Early Behavioral Interventions for Children with Autism Spectrum Disorder in Routine Clinical Care: A Systematic Review and Meta-analysis

Gro Janne H. Wergeland¹,², Maj-Britt Posserud¹,², Krister Fjermestad³, Urdur Njardvik⁴, & Lars-Göran Öst⁵,⁶

1) Department of Child and Adolescent Psychiatry, Division of Psychiatry, Haukeland University Hospital, Bergen, Norway
2) Department of Clinical Medicine, Faculty of Medicine, University of Bergen, Norway
3) Department of Psychology, University of Oslo, Norway
4) Department of Psychology, University of Iceland, Reykjavik, Iceland
5) Department of Psychology, Stockholm University, Stockholm, Sweden.
6) Bergen Center for Brain Plasticity, Haukeland University Hospital, Bergen, Norway

Author note

Conflict of interest: Maj-Britt Posserud holds a position at the scientific advisory board for Takeda regarding a slow-release formulation of melatonin for children with ASD. No other potential conflict of interest was reported for any of the other authors.

Corresponding concerning this article should be addressed to: Gro Janne Wergeland, Department of Child and Adolescent Psychiatry, Haukeland University Hospital, N-5021 Bergen, Norway. Email: gjwergeland@gmail.com

Phone: +47 48176828
Funding

No external funding was obtained for this meta-analysis.

Contributors

LGÖ designed the meta-analysis and together with GJW wrote the protocol. GJW conducted literature searches in collaboration with an academic librarian. KWF, MBP, and UN contributed in the screening process, the extraction of data, and writing of the paper. LGÖ wrote the coding scheme, meta-analyzed the included studies, and wrote the first draft of methods and results. GJW and LGÖ rated the methodology, and Risk of Bias of the included studies. GJW wrote the first draft of the introduction and discussion. All authors have approved the final manuscript.

Acknowledgements

The authors wish to thank academic librarian Regina Küfner Lein for assistance with the literature search.

Conflict of interest

Maj-Britt Posserud holds a position at the scientific advisory board for Takeda regarding a slow-release formulation of melatonin for children with ASD. No other potential conflict of interest was reported for any of the other authors.
Abstract

The current paper systematically reviews and meta-analyses the effectiveness of early behavioral interventions (BI) for children with autism spectrum disorder (ASD) in routine clinical care. The effectiveness of BI, methodological study quality, and moderators of treatment outcome were examined and benchmarked with efficacy studies. The quality of the evidence was assessed with the Cochrane risk of bias tool. Twenty-nine studies were included, comprising 1422 participants. Medium to large within-group effect sizes ($g = 0.76-1.27$) were found post-treatment for the outcome domains adaptive behavior, cognition, communication, and socialization, with large average effect size at post ($g = 0.94$) and at follow-up ($g = 1.08$). Comparison of effectiveness and efficacy studies showed that evidence-based early BI in routine clinical care yielded effects comparable to university research settings. The limitations include potential language and publication bias. The findings support evidence-based behavioral treatments delivered in routine clinical care as efficacious in reducing ASD symptoms. PROSPERO registration: ID CRD42020212833.

**Keywords:** Autism spectrum disorders, routine care; effectiveness, early behavior interventions, children, meta-analysis

*Public Health Statements:* Early behavioral interventions for autism spectrum disorders in children treated in routine clinical care was found efficacious in reducing symptoms within the outcome domains adaptive behavior, cognition, communication, and socialization, with medium to large within-group effect sizes at post-treatment and follow-up. The outcome of effectiveness studies was similar to that of efficacy studies. Our findings suggest that clinicians and patients can be confident about the effectiveness of early behavioral interventions with already established efficacy when delivered in routine clinical care. As treatment effects are not lost when evidence-based treatment programs are
transported from research clinics to routine clinical care, further implementation of evidence-based interventions is needed in routine clinical care for children with autism disorders.
Early Behavioral Interventions for Children with Autism Spectrum Disorder in Routine Clinical Care: A Systematic Review and Meta-analysis

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by challenges in social interaction, communication, as well as restricted and repetitive interests or behaviors (American Psychiatric Association, 2013). ASD usually emerges in childhood, and prevalence estimates indicate that ASD affects more than 0.6% of children (Elsabbagh et al., 2012). However, more recent studies have reported prevalence rates between 1.0-1.9% (Maenner et al., 2020; Posserud et al., 2021). Individuals with ASD are highly heterogeneous with diverse clinical manifestations, behavioral phenotypes, and cognitive functioning levels (Baio et al., 2018; Lord et al., 2018). A significant proportion of individuals with ASD have challenges in social function, communication, and adaptive functioning throughout life (Magiati et al., 2014). However, it is commonly acknowledged that early intervention can alleviate these challenges and mitigate core and associated ASD features (Smith & Iadarola, 2015).

Early intervention programs targeting ASD aim to enhance functioning in daily life in children with ASD. Several systematic reviews and meta-analyses of studies on early intervention programs for ASD have been published (e.g., Nevill et al., 2018; Rodgers et al., 2021; Sandbank et al., 2020). Because the field is diverse, these evaluation studies have varied in the key methodological features they considered, including treatment format, dose and intensity, and outcome measures. However, this variability challenges the interpretation and generalization of findings. A key challenge to summarizing existing evidence is that the reviewed intervention programs vary considerably in terms of theoretical frameworks, and of particular importance, in levels of scientific evidence. Some early intervention programs have been associated with
subsequent improvements in communication, language, cognition, and/or adaptive behavior, and may also affect long-term outcomes (Orinstein et al., 2014). There have been substantial advances in early detection and diagnosis of ASD in recent years, and the need for evidence-based early interventions has increased. Furthermore, the increasing empirical support for some early interventions highlights the necessity of a meta-analysis of evidence-based early interventions for children with ASD.

Systematic reviews and meta-analyses have identified applied behavior analysis (ABA), and developmental social-pragmatic (DSP) models as the best supported interventions for children with ASD (Rogers & Vismara, 2008; Smith & Iadarola, 2015). Individual, comprehensive ABA is commonly referred to as early intensive behavioral intervention (EIBI). EIBI is a structured teaching approach for children with ASD that is recommended to start before the age of five years (Lovaas, 1987; Smith, 2010). Key elements include highly specialized, individualized services for two or more years, the use of specific teaching procedures and behavioral strategies, supervision by ABA trained personnel, a 1:1 child to trainer ratio, and implementation in settings like home or school for 25+ hours/week (Lovaas, 1987; Smith, 2011). Complex tasks and skills are broken down into small steps and taught systematically. The structure of the intervention strategies is gradually reduced with the goal to improve the child’s functioning in everyday life situations. In developmental social-pragmatic models (DSP), the main theoretical principle is that the core ASD feature is difficulties in reciprocal social interaction, leading to a multitude of downstream challenges with communication, functioning, and development (Mundy & Crowson, 1997). Studies using EIBI and DSP interventions have reported efficacy and clinical utility, with reductions in symptom severity, and improvement in adaptive behavior, cognitive functioning, and communication (Smith & Iadarola, 2015).
In the most recent Evidence Based Update for Autism Spectrum Disorder, Smith and Iadarola (2015) evaluated psychological and behavioral interventions for children with ASD, under the age of five years. Two interventions were identified as well-established and three as probably efficacious. Interventions with well-established efficacy were individual, comprehensive ABA, and teacher-implemented, focused ABA+DSP. The interventions evaluated as probably efficacious were individual, focused ABA for alternative communication systems, individual, focused ABA in combination with DSP, and focused DSP parent training. It is encouraging that an increasing number of methodologically sound early intervention studies are being carried out for children with ASD, and that several interventions have empirical support. However, more attention should be directed at investigations of clinical effectiveness, i.e., how the interventions with established efficacy such as early behavioral interventions perform when delivered in routine clinical care (Lake et al., 2020; Nahmias et al., 2019; Wood et al., 2015).

The majority of the studies evaluating early interventions for children with ASD have been efficacy studies conducted in controlled research settings (Nahmias et al., 2019). Important differences between research clinics and routine clinical care regarding patients, therapists, and treatment contexts may impact the generalizability of results from efficacy studies to routine clinical care (Weisz, Ng, et al., 2013; Weisz, Ugueto, et al., 2013). For example, in efficacy research the intervention is tested with a methodologically stringent procedure to ensure high internal validity. Furthermore, the clinicians are often well trained and supervised and deliver the intervention with high fidelity (Bauer et al., 2015; Weisz, Ng, et al., 2013). Effectiveness trials, on the other hand, evaluate whether an intervention produces good effects under “real world” conditions such as routine clinical care. In these studies, the samples are often more heterogeneous, the training and supervision of clinicians vary
more than in efficacy trials, and the delivery of the intervention provided is not typically implemented and monitored with the same level of fidelity as in university research settings (Hunsley, 2007; Hunsley & Lee, 2007). Such differences between university settings and routine clinical care suggest that findings obtained in university research settings may not be directly transferable to clinical practice. Routine clinical care is a crucial service site where the majority of children with ASD will receive their mental health services (Brookman-Frazee et al., 2010). It is important for clinicians to know what outcomes to expect from empirically supported treatments of ASD when they are delivered in routine clinical practice, and how results compare with outcomes obtained in specialized university research settings.

Previous meta-analyses comparing studies from routine clinical care to efficacy studies have shown different results for different disorders. Two recent meta-analyses of effectiveness studies for children and adolescents with internalizing disorders (Wergeland et al., 2021) and externalizing disorders (Riise et al., 2021), respectively, reported treatment outcomes in routine clinical care that were comparable to those in university settings. Contrary to this, for ASD, a recent meta-analysis of studies of community-based early intervention programs found smaller effect sizes for community-based studies than for university-based clinical trials (Nahmias et al., 2019), indicating a gap between research settings and routine clinical care for ASD interventions. However, an updated meta-analysis is warranted due to some limitations in the Nahmias et al., (2019) review. Nahmias et al. (2019) included a wide range of interventions, not limiting the meta-analysis to interventions with established evidence. Furthermore, the comparison between university-based trials and routine clinical care in Nahmias et al. (2019) is questionable as four of the five university-based trials (Cohen et al., 2006; Howard et al., 2005; Magiati et al., 2007; Remington et al., 2009) in the systematic review by Reichow et al. (2012) that were used as comparison were included as “community-based early intervention studies” in
the Nahmias meta-analysis. This means that the comparison is flawed since these studies are counted in both categories of studies. Finally, Nahmias et al. used within-group effect size whereas Reichow et al. used between-group effect size, making the comparison of results challenging. These factors make an update about the current state of effectiveness of evidence-based early behavioral interventions for children with ASD in routine clinical care warranted.

The present study aims to provide a meta-analysis of the effectiveness of early behavioral interventions considered well-established (level 1) or probably efficacious (level 2) according to Smith and Iadarola (2015) for children with ASD treated in routine clinical care. We have included studies investigating the effectiveness of empirically supported treatment programs, delivered by practicing clinicians in routine clinical care, to patients referred for treatment through usual clinical routes. In the present meta-analysis, efficacy studies were directly meta-analyzed in comparison with effectiveness studies, using the same effect size measure. Our aims were threefold. First, to examine the effectiveness of evidence-based early behavioral interventions for ASD in young children (i.e., samples with a mean age below 5 years). Second, to evaluate the methodological quality of the effectiveness studies, and investigate potential moderators of treatment outcome. Third, to examine how the treatments delivered in routine clinical care fare in comparison to efficacy studies, in order to evaluate if early behavioral interventions in effectiveness studies lead to equivalent effect sizes as in efficacy studies.

Methods

The protocol for this meta-analysis was pre-registered at PROSPERO with ID CRD42020212833. The meta-analysis was conducted according to the PRISMA guidelines (Liberati et al., 2009), and reported according to AMSTAR 2 (Shea et al., 2017), see Supplement S8 and
Two independent raters were involved in the process of selecting studies, data extraction and categorization. The meta-analysis was designed according to the PICOS acronym in the following way:

- **Population**: children with ASD.
- **Intervention**: Behavioral intervention (BI) for ASD delivered in routine clinical care. For the search, the umbrella construct CBT was used, to include cognitive, behavioral, and combined treatments (CBT and BI).
- **Comparison**: within-group change, i.e., pre vs. post-data (and pre vs. follow-up data).
- **Outcome**: primary continuous measure.
- **Study design**: randomized controlled trials (RCTs) and open trials.

**Literature Search**

Studies were identified by a systematic and comprehensive literature search of electronic databases and scanning of the included articles’ reference lists. The search was applied to Ovid MEDLINE, Embase OVID, PsycINFO, ERIC, and Web of Science from the start of the databases to September 1st, 2020. An updated search was done April 7th, 2021. The list of search terms utilized to identify potential studies were generated by all five authors in collaboration with a university librarian, who conducted the database searches. For search strategy for Ovid MEDLINE, Embase OVID, PsychINFO, ERIC, and Web of Science, see the Supplement S1.

Three author pairs read the titles and abstract of all the papers from this initial search. When there was an indication of a group of patients receiving early BI in a non-university setting, the full-text was retrieved. The reference lists in the retrieved articles were then checked. In total,
267 full-text articles were considered for inclusion. The final decision for article inclusion was made using a stricter set of inclusion and exclusion criteria. The full text articles were read by author pairs and disagreements were resolved by consensus discussion. It was determined that 29 studies (presented in 33 articles) could be included.

**Inclusion Criteria**

In order to be included in the review and meta-analysis a study had to:

1. Be published, or in press, in an English language journal.
2. Have participants diagnosed with some form of ASD according to DSM or ICD.
3. Be testing a form of early BI for ASD classified as well-established or probably efficacious in the evidence base update review of Smith and Iadarola (2015).
4. Have participants referred for treatment through usual clinical routes (i.e., not recruitment via advertisements).
5. Be an effectiveness study, i.e., carried out in a non-university setting such as routine clinical care, preschools, or homes.
6. Have therapists/supervisors who are practicing clinicians with regular caseload for whom provision of service is a substantial part of the job (Shadish et al., 2000).
7. Have a treated sample consisting of at least 10 participants.
8. Have a maximum mean participant age of 60 months (i.e., 5 years).
9. Provide a continuous measure of the principal disorder treated.
**Exclusion Criteria**

1. The study is a secondary analysis of a previously published study. Separate follow-up studies to the basic study are included to provide follow-up data.

2. The study is an evaluation of a service where the results for individual disorders cannot be extracted.

3. The study is testing a combination of early behavioral intervention and pharmacological treatment and 100% of the participants in that condition receive both treatments.

   Figure 1 shows a flowchart of the inclusion of studies in the present meta-analysis. For references to included studies, these are marked with an asterisk in the reference list. For references to studies excluded in the meta-analysis, see Supplement S2.

**Potential Categorical Moderators**

   To include any potential categorical or continuous moderator in the analysis we required that at least 70% of the studies provided information on that variable. With lower proportions it is questionable if the information extracted is representative of the entire body of studies.

**Type of Study and Statistical Analysis**

   Type of study was either RCT (when an early BI-condition was compared with some kind of control/comparison condition) or open trial (when only an early BI-condition was used in the study). Statistical analysis was categorized as intent-to-treat if all randomized or starting participants were included in the statistical analysis or completers if dropouts were deleted.

**Target of Treatment and Setting**
Treatments were classified according to the description in the study (i.e., ABA, EIBI, Joint Attention Mediated Learning). However, it was not possible to classify the various components included in the interventions as the information in the studies most often did not provide this level of detail. Instead, we categorized the primary recipient of treatment as: child, or child and parent (if parents were trained in applying various treatment techniques), and treatment setting as: clinic, community center, home, preschool, or home + preschool. We also categorized the different target areas of the intervention (Supplement S6).

**Continent**

The country in which the study was carried out was categorized as North America, South America, Europe, Asia, Australia, or Africa.

**Potential Continuous Moderators**

The following continuous measures on which at least 70% of the studies provided information were used as potential moderators: year of publication, number of participants in the study, percent boys, mean age (in months), pre-treatment severity (calculated as percentage of the maximum score of the rating scale applied), methodology score (see below), risk-of-bias score (see below), months of treatment, hours of treatment/week, and percentage attrition in the treatment condition.

We also extracted information on a few other categorical and continuous variables, but these did not reach the 70% criterion and were excluded. A coding scheme and manual including the variables of interest were developed. The data extraction and categorizations were done independently by author pairs and any disagreements were solved after discussion.

**Methodological Quality**
The Psychotherapy Outcome Study Methodology Rating Scale (POMRS)

Methodological quality was assessed using the POMRS (Öst, 2008). The scale consists of 22 items covering various important aspects of the methodology in psychotherapy outcome research. Each item is rated as 0 = poor, 1 = fair, and 2 = good, and each step has a short qualitative description. Possible range is 0-44 points. Since not all items were applicable to all studies the total score was recalculated as a percentage of the maximum score possible for the individual study. The internal consistency of the scale was good with a McDonald’s ω of .80. The inter-rater reliability of the scale (between the first and the last author), based on 20% randomly selected and blindly rated studies was ICC = 0.92 (95% CI 0.49-0.99, p = 0.009) which is excellent according to Cicchetti (1994).

Risk-of-Bias

The Cochrane Collaboration tool for assessing risk-of-bias (Sterne et al., 2019) was used, and the following domains were rated: the randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. A high risk-of-bias in a domain was given 1 point, an unclear risk 0.5, and a low risk 0 point. Across the five domains the total score could vary between 0 and 5, with higher scores indicating higher risk-of-bias. Inter-rater reliability was assessed between the first and the last author based on 20% randomly selected and masked rated studies. This yielded an intra-class correlation, ICC = 0.96 (95% CI 0.32-0.99, p < 0.001) which is excellent according to Cicchetti (1994).

Effect Size Measures
In accordance with previous reviews and meta-analyses (Nahmias et al., 2019; Smith & Iadarola, 2015) we extracted data on the following constructs commonly assessed in ASD studies:


2. *Cognitive* (e.g., Bayley Scales of Infant Development, Second Edition; Bayley, 1993; Mullen Scales of Early Learning; Mullen 1995; Wechsler Preschool and Primary Scale of Intelligence-Revised; Wechsler, 1990).

3. *Communication* (e.g., Reynell Developmental Language Scales; Reynell, 1990; Vineland Adaptive Behavior Scales-Communication domain; Sparrow, Balla, & Cicchetti, 1985; Sparrow, Cicchetti, & Balla, 2005).


A list of the various instruments used in the effectiveness studies and that provided data to calculate effect sizes (ES) is presented in Supplement S3. When a study named its primary outcome measure among rating scales we used that. If no primary outcome was pinpointed we selected measures according to the following hierarchy: independent assessor or observer rating, teacher report scale, and parent report scale.

Ten (34.5%) of the 29 studies provided outcome data for all four constructs, 11 (37.9%) had data for three, five (17.2%) had for two, and three (10.3%) had data for only one of the constructs. On average, the studies had data on three of the four constructs. The basic meta-analytic
statistic uses the mean ES across the constructs each study provides data on. However, to obtain a tentative indication of possible differences in ES between the constructs we also compared the ES across the constructs, aware of the fact that these data are not independent.

**Meta-Analysis**

To obtain as large as possible a body of effectiveness studies we included both RCTs and open trials in the meta-analysis since within-group ES can be calculated from both types of studies. Within-group ES was calculated as \((M_{\text{pre}} - M_{\text{post}})/SD_{\text{pre}}\) according to recommendation by Lakens (2013) as the interventions may influence not only the means but also the standard deviations. The mean ES was computed by weighting each ES by the inverse of its variance. When a study presented intent-to-treat data (34.0%) these were used, if not, completer data (66.0%) were used.

Before pooling the effect sizes, we screened for statistical outliers, defined as outside \(M \pm 2SD\). At the post-treatment assessment, five (5.2%) of the ESs were outliers, and at follow-up assessment there was one (7.7%). For these ESs *winsorizing* (Lipsey & Wilson, 2001) was used by reducing outliers to the exact value of \(M+2D\). The software *Comprehensive Meta-Analysis v.3* (CMA; Borenstein et al., 2013) was used for all analyses and to correct for small sample sizes Hedges’s \(g\) was calculated. A random effects model was used since it cannot be assumed that the ESs come from the same population.

Heterogeneity among ESs was assessed with the \(Q\)- and the \(I^2\)-square statistic. The possibility of publication bias was analyzed with the trim-and-fill method of Duval and Tweedie (2000) as well as Egger’s regression intercept (Egger et al., 1997). Moderator analyses of continuous variables were carried out with meta-regression and for categorical variables with subgroup analysis using the mixed effect model.
Efficacy Studies for Comparison

To obtain the efficacy studies to be used in comparison of the effect of early BI in effectiveness studies we consulted the most recent evidence base update review of psychosocial treatments for ASD published in the Journal of Clinical Child and Adolescent Psychology. We chose this journal because it provides regular updates classifying treatment methods for different youth disorders according to the APA system for evaluating the evidence base. This was the Smith and Iadarola (2015) review, but we also checked the previous review of ASD (Rogers & Vismara, 2008) to get as comprehensive a list of efficacy RCTs as possible. From these reviews we listed the RCTs of some kind of BI evaluated as well-established or probably efficacious according to the criteria adopted by the Society of Clinical Child and Adolescent Psychology (Southam-Gerow & Prinstein, 2014). We then removed the RCTs which were already included among the effectiveness studies. This resulted in 16 efficacy RCTs for our comparison and these references are listed in the Supplement S4.

As for the effectiveness studies we extracted data for the primary continuous outcome measure, separately at post-treatment and follow-up assessment. To compare the two categories of studies on background variables we also extracted data on proportion of boys, mean age (in months), pre-treatment severity (calculated as percent of maximum score on the continuous measure), treatment duration (in months), treatment time (hours/week), and attrition rate. Other variables were not reported systematically, or not at all in a sufficient proportion of studies, which precluded inclusion as a background variable.

Power Analysis
In the overall comparison of effectiveness and efficacy studies we have the following number of studies and treatment conditions, which is the unit of analysis: effectiveness studies 29/32 and efficacy studies 16/18. This gives a total of 45 studies and 50 conditions with an average of 44 participants per condition. According to the formulas for power analysis in meta-analyses by Valentine et al. (2010) we would have 91% power to detect an effect size of 0.20, when assuming a high heterogeneity of effect sizes.

**Results**

**Description of the Studies**

**Study Characteristics**

Background data for the included studies are presented in Table 1. The majority of the 29 studies were conducted in North America \(n = 13\) or Europe \(n = 10\), whereas four came from Asia and two from Australia. Five studies \(17.2\%\) provided information on ethnicity, see Supplement S7. Only four \(13.8\%\) of the studies were RCTs whereas 25 \(86.2\%\) were uncontrolled open trials. The total number of participants receiving early BI in these studies was 1422. There was an overall majority of boys in the studies; with a mean of 84.5\% and a range from 71.4\% to 96.4\%. Five studies \(17\%\) did not provide information on participant gender. The mean age when starting treatment across all studies was 38.7 \(SD 7.9\) months, varying from 25.1 (Zachor & Ben Itzchak, 2010) to 54 months (Perry et al., 2008).

**Treatment Data**

Treatment data for the included studies are presented in Table 2. In labelling the programs we used the same names as provided in the original studies (second column). We classified the treatments according to the definitions provided by Smith and Iadarola (2015) as either
“comprehensive, aiming to address all areas of need, or focused, having a more circumscribed set of goals” (p. 902). Almost all (93.1%) studies used EIBI or ABA or modifications thereof and were comprehensive interventions. The child was always the target of treatment with varying degree of parental participation. Various clinicians (e.g., preschool teachers, counselors) delivered the treatment working at the center in question. These were supervised by professionals (e.g., psychologists) who usually were Board Certified Behavioral Analysts. In 15 studies (51.7%) the parents of the child with ASD also received training in how to apply ABA treatment in the home and other settings. Treatments were carried out in community centers (n = 8), preschools (n = 7), in patients’ homes (n = 7), in home + preschool (n = 6), at outpatient clinics (n = 3). Treatments were carried out over 18.6 (SD 11.4) months on average (range 3-48) and the mean hours of treatment/week was 24.1 (SD 9.6). The mean attrition rate was 5.7% (SD 6.3) with a range of 0-20.3%. Only four studies (13.8%) presented follow-up data, either in the original or a separate article, giving a total number of 33 papers included. The mean time since the end of treatment was 58.0 (SD 56.8) months, with a range from 12 to 138 months.

**Methodological Data**

**Methodology Ratings**

The research methodology score (% of maximum possible score for the instrument used in the individual study) had an overall mean of 45.5% (SD 6.9), which corresponds to a raw score of 19.4 points (see Table 1). RCTs (M 48.6%, SD 3.8) had a nonsignificant higher methodology score (p = 0.28) than the open trials (M 44.9%, SD 7.2).

**Risk of Bias**
The risk of bias categorization is presented in the Supplement S5. The different factors had the following proportions of a high risk-of-bias: the randomization process 73%, deviations from intended interventions 100%, missing outcome data 24%, measurement of the outcome 59%, and selection of the reported results 0%. In order to score the risk-of-bias a low risk was given 0, an unclear risk (some concerns) 0.5, and a high risk 1 point, which means that the total score could vary from 0 to 5 points. The total mean score was 2.94 (SD 0.90) and the RCTs (M 1.60, SD 0.42) had a significantly lower risk-of-bias (t(30) = 4.67, p < 0.0001) than the open trials (M = 3.19, SD = 0.74).

**Meta-Analysis**

**Attrition**

Eighteen studies (62.1%) provided information on the number of participants who dropped out of treatments. Using treatment condition (k = 21) as the unit of analysis the overall attrition rate was 8.8% (95% CI 6.4–12.1, z = 13.03, p < 0.0001). The difference between RCTs (6.4%) and open trials (9.0%) was not significant (Q_{between} = 0.42, df. =1, p = 0.52).

**Primary Outcome Measure**

The mean effect sizes of the primary continuous measure for all studies at post-treatment and follow-up assessment, which was done on average 50 months after the end of therapy, are presented in Table 3. The mean ES at post-treatment was large (g = 0.94) and significantly different from zero. As indicated by the Q- and I²-values heterogeneity was significant and large. At follow-up, the mean ES (g = 1.08) was also significantly different from zero and significantly heterogeneous.
The effect sizes for the four constructs we extracted data on are presented in Table 4. Since these data usually emanate from the same informant (e.g., an independent assessor or a parent) often using the same instrument (e.g., the Vineland Adaptive Behavior Scales) they are not independent. This lack of independence needs to be considered for the results presented in Table 4, which shows that the mean ES varies between 0.76 and 1.27, are all significantly different from zero, and significantly heterogeneous. There was a significant difference in ES between the constructs ($Q_{\text{between}} = 9.78$, df. = 3, $p = 0.021$), which was followed by pairwise Q-tests. These showed that the ES for communication (1.27) was significantly higher than the 0.76 for cognitive ($Q_{\text{between}} = 8.82$, df. = 1, $p = 0.003$), and the 0.81 for adaptive measures ($Q_{\text{between}} = 4.64$, df. = 1, $p = 0.021$). None of the other differences were statistically significant.

**Publication Bias**

The possibility of publication bias in the post-treatment data was investigated using Duval and Tweedie’s trim-and-fill method and Eggers regression intercept. The trim and fill method suggested that 13 studies should be trimmed, which would reduce the ES to 0.58 (95% CI 0.37-0.79). Egger’s regression intercept yielded a $t$-value of 2.13 ($p = 0.04$). Thus, publication bias may be an issue for these studies.

**Moderator Analyses**

Since the mean post-treatment ES was significantly heterogeneous we followed up with moderator analyses. The results for categorical variables using subgroup analysis are presented in Table 5. There was no significant difference between RCTs and open trials, indicating that participants in RCTs improved as much as (or non-significantly more than) patients in open trials. Regarding statistical analyses, studies with intent-to-treat analysis yielded non-significantly lower ES than studies using completer analysis. There was no significant difference in ES
depending on target of treatment. However, the setting yielded a significant $Q_{\text{between}} = 12.55$, df$ = 4$, $(p = 0.014)$, which was followed by pairwise Q-tests. The ES for home treatments (1.34) was significantly higher than the 0.69 for community centers $(Q_{\text{between}} = 9.41$, d.f. $= 1$, $p = 0.002$), the 0.72 for preschools $(Q_{\text{between}} = 7.12$, df$ = 1$, $p = 0.008$), and the 0.84 for home + preschool $(Q_{\text{between}} = 4.42$, df$ = 1$, $p = 0.026$). None of the other differences were significant. The continent at which the study was carried out was not associated with a significant difference. However, since so few studies originated from Asia (n = 4) or Australia (n = 2) we only compared the ESs for North America (n = 15) and Europe (n = 12), which yielded a non-significant difference $Q_{\text{between}} = 1.81$, df$ = 1$, $(p = 0.18)$.

Continuous variables on which at least 70% of the studies provided information were analyzed with the meta-regression module in the CMA program using the random effects analysis (see Table 6). Since 10 variables were included we used the Holm-Bonferroni correction (see Jaccard & Guilamo-Ramos, 2002). There was one positive moderator; hours of treatment/week, i.e., more intensive treatment was associated with higher ES. There was one negative moderator; year of publication, i.e. later publication year was associated with lower the ES. None of the other moderators was significant.

Efficacy-Effectiveness Comparison

In the following section (see Tables 7-8) data for the effectiveness studies reviewed so far were compared with data for the efficacy studies obtained from the evidence base update reviews on ASD (Rogers & Vismara, 2008; Smith & Iadarola, 2015).

**Background and Treatment Variables**
Table 7 displays comparisons between effectiveness and efficacy studies on some background variables and treatment variables. Since there are six variables tested, the Holm-Bonferroni correction was used. The only significant differences were on treatment period in months and treatment hours/week. Effectiveness studies had significantly longer treatments with higher intensity. This is due to the fact that among the effectiveness studies 93% were comprehensive, whereas only 6% of the efficacy studies were categorized as comprehensive. This yielded a significant difference using Fisher’s exact probability test (2-tailed) $p = 0.0001$. There were no significant differences between the two types of studies regarding mean age, proportion of boys, pre-treatment severity, and percent attrition. Thus, judging from the background and treatment variables which could be extracted the effectiveness studies do not comprise participants who are easier to treat than do the efficacy studies.

**Effect Size on Primary Outcome Measure**

Table 8 presents the subgroup analyses comparing the within-group effect size for effectiveness and efficacy studies within each outcome. Neither at post-treatment assessment (upper part) nor at follow-up (middle part) were there any significant differences between the two types of studies. For both types the ESs were significantly different from zero and the effects were maintained, or somewhat higher, at follow-up, which was done on average 58 months after treatment for effectiveness and 8.1 months for efficacy studies.

**Comparison of RCTs Only**

Since the outcomes presented in Table 8 may have been unduly influenced by open trials we repeated the analyses using only RCT effectiveness studies. The lower part of Table 8 shows that the effectiveness studies had a nominally higher ES than the efficacy studies, however the difference was not significant.
Discussion

The primary aim of this meta-analysis was to examine the effectiveness of early BI considered well-established or probably efficacious (Smith & Iadarola, 2015) for ASD in children when delivered in routine clinical care. The overall within-group effect size was large and significant, and medium to large and significant across the domains of adaptive behavior, cognition, communication, and socialization. Furthermore, the results showed that the outcomes were maintained at follow-up, and that a mean of 91% of the participants with ASD completed the intervention. The comparisons showed no significant differences in effect sizes between effectiveness and efficacy studies. Our results suggest that early BI for ASD are effective in routine clinical care, have a low attrition rate, and that the outcomes are comparable with how effective these interventions are in university research settings.

Our findings are in contrast with what was reported in a recent meta-analysis on the effectiveness of community-based early intervention for children with ASD (Nahmias et al., 2019; ESs [0.21-0.32]). The contrasting results may be due to several important differences between Nahmias et al. (2019) and the present meta-analysis. First, some inclusion and exclusion criteria were different yielding only 42% overlap in studies published during the same time period used in Nahmias et al. Importantly, whereas the present meta-analysis included only effectiveness and efficacy studies of early behavioral treatment methods evaluated as well-established or probably efficacious (Smith & Iadarola, 2015), Nahmias et al. (2019) also included studies of methods with a lower evidence-base level. The latter difference is important, as Nahmias et al. found variability across program outcomes with an indication of stronger results for programs based on evidence-based interventions. Also, Nahmias et al. (2019) employed different effect size measures when comparing effectiveness studies and efficacy studies, whereas the same
effect size measure for both categories of studies were used in the present meta-analysis. These important differences between the Nahmias et al. (2019) and the current meta-analysis in the body of studies included and statistical methods used, impedes the direct comparison of results.

Regarding the magnitude of effect sizes across the domains of adaptive behavior, cognition, communication, and socialization, our findings are difficult to compare as we are not aware of any other meta-analysis on the effectiveness of early interventions for ASD in children besides Nahmias et al. (2019).

We found that over 72% of the studies assessed outcomes across three or four domains. The variability probably reflects differences in the intervention targets across the studies. Furthermore, the tools utilized to measure outcome domains also varied across the studies. Currently, no uniform and common standard of outcome measures exists to assess the effects of interventions for children with ASD (Smith & Iadarola, 2015). As such, there is no consensus regarding type of outcome measure regarded as most clinically meaningful and psychometrically sound, challenging the comparison of outcomes across interventions.

Due to significant heterogeneity in the effect sizes, we examined some characteristics of the patient sample and treatment variables as potential moderators influencing treatment outcome. The number of hours/week was a significant positive moderator of effect size (i.e., more hours yielded higher effect size), and there was a nonsignificant trend that months of treatment moderated ES. The majority of the interventions were comprehensive with a mean hours of treatment per week of 24 across a mean of 19 months. The different EIBI interventions are highly intensive with up to 40 hours per week (Eldevik et al., 2009), with subsequent adaptions incorporating the EIBI techniques with varying treatment intensity (Smith & Iadarola, 2015). ASD is highly heterogeneous, and there is limited information on what degree of treatment
intensity is more effective for children and families. Previous meta-analyses and a recent study have demonstrated similar findings of a dose-response relationship between treatment intensity and outcome (Eldevik et al., 2020; Reichow, 2012; Virues-Ortega, 2010).

Later publication year was associated with lower effect sizes. The studies were conducted over a period of 33 years, during which the methodology of research designs, identification, and diagnosis of children with ASD, interventions, and inclusive educational support have evolved (Rodgers et al., 2021). It is likely that older studies may have observed larger effects, and that there are important differences in the contexts in which the interventions have been delivered during these years that may impact the effects. However, others have not found year of publication being related to outcome (e.g. Nahmias et al., 2019).

Several interesting patterns emerged in the moderator analyses of categorical variables, though mainly not to the point of statistical significance. There was no difference in effect sizes between RCTs and open trials. Only a small proportion (<13%) included a control group in the study design, and a similar small proportion (14%) included a follow-up assessment. In routine clinical care, it may be more challenging to include a control group and follow-up assessments for ethical, logistical, or financial reasons (Lake et al., 2020). For example, it may be a particular challenge to delay or withhold treatment for ASD as most of the interventions have shown efficacy for children aged five years or younger (Smith & Iadarola, 2015; Warren et al., 2012). Also, early intervention is suggested to have a larger impact as young children with ASD have not yet fallen as far behind compared to typically developing peers, and may be more amenable to change (Myers & Johnson, 2007). There is little data for comparison of follow-up across studies. Few studies report outcome beyond treatment termination, and with few exceptions there is a lack of information on long term effects (Rodgers et al., 2021). A follow-up of the children over extended periods of time in routine clinical
care could also be a challenge due to organizational and financial reasons (Lake et al., 2020). Thus, there could be several reasons for the low proportion of studies employing a control group and including follow-up assessments for children with ASD in routine clinical care. It is therefore encouraging that the effect sizes were similar across the study designs, even if the finding needs to be interpreted with caution given the small number of RCTs.

There were no differences in effect size across treatment targets (i.e., child and parent vs. child only). However, regarding setting our results showed that interventions delivered in the home resulted in a significantly higher effect size compared to the other settings. Best practice guidelines consistently identify active parent participation as an important intervention for children with ASD (Stadnick et al., 2015). Commonalities across interventions are the focus on systematically reinforcing target behaviors and promoting social communication and interacting, helping acquisition of skills, and interacting with others. Parents generally represent the most proximal and powerful environmental influence during early childhood. Family involvement is suggested to facilitate generalization and maintenance of acquired skills and to promote consistency across settings, by overcoming difficulties a child with ASD may have in conveying information across various settings (Smith & Iadarola, 2015; Stadnick et al., 2015). Thus, by involving the parents and guiding them to customize the intervention flexibly to their child’s individual learning style, implement and deliver the intervention in everyday situations, one may surpass what clinicians can provide (Smith & Iadarola, 2015). The possibility of a higher degree of family involvement when interventions are delivered in homes compared to the other settings could therefore explain the higher effect size. The finding of no differences across treatment targets, however, could be explained by the variation in parental participation across the different interventions. Furthermore, it could be explained by the smaller unique contribution of an
added parental component in addition to child only, which would have required a much larger body of studies to detect. This suggestion corresponds with studies reporting overall small effects of parent-mediated interventions (Nevill et al., 2018). Overall, an important implication is that positive outcomes for the domains of adaptive behavior, cognition, communication, and socialization can be obtained across treatment targets for children with ASD in routine clinical care.

Regarding methodological aspects, all studies were evaluated by using the psychotherapy outcome study methodology rating scale developed by Öst (2008). The results showed an overall mean of 19.2 points, which is encouraging with such a high proportion of open trials. This result is comparable to a recent meta-analysis on CBT for externalizing disorders in children and adolescents (Riise et al., 2021). Methodological flaws were noted in several of the studies, with RCTs having a significantly lower risk of bias. Overall, the effect size was not moderated by risk-of-bias score or methodology score. These results provide confidence in the overall findings of the meta-analysis.

We statistically compared our outcomes to efficacy studies to evaluate whether the magnitude of improvement achieved in routine clinical care is at the same level as randomized controlled trials from specialized research settings. The most encouraging and important finding from the present meta-analysis was that the effectiveness-efficacy comparison demonstrated no significant differences in effect size at post-treatment and follow-up. Importantly, there were no differences between the effectiveness and efficacy studies on effect sizes when only RCTs were analyzed, providing confidence in the findings. The only difference in the background and treatment variables was a higher number of treatment months, and higher number of treatment hours per week in the effectiveness studies, which is explained by the fact that 93% of the effectiveness studies were comprehensive, whereas only 6% of the efficacy studies had such a focus. Studies of treatments which are comprehensive will be more
intense and have a longer duration than studies focusing on one domain in ASD. The difference between comprehensive and focused treatments is also reflected in the measures used to evaluate treatment outcome across the four domains (i.e., adaptive, cognitive, communication, and social functioning). On average, the effectiveness studies provided data on three of the four domains, whereas the efficacy studies provided data on a mean of 1.6 domains, which aligns with a focused treatment study primarily needs to assess the domain covering the treatment’s focus. Overall, the results provide positive evidence for the transportability of empirically supported treatments for children with ASD from university research settings to routine clinical care.

Our meta-analysis contained several strong methodological elements; a power analysis indicating a high power to detect a small effect size; screening of abstract and extractions of information from the included studies in pairs of researchers where disparities were solved in consensus with all authors; ratings of methodological quality and risk-of-bias was done by one of the authors and independently by another, yielding excellent inter-rater reliability (ICC 0.92; 0.96). However, there are some limitations to consider. We only included peer-reviewed published or in press studies in English language journals. Studies published in other languages could have provided additional information about the effectiveness of early BI for ASD in children. Furthermore, the inclusion of only published studies could be viewed as a limitation. However, our pool of studies spanned four decades. Including unpublished studies could have introduced bias as it could have been easier to identify unpublished studies from more recent compared to earlier decades. For the included studies, only a small number (17%) provided information on ethnicity, which is too low to draw any conclusions. Furthermore, the difference in the percentage of comprehensive compared to focused interventions in the effectiveness versus efficacy studies complicated the comparison. It should also be noted that the review of efficacy studies
used for comparisons was published in 2015, and the early interventions for children with ASD have developed since then. Finally, it may be claimed that the effectiveness of an intervention is demonstrated when it exceeds the effects of the treatment youth and families usually received in the clinic, i.e. usual clinical care. However, our aim was to examine the degree of improvement that can be expected following early behavioral interventions with established evidence for ASD when delivered in routine clinical care. Thus, we included both open trials and RCTs to better capture all research conducted in routine clinical care contexts and be as comprehensive as possible.

The current meta-analysis suggests that early BI for ASD are effective in routine clinical care. However, it also highlights areas of improvements and opportunities, including a more detailed classification of the programmes used based on the core components, and a common standard of outcome measures for assessing the effectiveness of interventions. With the heterogeneity of ASD, more careful characteristics of children and their families would be beneficial to include in future studies so that clinicians may have a better understanding of whether a particular interventions is appropriate for the child and family.

In conclusion, our findings demonstrate the effectiveness of behavioral interventions considered well-established or probably efficacious according to the Smith and Iadarola (2015) update review for ASD in children, and suggest that clinicians can be confident about the effectiveness of these interventions in routine clinical care. Effect sizes are comparable to those in university research clinic settings. As treatment effects are not lost when these evidence-based treatment programs are transported from research clinics to routine clinical care, there is a need to further implement effective and evidence-based interventions in routine clinical care for children with ASD.


doi.org/10.1016/j.rasd.2014.05.010


doi.org/10.1016/j.rasd.2006.12.001
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Continent</th>
<th>Type of study</th>
<th>Analysis</th>
<th>Methodology score (%)</th>
<th>Severity</th>
<th>N</th>
<th>% boys</th>
<th>Mean age in months</th>
<th>Age range in months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson, 1987</td>
<td>USA</td>
<td>North America</td>
<td>Open trial</td>
<td>ITT</td>
<td>35.0</td>
<td></td>
<td>14</td>
<td></td>
<td>43.0</td>
<td>18-64</td>
</tr>
<tr>
<td>Ben-Itzchak, 2007</td>
<td>Israel</td>
<td>Asia</td>
<td>Open trial</td>
<td>ITT</td>
<td>40.0</td>
<td></td>
<td>25</td>
<td>92.0</td>
<td>26.6</td>
<td>20-32</td>
</tr>
<tr>
<td>Ben-Itzchak, 2014</td>
<td>Israel</td>
<td>Asia</td>
<td>Open trial</td>
<td>Compl.</td>
<td>42.5</td>
<td>0.760 [a]</td>
<td>46</td>
<td>84.8</td>
<td>25.5</td>
<td>17-33</td>
</tr>
<tr>
<td>Bibby, 2002</td>
<td>UK</td>
<td>Europe</td>
<td>Open trial</td>
<td>Compl.</td>
<td>40.0</td>
<td>0.613 [b]</td>
<td>66</td>
<td>83.3</td>
<td>45.0</td>
<td></td>
</tr>
<tr>
<td>Cohen, 2006</td>
<td>USA</td>
<td>North America</td>
<td>Open trial</td>
<td>Compl.</td>
<td>55.0</td>
<td>0.463 [c]</td>
<td>21</td>
<td>85.7</td>
<td>30.2</td>
<td>20-41</td>
</tr>
<tr>
<td>Eikeseth, 2012</td>
<td>Sweden</td>
<td>Europe</td>
<td>Open trial</td>
<td>Compl.</td>
<td>36.4</td>
<td>0.620 [d]</td>
<td>35</td>
<td>82.9</td>
<td>47.0</td>
<td>25-76</td>
</tr>
<tr>
<td>Eldevik, 2012</td>
<td>Norway</td>
<td>Europe</td>
<td>Open trial</td>
<td>ITT</td>
<td>52.3</td>
<td>0.519 [c]</td>
<td>31</td>
<td>80.6</td>
<td>42.2</td>
<td>24-72</td>
</tr>
<tr>
<td>Eldevik, 2020a</td>
<td>Norway</td>
<td>Europe</td>
<td>Open trial</td>
<td>ITT</td>
<td>50.0</td>
<td>0.515 [d]</td>
<td>21</td>
<td>90.5</td>
<td>46.8</td>
<td>32-69</td>
</tr>
<tr>
<td>Eldevik, 2020b</td>
<td>Norway</td>
<td>Europe</td>
<td>Open trial</td>
<td>ITT</td>
<td>50.0</td>
<td>0.677 [d]</td>
<td>36</td>
<td>80.6</td>
<td>44.1</td>
<td>21-77</td>
</tr>
<tr>
<td>Goin-Kochel, 2007</td>
<td>USA</td>
<td>North America</td>
<td>Open trial</td>
<td>Compl.</td>
<td>35.0</td>
<td>0.499 [c]</td>
<td>29</td>
<td>93.1</td>
<td>45.7</td>
<td>30-61</td>
</tr>
<tr>
<td>Hayward, 2009a</td>
<td>UK</td>
<td>Europe</td>
<td>Open trial</td>
<td>Compl.</td>
<td>54.5</td>
<td>0.521 [c]</td>
<td>23</td>
<td>82.6</td>
<td>37.4</td>
<td>24-42</td>
</tr>
<tr>
<td>Hayward, 2009b</td>
<td>UK</td>
<td>Europe</td>
<td>Open trial</td>
<td>Compl.</td>
<td>54.5</td>
<td>0.474 [c]</td>
<td>21</td>
<td>71.4</td>
<td>34.2</td>
<td>24-42</td>
</tr>
<tr>
<td>Howard, 2005; 2014</td>
<td>USA</td>
<td>North America</td>
<td>Open trial</td>
<td>Compl.</td>
<td>40.9</td>
<td>0.585 [c]</td>
<td>29</td>
<td>86.2</td>
<td>30.9</td>
<td></td>
</tr>
<tr>
<td>Janson, 2020</td>
<td>Sweden</td>
<td>Europe</td>
<td>RCT</td>
<td>Compl.</td>
<td>47.7</td>
<td>0.760 [c]</td>
<td>19</td>
<td>78.9</td>
<td>40.1</td>
<td>30-48</td>
</tr>
<tr>
<td>Luiselli, 2000</td>
<td>USA</td>
<td>North America</td>
<td>Open trial</td>
<td>Compl.</td>
<td>42.5</td>
<td></td>
<td>16</td>
<td>93.8</td>
<td>39.7</td>
<td>26-57</td>
</tr>
<tr>
<td>Study</td>
<td>Location</td>
<td>Region</td>
<td>Trial Type</td>
<td>Analysis</td>
<td>Severity</td>
<td>% Improvement</td>
<td>% Severity Change</td>
<td>Study Duration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>----------</td>
<td>----------</td>
<td>---------------</td>
<td>------------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magiati, 2007, 2011</td>
<td>UK</td>
<td>Europe</td>
<td>Open trial</td>
<td>Compl.</td>
<td>54.5</td>
<td>0.542 [c]</td>
<td>28</td>
<td>96.4</td>
<td>38.0</td>
<td>23-54</td>
</tr>
<tr>
<td>Makrygianni, 2018</td>
<td>Greece</td>
<td>Europe</td>
<td>Open trial</td>
<td>Compl.</td>
<td>38.6</td>
<td>0.424 [e]</td>
<td>10</td>
<td>54.0</td>
<td>30-78</td>
<td></td>
</tr>
<tr>
<td>Perry, 2008</td>
<td>Canada</td>
<td>North America</td>
<td>Open trial</td>
<td>Compl.</td>
<td>40.0</td>
<td>0.602 [d]</td>
<td>332</td>
<td>83.1</td>
<td>54.0</td>
<td>20-86</td>
</tr>
<tr>
<td>Rad, 2019</td>
<td>Romania</td>
<td>Europe</td>
<td>Open trial</td>
<td>Compl.</td>
<td>47.7</td>
<td>0.529 [a]</td>
<td>48</td>
<td>81.3</td>
<td>42.7</td>
<td>24-60</td>
</tr>
<tr>
<td>Remington, 2007; 2011</td>
<td>UK</td>
<td>Europe</td>
<td>Open trial</td>
<td>Compl.</td>
<td>54.8</td>
<td>0.568 [f]</td>
<td>23</td>
<td>35.7</td>
<td>30-42</td>
<td></td>
</tr>
<tr>
<td>Rivard, 2014; 2019</td>
<td>Canada</td>
<td>North America</td>
<td>Open trial</td>
<td>ITT</td>
<td>47.7</td>
<td>0.587 [d]</td>
<td>93</td>
<td>75.3</td>
<td>46.0</td>
<td>33-57</td>
</tr>
<tr>
<td>Sallows, 2005a</td>
<td>USA</td>
<td>North America</td>
<td>RCT</td>
<td>Compl.</td>
<td>45.5</td>
<td>0.542 [c]</td>
<td>13</td>
<td>84.6</td>
<td>35.0</td>
<td>35-37</td>
</tr>
<tr>
<td>Sallows, 2005b</td>
<td>USA</td>
<td>North America</td>
<td>RCT</td>
<td>ITT</td>
<td>45.5</td>
<td>0.532 [c]</td>
<td>10</td>
<td>80.0</td>
<td>37.1</td>
<td>35-37</td>
</tr>
<tr>
<td>Schertz, 2013</td>
<td>USA</td>
<td>North America</td>
<td>RCT</td>
<td>Compl.</td>
<td>54.5</td>
<td>0.650 [a]</td>
<td>11</td>
<td>24.6</td>
<td>25-28</td>
<td></td>
</tr>
<tr>
<td>Sheinkopf, 1998</td>
<td>USA</td>
<td>North America</td>
<td>Open trial</td>
<td>Compl.</td>
<td>36.4</td>
<td>0.546 [g]</td>
<td>11</td>
<td>33.8</td>
<td>23-47</td>
<td></td>
</tr>
<tr>
<td>Smith, 2015</td>
<td>USA</td>
<td>North America</td>
<td>Open trial</td>
<td>ITT</td>
<td>52.5</td>
<td>0.850 [a]</td>
<td>71</td>
<td>84.5</td>
<td>39.2</td>
<td>24-58</td>
</tr>
<tr>
<td>Smith, 2019</td>
<td>Canada</td>
<td>North America</td>
<td>Open trial</td>
<td>Compl.</td>
<td>31.8</td>
<td>0.382 [c]</td>
<td>147</td>
<td>76.5</td>
<td>41.0</td>
<td></td>
</tr>
<tr>
<td>Tonge, 2014</td>
<td>Australia</td>
<td>Australia</td>
<td>RCT</td>
<td>ITT</td>
<td>50.0</td>
<td>0.687 [d]</td>
<td>35</td>
<td>74.3</td>
<td>46.0</td>
<td>24-65</td>
</tr>
<tr>
<td>Waters, 2020</td>
<td>USA</td>
<td>North America</td>
<td>Open trial</td>
<td>ITT</td>
<td>50.0</td>
<td>0.461 [c]</td>
<td>48</td>
<td>93.8</td>
<td>37.8</td>
<td></td>
</tr>
<tr>
<td>Wood, 2018</td>
<td>Australia</td>
<td>Australia</td>
<td>Open trial</td>
<td>Compl.</td>
<td>42.5</td>
<td>0.603 [d]</td>
<td>45</td>
<td>78.0</td>
<td>40.7</td>
<td>24-64</td>
</tr>
<tr>
<td>Zachor, 2007</td>
<td>Israel</td>
<td>Asia</td>
<td>Open trial</td>
<td>ITT</td>
<td>40.9</td>
<td>0.20</td>
<td>95.0</td>
<td>27.7</td>
<td>22-34</td>
<td></td>
</tr>
<tr>
<td>Zachor, 2010</td>
<td>Israel</td>
<td>Asia</td>
<td>Open trial</td>
<td>Compl.</td>
<td>45.5</td>
<td>0.45</td>
<td>91.0</td>
<td>25.1</td>
<td>17-35</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.

Treatment data of the included studies.

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatment method *</th>
<th>Treatment target</th>
<th>Setting</th>
<th>Treatment hrs/week</th>
<th>Months of treatment</th>
<th>Percent attrition</th>
<th>F-up months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson, 1987</td>
<td>EIBI*</td>
<td>I+P</td>
<td>H</td>
<td>20.0</td>
<td>12.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Ben-Itzchak, 2007</td>
<td>ABA*</td>
<td>I+P</td>
<td>CC</td>
<td>35.0</td>
<td>12.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Ben-Itzchak, 2014</td>
<td>EIBI*</td>
<td>I+P</td>
<td>CC</td>
<td>20.0</td>
<td>24.0</td>
<td>138</td>
<td></td>
</tr>
<tr>
<td>Bibby, 2002</td>
<td>EIBI*</td>
<td>I+P</td>
<td>H</td>
<td>30.3</td>
<td>32.8</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Cohen, 2006</td>
<td>EIBI*</td>
<td>I+P</td>
<td>H+P</td>
<td>37.5</td>
<td>36.0</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>Eikeseth, 2012</td>
<td>EIBI*</td>
<td>I</td>
<td>P</td>
<td>23.0</td>
<td>12.0</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Eldevik, 2012</td>
<td>EIBI*</td>
<td>I</td>
<td>P</td>
<td>13.6</td>
<td>25.1</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Eldevik, 2020a</td>
<td>EIBI-lower*</td>
<td>I</td>
<td>P</td>
<td>11.1</td>
<td>12.5</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Eldevik, 2020b</td>
<td>EIBI-higher*</td>
<td>I</td>
<td>P</td>
<td>18.2</td>
<td>12.1</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Goin-Kochel, 2007</td>
<td>ABA*</td>
<td>I</td>
<td>P</td>
<td>30.0</td>
<td>23.7</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Hayward, 2009a</td>
<td>EIBI-clinic*</td>
<td>I+P</td>
<td>C</td>
<td>37.4</td>
<td>12.0</td>
<td>13.0</td>
<td></td>
</tr>
<tr>
<td>Hayward, 2009b</td>
<td>EIBI-parent*</td>
<td>I+P</td>
<td>H</td>
<td>34.2</td>
<td>12.0</td>
<td>9.5</td>
<td></td>
</tr>
<tr>
<td>Howard, 2005</td>
<td>IBT*</td>
<td>I</td>
<td>H+P</td>
<td>32.5</td>
<td>14.2</td>
<td>9.8</td>
<td></td>
</tr>
<tr>
<td>Janson, 2020</td>
<td>IBT*</td>
<td>I+P</td>
<td>P</td>
<td>14.4</td>
<td>3.0</td>
<td>9.5</td>
<td></td>
</tr>
<tr>
<td>Luiselli, 2000</td>
<td>EIBI*</td>
<td>I</td>
<td>H</td>
<td>13.7</td>
<td>9.4</td>
<td>9.4</td>
<td></td>
</tr>
<tr>
<td>Magiati, 2007</td>
<td>EIBI*</td>
<td>I</td>
<td>H</td>
<td>32.8</td>
<td>24.0</td>
<td>15.2</td>
<td>58</td>
</tr>
<tr>
<td>Makrygianni, 2018</td>
<td>ABA*</td>
<td>I</td>
<td>P</td>
<td>23.8</td>
<td>9.0</td>
<td>9.0</td>
<td></td>
</tr>
<tr>
<td>Perry, 2008</td>
<td>IBI*</td>
<td>I</td>
<td>CC+H</td>
<td>30.0</td>
<td>18.0</td>
<td>18.0</td>
<td></td>
</tr>
<tr>
<td>Rad, 2019</td>
<td>ABA*</td>
<td>I</td>
<td>C</td>
<td>10.0</td>
<td>12.0</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>Remington, 2007</td>
<td>EIBI*</td>
<td>I+P</td>
<td>H+P</td>
<td>18.1</td>
<td>23.0</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Rivard, 2014</td>
<td>EIBI*</td>
<td>I+P</td>
<td>CC</td>
<td>18.0</td>
<td>12.0</td>
<td>0.0</td>
<td>12</td>
</tr>
<tr>
<td>Sallows, 2005a</td>
<td>EIBI-clinic*</td>
<td>I+P</td>
<td>C</td>
<td>37.6</td>
<td>48.0</td>
<td>7.7</td>
<td></td>
</tr>
<tr>
<td>Sallows, 2005b</td>
<td>EIBI-parent*</td>
<td>I+P</td>
<td>H</td>
<td>35.7</td>
<td>48.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Schertz, 2013</td>
<td>JAML</td>
<td>I+P</td>
<td>H</td>
<td>9.4</td>
<td>7.0</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>Sheinkopf, 1998</td>
<td>EIBI*</td>
<td>I+P</td>
<td>H+P</td>
<td>27.0</td>
<td>15.7</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>Smith, 2015</td>
<td>EIBI*</td>
<td>I</td>
<td>CC</td>
<td>18.4</td>
<td>24.0</td>
<td>9.9</td>
<td></td>
</tr>
<tr>
<td>Smith, 2019</td>
<td>EIBI*</td>
<td>I</td>
<td>H+P</td>
<td>20.0</td>
<td>12.0</td>
<td>12.0</td>
<td></td>
</tr>
<tr>
<td>Tonge, 2014</td>
<td>PEBM</td>
<td>I+P</td>
<td>CC</td>
<td>7.3</td>
<td>5.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Waters, 2020</td>
<td>EIBI*</td>
<td>I+P</td>
<td>H+P</td>
<td>37.5</td>
<td>36.0</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>Treatment</td>
<td>Setting</td>
<td>Age</td>
<td>Duration</td>
<td>Cost</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
<td>-----------</td>
<td>---------</td>
<td>-----</td>
<td>----------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>Wood, 2018</td>
<td>EIBI*</td>
<td>I</td>
<td>CC</td>
<td>20.3</td>
<td>24.3</td>
<td>20.3</td>
<td></td>
</tr>
<tr>
<td>Zachor, 2007</td>
<td>ABA*</td>
<td>I</td>
<td>CC</td>
<td>35.0</td>
<td>12.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Zachor, 2010</td>
<td>ABA*</td>
<td>I+P</td>
<td>CC</td>
<td>20.0</td>
<td>12.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3
Within-group effect size (Hedges’ $g$) for all studies (RCTs and open trials).

<table>
<thead>
<tr>
<th>Time point</th>
<th>$k$</th>
<th>$g$-value</th>
<th>95% CI</th>
<th>$z$-value</th>
<th>Q-value</th>
<th>$I^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>33</td>
<td>0.94</td>
<td>0.76-1.13</td>
<td>9.93$^b$</td>
<td>199.9$^b$</td>
<td>84.0</td>
</tr>
<tr>
<td>Follow-up</td>
<td>4</td>
<td>1.08</td>
<td>0.17-1.98</td>
<td>2.34$^a$</td>
<td>42.1$^b$</td>
<td>92.9</td>
</tr>
</tbody>
</table>

Note. $k$ = number of treatment conditions. $^a p < 0.05, ^b p < 0.0001.$
**Table 4**

Within-group effect size (Hedges’ *g*) at post-treatment for all studies divided on construct.

<table>
<thead>
<tr>
<th>Construct</th>
<th>k</th>
<th>g-value</th>
<th>95% CI</th>
<th>z-value</th>
<th>Q-value</th>
<th>$I^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive</td>
<td>21</td>
<td>0.81</td>
<td>0.50-1.12</td>
<td>5.10a</td>
<td>281.3a</td>
<td>92.9</td>
</tr>
<tr>
<td>Cognitive</td>
<td>24</td>
<td>0.76</td>
<td>0.57-0.96</td>
<td>7.61a</td>
<td>129.1a</td>
<td>82.2</td>
</tr>
<tr>
<td>Communication</td>
<td>27</td>
<td>1.27</td>
<td>1.00-1.54</td>
<td>9.20a</td>
<td>232.0a</td>
<td>88.8</td>
</tr>
<tr>
<td>Socialization</td>
<td>24</td>
<td>1.01</td>
<td>0.76-1.26</td>
<td>7.91a</td>
<td>216.0a</td>
<td>89.4</td>
</tr>
</tbody>
</table>

*Note. k = number of treatment conditions. *a* *p* < 0.0001.*
### Table 5

Subgroup analysis of the effect size for all studies at post-treatment.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect size</th>
<th>k</th>
<th>g-value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of study</strong> (Q_b = 0.44, p = 0.51)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCT</td>
<td></td>
<td>5</td>
<td>1.12</td>
<td>0.55-1.70</td>
</tr>
<tr>
<td>Open trial</td>
<td></td>
<td>28</td>
<td>0.92</td>
<td>0.72-1.12</td>
</tr>
<tr>
<td><strong>Statistical analysis</strong> (Q_b = 0.74, p = 0.39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intent-to-treat</td>
<td></td>
<td>12</td>
<td>0.83</td>
<td>0.57-1.10</td>
</tr>
<tr>
<td>Completers</td>
<td></td>
<td>21</td>
<td>0.99</td>
<td>0.75-1.23</td>
</tr>
<tr>
<td><strong>Treatment target</strong> (Q_b = 0.10, p = 0.76)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child</td>
<td></td>
<td>15</td>
<td>0.90</td>
<td>0.62-1.19</td>
</tr>
<tr>
<td>Child + parent</td>
<td></td>
<td>18</td>
<td>0.96</td>
<td>0.72-1.20</td>
</tr>
<tr>
<td><strong>Setting</strong> (Q_b = 12.55, p = 0.014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home</td>
<td></td>
<td>7</td>
<td>1.34</td>
<td>1.03-1.66</td>
</tr>
<tr>
<td>Clinic</td>
<td></td>
<td>3</td>
<td>1.33</td>
<td>0.63-2.02</td>
</tr>
<tr>
<td>Home + preschool</td>
<td></td>
<td>7</td>
<td>0.84</td>
<td>0.49-1.19</td>
</tr>
<tr>
<td>Preschool</td>
<td></td>
<td>7</td>
<td>0.72</td>
<td>0.39-1.05</td>
</tr>
<tr>
<td>Community centre</td>
<td></td>
<td>8</td>
<td>0.69</td>
<td>0.41-0.96</td>
</tr>
<tr>
<td><strong>Continent</strong> (Q_b = 1.98, p = 0.58)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td></td>
<td>4</td>
<td>0.93</td>
<td>0.36-1.50</td>
</tr>
<tr>
<td>Australia</td>
<td></td>
<td>2</td>
<td>0.80</td>
<td>0.32-1.29</td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td>12</td>
<td>0.82</td>
<td>0.58-1.05</td>
</tr>
<tr>
<td>North America</td>
<td></td>
<td>15</td>
<td>1.09</td>
<td>0.76-1.43</td>
</tr>
</tbody>
</table>

*Note. k = number of treatment conditions, Q_b = Q between subgroups. The statistic in parenthesis tests if the subgroups within the individual category differ significantly from each other.*
Table 6

Meta-regression analysis of the effect size for all studies at post-treatment.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$k$</th>
<th>Point estimate</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours of tx/week</td>
<td>33</td>
<td>0.032</td>
<td>3.70</td>
<td>0.0002</td>
</tr>
<tr>
<td>Year of publication</td>
<td>33</td>
<td>-0.043</td>
<td>-3.60</td>
<td>0.0003</td>
</tr>
<tr>
<td>Months of treatment</td>
<td>33</td>
<td>0.017</td>
<td>1.83</td>
<td>0.07</td>
</tr>
<tr>
<td>Percent boys</td>
<td>28</td>
<td>0.020</td>
<td>1.32</td>
<td>0.19</td>
</tr>
<tr>
<td>Severity</td>
<td>28</td>
<td>-0.984</td>
<td>-1.03</td>
<td>0.30</td>
</tr>
<tr>
<td>Age</td>
<td>33</td>
<td>-0.012</td>
<td>-0.98</td>
<td>0.33</td>
</tr>
<tr>
<td>Percent attrition</td>
<td>22</td>
<td>0.018</td>
<td>0.95</td>
<td>0.34</td>
</tr>
<tr>
<td>Risk-of-bias score</td>
<td>33</td>
<td>-0.048</td>
<td>-0.43</td>
<td>0.66</td>
</tr>
<tr>
<td>Methodology score</td>
<td>33</td>
<td>-0.005</td>
<td>-0.38</td>
<td>0.70</td>
</tr>
<tr>
<td>Number of participants</td>
<td>33</td>
<td>0.0002</td>
<td>0.16</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Note. $k$ = number of treatment conditions. tx=treatment.
### Table 7

Some background and treatment data (M and SD) for effectiveness and efficacy studies.

<table>
<thead>
<tr>
<th>Study type</th>
<th>$k$</th>
<th>Boys (%)</th>
<th>Age (months)</th>
<th>Severity (%)</th>
<th>Treatment (months)</th>
<th>Treatment hours/week</th>
<th>Attrition (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectiveness</td>
<td>32</td>
<td>84.5 (6.8)</td>
<td>38.7 (7.9)</td>
<td>57.4 (10.7)</td>
<td>18.6 (11.4)</td>
<td>24.1 (9.6)</td>
<td>6.1 (6.2)</td>
</tr>
<tr>
<td>Efficacy</td>
<td>18</td>
<td>85.7 (6.4)</td>
<td>43.0 (8.8)</td>
<td>56.4 (16.6)</td>
<td>7.2 (7.6)</td>
<td>8.6 (9.1)</td>
<td>5.7 (8.7)</td>
</tr>
</tbody>
</table>

$p$: 0.60 0.09 0.80 0.0004 0.0001 0.85

*Note. $k =$ number of treatment conditions, Severity = percentage of the maximum score on the primary outcome measure. Attrition (%) = proportion dropping out of those participating in at least one session.*
Table 8

Effect sizes (Hedges’ g) for effectiveness and efficacy studies.

<table>
<thead>
<tr>
<th>Time point</th>
<th>Study type</th>
<th>k</th>
<th>g-value</th>
<th>95% CI</th>
<th>z-value</th>
<th>Q between</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>Effectiveness</td>
<td>33</td>
<td>0.94</td>
<td>0.76-1.13</td>
<td>9.93(^b)</td>
<td></td>
<td>3.16</td>
</tr>
<tr>
<td></td>
<td>Efficacy</td>
<td>18</td>
<td>0.69</td>
<td>0.48-0.90</td>
<td>6.39(^b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td>Effectiveness</td>
<td>4</td>
<td>1.08</td>
<td>0.17-1.98</td>
<td>2.34(^a)</td>
<td></td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Efficacy</td>
<td>6</td>
<td>1.54</td>
<td>0.84-2.25</td>
<td>4.29(^b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCTs only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>Effectiveness</td>
<td>5</td>
<td>1.12</td>
<td>0.55-1.70</td>
<td>3.84(^b)</td>
<td></td>
<td>1.97</td>
</tr>
<tr>
<td></td>
<td>Efficacy</td>
<td>18</td>
<td>0.69</td>
<td>0.48-0.90</td>
<td>6.39(^b)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. \(k\) = number of comparisons. \(^a\) \(p < 0.05\), \(^b\) \(p < 0.0001\). Q between = comparison Effectiveness vs. Efficacy at the respective time points.
Figure 1. Flowchart of the inclusion of studies.

Note. *29 unique studies across 34 articles.
Supplement to

S1. Literature search strategies.

S2. References to excluded studies.

S3. Table of rating scales used for primary continuous measure.

S4. References to efficacy studies used for comparison.

S5. Table of risk-of-bias evaluation.

S6. Table of ethnicity.

S7. Table of focus of treatment.

S8. PRISMA Checklist.

S9. AMSTAR2.
SI. Literature search strategies

Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily <1946 to April 06, 2021> 07-04-21

1. exp Behavior Therapy/ 78488
2. (CBT or behavio?r* therapy or behavio?r* intervent* or behavio?r* modification or behavio?r* management or EIBI or behavio?r* analys* or ABA).ti,ab,kw. 60569
3. (Discrete-trial adj1 (instruction or training or teaching or intervention)).ti,ab,kw. 107
4. token economy/ or Conditioning, psychological/ or exp Reinforcement, Psychology/ or exp "Task performance and analysis"/ 102411
5. (contingency management or token economy or fading conditioning or omission conditioning or reinforcement or functional communication or picture exchange communication system or PECS or task analysis or response interruption or redirection).ti,ab,kw. 42988
6. 1 or 2 or 3 or 4 or 5 244155
7. exp Child Development Disorders, Pervasive/ 37499
8. (autism or (autistic adj2 disorder*) or asperger* or pervasive development* disorder* or childhood disintegrative disorder* or ASD or PDD or PDD-NOS).ti,ab,kw. 60779
9. 7 or 8 66020
10. adolescent/ or exp child/ 3092326
11. (pediatric* or paediatric* or infant* or infancy or toddler* or preschool* or pre-school* or child* or adolescen* or youth).ti,ab,kw. 2135035
12. 10 or 11 3948915
13. 6 and 9 and 123113
14. Outpatient Clinics, Hospital/ 15725
15. Community Mental Health Services/ 18718
16. Schools/ 40428
17. (effectiveness or community clinic* or outpatient clinic* or routine care or regular care or daycare or day-care or school* or kindergar?en).ti,ab,kw. 810070
18. 14 or 15 or 16 or 17 844685
19. 6 and 9 and 12 and 18 660

Database: Embase (Ovid) <1974 to 2021 April 06>
07-04-2021

1. exp behavior therapy/ or exp cognitive behavioral therapy/ 60460
(CBT or behavio?r* therapy or behavio?r* intervent* or behavio?r* modification or behavio?r* management or EIBI or behavio?r* analys* or ABA).ti,ab,kw. 80865

(Discrete-trial adj1 (instruction or training or teaching or intervention)).ti,ab,kw. 132

reinforcement/ or exp conditioning/ or task performance/ 230105

(contingency management or token economy or fading conditioning or omission conditioning or reinforcement or functional communication or picture exchange communication system or PECS or task analysis or response interruption or redirection).ti,ab,kw. 51093

1 or 2 or 3 or 4 or 5 366684

exp autism/ 75407

7 or 8 102686

child/ or preschool child/ or school child/ or adolescent/ 2882587

(pediatric* or paediatric* or infant* or infancy or toddler* or preschool* or pre-school* or child* or adolescent* or youth).ti,ab,kw. 2632354

10 or 11 3891767

6 and 9 and 125286

outpatient department/ 68138

community care/ 55952

exp community mental health service/ 364

exp school/ 364703

(effectiveness or community clinic* or outpatient clinic* or routine care or regular care or daycare or day-care or school* or kindergar?en).ti,ab,kw. 1075040

14 or 15 or 16 or 17 or 18 1399995

13 and 19 1076

limit 20 to conference abstracts 123

20 not 21 953

Database: APA PsycInfo (Ovid) <1806 to March Week 5 2021>
07-04-21

exp behavior therapy/ or exp cognitive behavior therapy/ 42514

(CBT or behavio?r* therapy or behavio?r* intervent* or behavio?r* modification or behavio?r* management or EIBI or behavio?r* analys* or ABA).tw. 67160

(Discrete-trial adj1 (instruction or training or teaching or intervention)).tw. 343
exp contingency management/ 3112
exp classical conditioning/ or exp reinforcement/ or communication skills training/ or "extinction (learning)"/ or exp task analysis/ 85139
(contingency management or token economy or fading conditioning or omission conditioning or reinforcement or functional communication or picture exchange communication system or PECS or task analysis or response interruption or redirection).tw. 50799
1 or 2 or 3 or 4 or 5 or 6 187365
autism spectrum disorders/ 46071
(autoism or (autistic adj2 disorder*) or asperger* or pervasive development* disorder* or childhood disintegrative disorder* or ASD or PDD or PDD-NOS).tw. 56059
8 or 9 58659
7 and 10 5737
limit 11 to (100 childhood <birth to age 12 yrs> or 200 adolescence <age 13 to 17 yrs>) 3516
(pediatric* or paediatric* or infant* or infancy or toddler* or preschool* or pre-school* or child* or adolescen* or youth).tw. 980226
11 and 13 4202
12 or 14 4716
exp outpatient treatment/ 7137
exp schools/ 71546
(effectiveness or community clinic* or outpatient clinic* or routine care or regular care or daycare or day-care or school* or kindergar?en).tw. 586336
16 or 17 or 18 605990
15 and 19 1400
limit 20 to dissertation 240
20 not 21 1160


----------------------------------------------
S2  DE "Task Analysis"  12,079
S3  TI ( CBT or "behavior* therapy" or "behavior* intervention" or "behavior* modification" or "behavior* management" or EIBI or "behavior* analysis" or ABA ) OR AB ( CBT or "behavior* therapy" or "behavior* intervention" or "behavior* modification" or "behavior* management" or EIBI or "behavior* analysis" or ABA )  9,120
S4  TI ( (Discrete-trial or "discrete trial") N0 (instruction or training or teaching or intervention) ) OR AB ( (Discrete-trial or "discrete trial") N0 (instruction or training or teaching or intervention) )  156
S5  TI ( "contingency management" or "token economy" or "fading conditioning" or "omission conditioning" or reinforcement or "functional communication" or "picture exchange communication system" or PECS or "task analysis" or "response interruption" or redirection ) OR AB ( "contingency management" or "token economy" or "fading conditioning" or "omission conditioning" or reinforcement or "functional communication" or "picture exchange communication system" or PECS or "task analysis" or "response interruption" or redirection )  10,533
S6  S1 OR S2 OR S3 OR S4 OR S5  45,813
S7  DE "Pervasive Developmental Disorders" OR DE "Asperger Syndrome" OR DE "Autism"  16,247
S8  TI ( autism or (autistic N1 disorder*) or asperger* or "pervasive development* disorder*" or "childhood disintegrative disorder*" or ASD or PDD or PDD-NOS ) OR AB ( autism or (autistic N1 disorder*) or asperger* or "pervasive development* disorder*" or "childhood disintegrative disorder*" or ASD or PDD or PDD-NOS )  15,264
S9  S7 OR S8  16,906
S10  DE "Adolescents" OR DE "Children"  88,938
S11  TI ( pediatric* or paediatric* or infant* or infancy or toddler* or preschool* or pre-school* or child* or adolescent* or youth ) OR AB ( pediatric* or paediatric* or infant* or infancy or toddler* or preschool* or pre-school* or child* or adolescent* or youth )  376,716
S12  S10 OR S11  389,140
S13  S6 AND S9 AND S12  2,147
S14  DE "Schools" OR DE "Bilingual Schools" OR DE "Boarding Schools" OR DE "Colleges" OR DE "Community Schools" OR DE "Consolidated Schools" OR DE "Correspondence Schools" OR DE "Day Schools" OR DE "Disadvantaged Schools" OR DE "Elementary Schools" OR DE "Experimental Schools" OR DE "Folk Schools" OR DE "Free Schools" OR DE "International Schools" OR DE "Laboratory Schools" OR DE "Magnet Schools" OR DE "Middle Schools" OR DE "Military Schools" OR DE "Montessori Schools" OR DE "Multiunit Schools" OR DE "Neighborhood Schools" OR DE "Nursery Schools" OR DE "Open Plan Schools" OR DE "Private Schools" OR DE "Professional Development Schools" OR DE "Public Schools" OR DE "Racially Balanced Schools" OR DE "Regional Schools" OR DE "Rural Schools" OR DE "Schools of Education" OR DE "Secondary Schools" OR DE "Single Sex Schools" OR DE "Slum Schools" OR DE "Small Schools" OR DE "Special Schools" OR DE "State Schools" OR DE "Suburban Schools" OR DE "Summer Schools" OR DE "Traditional Schools" OR DE "Urban Schools" OR DE "Vocational Schools" OR DE "Year Round Schools"  142,134
S15  DE "Child Care"  4,623
S16  TI ( effectiveness or community clinic* or "outpatient clinic*" or "routine care" or "regular care" or daycare or day-care or school* or kindergar#en ) OR AB ( effectiveness or
community clinic* or "outpatient clinic*" or "routine care" or "regular care" or daycare or daycare or school* or kindergar#en )  

S17  
S14 OR S15 OR S16  

S18  
S6 AND S9 AND S12 AND S17  

597,381  

629,068  

708  

Web of science (Clarivate)  
Indexes=SCI-EXPANDED, SSCI, A&HCI, ESCI  
Timespan=All years, search date 7th April 2021.

# 1  
102,757 - TOPIC: (CBT or "behavio$r* therapy" or "behavio$r* intervent*" or "behavio$r* modification" or "behavio$r* management" or EIBI or "behavio$r* analys*" or ABA)

# 2  
237 - TOPIC: ("Discrete-trial" or "discrete trial") NEAR/0 ("instruction" or "training" or "teaching" or "intervention")

# 3  
130,278 - TOPIC: ("contingency management" or "token economy" or "fading conditioning" or "omission conditioning" or "reinforcement" or "functional communication" or "picture exchange communication system" or PECS or "task analysis" or "response interruption" or "redirection")

# 4  
230,931 - #3 OR #2 OR #1

# 5  
88,661 - TS=("autism" or ("autistic" NEAR/1 disorder*) or asperger* or "pervasive development* disorder*" or "childhood disintegrative disorder*" or ASD or PDD or PDD-NOS)

# 6  
2,650,572 - TOPIC: (pediatric* or paediatric* or infant* or infancy or toddler* or preschool* or pre-school* or child* or adolescen* or "youth")

# 7  
1,455,145 - TOPIC: ("effectiveness" or "community clinic*" or "outpatient clinic*" or "routine care" or "regular care" or "daycare" or "day-care" or school* or "kindergar$en")

# 8  
1,056 - #7 AND #6 AND #5 AND #4
S2. References to excluded studies.

Studies could be excluded for more than one reason and could thus potentially be listed in several of the categories below. However, reason for exclusion was categorized as the first exclusion criteria that became evident to the authors when reading the study.

Treatment at research clinics:


*Age > 5 years:*

Brookman-Frazee, L. I., Drahota, A., & Stadnick, N. (2012). Training community mental health therapists to deliver a package of evidence-based practice strategies for school-age children with...


Follow-up study of Eikeseth et al., 2002:


Practicing clinicians:


Participants not clinically referred:


*Dissertations, conference proceedings, book chapters, Reviews:*


Gaden, G. G. (2012). The Impact of a Year-Long, Same School Social Skills Instruction Program on Students' with Verified Behavioral Disorders, Autism Spectrum Disorders, and Attention Deficit Hyperactivity Disorders Perceptions of Program Effectiveness, ProQuest LLC.


*Comb. of treatment and drug:*


*<10 participants in treatment condition:*


**Not targeted diagnosis:**


*No measure of primary disorder:*


Secondary analyses:


*Not in English language journal*


*Participants not diagnosed:*


*Not testing a form of BI, which are at level 1 or 2 according to the evidence base update review of Smith and Iadarola (2015), Table 6, p. 913.*


McDaniel, J., Yoder, P., Crandall, M., Millan, M. E., Ardel, C. M., Gengoux, G. W. et al. (2020). Effects of Pivotal Response Treatment on Reciprocal Vocal Contingency in a Randomized Controlled


S3. Rating scales used for primary continuous measure.

1. Cognitive

Early Learning Accomplishments Profile (ELAP) (Glover et al., 1988)

Learning Accomplishments Profile (LAP) (Sanford and Zelman, 1981)

Leiter International Performance Scale–Revised (Leiter-R; Roid & Miller, 1997)

Merrill-Palmer Scale of Mental Tests (MPS) (Roid, G.H., & Sampers, J.L., 2004).
Mullen Scales of Early Learning (MSEL; Mullen 1995)
Psychoeducational Profile–Revised (PEP-R) (Schopler et al., 1990)

Stanford-Binet Intelligence Scale Fourth Edition (Thorndike, Hagen, & Sattler, 1986)

Wechsler Preschool and Primary Scale of Intelligence-Revised (Wechsler, 1990)

Wechsler Intelligence Scale for Children. 3rd ed. (Wechsler, 1992)

2. Communication
Assessment of Basic Language and Learning Skills (ABLLS) (Partington & Sundberg, 1998)

Autism Diagnostic Observation Schedule (ADOS) (Lord, Rutter, DiLavore, & Risi, 1999).

Developmental-behavioral scales (Alpern, Boll, & Shearer, 2000)

Early Learning Accomplishments Profile (ELAP) (Glover et al., 1988)

Learning Accomplishments Profile (LAP) (Sanford and Zelman, 1981)

Reynell Developmental Language Scales (Reynell, 1990)
Uniform Performance Assessment System (UPAS) (White, Edgar, Haring, Afflick, & Hayden, 1978)

Vineland Adaptive Behavior Scales (VABS) (Sparrow, Balla, & Cicchetti, 1985; Sparrow, Cicchetti, & Balla, 2005)

3. Socialization
Adaptive Behavior Assessment System-II (ABAS-II; Harrison & Oakland, 2003).

Assessment of Basic Language and Learning Skills (ABLLS) (Partington & Sundberg, 1998)

Autism Diagnostic Observation Schedule (ADOS) (Lord, Rutter, DiLavore, & Risi, 1999).

Early Learning Accomplishments Profile (ELAP) (Glover et al., 1988)

Learning Accomplishments Profile (LAP) (Sanford & Zelman, 1981)

Social Skills Improvement System (SSIS) (Elliott & Gresham, 2007)

Uniform Performance Assessment System (UPAS) (White, Edgar, Haring, Afflick, & Hayden, 1978)

Vineland Adaptive Behavior Scales (VABS) (Sparrow, Balla, & Cicchetti, 1985; Sparrow, Cicchetti, & Balla, 2005)

4. Adaptive
Adaptive Behavior Assessment System-II (ABAS-II; Harrison & Oakland, 2003).

Assessment of Basic Language and Learning Skills (ABLLS) (Partington & Sundberg, 1998)

Vineland Adaptive Behavior Scales (VABS) (Sparrow, Balla, & Cicchetti, 1985; Sparrow, Cicchetti, & Balla, 2005)
References to the instruments


S4. References to efficacy studies used for comparison.


### S5. Table of risk-of-bias evaluation.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson, 1987</td>
<td>NA</td>
<td>H</td>
<td>S</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Ben-Itzchak, 2007</td>
<td>NA</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Ben-Itzchak, 2014</td>
<td>NA</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Bibby, 2002</td>
<td>NA</td>
<td>H</td>
<td>S</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Cohen, 2006</td>
<td>NA</td>
<td>H</td>
<td>S</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Eikeseth, 2012</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Eldevik, 2012</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Eldevik, 2020</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>S</td>
</tr>
<tr>
<td>Goin-Kochel, 2007</td>
<td>NA</td>
<td>H</td>
<td>S</td>
<td>H</td>
<td>S</td>
</tr>
<tr>
<td>Hayward, 2009</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Howard, 2005; 2014</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>S</td>
<td>H</td>
</tr>
<tr>
<td>Janson, 2020</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Luiselli, 2000</td>
<td>NA</td>
<td>H</td>
<td>S</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Magiati, 2007; 2011</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Makrygianni, 2018</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Perry, 2008</td>
<td>NA</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Rad, 2019</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Remington, 2007</td>
<td>NA</td>
<td>H</td>
<td>L</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Rivard, 2014; 2019</td>
<td>NA</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Sallows, 2005</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Schertz, 2013</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Sheinkopf, 1998</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Smith, 2015</td>
<td>NA</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Smith, 2019</td>
<td>NA</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Tonge, 2014</td>
<td>L</td>
<td>H</td>
<td>L</td>
<td>S</td>
<td>L</td>
</tr>
<tr>
<td>Waters, 2020</td>
<td>H</td>
<td>H</td>
<td>S</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>Wood, 2018</td>
<td>NA</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Zachor, 2007</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>L</td>
</tr>
<tr>
<td>Zachor, 2010</td>
<td>H</td>
<td>H</td>
<td>S</td>
<td>H</td>
<td>L</td>
</tr>
</tbody>
</table>

Note: NA = Not applicable, H = high risk, L = low risk, S = some concerns.
**Table of ethnicity.**

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>White/ Caucaustian/ European American</th>
<th>African American, Black British</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson, 1987</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Ben-Itzchak, 2007</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Ben-Itzchak, 2014</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Bibby, 2002</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Cohen, 2006</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Eikeseth, 2012</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Eldevik, 2012</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Eldevik, 2020a</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Eldevik, 2020b</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Goin-Kochel, 2007</td>
<td>29</td>
<td>86.2</td>
<td>10.3</td>
<td>3.4</td>
</tr>
<tr>
<td>Hayward, 2009a</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Hayward, 2009b</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Howard, 2005; 2014</td>
<td>29</td>
<td>72.4</td>
<td></td>
<td>27.6</td>
</tr>
<tr>
<td>Janson, 2020</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Luiselli, 2000</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Magiati, 2007, 2011</td>
<td>28</td>
<td>75.0</td>
<td>4.6*</td>
<td>21.4</td>
</tr>
<tr>
<td>Makrygianni, 2018</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Perry, 2008</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Rad, 2019</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Remington, 2007; 2011</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Rivard, 2014; 2019</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Sallows, 2005a</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Sallows, 2005b</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Schertz, 2013</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Sheinkopf, 1998</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Smith, 2015</td>
<td>71</td>
<td>88.7</td>
<td>7.0</td>
<td>4.2</td>
</tr>
<tr>
<td>Smith, 2019</td>
<td>185</td>
<td>95.0</td>
<td></td>
<td>5.0</td>
</tr>
<tr>
<td>Tonge, 2014</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
<tr>
<td>Waters, 2020</td>
<td></td>
<td></td>
<td></td>
<td>NI</td>
</tr>
</tbody>
</table>
Wood, 2018  
Zachor, 2007  
Zachor, 2010  

Note. NI= No information provided. *= Black British
### Table of focus of treatment.

<table>
<thead>
<tr>
<th>Study</th>
<th>Method</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson, 1987</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Ben-Itzchak, 2007</td>
<td>ABA*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Ben-Itzchak, 2014</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Bibby, 2002</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Cohen, 2006</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Eikeseth, 2012</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Eldevik, 2012</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Eldevik, 2020a</td>
<td>EIBI-lower*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Eldevik, 2020b</td>
<td>EIBI-higher*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Goin-Kochel, 2007</td>
<td>ABA*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Hayward, 2009a</td>
<td>EIBI-clinic*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Hayward, 2009b</td>
<td>EIBI-parent*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Howard, 2005; 2014</td>
<td>IBT*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Janson, 2020</td>
<td>IBT*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Luiselli, 2000</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Magiati, 2007, 2011</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Makrygianni, 2018</td>
<td>ABA*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Perry, 2008</td>
<td>IBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Rad, 2019</td>
<td>ABA*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Remington, 2007; 2011</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Rivard, 2014; 2019</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Sallows, 2005a</td>
<td>EIBI-clinic*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Sallows, 2005b</td>
<td>EIBI-parent*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Schertz, 2013</td>
<td>JAML</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Sheinkopf, 1998</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Smith, 2015</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Smith, 2019</td>
<td>EIBI*</td>
<td>+ + + + + +</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>+</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------</td>
<td>---</td>
</tr>
<tr>
<td>Tonge, 2014</td>
<td>PEBM</td>
<td>+</td>
</tr>
<tr>
<td>Waters, 2020</td>
<td>EIBI*</td>
<td>+</td>
</tr>
<tr>
<td>Wood, 2018</td>
<td>EIBI*</td>
<td>+</td>
</tr>
<tr>
<td>Zachor, 2007</td>
<td>ABA*</td>
<td>+</td>
</tr>
<tr>
<td>Zachor, 2010</td>
<td>ABA*</td>
<td>+</td>
</tr>
</tbody>
</table>

*Note:* + = Indicated as a focus for the intervention; - = No information is provided