsletter and recruitment letters sent to eligible families attending a PDD clinic. 24/27 had received diagnoses from the PDD clinic. Sample represented a range verbal abilities from nonverbal (n=4) to relatively fluent. Most made 1-5 word non-spontaneous requests. 2 drop-outs (1 each group) after 2nd session; one as parents felt child was relatively higher functioning than peers, the other due to practical issues in attendance. Inclusion criteria included being aged between 4 and 6 years; and diagnosed with autistic disorder. Excluded children with other autism spectrum disorders (Asperger's disorder, Rett's disorder, childhood disintegrative disorder, and PDD-NOS).	Eligible children were allocated to one of six groups (3 in DT and 3 in PA) based on which time-slot dates they were able to attend. Days remained the same between groups. Parents were blind to intervention groups. Matched random assignment was used to ensure children of similar levels of functioning were assigned to each group, based on the child's Autism Quotient standard score derived from the Gilliam Autism Rating Scale (GARS). **Follow up**: 5 weeks from baseline (during last intervention session) **Intervention intensity:** Groups met for 5 weeks, three times per week, 1 hour each time (M=14 sessions/hours attended; no group difference). Study period spanned 5 months (interventions not concurrent). **Statistics** Multivariate analysis of	Data were derived and coded from videotapes of first 30 minutes of free play of the first session, and last 30 minutes of free play of the last/fifth session, representing "pre- and post-" intervention assessments. These periods were unstructured with access to play materials but with no prompting or intervention from facilitators, apart from when a child initiated an interaction or there was disruptive behaviour. Social Interaction Observation Code (SIOC) measured frequency, duration and nature (positive or negative) of the videotaped social interactions. Assessment of Basic Language and Learning Skills (ABLLS) - revised "Group Instruction" cluster indexed "learning readiness" for small group learning goals such as sitting, attending, learning, answering questions. Assessed during first and last session circle	**Pre-treatment:** No significant difference at baseline on GARS, average age or social behaviours (SIOC). **Post-treatment:** Both groups improved in their pro-social behaviors. A group interaction revealed that the DT group made more gains in social skills than PA group for these behaviours (with effect sizes): ▪ initiating behaviors; $F(1, 23) = 6.287$, $p= 0.020$ ($\eta=0.215$), ▪ responding behaviors, $F(1, 23) = 11.243$, $p=0.003$ ($\eta=0.328$), ▪ interacting behaviors, $F(1, 23) = 9.324$, $p=0.006$ ($\eta=0.288$). Both DT and PA groups improved in their learning readiness and group orienting behaviors (ABLLS): $F(1,23) =14.843$, $p=0.001$. However, the DT group did not show more improvement over time than the PA group: $F(1,23) = 3.270$, $p=0.084$. Parental satisfaction	▪ not randomised, but allocated based on whether time slots suited parents (may be open to subtle biases). ▪ not reported how many people from PDD clinic sent letter to determine real response rate ▪ intention to treat not used as 2 drop-outs excluded from all analyses ▪ intervention described in detail ▪ no effective control as both interventions involved ABA and were similar in many respects ▪ outcomes narrowly defined ▪ blind coding of SIOC but not ABLLS ▪ inter-rater reliability for SIOC high ▪ groups did not vary at baseline ▪ limited outcome and potential confounder variables measured ▪ treatment fidelity not measured. Facilitators were University students who had received over 60 hours of training ▪ generalisation and maintenance not assessed, with follow-up occurring during last

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
	behavior management for inappropriate behaviors (ignoring, differential reinforcement, and brief timeout), and introduction of identical toys. **Study setting**: Group sessions, clinic/University setting (not described)	**Sample characteristics (of children with ASD)**: Sex: 80% male Age range: 4-6 years 2 children in each group were nonverbal (and had no augmentative communication systems) DT: N=13/14, 9 boys, 4 girls; 9 Caucasians, 4 African-Americans PA: N=12/13, 11 boys, 1 girl, all Caucasians **Diagnosis**: All diagnosed with autistic disorder, by standardised assessments from a relevant professional.	variance for repeated measures using the General Linear Model statistic.	times. SIOC data coded by two trained raters. Inter-rater reliability=98.4%. Forced consensus conducted for any disagreements. ABLLS tool administered by group facilitators. High internal consistency. Parent satisfaction also assessed post intervention.	was high (M=5.7 for both groups), on a 1-7 scale. **Authors' conclusions:** "Findings indicated that while members of both groups increased pro-social behaviors, the direct teaching group made more gains in social skills."	part of intervention. **Quality score**: fair (-)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Machalicek et al. 2007) USA, Australia and Italy SR of SCED studies	**Evidence level:** III-1 **Aim:** To evaluate research on the treatment of challenging behaviour in school settings for students with ASD. **Search period**: 1995 - 2006. **Databases:** Medline PsycInfo ERIC (search terms not given) The authors also examined the reference sections of included studies.	**Inclusion criteria:** ▪ included participants ages 3 – 21 years with a diagnosis of ASD ▪ utilised a single subject design ▪ published in a peer reviewed journal between 1995 and 2005 ▪ applied an intervention in an effort to reduce challenging behaviour ▪ took place within the context of a classroom. **Exclusion criteria:** none stated.	26 studies met the inclusion criteria and were divided into 4 categories, depending on study procedures: 1) antecedent manipulations (procedure focused on addressing environmental conditions occurring prior to the occurrence of challenging behaviour) – 10 studies (5 of which used social stories which is regarded as 'borderline' ABA and excluded from this report) 2) change in instructional context (changes in the instructional context to reduce challenging behaviour) – 5 studies 3) differential reinforcement (procedures that differentially reinforced alternative or other behaviour) – 8 studies 4) self management (procedures that attempted to decrease challenging behaviour by increasing a students' independent task completion, their independent transition through a classroom schedule or focused on teaching students to self monitor their challenging behaviour) – 3 studies. 85% of all studies reported decreases in challenging behaviour and attributed these to the intervention. Of these studies, 10 studies reported elimination of the challenging behaviour(s) of at least 1 student during intervention in at least 1 condition. Each of the categories appeared to be effective (data not given). The top 4 groups of challenging behaviour most often targeted for intervention were: 1) screaming, yelling, shouting or crying – 15 studies 2) stereotypy – 14 studies 3) non compliance – 11 studies 4) aggression – 11 studies 13/26 studies did not conduct functional behavioural assessment (FBA) prior to choosing an intervention to treat the challenging behaviour yet most of the interventions (73%) reported equally positive findings (contrary to the results from past research which has indicated that the use of FBA increases the likelihood of Rx success) Most of the interventions were implemented in a special education classroom with a teacher and 1 or more paraprofessionals present. 2 studies demonstrated the ability of teachers to carry out an experimental FBA in the classroom with limited training and	▪ most of the included studies very small ▪ table provided that gave details of the number in each study, the age of participants, the intervention, study design, target behaviour(s) and findings ▪ no data given – results for each study reported as either "positive" , "mixed" or "inconclusive" ▪ difficult to determine whether treatment gains long lasting as many studies did not have follow up or maintenance assessment ▪ some studies had mixed findings and there is concern about the variability in the data of some studies ▪ studies within each category discussed in narrative format but conclusions are imprecise and very general ▪ details of each intervention not described fully – e.g. duration, intensity, follow up – so difficult to reach any conclusions about what aspects of SS interventions are effective ▪ authors suggest that future research should assess ways in which social validity of treatments be incorporated and assessed in studies. **Quality score:** fair (-)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
			support. Formal measures of the social validity of interventions were reported by a minority of studies. **Authors' conclusions:** The authors concluded that "this review suggests that the treatments utilised have effectively decreased or eliminated a variety of challenging behaviours across many ages (and that) . the studies . . point to the feasibility of conducting challenging behaviour intervention research in classrooms". Half the interventions did not carry out any type of FBA prior to designing the intervention, but most of the interventions reported equally positive findings.	

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Mancil 2006) USA SR of SCED studies	**Evidence level:** III-1 **Aim:** To examine functional communication training (FCT) and the environments and individuals involved in the training and the effectiveness of FCT with children who have a diagnosis of ASD. **Search period**: 1985 – "present" **Databases**: ERIC, Education, PsycINFO, Academic Search Premier (search terms provided), hand searching of four Journals.	**Inclusion criteria:** ▪ at least one participant was a child with ASD ▪ functional behaviour assessment (FBA) used to determine the function of the challenging behaviour ▪ the primary intervention was FCT. **Exclusion criteria:** none reported	Retrieved 30 published research articles, 8 of which were reported as including only subjects with diagnoses of autism. All were SCED (reversal and multiple baseline) studies of 22 children aged 2.7-13 years (mean=8). Researchers and research assistants implemented most of the studies and all but two were conducted in separate clinic rooms, usually containing only a table and chairs. For reversal studies, the trainer reinforced relevant communicative responses, alternating to reinforcing irrelevant responses. Main findings: ▪ All studies reported success of FCT in decreasing challenging behaviours regardless of topography (ie, aggression, tantrums, self-injurious behaviour) with a corresponding increase in communication mands, though these mands were limited in range and scope. ▪ Suggestion from one study that communication response must match the challenging behaviour's function. ▪ All studies reported high treatment fidelity and high inter-rater reliability for coding of outcomes. ▪ No study conducted long term follow-up or assessed whether children developed independence ▪ Most researchers did not consider generalisation and only one study (of 8) occurred in a natural environment ▪ The majority of research continues to conduct FCT similarly to the first published article in 1985, typically conducting the research themselves, in small, isolated rooms, and focusing on one communication mand. Authors argue that this limitation poses a critical threat to maintenance and generalisability. Natural environments pose sensory issues (noise, lighting, visual stimuli) not present in bare clinic rooms. **Authors' conclusions:** "FCT consistently reduces challenging behaviour and increases communication; however, the majority of research is clinically based and focuses on one communication mand. Future research teams should address maintenance and generalisation by training teachers in classrooms and parents in homes while collecting data across time."	▪ focus on functional behaviour assessment for children with autism ▪ moderate range of databases searched ▪ no details of coding and appraisal criteria given ▪ detailed description of individual study methods and results given in Tables ▪ findings summarised briefly in text, with some methodological critique. ▪ future research directions briefly outlined ▪ conclusions based on 8 primary studies ▪ no appraised studies eligible for inclusion in current review (all SCED). **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Matson et al. 2007) USA Mostly SR of SCED studies	**Evidence level:** III-1 **Aim:** To review the treatments that target social skills development in children with ASD. **Search period**: not described. **Databases:** Medline Google Scholar (search terms not given). Handsearching of following journals: ▪ Research in Developmental Disabilities ▪ Research in Autism Spectrum Disorders ▪ Autism ▪ Focus on Autism and other Developmental Disabilities ▪ Journal of Applied Behaviour Analysis ▪ Behaviour Modification ▪ Journal of Autism and Developmental Disorders ▪ Journal of Positive Behaviour Interventions Reference lists of each article that met the criteria for inclusion also searched.	**Inclusion criteria:** ▪ children described as having ASDs ▪ at least some portion of the sample being 12 years of age or younger ▪ study has a recognized controlled experimental design, single-case design or group design. **Exclusion criteria:** none stated.	79 studies identified overall. The authors grouped the interventions into 5 categories: ▪ Modelling and reinforcement (33 studies) ▪ Peer mediated interventions (20 studies) ▪ Reinforcement schedules and activities (8 studies) ▪ Scripts and stories (10 studies) ▪ Miscellaneous (8 studies). The first, second and third categories are relevant to this review of ABA. Other categories may contain ABA type interventions but generalisations from these categories may not be relevant to this review of ABA as it is difficult to separate out the ABA components from other approaches. Modelling and reinforcement: All studies identified were very small (range n=2 to 26) and most were not controlled (90% were SCED). Settings included home, clinic, school and treatment program. In these studies, modelling was undertaken, discrete target behaviours were rated, feedback was given with suggestions for improvement and additional practice, reinforcement was used for appropriate responding. Peer mediated interventions Methods in the studies varied – usually an experimenter worked to teach peers to model and/or prompt appropriate social behaviours. The settings in these studies were limited to school or clinic. Studies were mostly small with SCED design. Reinforcement schedules and activities: Studies were mostly small with SCED design. All but one study were undertaken at school and most were implemented by teachers or experimenters. Reinforcement was used and, in some instances, prompts or time delays between prompts and reinforcement. No advanced conceptual skills were required and edible reinforcers were commonly employed. **Authors' conclusions:** The authors concluded that modelling and reinforcement interventions are "efficient and effective but may be best for older children with better mental skills and experiences and the ability to generalize these skills to other settings with minimal additional training". The authors noted that peer mediated interventions were popular but the approach was limited by the setting and the age of	▪ no clearly focused question ▪ simple descriptions of each study, who implemented the intervention, study setting, number of participants and whether there was follow up and generalisation ▪ studies very small, mostly SCED design ▪ description of results for each category in narrative format very vague and unhelpful ▪ no statistics or numbers quoted to support vague conclusions ▪ some suggestions given for future research. **Quality score:** fair (-)

			the child acting as a peer. The authors concluded that reinforcement schedules and activities can be effective as children with autism have an aversion to physical contact and the "need for sameness can make schedules particularly salient and effective with this population". Authors claimed that there was an important need for group research designs as the majority of the included studies were very small.	

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Ayres and Langone 2005) USA SR of single case experimental designs (SCED) studies	**Evidence level:** III-1 **Aim:** To synthesise research findings on efficacy of video based instruction for people with autism in two primary areas: instruction of social skills, and instruction of functional skills. **Search period**: not reported **Databases**: ERIC and PsycINFO (search terms provided). In addition, hand searching of three relevant Journals and an ancestral search of references.	**Inclusion criteria:** ▪ empirical study in a peer reviewed Journal ▪ use of video as an intervention tool for students with autism ▪ results reported for individuals with autism separately from others ▪ article written in the English language. **Exclusion criteria:** none reported	15 articles eligible for inclusion: 9 used video to teach social skills (representing 31 people with autism) and 6 used video primarily to improve functional skills (representing 14 people with autism). Ages ranged from pre-schoolers to 2 adults aged 20. All were SCED studies. Videos varied in terms of whether they presented adult or peer models, non-human models or footage of the individual with autism themselves (video self modelling, where prompting and off-task behaviour is edited out). Methodological critique of potential flaws, skewed interpretation of results and hypotheses for findings. For example, ceiling effects, accelerated baselines, highly variable data, inability to separate effects based on different treatment elements. Highlighted results included: ▪ five studies demonstrated the power of video for teaching conversational skills. Students with autism were able to accurately imitate the models presented via video. ▪ studies demonstrated how video self modelling was used to reduce tantrum behaviour. and how videos were an instructional component for teaching daily living skills such as shopping, cleaning glasses and mailing a letter. ▪ video has successfully been used to teach acquisition and generalisation of functional skills to students with autism. ▪ discussed how videos can serve as a supplement and extension of the *in-vivo* instruction, and for recording of conversations. Videos can also isolate steps of a process and show reliably delivered and standardized repetitions whilst manipulating important exemplars. ▪ little progress has been made in identifying the critical components of video models and video based instructions. **Authors' conclusions:** "Researchers are successfully using video to teach a variety of social and functional skills." "whilst this area of research is expanding, more detailed studies are needed to better describe specific aspects of video based instruction".	▪ focus on use of video based instruction for people with autism, included video modelling and video priming without models. ▪ dates of search period not stated (may not have been limited) ▪ narrow and selective range of databases searched ▪ hand-searching of three Journals, and ancestral searching of reference lists. ▪ no checklists, critical appraisal methods or coding described ▪ description of individual study methods and results given in Tables ▪ findings summarised in text. ▪ future research directions briefly outlined ▪ conclusions based on 15 primary studies ▪ no appraised studies eligible for inclusion in current review (all SCED). **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Campbell 2003) USA SR of single case experimental design (SCED) studies	**Evidence level:** III-1 **Aim:** To review the efficacy of behavioural interventions for problem behaviour in persons with autism. **Search period**: 1966 – 1998 **Databases**: PsycLit, ERIC, Medline. (Search terms provided). In addition, extensive hand-searching of pertinent journals.	**Inclusion criteria:** ▪ diagnosed with autistic disorder ▪ SCED studies only ▪ baseline and treatment phases present, repeated data points reported ▪ treatment targeted SIB, stereotypy, aggression, or property destruction. **Exclusion criteria:** Less than two baseline data points. Detailed description of coding and statistics. Three effect sizes were calculated for each article. Inter-coder agreement > 80% for categorical and good for continuous variables except length of treatment (Pearson's *r*=.703).	117 articles (representing 181 participants) met selection criteria. Behavioural treatments were significantly effective in reducing problem behaviour in individuals with autism. Treatment significantly reduced problem behaviour by 76% (mean baseline reduction effect size), significantly different from zero (p<.001). Across all studies, percentage of zero data (PZD) averaged 43 (SD=36) (significantly different from zero, p<.001). Hierarchical multiple regressions were performed. Type of problem behaviour and type of technique used (eg; punishment, positive reinforcement, sensory extinction) did not account for variance in effect size. In ANOVA analyses, the presence of pre-treatment functional assessment resulted in significantly higher average PZD scores in studies than those without (p=.012). Of studies with pre-treatment functional assessment, those employing experimental functional analysis (EFA) showed significantly greater behavioural suppression than those with other functional assessment methods. Higher reliability (p<.01) and a greater number of treatment observations (p<.05) were related to higher PZD treatment outcome scores. No other participant, treatment or experimental variables contributed to efficacy in the regression equation. **Authors' conclusions:** "Behavioural treatments are effective in reducing problem behaviours in individuals with autism." Results "match consistent findings that document the general lack of influence that participant variables, such as age, gender or level of mental retardation, exert within the context of the general effectiveness of behavioural treatment". "The most salient clinical implication is that behavioural treatments are more effective when preceded by a functional assessment". The review "points to the benefits of conducting EFA as opposed to indirect or descriptive methods of functional assessment" when treating problematic behaviours.	▪ explicit clinical questions described ▪ reasonable range of databases searched and extensive hand-searching ▪ thorough account of coding and statistical analyses ▪ used Bonferoni's technique to adjust p value for multiple tests ▪ acceptable inter-rater reliability for coding ▪ study results synthesised and moderating variables investigated statistically in the text and detailed tables ▪ PZD suggested as more stringent indicator of treatment efficacy as it indicates behavioural suppression, not just reduction ▪ all studies were SCEDs and therefore no primary study met criteria for inclusion in current review. **Quality score:** very good (++)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Horner et al. 2002) USA SR of single case experimental design (SCED) studies	**Evidence level:** III-1 **Aim:** To examine the effectiveness, efficiency and relevance of behavioural interventions for children 8 years of age or younger who have a formal diagnosis of autism and engage in problem behaviours. **Search period**: 1996-2000 **Databases**: ERIC, EXCEPTIONAL CHILD, PsycINFO. (Search terms provided). In addition, hand searching of many Journals.	**Inclusion criteria:** ▪ people with autism aged less than 97 months old (<8 years) ▪ used problem behaviour as dependent variable ▪ employed experimental design that allowed causal relationship to be investigated ▪ provided data for individuals ▪ included at least 3 data points for outcomes **Exclusion criteria:** none reported Inter-rater agreement on selection (100%) and appraisal (90%) assessed. Details given of data coded. NAS criteria for assessing intervention studies employed.	Retrieved 41 published research articles, 9 of which were eligible for inclusion, representing 24 participants and 37 comparisons. All were SCED studies with non-blinded observation. Conclusions based on 9 primary (mentioned above), as well as 5 secondary studies identified by hand-searching Journals. ▪ Nearly 60% of the comparisons reported 90% reduction in problem behaviour. ▪ 6 of 9 primary studies reported on maintenance, all showing that the behavioural reduction remained within 15% of initial levels at follow-up of a mean of 12 weeks. ▪ Only two studies examined generalisation. ▪ Interventions developed from functional assessment appear more likely to result in significant behaviour reduction. ▪ Authors conclude that behavioral intervention should (a) emphasise control of stimulus-based events that make the problem behavior irrelevant (eg, reduce access to aversive events, present regular access to preferred events, minimise rewarding problem behaviour); (b) teach socially appropriate behaviours that both make the child more competent in the environment and produces the same environmental effect as the problem behaviors; (c) organise consequences to prevent reinforcement of problem behaviour; (d) organise consequences to maximise reinforcement of competing, appropriate behaviours; (e) ensure that the specific procedures employed are within the skills, resources and values of those who must implement them; and (f) maintain systems of data collection to ensure that the effects of the intervention may be assessed. **Authors' conclusions:** "There is reason for significant optimism that behavioural interventions can result in reductions of problem behaviour of 80 to 90%." "When problem behaviors are identified, conduct a functional assessment." "There is insufficient information about generalisation and maintenance of effects in the reviewed literature".	▪ focus on early behavioural interventions for problem behaviour for children with autism ▪ described as a "targeted review". A narrow publication period considered ▪ narrow and selective range of databases searched ▪ details of coding and appraisal criteria given ▪ inter-rater agreement very high. The two items where it wasn't high were excluded – lifestyle change and impact on non-problem behaviour ▪ detailed description of individual study results given in tables, and table of appraisal criteria given in Appendices ▪ individual studies contributed more than one comparison and one study contributed 32% of the comparison data ▪ findings summarised in text, though with minimal methodological critique. ▪ future research directions briefly outlined ▪ conclusions based on 9 primary studies systematically identified, as well as 5 secondary studies identified by hand-searching Journals (not reported here as not meeting criteria for systematic search) ▪ no appraised studies eligible for inclusion in current review (all SCED). **Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(McConnell 2002) USA SR of SCED studies	**Evidence level:** III-1 **Aim:** To review the knowledge available from aggregated research on the characteristics of social interactions and social relationships among young children with autism, with special attention to strategies and tactics that promote competence or improved performance in this area. **Search period**: earlier date not specified, later date May 2000, later extended to October 2001. **Databases:** PsycINFO (search terms given) Also manual searches of journals and systematic searches for antecedent and descendant studies using references from published works and online citation indices.	**Inclusion criteria:** ▪ focus on teaching, training or providing intervention on social functioning of children with autism ▪ no other specific inclusion criteria given. **Exclusion criteria:** ▪ all or a large proportion of the subjects >8 years of age ▪ reports of case studies or program descriptions ▪ investigations published before 1979.	55 studies were identified. The authors divided the studies into 5 general categories: ▪ ecological variations (promotion of social interaction through manipulations/arrangements or general features of the physical/social environment) – 11 studies ▪ collateral skills interventions (where children with autism demonstrate increases in social interaction as a function of training in other skills) – 9 studies ▪ child-specific interventions (instructional and/or reinforcement procedures designed specifically to increase the skill, frequency or quality of social behaviours) – 15 studies ▪ peer behaviour (provision of social skills training and other manipulations to other children that are designed to change social interactions/skills for children with autism) – 30 studies ▪ comprehensive interventions (those that include 2 or more of the previous interventions) – 7 studies Ecological: Authors state that ecological variations can under some conditions produce weak to moderate effects on the social interaction of young children with autism but the effects are variable. Ecological interventions are thus likely to be necessary but not sufficient for producing changes in social interaction and development. Collateral: These interventions may increase social interaction by bringing children with autism into contact with typically developing peers thus activating natural processes for social development. Child-specific: These interventions can increase social interaction both as direct effects of intervention and through promotion of generalisation or maintenance. But these interventions in isolation have limited potential, especially in relation to long term effectiveness. Peer-mediated: These interventions have demonstrated powerful and robust treatment effects across a number of children, investigators	▪ search systematic but only one database accessed ▪ included studies appraised for internal and external validity and generalisation ▪ variations in the diagnostic procedures used for identifying children with autism in the included studies ▪ different ages of identification of children with autism may have confounded results ▪ difficult to separate out the changing patterns of the social environment and the ways this environment is influenced by a variety of social and policy issues ▪ clear descriptions of participants, methodology and quality of the included studies ▪ useful taxonomy developed by the authors for synthesizing the studies ▪ narrative synthesis of results in different categories of study ▪ effect sizes not estimated as this approach not considered by authors suitable for SCED studies ▪ because there are no quantitative results of effectiveness, it is difficult to determine the setting, participant and procedural features that lead to the most beneficial outcomes and there are no relative comparisons of treatment effectiveness across different intervention strategies ▪ most of the studies conducted in classroom settings with teachers or other professionals as interventionists and classmates as interactive partners ▪ no evidence for specific packages of social intervention procedures but some support in this review for intervention components. ▪ suggestions for future research provided

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
			and intervention variations. However they are limited as it is not clear that effects generalize to untrained peers and situations. Comprehensive: These interventions can produce effects on social interaction in intervention settings with some evidence of generalisation to other settings during the same time period. **Authors' conclusions:** "Social interaction development. . .should be a routine component (when needed) of any educational treatment program". However, the authors acknowledged that questions remained about the effects of different treatment components and about the short and long term generalisation effects of intervention or whether children with autism need different interventions compared to children with other disabilities or deficits.	**Quality score:** good (+)

Reference, country, design	Evidence level, aim and search method	Inclusion and exclusion criteria	Results and authors' conclusions	Comments and quality score
(Smith 1999) USA SR of non-randomised experimental studies	**Evidence level:** III-1 **Aim:** Not explicitly stated, but focus is to critique peer-reviewed outcome studies of early intervention for children with autism. **Search period**: 1980 – Dec 1996 **Databases**: ERIC, Medline, and PsycLit (search terms provided). In addition, searched citations in five "recently published" review papers.	**Inclusion criteria:** ▪ children average age of 5 or younger at treatment onset ▪ children received direct services ▪ services were comprehensive, aimed at addressing multiple problems ▪ provided data on treatment outcome ▪ published in peer reviewed Journals **Exclusion criteria:** none reported	13 publications reviewed: 10 related to ABA, 1 to TEACCH, and 2 to Colorado Health Sciences. The results from the ABA studies reported here, relating to 7 programmes: 4 home-based (May, Murdoch, UCLA, and UCSF) and 3 school-based (Douglass, LEAP, PCDI). Mean IQ gains of 7-28 points were reported. However there were mixed results between studies, and no clear correlations among treatment intensity, treatment model, and outcome. The UCLA home-based Young Autism Project programme (Lovaas, 1987) provided the most intensive treatment of 40 hours/week and reported the highest gains. The intensive treatment group averaged 22-31 higher IQ points than two control groups. Whilst relatively strong methodologically, limitations included group assignment based on therapist availability, and differing IQ assessment tools at intake. Individual children varied substantially in treatment responsivity and though children who were relatively high-functioning at intake tended to improve the most, exceptions were numerous. Key limitations across most ABA studies included: lack of random assignment to treatment group, lack of standardised assessment of diagnosis of autism, lack of blind independent clinical assessments, lack of replicable description of intervention, lack of follow-up beyond treatment (except for McEachin et al, 1993), small sample sizes, reliance on a single outcome measure, and use of inexperienced and poorly trained therapists. Home-based studies were particularly flawed. **Authors' conclusions:** "Though all investigators reported substantial improvements, the nature of improvements varied substantially across studies." Problems "hinder drawing conclusions from existing early intervention studies." "Further progress is likely to require more rigorous clinical trials, more comprehensive pre-treatment and follow-up assessments, and greater attention to brain-behaviour relationships."	▪ focus on early intervention for children with autism ▪ moderate range of databases searched, with reference checking of selected recent reviews ▪ no appraisal checklists mentioned ▪ studies described critically in the text, and detailed tables with description of interventions, methods and study results ▪ detailed, balanced discussion of individual studies and their strengths and limitations ▪ detailed discussion of areas for future research ▪ conclusions based on 13 studies (10 ABA) ▪ for ABA, described range of non-randomised experimental (where allocation methods were open to biases) or cohort studies ▪ 1 study met criteria for inclusion in review - (Sheinkopf and Siegel 1998) **Quality score:** good (+)

4. Level III-2 studies

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Carr and Felce 2007) Wales, UK Non randomised experimental study	**Evidence level**: III-2 **Interventions:** PECS + eclectic teaching approach (ECI) **Comparator:** Eclectic teaching approach (ECI) **Study setting**: Interventions took place in the child's own classroom which had a permanent teacher and 2 classroom aides with a range of 6 to 8 students per class.	**Participants:** N=24 in PECS group and N=17 in control group, recruited from special education classrooms or specialists units for autism across South Wales. **Inclusion:** ▪ children aged between 3 and 7 years ▪ previous diagnosis of autism from a clinical practitioner (verified by classroom teacher ▪ no previous PECs teaching beyond Phase 1. **Sample characteristics**: Age (mean): PECS: 5.5 years Control: 5.9 years Language (mean) (PLS-3UK): PECS: 7.7 months Control: 9.4 months Adaptive behaviour (mean) (VABS): PECS: 14.7 months Control: 14.8 months **Diagnosis**: 100% with autism diagnosed by a clinical practitioner and verified through classroom	Children were allocated to PECS if they lived within 50 miles of the researcher's base and children were allocated to control if they lived outside this range. All children received an eclectic education program with no particular specialized approach to intervention. Children in the PECS group had a personal communication folder and accumulated a set of Velcro-backed coloured pictures individualised to his/her range of preferred items. The researchers were trained to deliver PECS teaching. Reliability between the 2 researchers was measured for the scoring of outcomes. **Intensity of intervention:** The PECS children had a total of 15 hours of PECS teaching up to Phase III over a period of 4 to 5 weeks. **Timing of assessment**: The PECS children had	An observation instrument was designed to record communicative interactions between the children and their classroom teachers and assistants. There were 5 categories: ▪ total number of child to adult initiations ▪ no of child to adult initiations with a response from the adult ▪ total no of adult initiations with opportunity for the child to respond ▪ no of adult initiations with opportunity for the child to respond and with a response from the child ▪ no of adult initiations with no opportunity for the child to respond. 56% of assessments undertaken by both researchers independently; presumably the other assessments scored singly by either researcher.	**Pre-treatment:** No significant differences between groups on chronological age, or VABS or PLS-3UK assessments. **Post-treatment:** There was no evidence of a difference in communicative outcomes between T0 and T1 for PECS children. Frequency of total child to adult initiations was significantly higher for PECS than control children at T2 ($z=5.3$, $p<0.00003$) The frequency of linguistic communications was significantly higher for PECS than for control children at T2 ($z=6.93$, $p<0.00003$) The percentage of adult responses given to child initiated communications was significantly higher for PECS than control children at T2 ($z=2.8$, $p<0.0026$) There was no significant difference in frequency of initiations giving opportunity for child response between groups at T2. There was a significantly higher %age of child	▪ no significant differences at baseline between groups in chronological age, VABS scores and PLS-3UK scores ▪ group assignment was not randomised but allocated on the basis of place of residence (children were assigned PECS if they lived within 50 miles of the researcher's base and children were assigned control if they lived outside this range) ▪ numbers not equivalent in each group ▪ within group measures were undertaken for the PECS group to control for the effects of maturation ▪ reliability measures indicated good inter-observer agreement in scoring ▪ results of effectiveness with PECS do not

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		teachers	a 2 hour evaluation at Time 0 to control for the effects of maturation which was followed by 5 weeks of no intervention. All children had 2 hour observations at Time 1 (1 week prior to the commencement of teaching) with measurement of VABS and PLS-3UK (baseline). The PECS children then had 4-5 weeks of intervention and the control children had no intervention other than the regular program. All children had 2 hour observations at Time 2.		responses to adult initiated communications for PECS than for control children at T2 ($z=2.3$, $p<0.0107$) The frequency of adult to child initiations with no opportunity to respond was significantly lower for PECS than for control children ($z=-1.65$, $p<0.0495$) There were no differences between outcomes at T1 between groups. Also, for all outcomes except adult to child initiations with opportunity for child response, there was a significant change in outcomes between T1 (baseline) and T2 for PECS children (within group difference). There was a significant difference between T1 (baseline) and T2 for control children only for the outcome: adult to child initiations with opportunity for child response $T(17)=35$; $p<0.05$. **Authors' conclusions:** "Communicative initiations and dyadic interactions increased significantly between the children and teachers in the PECS group but not for the control group".	necessarily translate to maintenance over the long term ▪ the PECS intervention required 2 extra teachers in the classroom. The study has not been able to determine whether the results are due to extra teacher input/presence or the procedures themselves ▪ the presence of the researchers during T2 observations may also have affected the assessment of outcomes as teachers may have been aiming for optimum outcomes with PECS. **Quality score:** fair (-)-

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Eikeseth et al. 2002) (Eikeseth et al. 2007) Norway Non randomised experimental study	**Evidence level**: III-2 **Intervention:** Intensive behavioural treatment based on the UCLA treatment model (based on 1981 Lovaas manual, except that aversives not used) (EIBI) **Comparator:** Eclectic care intervention (ECI) treatment designed to reflect best practices for serving children with autism (incorporated elements from TEACCH, sensory motor therapies and ABA as well as personal experience) **Study setting**: Public kindergartens and elementary schools for typically developing children in Norway with no two participants enrolled in the same class (ECI).	**Participants:** N=25 children (13 in EIBI group, 12 in ECI) who were referred during Nov 1995 and Nov 1998. There were 2 dropouts/1 from each group (it is not clear, but it appears that these were in addition to the 25 randomised). **Inclusion criteria:** ▪ diagnosis of childhood autism (ICD-10) from both the ADI-R and an independent child psychologist ▪ aged 4 to 7 years at enrolment ▪ deviation/ratio IQ of 50 or above on WPPSI-R or BSID-R ▪ absence of medical conditions other than autism. **Sample characteristics**: Gender (male/female): EIBI: 8/5 ECI: 11/1 Age (mean): not reported	Allocation based on availability of personnel to supervise the ABA treatment. EIBI: Prior the study, none of the therapists had any supervised experience in the implementation of behavioural treatment for children with autism. They received 10 hours/week of supervision during the study by trained supervisors and weekly meetings with project directors (psychologists with 10 years experience implementing the treatment). Parental worked alongside the therapists at school for the first 3 months of Rx with the purpose of extending the Rx into the home and community. Weekly 2 hour meetings held with each child (with child, caregiver, therapists, supervisor and director) to modify Rx. ECI: Treatment was a combination of approaches and was individually selected for each child based on recommendations of team of school personnel.	Assessments by licensed clinical psychologist using: ▪ intellectual functioning - standardised intelligence tests (either WPPSI-R or the WISC-R) ▪ visual spatial skills (Merrill-Palmer Scale of Mental Tests (MPSMT) or performance subscale of WPPSI-R or WISC-R for older children) ▪ language functioning (Reynell Developmental Language Scales (RDLS) or verbal subscale of WPPSI-R or WISC-R for older children) ▪ adaptive behaviours (Vineland Adaptive Behaviour Scales) (VABS)	**Pre-treatment:** No differences between groups in chronological age, IQ, language or VABS. **1 year follow up:** EIBI: m=28 Rx hours/wk ECI: m=29.1 Rx hrs/wk Significant differences: full scale IQ, language comprehension, expressive language and Vineland communication domain Full scale IQ: EIBI: +17 points ECI: +4 points p<0.01 Language comprehension: EIBI: +17 points ECI: +0 points p<0.05 Expressive language: EIBI: +17 points ECI: -2 points p<0.05 Vineland communication domain: EIBI: +15 points ECI: -1 point p<0.01. Composite VABS: EIBI: +11 points ECI: +0 p<0.05 Also children in EIBI group more likely to have IQs in average range compared to control (7/13 vs 2/12). No significant differences between groups for: daily living and socialization (VABS subscales).	▪ interventions took place at school rather than at home ▪ children received an average of 28 hours per week (instead of the 40 recommended by Lovaas) ▪ children were more able than those in some of the other RCTs of ABA, thus achieving more gains ▪ no differences between year 1 and year 2 ½ follow up for most measures, except for adaptive behaviour and socialization (suggesting that it may be important to extend ABA treatment beyond 1 year to achieve reliable gains in social behaviour and daily living skills) ▪ large individual differences in gains were made by children in the EIBI group (no variables at intake predicted these differences). Thus a priority for future research is to determine which variables are associated with individual differences in response to Rx (age was not a predictor) ▪ note: 40% of treatment in the EI arm was ABA techniques Limitations:

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		Diagnosis: 100% diagnosed with childhood autism (ICD10) (recent, ie, <6 months earlier) from both the ADI-R and an independent child psychologist	Implemented one to one in a separate room by therapists. Therapists had weekly 2 hour consultations with the supervisors and directors who oversaw the behavioural treatment. **Intensity:**: Each participant had a minimum of 2 therapists and a minimum of 4 to 6 hours per week of treatment by a special education teacher, followed by treatment with aide(s) (Rx administered with the child alone with the therapist in a separate room). When not in the sessions, the child was mainstreamed with his/her classmates while being shadowed by the therapist. **Timing of assessment**: At intake into the study, 1 year after treatment began and approximately 2 ½ years after treatment began (2007 publication). **Statistics:** T tests and chi square tests with Dunn Bonferroni corrections.		**2 ½ year follow up** (ie, at average 8 years 2 months of age): EIBI: m=18 Rx hrs/wk ECI: m=16 Rx hrs/wk Significant differences: Full scale IQ: EIBI: m=+25 points ECI: m=+7 points VABS adaptive functioning: EIBI: m=+12 points ECI: m=-10 points VABS communication: EIBI: m=+20 points ECI: m=-7 points VABS daily living skills: EIBI: m=+9 points ECI: m=-6 points VABS socialization skills: EIBI: m=+12 points ECI: m=-12 points For social emotional functioning, there were no differences between groups (except for lower rates of social problems and aggressive behaviour for children having EIBI (compared to chn having ECI)). Children in EIBI group more likely to have IQs in the average range compared to ECI (7/13 vs 2/12). **Authors' conclusions:** "The behavioural treatment group showed larger increases in IQ and adaptive functioning than did the eclectic group [and] fewer aberrant behaviours and social problems at followup."	▪ small sample size ▪ not randomised ▪ measures focussed more on cognitive than social development ▪ Strengths: ▪ comprehensive uniform assessment protocols administered by blind examiners ▪ manualized research-based interventions for the EIBI group ▪ treatment supervision by experienced personnel ▪ measures of the amount of treatment that the children received, skills addressed in treatment and education of the therapists ▪ group assignment performed by a professional who was independent of the study ▪ careful documentation of types of approaches used in each arm. **Quality score**: good (+)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Magiati et al. 2007) UK Cohort study	**Evidence level**: III-2 Cohort study **Intervention:** Home-based Early Intensive Behavioural Intervention (EIBI-H) programme offering 1:1 home teaching, all using trial teaching techniques. **Comparators:** Specialist autism-specific school-based nursery (School) programme sourced from 10 different schools, including autism-specific nurseries or units or within generic special schools. Eclectic programme emphasising structure, visual cues, individualised teaching and close liaison with parents. Used mix of approaches including TEACCH, PECS, Makaton, and SPELL (ECI). **Study setting**: Community setting (schools or homes).	**Participants:** N=44/63 (70%) children with ASD participated. ▪ chronological age (CA) 22–54 months. ▪ independent diagnosis of ASD ▪ diagnosis additionally confirmed in the majority of cases on the ADI-R ▪ no additional major medical diagnoses ▪ English main language at home ▪ living within 3 hours of Central London. ▪ enrolled in EIBI or autism-specific nursery ▪ receiving no other intensive intervention. 19 excluded (7 from EIBI, 12 from school) for following reasons: ADI-R criteria not met, late intakes, additional medical diagnosis, 2 moved to other groups, 1 dropped out of EIBI, 3 followed EIBI in school	Naturalistic opportunistic study where familiies had already chosen the intervention they wished to pursue. Intention-to-treat approach. **Follow up**: 2 years (mean ranging 23-27 mths), no difference between groups. **Intervention intensity:** EIBI: Most attended an initial workshop, and employed a consultant, supervisor and/or therapist. The average number of therapists working for each family was 9 (range 3-18). ECI: 1. The average amount of 1:1 teaching was 6 hours per week (range 90 minutes to 20–25 hours per week). Average hours reported per week/per child were higher for EIBI children than ECI children at baseline (32.4 cf 25.6 hours; p<.001) and follow-up (33.2 cf 27.4 hours; p<.001).	Children assessed using: ▪ "Best test" mental age (MA) and IQ from one of: Merrill-Palmer Scale of Mental Tests (MPSMT), Bayley Scales of Infant Development (BSID), Weschler Preschool and Primary School Intelligence scales-Revised (WPPSI-R)., as appropriate ▪ Adaptive behaviour: Vineland Adaptive Behaviour Scales (VABS) ▪ Receptive and Expressive language: British Picture Vocabulary Scale-II, and the Expressive One-Word Picture Vocabulary Test-Revised (raw scores) ▪ Play: Symbolic Play Test-Second Edition (SPT-II) and Test of Pretend Play, depending on scores. ▪ ADI-R to assess autism severity. Questionnaire also assessed type, intensity and duration of interventions from parents.	**Pre-treatment:** Parents of EIBI children were more highly educated (p<.04), than those in ECI group. IQ and VABS did vary between groups but in terms of clinical significance the difference was said to be very small (4 points). Both were highly correlated (r=.84) and so IQ was used as a covariate in ANCOVA's. **Post-treatment:** There were no group differences in cognitive, play and language skills or severity of autism. There was a difference *approaching* significance in VABS Daily Living Skills standard scores (p =.06) such that EIBI children's scores decreased less compared to those of the ECI group children. This was clinically small (9 pts, and less than one SD). No child in either group was in mainstream school without 1:1 support. Tracking individual patterns, change was	▪ "ABA-naïve" classes not possible ▪ age range of participants narrow ▪ assessment not blinded ▪ treatment facilitators and parents not able to be blinded ▪ parents were more highly educated in EIBI group and may respond differently to their children at home ▪ treatment groups varied in IQ at baseline; baseline IQ entered as a covariate in ANCOVA's. ▪ EIBI children received more hours of intervention than those in ECI group, however the average hours per week were of a similar intensity for EIBI and school groups (32 vs 25, respectively) ▪ treatment fidelity not directly assessed ▪ inter-rater reliability acceptable ▪ maintenance not

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		setting, and 1 joined another intensive method. **Sample characteristics:** Sex: 89% male Age range: 23-53 months EIBI: N=28/35: 27 boys, 1 girl ECI: N=16/28: 12 boys, 4 girls **Diagnosis**: Independent, professional diagnosis of autism or ASD confirmed in majority of cases using Autism Diagnosis Interview Revised –(ADI-R).	Other treatments were used by the two groups. Significantly more EIBI group than School group children followed special dietary and other biological interventions, whilst more School children than EIBI children followed extra-curricular educational interventions.	Video-recorded assessments for IQ, play and language scored blind by another assessor to demonstrate interclass correlation coefficient of 0.99 for raw scores, and for a selection of ADI-R; inter-rater agreement averaging 84%.	generally small to moderate for most children; few made major improvements. In all areas assessed, the extent of individual variation in progress was evident in both groups. A hierarchical multiple regression indicated that baseline IQ and receptive language raw scores best predicted progress after 2 years. Baseline Vineland standard scores and ADI-R total raw scores also contributed to the model. **Authors' conclusions:** **"**Community-based EIBI is beneficial for some children with autism. However, *specialist*, relatively intensive nursery provision in the UK produced similar outcomes." "Children with initially high cognitive and language functioning tended to make more progress, regardless of intervention group." "Our data support the growing consensus that no one intervention for children with ASD is universally superior to all others."	investigated **Quality score**: good (+)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Reed et al. 2007) UK Cohort study	**Evidence level**: III-2 Children were analysed by whether a high intensity or low intensity programme offered. **Intervention** High intensity (HI) EIBI: 20-40 hours intervention per week from one of three home-based ABA programmes offering mostly 1:1 teaching provided by tutors and guided by an ABA supervisor: Lovaas style programme (n=4); Verbal behaviour approach (n=5); and the Comprehensive Application of Behaviour Analysis to Schooling (CABAS) (n=5). All programmes used an antecedent, behaviour (sometimes prompted) and consequence. As the sample size for each intervention programme type was under 6, we will not report group comparisons. **Comparator**: Low intensity (LI) EIBI: 10-20 hours of intervention per week, offering similar generic	**Participants:** N=27 children with autism participated, recruited from an existing early intervention programme. Children selected if aged between 2 years 6 mths and 4 years, had received no other major intervention, were at the start of their intervention, and had received a diagnosis of ASD. **Sample characteristics (of children with ASD)**: Sex: all male HI: N=14 males LI: N=13 males **Diagnosis**: All diagnosed with ASD, but no details given of testing schedule.	Eligible children in three intervention groups assigned based on what was offered to the child in their particular area (of similar socio-demographic profile), divided into high-intensity and low-intensity comparators. **Follow up**: 9-10 months from baseline (presumably immediately post/during intervention) **Intervention intensity:** Each session lasted 2-3 hours, involving 8-10 tasks of 5-10 min each, separated by a 5-10 minute break. HI: 20-40 hours intervention per week, mean=30.4 (SD=5.0). LI: 10-20 hours of intervention per week, mean=12.6 (SD=2.3), including up to 4 home-based 1:1 direct teaching sessions per week.	Educational psychologist administered the following: ▪ Gilliam Autism Rating Scale (GARS), comprising subtests: stereotyped behaviours, communication, social interaction, and developmental disturbances. ▪ British Abilities Scales (BAS II) – Early Years Battery - measures cognitive abilities which index educational achievement. Psychologist assisted parents in completing: ▪ Psycho-educational Profile-Revised (PEP-R) measuring developmental functioning in seven developmental domains. ▪ Adaptive skills: Vineland Adaptive Behaviour Scale (VABS) measures day-to-day adaptive functioning along four domains (socialization, communication, daily living skills, motor skills). Tutors and parents reported on intervention characteristics post	**Pre-treatment:** No significant difference on sample characteristics between intensity groups. **Post-treatment:** A significant slight decrease in GARS for overall autistic severity for both intensity groups, and no significant between-group difference. Whilst within-group pre-versus-post intervention increases found for intellectual functioning and educational functioning for HI group, and for educational functioning for LI group, the only significant group-difference in change scores was for educational functioning ($p<.01$), reflecting higher gains for high-intensity group. Mean gains in the three outcome measures was computed and the relationship between mean gain and the number of hours per week of intervention was represented as a regression equation, for those in the HI EIBI	▪ children allocated based on programme availability in area ▪ groups did not vary at baseline, but may have in unmeasured differences (eg, parent or socio-demographic factors reflecting any biases in availability to programmes by area) ▪ participation rates not reported and little detail on recruitment ▪ any drop-out rates not mentioned ▪ all groups received interventions which were based on ABA principles, differing in intensity/hours per week – but also in type of programme ▪ assessment was blinded ▪ parents and intervention staff not able to be blinded to intervention ▪ inter-rater reliability of coding not investigated ▪ power calculations not mentioned ▪ treatment fidelity not measured but all

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
	ABA programmes to those in the EIBI group but at a lower intensity. **Study setting**: Home-based setting			treatment.	group. This revealed a pattern of decreased mean outcome gains as the hours of intervention per week increased. **Authors' conclusions:** "The high-intensity group did better than the low-intensity group, but within the high-intensity group there was an inverse relationship between the temporal input and the gains. This finding implies that the suggested 40 hours/week input may not be optimal, and once over a certain level of temporal input, perhaps around 20 hours a week, there are diminishing returns for increasing the temporal input of a programme."	followed manuals ▪ maintenance not measured. **Quality score**: fair (-)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Remington et al. 2007) UK Cohort study	**Evidence level**: III-2 Cohort study **Intervention:** Home-based Early Intensive Behavioural Intervention (EIBI-H). Based on applied behavioural analysis (eg; shaping, chaining, prompting, fading, modelling, discrimination learning, task analysis, functional analysis) and using discrete trial training methods. Programs delivered by Local Education Authority (LEA) funded University-based clinics (n=13) or private sector clinics (n=12) with some variation in programme between providers and individualised for children . **Comparator:** Usual care (ECI) of standard treatment offered as statutory services by their LEA. **Study setting**: EIBI was home-based Both EIBI and usual care groups accessed	**Participants:** N=44 preschool children with autism who met study criteria, recruited through referrals, advertisements with autism society, parent groups and charities. Selection criteria included: diagnosis with autism, aged between 30 and 42 months at induction, free of any serious medical condition that would interfere with intervention or affect development, and living in the family home. 7 children, 2 in EIBI and 5 in ECI groups did not participate in video-taped assessment due to behavioural problems, inattention, or absence of parental permission. These children didn't differ from others at baseline. **Sample characteristics**: Participants' gender not reported. EIBI: N=23, Mean age: 35.7 mths (SD=4.4) UC: N=21, Mean	Groups were identified from recruitment strategy of those who had opted for EIBI and those not actively seeking behavioural intervention (usual care group) with no experimental assignment to group. **Follow up**: 12 months & 24 months post baseline **Intervention intensity:** EIBI: 2 years of home-based 1-to-1 therapy delivered by trained tutors and parents, (supervised by more experienced staff) for average of 25.6 hours per week (SD=4.8; range=18.4-34.0). Regular team meetings held with tutors and parents. Children in EIBI and UC groups attended mainstream or special needs schools, though EIBI attended less frequently due to demands of home programme. Schools frequently incorporated TEACCH and PECS methods.	Children outcomes: Administered by an author: ▪ Nonverbal communication: Early Social Communication Scales – video-taped observational instrument of rates of: "Initiating joint attention", & "responding to joint attention". ▪ Intellectual functioning: Bayley Scales or the Stanford Binet Intelligence Scale, depending on age of child. ▪ Language: Reynell Development Language Scales: expressive language and comprehension. ▪ Adaptive skills: Vineland Adaptive Behaviour Scale-Survey (1 week prior to assessment visit), measuring: socialization, communication, daily living skills, motor skills. Parent ratings of child behaviour: ▪ Positive Social subscale of the Nisonger Child Behaviour Rating	**Pre-treatment:** Children in UC group on average 3 months older than EIBI (p<.05). **Post-treatment:** Using ANCOVAs, with baseline scores and age as covariates: Significant main group effects at p<.05 level demonstrating an advantage for EIBI over UC at 12 mths, maintained at 24 mths for: responding to joint attention: IQ; Mental age; Vineland daily living skills, and a trend for motor skills (p=.057). At 24 mths, mothers' rated positive social behaviour higher in EIBI group than UC group (and a trend for fathers' ratings, p=.053). More children were able to obtain a score on Reynell Development Language scales in EIBI group than UC group at 12 and 24 mths. No significant group differences for observed initiating joint attention, Vineland Composite score, Motor skills, Socialization, or Communication subscales; for parental	▪ observational study, no assignment to group ▪ participation and drop-out rates not reported ▪ assessment "blinded", however assessment occurred in children's homes and "physical or behavioural cues may have signalled their treatment goup ▪ facilitators and parents not blinded ▪ inter-rater reliability of videotapes acceptable ▪ "ABA-naïve" classes not possible as some aspects employed in schooling ▪ treatment groups varied in age at baseline; adjusted for in some analyses. ▪ no adjustment of p for multiple tests ▪ researchers could not control course of therapy as services provided by range of providers, though some supervision provided. ▪ treatment fidelity variable as staff turnover common and new staff "slow to train". Average of 25.6 hours of home-based EIBI per week instead of target intensity of 40 p/wk.

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
	publicly funded educational facilities.	age=38.4 mths (SD=4.0) **Diagnosis**: All newly diagnosed with autism based on Autism Diagnostic Interview -Revised (ADI-R)	Some families in both groups also used other interventions including dietary interventions, prescription medication, high-dose vitamin injections, and homeopathic interventions. **Statistics** ANCOVAs employed, entering baseline outcome measures as covariates. Statistical measures of a *reliable* IQ response determined as having 24 mth follow-up IQ's which deviated by more than 23.94 IQ points from baseline (the SD of the variation from baseline to 24 mths in the UC group). Statistical measures of a *clinically significant* IQ response determined as having an IQ higher than 81.93 at 24 mth follow-up: halfway between mean baseline IQ of all 44 children and the typical population mean (of 100).	Form ▪ Developmental Behavior Checklist ▪ Autism Screening Questionnaire Parent self-report: ▪ Hospital Anxiety and Depression Scales ▪ Questionnaire on Resources and Stress-Friedrich-Parent and Family Problems. ▪ Kansas Inventory of Parental Perceptions-Positive Contributions Inter-rater reliability of videotapes determined for 25% of children and interclass correlations were between 0.95-0.99	ratings of children's developmental or autistic behaviour. Fathers in EIBI group reported more symptoms of depression at 12 and 24 mths than UC groups ($F(1,28)=5.19$, $p<.031$). No other group effects on parental outcome effects. 5 children in EIBI and 3 in UC responded both *reliably* and *clinically* in IQ by 24 mths follow-up (see Method). No reliable regression in IQ for EIBI group c.f. 3 regressing reliably in UC group. **Authors' conclusions:** **"**After 2 years, robust differences favoring intensive behavioral intervention were observed on measures of intelligence, language, daily living skills, positive social behavior, and a statistical measure of best outcome for individual children. Measures of parental well-being, obtained" … "produced no evidence that behavioral intervention created increased problems".	▪ parental self-report of children's behaviour is open to biases, as parents seeking EIBI and more personally involved in therapy may be more motivated to see and rate improvement. ▪ those achieving reliable and clinically significant change were identified (incorporating baseline IQ scores) for a small sample and the variation between groups was very small (3 versus 5). ▪ argued that increase in depressive symptoms for fathers in EIBI group may in part be "regression to the mean" as they had higher scores at baseline than UC. ▪ no measurement of maintenance ▪ power calculations not reported. **Quality score**: fair (-)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Zachor et al. 2007) Israel Non randomised experimental study	**Evidence level**: III-2 **Interventions:** ABA approach (EIBI): program based on ABA principles with a curriculum that included DTT, naturalistic and incidental teaching techniques (EIBI) **Comparator:** Eclectic-Developmental (ECI) approach: program based on principles from several approaches, mainly from the developmental oriented philosophy and the DIR model but also strategies from TEACCH and ABA as well. **Study setting**: Programs were based in 2 different centres (located in different counties), both included preschool routines and provided services for 8 hours per day and the children enrolled were under 3 years of age. Both programs received the same budget per child from the same national agencies.	**Participants:** N=20 received EIBI and n=19 received ECI. Children in each group were matched for age, autism severity and cognitive level. **Sample characteristics**: Gender: EIBI: 19:1 (male/female) ECI: 18:1 (male/female) Age (mean/range): EIBI: 27.7 months (22-34) ECI: 28.8 months (23-33) **Diagnosis**: All were diagnosed with autism using the ADI and met established criteria for autism/PDD-NOS according to DSM-IV criteria.	Both programs were center based, used preschool routines, provided services for 8 hours per day and the children enrolled were <3 yrs of age. Pre-intervention evaluation (baseline) was performed within the first month of enrolment with ADOS (only language and communication and reciprocal social interaction domains) and cognitive evaluation (BSID or SBIS according to children's language abilities). EIBI: One on one treatment was provided by skilled behaviour therapists for 35 hours/week – therapy was individualised for each child. Program included speech and occupational therapy and group activities. Children were included in supervised preschool program if they attained sufficient skills. Success was defined as accurate performance in 80% of the trials. Therapists supervised by trained	**Pre treatment:** ADOS test module 1 or 2 (only language/ communication and reciprocal social interaction domains) and cognitive evaluation (BSID-II or SBIS according to childrens' language abilities) **Post treatment**: ADOS Module 1 or 2 (only language/ communication and reciprocal social interaction domains)	**Pre-treatment:** No significant differences between groups in their ADOS scores (language and communication and reciprocal social interaction) and IQ. **Post-treatment:** Post treatment ADOS scores: ▪ Language and communication mean (SD): EIBI: 7.2 (4.1) ECI: 11.1 (6.7) ▪ Reciprocal social interaction mean (SD): EIBI: 11.1 (6.7) ECI: 13.3 (4.8) Changes in ADOS scores were assessed by MANOVA (intervention group X time) repeated measures. For language and communication, significant differences were found btwn EIBI and ECI, $p<0.01$. For reciprocal communication, differences btwn groups almost reached significance, $p=0.07$. Stability of diagnosis was also assessed. At baseline, 18/19 in ECI group and 19/20 of EIBI	▪ the authors claim stringent matching criteria in reference to the children's age, profile and intensity of intervention in both groups ▪ both intervention groups had the same number of weekly school-based hours and the same government allocated budget per child ▪ no difference was reported between groups in the pre-intervention evaluation (ADOS scores and IQ scores). However, no table provided to show other baseline characteristics of children between groups ▪ sample not randomised – children allocated by matching on characteristics ▪ only systematic bias claimed for group allocation is by place of residence (programs in 2 different centres) ▪ lack of blinding in the assessment of outcomes may have caused bias. Not clear

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
			behaviour analysts. ECI: Program included daily work in small group activities supervised by special education teachers with experience in autism. Children also received individual therapy from various therapists: speech and language, occupational and music therapies and structured cognitive teaching. Parents also involved with program. Children included in supervised preschool program if they had attained sufficient skills. **Intensity of Rx:** EIBI had 35 hours/wk; intensity not clear in the ECI group **Timing of assessment**: Baseline assessment and post treatment assessment after 1 year of treatment. **Statistics:** MANOVA and ANOVA		group had autism and 2 children (1 from each group) had ASD. Change of diagnostic classification significantly higher for the EIBI group (4 were no longer on the autism spectrum in EIBI group vs none in ECI group), $p<0.05$. Pre-treatment IQ scores were positively related to ADOS scores at pre and post intervention times but not to progress over time. **Authors' conclusions:** "Pre-post intervention differences in language and communication were significant only for the ABA group. Both groups showed significant improvement in reciprocal social interaction domain . . .[but] the effect size was greater for the ABA group." "Behavioural intervention is more effective than eclectic approach in improving autism core symptoms in young children with autism".	who did the assessments. **Quality score**: fair -

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Cohen et al. 2006) USA Non randomised experimental study	**Evidence level**: III-2 **Interventions:** Early intensive behavioural treatment (EIBI) **Comparator:** No EIBI treatment – eclectic services mainly from special day classes from local public schools (ECI) **Study setting**: Dependent on the age of the child. For EIBI group, initially 26 to 31 hours/wk of home instruction, 3 to 5 hours/wk of peer play and 6 to 9 hours/wk of preschool. As the study progressed, the balance of these components changed. Settings varied for children in the comparison group as these were chosen by parents (most were enrolled in special day classes at the age of 3 at local schools, operating for 3 to 5 days a week for up to 5 hours per day)	**Participants:** 42 children in 2 groups (21:21). 5 dropouts but it appears that these were additional to the 42 (3 in EIBI group and 2 in control group) **Inclusion:** ▪ primary diagnosis of ASD or PDD NOS based on an evaluation by an independent licensed psychologist and confirmed by ADI-R ▪ pre-treatment IQ above 35 on Bayley Scales of Infant Development – Revised (BSID-R) ▪ chronological age between 18 and 42 months at diagnosis and under 48 months at Rx onset ▪ no severe medical limitation or illness including motor or sensory deficits that would preclude a child from participating in 30 hours per week of Rx ▪ residence within 60 km of the Rx agency ▪ no more than 400 hours of behavioural intervention prior to intake ▪ parental agreement to participate actively in	Authors stated that legal and ethical issues precluded randomisation – children allocated to groups according to parental selection of programs. Children in comparison group were matched to those in EIBI group. **Intensity of intervention:** Chn in EIBI group had 35 to 40 hours of EIBI based on Lovaas' UCLA treatment model (17/21 remained in the program for the full 3 years but dropouts included in the final analysis). EIBI consisted of: in-home 1:1 instruction, peer play training and regular education classroom inclusion (proportion dependent on age of child). Chn in the comparison group (ECI) received community services that their families selected from the Matrix of Educational Options. **Timing of assessment**: Outcomes measured prior to treatment and at year 1, 2 and 3.	**Pre treatment:** Licensed psychologist independent of the study conducted: ▪ Standardised behaviour observation ▪ Parent interview ▪ Developmental tests (BSID-R, MPSMT, RDLS, VABS) **Post treatment** Primary: IQ (BSID-R or WPPSI) Secondary: ▪ Merrill Palmer Scale of Mental Tests (MPSMT) ▪ Reynell (RDLS) Language Comprehension ▪ Reynell (RDLS) Expressive Language ▪ Vineland Adaptive Behaviour Scales (VABS) ▪ Classroom placement	**Pre-treatment:** EIBI group had significantly greater proportion of children with ASD, parents had significantly more education and were significantly more likely to be married than comparison group. **Post-treatment:** The results shown below give the baseline/pre-treatment outcome measurement and the outcome measurement at year 3 IQ (mean): Baseline/year 3 EIBI: 62/87 (incr of 25) ECI: 59/73 (incr of 14) Difference btwn grps: p<0.05 MPSMT: No difference btwn grps – both grps had a mean increase of 13 points. RDLS Comprehension: No difference btwn grps, although trend, p=0.06. EIBI: 52/72 (incr of 20) ECI: 53/62 (incr of 9) RDLS Expressive: No difference btwn	▪ study not randomised as treatment was funded by public agencies that were required to offer free and appropriate services ▪ outcomes assessed independently by a trained evaluator who was not aware of the group allocation (however no evidence provided to ensure that they remained unaware) ▪ no significant differences at baseline between groups in chronological age, VABS scores and PLS-3UK scores However, the EIBI and comparison groups differed on other demographic variables at baseline (mother education, father education and diagnosis). The EIBI group had more children with autism and fewer with PDD NOS than comparison group. The EIBI group had more 2-parent families and better educated families than the comparison group. Statistical adjustments made but these are not a satisfactory solution to pre-existing group

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		parent training and generalisation and to have an adult present during home intervention hours. **Sample characteristics**: Gender (male:female) EIBI: 18:3 ECI: 17:4 Age (mean): EIBI: 30.2 months ECI: 33.2 months IQ: EIBI: 61.6 ECI: 59.4 **Diagnosis**: 100% with ASD or PDD NOS diagnosed by an independent clinical psychologist and confirmed by ADI-R ASD:PDD NOS: EIBI: 20:1 ECI: 15:6	**Statistics** Repeated measures analysis of covariance (ANCOVA) for each measure with pre-treatment score as the covariate and year 1, year 2 and year 3 scores as the repeated dependent measures.		grps, p=0.13. VABS – Composite EIBI: 70/79 (incr of 9) ECI: 71/67 (decr of 4) Classroom placement: EIBI: 17/21 were included in regular classroom settings at year 3 (6/17 fully included without assistance, 4/17 were fading the shadow tutor and 7/21 required full shadows) ECI: 1/21 was included in regular classroom setting. Analyses were undertaken to control for inequalities at pre-treatment and these changed the results. **Authors' conclusions:** "The UCLA/Lovaas model of early intenstive behavioural treatment can be implemented in a non-university community-based setting".	differences ▪ where data were missing, participants were removed from the analysis, with visual inspection to ensure that the missing data were random or unbiased ▪ the comparison group received diverse interventions and it was not possible to measure treatment fidelity ▪ diagnosis based on clinician report and ADI-R. Authors considered that the inclusion of the Autism Diagnostic Observation Schedule (ADOS) would have increased confidence in the initial diagnosis and inclusion of ADOS and Theory of Mind Test (TOM) at follow up assessments would have indicated whether chn continued to display ASD behaviours ▪ study designed to assess whether the UCLA Lovaas treatment was effective but the Rx protocol was expanded to include social skills training in year 3 using a discrete trial format ▪ group X time interactions were not significant. **Quality score:** fair (-)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Eldevik et al. 2006) Norway/USA Non randomised experimental study	**Evidence level**: III-2 **Interventions:** Minimal intensity one-to-one behavioural treatment) (MIBI) **Comparator:** Minimal intensity one-to-one eclectic treatment (MIECI) **Study setting**: All participants attended regular kindergartens or elementary school classes for typical children in their local community (for at least 20 hours each week). A separate Rx room was used for administering most of the treatment – some of the teaching sessions such as social skills training had typical children also working together with the study participant.	**Participants:** 28 children in 2 groups (13:15). Children recruited from one of 3 regional habilitation services in Norway between 1993 and 2001. **Inclusion:** ▪ diagnosis of autism and mental retardation (MR) according to ICD-10 criteria from a licensed psychologist and/or medical doctor ▪ chronological age <6 years at the start of treatment ▪ no medical conditions that could interfere with Rx ▪ treatment record indicating that the child had received 10 to 20 hours per week of one-to-one treatment ▪ assessments of intellectual functioning, adaptive behaviour and language conducted pre-treatment and after 2 years of Rx. **Sample characteristics**: Gender (male:female) MIBI: 10:3) MIECI: 14:1) Age (mean/range): MIBI: 53.0 months (36 to 68)	Allocation to groups made according to the treatment they had already received prior to enrolment – final determination managed by local professionals and parents. Supervisors, therapists and primary caregivers involved in MIBI were given training based on Lovaas treatment manuals and based on operant conditioning techniques (parents were expected to ensure generalisation and maintenance of skills to home and community settings). MIECI treatment included at least 2 of: alternative communication, ABA, total communication sensory motor therapies, programs based on TEACCH and other methods. Therapists and supervisors were also trained. Therapists for both groups served as 1:1 aides in the mainstream unit or classroom after individual treatment. **Intensity of intervention:**. Protocol specified less than 20 hours of one-to-one treatment per week and placement with typical peers for both treatment	Detailed information collected for each child's treatment program (weekly hours of one-to-one treatment etc) Assessments at intake and after 2 years: ▪ Intellectual functioning (BSID-II, SBIS, WPPSI-R, WISC-R) ▪ Language functioning (RDLS, PEP-R) ▪ Adaptive behaviour (VABS) ▪ Non verbal intelligence (MPSMT) ▪ Pathology data (from archival data, parental reports or a direct observation of the child). The measure for estimating pathology was adopted from Lovaas 1987 (score of 1=symptom present, 0=symptom absent). The behaviours assessed were: no words, not affectionate, no toy play, no peer play, stereotypical behaviours, severe tantrums and not toilet trained. Maximum score was 7 and minimum was 0	**Pre-treatment:** The MIECI group had higher scores at baseline on almost all variables but a significant difference was only found between groups on VABS socialization. **Post-treatment:** The MIBI group received 12.5 hours/wk of one-to-one treatment for 20 months and the MIECI group received 12 hours of one-to-one Rx for 21 months. There were no differences in the number of therapists per group or intensity of the Rx. The groups differed significantly in the type of Rx offered and treatment goals but there were no differences in the training experience of the therapists. Mean gains/losses (significant difference btwn groups): IQ ($p<0.001$): MIBI: +8.2 points MIECI: -2.9 points Language comprehension ($p<0.05$): MIBI: +6.8 points MIECI: -7.7 points Expressive language ($p<0.05$): MIBI: +11.0 points MIECI: -6.4 points Communication domain of	▪ retrospective analysis and no random assignment to groups. ▪ major possibilities of confounding: (1) because parents were involved in selecting the Rx for their children, this may have resulted in more motivated families seeking MIBI treatment because parental involvement is more central in this approach (2) parents may have had higher education and higher SES and this affected the choice of group for their children (3) the study did not reliably measure the portion of time spent on each treatment type in the MIECI group (4) quality of implementation of treatment in the MIECI group was not monitored (5) possible contamination as the MIECI group also used ABA therapy.

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		MIECI: 49 months (21 to 69) **Diagnosis**: 100% with autism and MR according to ICD-10 criteria from a licensed psychologist or medical doctor. ADI-R also administered to confirm the diagnosis for 24/28 of the children (only 13/24 by an independent professional, the remainder by one of the authors)	groups. **Timing of assessment**: Outcomes measured at intake and after 2 years. **Statistics** T tests and chi square tests.	(reliability checked for 30% of scoring). ▪ Degree of MR according to ICD-10 criteria. Data were collected by professionals independent of the study for 55% of tests (the authors undertook data collection for 37% and the remaining 8% were collected by both within 6 months of each other and reliability checked).	VABS ($p<0.01$): MIBI: +4.4 points MIECI: -4.5 points There were no differences between groups for non verbal intelligence, VABS composite, VABS daily living and VABS socialization. Significant differences in favour of the MIBI group on 4 of the 7 pathology symptoms (affection, toy play, peer play and toilet training). There were also significant differences in the proportion of the children changing their MR classification btwn groups. **Authors' conclusions:** "After 2 years of treatment, the [MIBI] group made larger gains than the [MIECI] group in most areas. However gains were more modest than those reported in previous studies with children receiving more intensive behavioural treatment and it is questionable whether they were clinically significant".	▪ detailed questionnaires of individual treatment were completed retrospectively (btwn 6 months and 3 years after the rest of the assessments) ▪ the 2 groups did not differ at baseline on most of the assessment variables but important demographic data not presented ▪ small sample size for the study indicates that there is likely to be a large margin of error and thus the clinical significance of the results is doubtful ▪ scores on IQ, language and adaptive behaviour declined in the MIECI group in spite of the fact that their individualised treatment was of similar intensity to that in the MIBI group ▪ assessments were partly conducted by the authors of the study and not by independent evaluators ▪ paragraph in discussion on future research. **Quality Score:** fair (-)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Wetherby and Woods 2006) USA Non randomised experimental study	**Evidence level**: III-2 **Interventions:** Early Social Interaction project (ESI) (incorporates the NRC recommendations within a family centered natural environments approach). Implemented by parents. **Comparator:** No treatment group (NTG). The scores for this group were directly compared with those of the ESI group after their intervention. **Study setting**: Home based sessions for the intervention group plus parent-child play groups. .	**Participants:** 17 children in ESI group recruited from an ongoing longitudinal screening study of children under the age of 2 years. These chn were matched with 18 children in the control group, recruited when they were first referred to community agencies on suspicion of having ASD **Inclusion:** ▪ diagnosis of autistic disorder or PDD NOS confirmed by ADOS falling at or above the cutoff for ASD (at 36 months of age) **Sample characteristics**: Gender (%male) ESI: 88% NTG: 78% Age (mean (SD)) months: ESI: 18.2 (3.9) months NTG: 31.6 (3.5) months **Diagnosis**: Children in the ESI group were recruited at a younger age and had significant red flags for ASD and had a provisional diagnosis of	No control intervention. Changes in pre and post intervention measures of social communication were analysed for the ESI group and the post test social communication measures of the ESI group were compared with a control group at the same age. ESI: Parents trained by 4 speech language pathologists and an early childhood education specialist to use a variety of intervention strategies within daily routines to increase opportunities to practice social communication goals. The strategies included: environmental arrangement, waiting, use of natural reinforcers, balanced turn taking, modelling, contingent imitation, requesting imitation and time delay – these strategies were individualised.	4 trained examiners scored the children on social communication measures: ▪ Communication and Symbolic Behaviour Scales Developmental Profile (CSBS DP)	**Pre-treatment:** Comparison of pre-treatment variables between groups not relevant as comparison group was on average about 1 year older than ESI group. **Post-treatment:** The within ESI group differences are not relevant to this review. Differences between ESI grp post ESI and control group: ESI>control: ▪ Social signals (gaze shifts, shared positive affect, gaze/point follow) P=0.001, p=0.005, p=0.014. ▪ Rate of communicating P=0.007 ▪ Communicative functions (behaviour regulation, social interaction, joint attention) P=0.005, 0.015, 0.021 ▪ Understanding P=0.030 There were no differences in the other measures of social communication: ▪ Communicative means (inventory of gestures, inventory of consonants, inventory of words) ▪ Symbolic capacity (inventory of actions, actions to others) Differences between ESI grp pre ESI and control grp: Control>ESI:	▪ major flaw of the study: not possible to determine whether the 2 groups were equivalent at equivalent ages as control group recruited at 1 year older than intervention grp ▪ parents were not required to document the time spent implementing the ESI program ▪ treatment fidelity monitored by self assessment checklist ▪ sample size relatively small for a group design ▪ assessors claimed to be blind to the children's "diagnosis" – query whether this is a misprint as all children were understood to have ASD ▪ inter rater reliability calculated for assessment of outcomes (on at least 20% of behaviour samples) ▪ authors commented that it was not ethical to withhold Rx from children with ASD so future RCTs would need to compare

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
		ASD. Children in control and children in ESI group (post ESI intervention) had a diagnostic evaluation at 36 months. They were required to have a diagnosis of ASD or PDD NOS (according to DSM-IV and confirmed by ADOS). Note: this was done post intervention for the ESI group.	**Intensity** Not clear. Treatment fidelity in the ESI group was monitored. **Timing of assessment**: ESI group had assessment at entry to study (18.2 months) and again (post intervention) at m=30.7 months. Control group had assessment only at entry to study (31.6 months). **Statistics** ANOVA with Welch correction for lack of homogeneous variances to evaluate group differences.		▪ Communicative means (inventory of gestures, consonants and words), p=0.043, p=0.002, p=0.002 ▪ Actions to others, p<0.001 (Note: control group 1 year older). No differences between groups on the other measures. Differences between groups on language stage: Percentage of group who were verbal: ESI group (pre ESI): 5.9% ESI group (post ESI): 76.5% Control group: 55.6% (no p value reported) **Authors' conclusions** [The study] "provides preliminary evidence suggesting that intervention beginning in the second year of life may have an effect on social communication and a secondary effect on developmental outcome". The authors concluded that the improvements that the ESI children showed in communicative means and symbolic play may have been due to maturation but that the improvement in social signals, rate of communicating, communicative functions and understanding may have been due to ESI. NOTE: Communicative means and symbolic play are often the focus of traditional interventions for chn with ASD.	parent implemented interventions with clinician implemented interventions. **Quality score**: fair (-)

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Howard et al. 2005) USA Non randomised experimental study	**Evidence level**: III-2 **Interventions:** Intensive behaviour analytic treatment (EIBI) **Comparator:** (1) Intensive 'eclectic' intervention in public special ed classrooms (combination of methods) -autism educational programming (ECI) (2) Non intensive public early intervention programs (combination of methods) - generic educational programming (GEP) **Study setting**: EIBI: multiple settings (including home, school and the community). ECI: children enrolled in public school classrooms designed for chn with autism GEP: children enrolled in local community special ed classrooms identified as early intervention or communicatively handicapped preschool programs serving chn	**Participants:** 61 children diagnosed with ASD or PDD NOS. **Inclusion criteria:** ▪ diagnosis of ASD or PDD NOS according to DSM-IV criteria by independent examiners before the child was 48 months of age ▪ entry into an intervention program before 48 months of age ▪ english as the primary language spoken in the child's home ▪ no significant medical condition other than ASD or PDD NOS ▪ no prior treatment of more than 100 hours. 78 children met these criteria but 17 were excluded (reasons given). **Sample characteristics**: Gender (male:female): EIBI: 25:4 ECI: 13:3 GEP: 16:0 Diagnosis (autism: PDD NOS): EIBI: 24:5 ECI: 12:4 GEP: 9:7 Age at intake (months): EIBI: 30.9 ECI: 37.4	Assignment to groups was determined by health professionals and parents. EIBI: Children had 50-100 learning opportunities per hour presented via discrete trial, incidental teaching and other ABA procedures. Therapists were trained and supervised. Parents were also trained and met with supervisors regularly ECI: A special ed teacher supervised the work of 4-8 paraprofessional aides using a variety of methods (discrete trial training, PECS, sensory integration therapy, TEACCH and other activities such as circle time and music). Teachers also consulted with ABA trainees and speech and language training was available. GEP: Each classroom had special ed teachers or certified speech and language pathologists who supervised 1-2 paraprofessional aides. Activities were described as 'developmentally appropriate' with emphasis on exposure to language, play activities and a variety of sensory	Conducted by experienced psychologists and speech and language contractors who were independent contractors. Took place in child's home, clinician's office or regional center. The following tests administered at intake and follow up, dependent on the chronological age of the child: ▪ Cognitive skills (BSID-II, WPPSI-R, DP-II, SBIS, DAS, DAYC, PEP-R) ▪ Non verbal skills (MPSMT, SBPT, Leiter-R) ▪ Receptive and expressive language (RDLS, RITLS, REELS-2, PLS-3, ITDA, PPVT-III, EVT, DP-II, SICD-R, EOWPVT, ROWPVT) ▪ Adaptive skills (VABS, DDST II, DP-II, RIDES)	**Pre-treatment:** Clear evidence of developmental delay for all children. Scores on tests equivalent, except GP group had higher mean age equivalent score than AP group. **Post-treatment:** There were no statistically significant differences between the mean scores of children in the ECI and GEP groups. The EIBI group had significantly higher mean scores on all domains than the other 2 groups, except for motor skills (which were close to normal for all groups at intake). Learning rates at follow up were also significantly higher for the EIBI group than either the ECI or GEP groups – they were normal or above normal for the EIBI group in all skill domains. By contrast, only non verbal skills were close to normal rates for children in the ECI and GEP groups. Results from analysis of change scores mirrored those of final scores.	▪ study not randomised – assignment to groups by parental preference ▪ groups not equivalent at baseline (children in EIBI group were younger than those in other groups and parental level of education in EIBI group higher) but authors attempted to control for this in the analysis ▪ the examiners who conducted assessments were not blind to the group's assignment at follow up testing ▪ results were analysed only in terms of performances on standardised norm-referenced assessments conducted in formal testing situations, rather than the repeated direct observational measurement of behaviour in situ that characterises ABA ▪ the analyses compared group mean scores statistically (thus not capturing the performance of

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
	with a variety of disabilities.	GEP: 34.6 Parent years of education (mean): EIBI: 14.4 ECI: 13.1 GEP: 13.0 **Diagnosis**: Chn were diagnosed with ASD or PDD NOS according to DSM-IV criteria by independent examiners before the child was 48 months of age.	experiences. Speech and language therapy were also available. **Intensity of intervention:**. EIBI: 1:1 adult:child ratio, 25-40 hours/week. ECI: 1:1 ratio or 1:2 ratio, 30 hours/week GEP: small groups (1:6 ratio), 15 hours/week **Timing of assessment**: Intake within 2 months of study entry and follow up an average of 14 months after treatment entry. **Statistics** Multiple regression to compare the 3 groups of children, controlling for individual differences in age at diagnosis and parental education.		**Authors' conclusions** "Young children with autism of PDD NOS who received IBT [EIBI] for about 14 months outperformed comparable children who received 'eclectic' intervention services on virtually every follow up measure".	individuals over time) ▪ treatment integrity was not measured. **Quality score:** fair **(-)**

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
(Sheinkopf and Siegel 1998) USA Non randomised experimental study	**Evidence level**: III-2 **Interventions:** Intensive home based behavioural therapy (EIBI-H) **Comparator:** Matched control group receiving standard school based interventions and brief one to one interventions (ECI) **Study setting**: Community/home and school based services in the San Francisco bay area.	**Participants:** N=22 (11 in each group). **Sample characteristics**: Sex: not reported Mean age (range): 33 (29 to 47) months Mean IQ: 63 **Diagnosis**: 10 pairs with diagnosed with autism and 1 pair with PDD-NOS. Diagnosis based on both developmental information derived from parent interviews and standardised behavioural observations; made by consensus of a child psychologist and child psychiatrist with experience in the field of autism. **Statistics** T tests	The experimental group had treatment that was said to be based on Lovaas therapy by the parents of the included children (but it was not observed directly). Parents received a detailed manual of treatment methods and they acted as therapists together with a variety of paraprofessionals with supervision from behaviourally trained clinicians in the community. Punishment was limited to mild aversives. Matching of children in the control group was accomplished by reverse serial selection of any case matching an index case on (1) chronological age, (2) mental age; (3) the interval between pre and post treatment assessments; and (4) diagnosis. Information on the treatments received by both groups was gathered via telephone interviews with parents – there was no direct observation to account for Rx fidelity.	**Pre and post treatment:** Cognitive data: ▪ Merrill Palmer Scale of Mental Tests (MPSMT) ▪ Bayley Scales of Infant Development (BSID) ▪ Wechsler Preschool and Primary Scale of Intelligence (WPPSI) ▪ Cattell Infant Intelligence Scale (CIIS) Behavioural data: ▪ Number of positive DSM-III-R symptoms as a percentage of the number of age appropriate scorable symptoms ▪ Measure of symptom severity	**Pre-treatment:** No differences between groups on pre treatment assessments. **Post-treatment:** At post treatment follow up, 2 pairs were not included in the assessment of group differences in IQ. Mean number (SD) of hours treatment: EIBI-H: 27.0 (7.9) hrs ECI: 11.1 (5.3) hrs (but no correlation btwn final IQ and no of weekly hours of Rx) Mean IQ: ▪ EIBI-H: 89.7 ▪ ECI: 64.3 (paired t test, difference (25 IQ points) btwn groups, p=0.01) Proportion with IQ >65: ▪ EIBI-H: 100% ▪ ECI: 54.5% (no test of difference reported) Percentage of DSM-III-R symptoms rated positive: ▪ EIBI-H: 46.6% ▪ ECI: 54.0% (no stat diff btwn grps)	Limitations: ▪ parents chose to instigate Lovaas therapy and researchers 'selected' the children retrospectively for inclusion from a larger longitudinal study ▪ although experimental Rx was effective, it is not possible to determine whether it is more effective than anything other than standard low intensity school based interventions (this was the only comparator and was poorly defined) ▪ the relative contributions of treatment intensity and specific therapeutic techniques were unclear in the study. There was no relationship between Rx intensity and IQ change but other research should be designed to "disentangle the active ingredients of therapy" ▪ therapy sessions were not observed directly – results based on parental reports – treatment fidelity not assessed ▪ the IQ tests used in the study varied between individuals with different tests weighted towards different types of skills (however, the proportions of children receiving various tests did not differ across groups) ▪ sample not randomised but groups equally matched on age, diagnosis, initial IQ and other variables. However, parents chose the treatment and were not blind to treatment type ▪ possibility of other factors related to parental choice of Rx that could

Reference, country, design	Evidence level, comparators, study setting	Sample characteristics	Methods	Outcome measures	Results and authors' conclusions	Comments and quality score
			Intensity: Both groups received school based services but the control group spent more time per week in school than children in the EIBI-H group. All children were placed in special education classrooms (a subset of children in the EIBI-H group was subsequently placed in regular classrooms). One on one services were provided to children in both groups. Because of the nature of the intervention, EIBI-H children spent twice the number of weekly hours of intervention than children in the control group. **Timing of assessment**: Assessments made pre-treatment and post treatment (an average of 20.4 months for the EIBI-H group and an average of 18.1 months for the ECI group).		Mean (SD) symptom severity: ▪ EIBI-H: 18.6 (7.0) ▪ ECI: 24.6 (8.0) (p=0.014) Diagnosis of autism at the end of the study ▪ EIBI-H: 10/11 ▪ ECI: 11/11 **Authors' conclusions:** "Intensive home-based behavioural treatment can be implemented successfully in the field, without the support of an academic center".	have confounded the results, eg, availability of social support, marital status, financial resources ▪ no measurement of other outcomes such as language development, adaptive behaviour, academic or social functioning ▪ no indication whether outcome assessment blinded ▪ no documentation of Rx fidelity. Other Comments ▪ implications of study are that intensive home-based behavioural Rx can be implemented successfully in the field without the support of an academic centre ▪ study failed to find more than modest effects on symptomatic behaviour despite the findings of large effects on cognitive functioning – effects on IQ were the largest in magnitude. Implications for future research: The study found that home based behavioural therapy is a ***good*** option for children with autism but did not indicate whether this Rx is ***better*** than other Rxs of similar intensity and/or structure – future research needs to consider a comparison of various Rx approaches **Quality score**: fair (-)