Clinical Psychology Review 33 (2013) 940–953

Contents lists available at ScienceDirect

Clinical Psychology Review

The TEACCH program for children and adults with autism: A meta-analysis of intervention studies☆

Javier Virues-Ortega [a,*], Flávia M. Julio [b], Roberto Pastor-Barriuso [c]

[a] St.Amant Research Centre & University of Manitoba, P518 Duff Roblin Bldg, 190 Dysart Rd, R3T 2N2 Winnipeg, MB, Canada
[b] University of Manitoba, P260 Duff Roblin Bldg, 190 Dysart Rd, R3T 2N2 Winnipeg, MB, Canada
[c] Research Network in Epidemiology and Public Health (CIBERESP), National Center for Epidemiology, Carlos III Institute of Health, Sinesio Delgado 4, 28029 Madrid, Spain

HIGHLIGHTS

- TEACCH had small effects on perceptual, motor, verbal, and cognitive skills.
- Effects on communication, ADL, and motor functioning were negligible or small.
- Gains in social and maladaptive behavior were larger but need further replication.
- Our analysis should be considered exploratory owing to the limited data available.

ARTICLE INFO

Article history:
Received 6 July 2012
Revised 28 April 2013
Accepted 16 July 2013
Available online 24 July 2013

Keywords:
TEACCH
Meta-analysis
Outcome research
Autism

ABSTRACT

The intervention program for autism known as Treatment and Education of Autistic and Related Communication Handicapped Children (TEACCH) is considered an emerging practice for autism. In the present study we used state-of-the-art meta-analytical procedures to examine the pooled clinical effects of TEACCH in a variety of outcomes. A total of 13 studies were selected for meta-analysis totaling 172 individuals with autism exposed to TEACCH. Standardized measures of perceptual, motor, adaptive, verbal and cognitive skills were identified as treatment outcomes. We used inverse-variance weighted random effects meta-analysis supplemented with quality assessment, sensitivity analysis, meta-regression, and heterogeneity and publication bias tests. The results suggested that TEACCH effects on perceptual, motor, verbal and cognitive skills were of small magnitude in the meta-analyzed studies. Effects over adaptive behavioral repertoires including communication, activities of daily living, and motor functioning were within the negligible to small range. There were moderate to large gains in social behavior and maladaptive behavior. The effects of the TEACCH program were not moderated by aspects of the intervention such as duration (total weeks), intensity (hours per week), and setting (home-based vs. center-based). While the present meta-analysis provided limited support for the TEACCH program as a comprehensive intervention, our results should be considered exploratory owing to the limited pool of studies available.

Contents

Abbreviations: AA-PEP, Adolescent and Adult Psychoeducational Profile; ASD, Autism spectrum disorders; CI, Confidence interval; *ES*, Effect size; M-P-R, Merrill-Palmer—Revised Scales of Development; PEP-R, Psychoeducational Profile Revised; TEACCH, Treatment and Education of Autistic and Related Communication Handicapped Children; VABS, Vineland Adaptive Behavior Scale.

☆ Authors would like to show their appreciation to the authors of the meta-analyzed studies for providing additional details on their studies. Kazunari Hashimoto helped to screen references published in Japanese. Christine van Loon provided assistance with the English style. This study has not been subject to pre-submission approval by any organization.

* Corresponding author. Tel.: +1 204 898 2495.
E-mail addresses: javier.virues@ad.umanitoba.ca (J. Virues-Ortega), juliof@cc.umanitoba.ca (F.M. Julio), rpastor@isciii.es (R. Pastor-Barriuso).

http://dx.doi.org/10.1016/j.cpr.2013.07.005

1. Introduction

Autism spectrum disorders (ASD) are pervasive developmental disabilities that affect 1 to 2.5% of children in the general population (Baio & Baio, 2012; Kim et al., 2011). Various approaches to treatment attempt to palliate the symptoms of autism and to establish more typical academic, daily living, social, and verbal repertoires. Over the last decades a myriad of psychosocial and educational interventions have been developed. For instance, the National Institutes of Health and Autism Speaks®, the largest funders of autism research, have provided support to Treatment and Education of Autistic and Related Communication Handicapped Children (TEACCH), animal-assisted therapy, floortime, sensory integration therapy, Early Start Denver Model, Picture Exchange Communication System, Relationship Development Intervention, parent-training models, and applied behavior analysis, to mention only a few approaches to treatment (Autism Speaks, 2012; National Institutes of Health, 2011). These interventions vary in their conceptual basis, procedures, targeted repertoires and skills, and in the quality, magnitude and relevance of the empirical evidence they have accrued over the years.

Among these various approaches, only a few have generated a sufficient number of clinical trials to allow for meta-analyses and systematic reviews. For instance, four meta-analyses have been published on interventions based on applied behavior analysis (Eldevik et al., 2009; Reichow, 2011; Reichow & Wolery, 2009; Virués-Ortega, 2010), while systematic reviews are available for Picture Exchange Communication System (Fillipin, Reska, & Watson, 2010; Preston & Carter, 2009; Tien, 2008) and sensory integration therapy (Sinha, Silove, Williams, & Hayen, 2004).

An intervention paradigm that has generated significant empirical research and service demand is the TEACCH program (Mesibov & Shea, 2010). A multi-national survey completed by parents of children with autism indicated that over 30% of families currently use or had used the TEACCH program (Green et al., 2006). Similarly, an epidemiological survey conducted in Finland indicated that over 40% of children with autism were receiving services based on TEACCH (Kielinen, Linna, & Moilanen, 2002). In addition, TEACCH is used frequently as a specialized education service in schools (Hess, Morrier, Heflin, & Ivey, 2008).

The TEACCH program (a) emphasizes a close working relationship between parents and practitioners, (b) adapts the intervention to the particular characteristics of the individual client, and (c) makes use of structured teaching experiences (Van Bourgondien & Schopler, 1996). In a typical TEACCH intervention, the individuals' abilities are assessed through standardized tests (e.g., Psychoeducational Profile). The results of the assessment will provide the basis for the development of a curriculum that will be consistent with the individual needs of the client (Mesibov, 1997). The TEACCH specialist would use structured teaching procedures to facilitate the acquisition of the learning goals composing the individual's curriculum. The structured teaching component requires the individual's environment and activities to be organized in ways that would optimize learning and avoid frustration. Three factors are reportedly essential in this connection: (a) organization of the physical environment in a way that is consistent with the needs of the child (e.g., minimizing possible distractions), (b) arrangement of activities in a predictable fashion (e.g., use of visual schedules of daily routines), and (c) organization of the materials and tasks to promote independence from adult directions/prompts (e.g., use visual materials if the student is more able to benefit from them).

The studies that have been conducted on TEACCH vary in the set of TEACCH components utilized, the intensity and duration of the intervention, the participants' characteristics (age, diagnosis, pre-intervention functioning), the assessment methods used, and the study design (pre–post, between-group) (e.g., Eikeseth, 2009; Ospina et al., 2008; Simpson et al., 2005). Such variability precludes a straightforward appreciation of treatment effectiveness. Although syntheses do exist in the literature, they do not meet the standards of a systematic review or a meta-analysis (e.g., Probst, Jung, Micheel, & Glen, 2010; Simpson et al., 2005) and they do not use standard methods to extract and combine data (Moher, Liberati, Tetzlaff, & Altman, 2009). This scenario further increases the complexity of appraising the magnitude and quality of the empirical evidence associated with the TEACCH program. A more clear appreciation of the evidence in support of the TEACCH program could be achieved by way of meta-analysis. Only recently, a sufficient number of studies have been published following a relatively consistent methodology to allow for a meta-analysis with sufficient power. The TEACCH program is among the few nonpharmacological approaches to meet this threshold (see another exception in Virués-Ortega, 2010).

Meta-analysis methodology based on randomized controlled trials and standardized assessments has been recognized as a valuable albeit limited approach for program evaluation. For example, Kazdin (2008) expanded appraisal of evidence-based practices includes mechanisms of change, moderators of change, and qualitative research, which cannot be fully explored with meta-analysis or other traditional approaches to clinical research (see also Mesibov & Shea, 2010). However, several of the guiding principles of the TEACCH program emphasize particular targets that are potentially accountable through standardized assessment and therefore by traditional methods of clinical research. Particularly, as part of the principles of *improved adaptation*, *structured teaching*, and *skill enhancement*, TEACCH targets key skills in activities of daily living, communication, language, social skills, executive functioning, attention, and engagement (Schopler, 2005). Skill acquisition across all these domains is expected to have a comprehensive effect of the individual's intellectual functioning, which could in turn be assessed through measures of developmental age and intellectual functioning. Also, improved adaptation, by way of skill acquisition and environmental accommodation, is expected to reduce problem behavior. The TEACCH program is defined as an approach that supports the "individual's ability to learn, comprehend, and apply learning across situations" (Mesibov, Shea, & Schopler, 2005, p. 216). Therefore, transversal skills expanding independent learning and generalization are a critical outcome of the program. Skills within this broad category may include imitation, as well as verbal and social skills. Finally, TEACCH takes advantage of the individual's relative strength in visual processing by adapting the physical environment and using visual structures for organizing space and activities. These adaptations may have a favorable impact on the individual's performance in perceptual tasks and in activities requiring visual and motor coordination.

The purpose of the present study was to provide a preliminary and comprehensive summary of the evidence in support of the TEACCH program using current meta-analytical methods. The specific purposes of this study were: (a) to conduct a meta-analysis of studies evaluating the TEACCH program effect over a variety of standardized outcomes including perceptual and motor skills, activities of daily living, behavioral adaptive skills, cognition, and language, and (b) to identify specific characteristics of the sample, the intervention and the study methodology that could be reliably associated with increased intervention effectiveness. This information may be valuable to clinicians and administrators for the purposes of treatment planning and resource allocation.

2. Methods

2.1. Literature search and study selection

The intervention studies analyzed in the present meta-analysis were selected from unrestricted searches conducted in MEDLINE (PubMed search engine), PsycINFO (ProQuest search engine), and the Cochrane Central Register of Controlled Trials (Cochrane Library, 2012) using "TEACCH" or "Treatment and Education for Children," as search terms (search date: December 1, 2012). Related search terms (e.g., structured teaching, structured instruction, structured training) that did not result in increased sensitivity to studies appropriate for meta-analysis were not included in our final search strategy. More details on the search strategy are available in the online supplementary material (Appendix A). The search was not restricted by language or time of publication. References were screened based on the following set of inclusion/exclusion criteria: (a) published/in press peer-reviewed or non-peer reviewed journal article reporting an empirical intervention (unpublished studies, dissertations, books, editorials, reviews, case reports, and essays were excluded), (b) the study reports a TEACCH intervention trial for individuals with ASD, and (c) the intervention group of the study is composed of five individuals or more. Criterion (a) was screened on the basis of the title and abstract of the study. All other criteria were screened after examining the full report of the study. During the screening of references we consulted with specialized translators regarding studies published in languages other than English, Spanish, Portuguese, and French (spoken by the authors). Among the studies potentially appropriate for meta-analysis, we excluded those that failed to report pre- and post-test mean and variability estimates after one attempt was made to contact the authors. Studies reporting outcomes that were not present in at least two other studies were used to compute mean effect sizes of the intervention across all studies, but isolated outcomes were not reported individually. Two judges with graduate training on outcome research screened references independently. The overall percentage of agreement on the exclusion/inclusion status of references screened was 95%. Disagreements were settled by consensus. The current study is in compliance with the Meta-Analysis Reporting Standards (MARS; American Psychological Association, 2009, pp. 251–252) and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (PRISMA guidelines; Moher et al., 2009).

2.2. Data extraction

We retrieved the following information from the studies included in the meta-analysis: participant information (mean pre-intervention age, mean pre-intervention developmental/mental age, percentage of male participants), aspects of the intervention (intensity in weekly hours, duration in weeks, intervention setting), methodological features of the study (study design, sample size, assessment instruments, characteristics of the control group), and the information required to compute effect sizes (pre- and post-test means and variability estimates of all relevant outcomes).

Methodological quality of each of the studies selected for meta-analysis was assessed by means of an adapted version of the Downs and Black checklist of methodological quality of randomized and non-randomized studies of health care interventions (Downs & Black, 1998, see online supplementary material, Appendix B). Two individuals with graduate training in outcome research methodology screened the studies independently. Disagreements were settled by consensus. Items on randomization (items 23, 24) were considered non-applicable for pre–post studies.

2.3. Statistical analysis

We used effect sizes as standardized outcomes of the effect of TEACCH on specific dependent variables. We incorporated studies with metrics derived from between-group and pre–post designs. Effect sizes of between-group studies were computed as the difference in outcome change overtime (post- minus pre-test mean scores) between the intervention and control groups, divided by the pre-test standard deviation pooled across intervention and control groups. Intervention group was composed of all individuals in the study undergoing TEACCH, while individuals in the control group were those not receiving TEACCH. Effect sizes of interventions reported as pre–post studies were obtained by dividing the mean difference between post- and pre-test measures by the pre-test standard deviation. For the purposes of interpreting our results we considered effect sizes below 0.44 as small. Moderate effects ranged from 0.45 to 0.79. Finally, effect sizes above 0.80 were considered large effects (Cohen, 1988).

The combination of studies with different metrics (pre–post and between-group studies) in the present meta-analysis, rests on the assumption that changes at post-test are the result of treatment. According to this assumption the effect sizes from pre–post studies could be interpreted as the effect of the intervention on the dependent variable established in standard deviation units of the dependent variable before the intervention (Morris & DeShon, 2002). The small number of pre–post studies found, relative to the between-group studies, allowed us to conduct sensitivity analyses for most outcome measures by restricting the meta-analysis to between-group studies. Sensitivity analyses allowed us to test the consistency of the meta-analysis across study metrics (e.g., Chootrakool, Shi, & Yue, 2011). Although only two pre–post studies implemented interventions of over 12 months in duration (Persson, 2000; Siaperas, Higgins, & Proios, 2007), we conducted sensitivity analyses to account for the potential bias induced by pre–post designs, which are insensitive to developmental or other time-dependent factors.

In order to obtain effect sizes we retrieved the mean and standard deviations directly from the research reports or through fully reported datasets of all individuals exposed to the intervention. We computed measure-specific effect size variances corrected for small-sample bias (Becker, 1988; Morris, 2008). Pretest–posttest correlations are necessary to compute effect size variances. We requested the authors of all the studies selected for meta-analysis to provide us with their original datasets in order to compute pretest–posttest correlations. Usable datasets were provided for Panerai, Ferrante, and Zingale (2002) and Panerai et al. (2009), which was sufficient to obtain pooled pretest–posttest correlations of all outcomes (Morris & DeShon, 2002). The pooled value for a given outcome was used to obtain the variance of all effect sizes of studies reporting that outcome.

We calculated pooled effect sizes and 95% confidence intervals for all dependent variables by means of inverse-variance weighted random effects meta-analysis (Cottrell, Drew, Gibson, Holroyd, & O'Donnell, 2007). The random effects model was selected on the basis of the relatively variable sample age, design, and treatment duration found in this literature (Table 1).

Table 1
Characteristics of the studies included in the meta-analysis of treatment outcomes of the TEACCH program.

Study	Country	Participants				Methods			Intervention		
		Diagnosis	Male (%)	Mean age (years)	Pre-test mental age	Sample size	Assessment methods	Quality index	Setting[a]	Intensity (h/week)	Duration (weeks)
Aoyama (1995)	Japan	Autism & ID	100	School age	41 (IQ)	5	BO	44	Center	–	1
Braiden et al. (2012)[b]	Ireland	Autism	94	3.2	_	18	PEP-3	50	Home-parent	4.5	10
McConkey et al. (2010)[c]	Ireland	Autism	90	2.8	20.1	35/26	PEP-R; VABS; GARS	70	Home	1.5	10
Ozonoff and Cathcart (1998)	US	Autism	82	4.4	21.4	11/11	PEP-R	60	Home-parent	1.0	4
Panerai et al. (2002)[d]	Italy	Autism & ID	100	9.1	17.9	13/10	PEP-R; VABS	73	Center	30.0	12
Panerai et al. (2009)[e]	Italy	Autism & ID	100	8.7	23.7	11/10	PEP-R; VABS	63	School, home-parent	–	36
Panerai et al. (2009)[e]	Italy	Autism & ID	100	9.7	20.6	8/8	PEP-R; VABS	63	Center	12.5	36
Persson (2000)	Sweden	Autism & ID	100	32.3	–	7	AA-PEP	53	Group home	–	30
Probst and Leppert (2008)	Germany	Autism	70	10.0	–	10	CCBSS	50	School	–	9
Siaperas and Beadle-Brown (2006)	Greece	Autism	67	21.3	–	10	BO	63	Center	–	6
Siaperas et al. (2007)	Greece	Autism	70	21.3	–	10	BO	63	Center	–	18
Tsang et al. (2007)	China	Autism & PDD NOS	94	4.6	–	18/16	C-PEP; HK-VABS; M-P-R	73	Center	35.0	12
Van Bourgondien et al. (2003)[f]	US	Autism & ID	100	23.7	28.1	6	ERS	47	Center	–	52
Welterlin et al. (2012)	US	Autism	90	2.5	57.1	10/10	MSEL, SIB-R	77	Home-parent	_	14

Notes. Sample size reported as total subjects for pre–post designs, or number of subjects in intervention/control groups for between-group studies. Quality computed as percent of total score of Downs and Black (1998) quality checklist. AA-PEP = Adolescent and Adult Psychoeducational Profile; BO = behavioral observation; CCBSS = Classroom Child Behavioral Symptom Scale; C-PEP = Chinese Psychoeducational Profile; ERS = Environmental Rating Scale; GARS = Gilliam Autism Rating Scale; HK-VABS = Hong Kong Based VABS; ID = comorbid intellectual disability; M-P-R = Merrill-Palmer—Revised; MSEL = Mullen Scales of Early Learning; PDD NOS = pervasive developmental disability not otherwise specified; PEP-R = Psychoeducational Profile Revised; SIB-R = Scales of Independent Behavior—Revised; VABS = Vineland Adaptive Behavior Scales.

[a] Home: delivered at home by specialized staff with parental support; home-parent: delivered at home by trained parents.

[b] Intensity computed as weekly hours of training that parents received.

[c] VABS outcomes reported in both control and experimental groups; PEP-R outcomes reported only in the experimental group ($n = 32$).

[d] Intensity according to S. Panerai, personal communication, December 17, 2011.

[e] This study was analyzed as two separate between-group studies with different treatment settings (school plus home vs. center-based).

[f] This study was analyzed as a within-subjects trial (pre-test estimates of the control groups were not reported).

Due to the limited number of studies, and the marginal effect of age on treatment outcomes (according to preliminary sensitivity analysis) we obtained pooled effects across all ages. Analyses by common age groupings strengthen the potential practicality of meta-analysis findings owing to the diversity of needs of individuals with developmental disabilities across the life span. Therefore, we supplemented the omnibus analysis with a thorough sensitivity analysis of all outcomes across three age groups (5 years or less, 6 to 17 years, and 18 years or older). Studies were assigned to each age group on the basis of participants' average age. This analysis provided a detailed account of the potential variations of treatment effects across common age groups.

We computed the I^2 statistic expressed as a percentage to inform the variation of outcomes across studies independent of the number of studies. I^2 can be interpreted as the extent to which outcome variability is inconsistent across studies, whether as the result of chance or true differences (heterogeneity) (Higgins & Thompson, 2002). I^2 values below 25% suggest a low level of heterogeneity, while values above 75% would suggest high levels of heterogeneity (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006).

We conducted a random effects meta-regression of all outcomes using the following continuous moderators: gender (percentage of male participants), developmental/mental age as established by a valid standardized assessment, intervention intensity (weekly hours), intervention duration (total weeks), and study quality (0–100 quality index; Downs & Black, 1998). We also tested intervention setting (home-based vs. center-based) as a categorical moderator. For the purposes of the meta-regression, interventions delivered partially or totally at a school, center, residence, or group home were considered center-based (e.g., the group reported by Panerai et al. (2009), receiving parental training and school support was classified as home-based). Interventions delivered exclusively in a home-setting delivered either by trained parents (e.g., Braiden, McDaniel, McCrudden, Hanes, & Crozier, 2012) or trained professionals (McConkey et al., 2010) were considered home-based. Meta-regression analyses were accompanied by additional sensitivity analyses, which were conducted by restricting random effect meta-analysis to two levels of the predictor (for continuous variables we used the median value of the predictor as the cut-off point). For instance, we conducted two meta-analyses of cognitive functioning restricted to interventions of less than three months in duration (median intervention duration). Subsequently, we replicated the analysis with interventions of more than three months in duration. Meta-regression would indicate whether the continuous predictor was statistically significant. Publication bias and statistical outliers (small-study effects) were monitored by means of the Egger's test (Egger, Smith, Schneider, & Minder, 1997; Thompson & Sharp, 1999).

In order to avoid an excessive loss of information due to the selection of specific outcomes within a study, we computed an overall mean effect size per study based on all psychologically and behaviorally relevant measures included in each study. Outcomes not directly assessing the participants' psychological or behavioral functioning were excluded (e.g., parental stress). The sign of effect sizes was reversed if a negative effect denoted a favorable outcome (e.g., problem behavior). We then conducted unweighted and weighted random effects meta-analyses of the mean effect sizes across all studies and outcomes (Shadish & Haddock, 2009). For this analysis we computed a mean effect size for each study appropriate for meta-analysis. The mean effect size incorporated all the outcomes, isolated or otherwise, reported in each study.

According to a preliminary review of study abstracts, we expected to identify approximately five studies per outcome meeting inclusion criteria with a typical sample size of about 30 participants (15 participants in each intervention arm). Based on this preliminary information and the methods proposed by Hedges and Pigott (2001), we conducted a prospective power analysis of our random-effects meta-analysis under different plausible values for the underlying mean effect size and the between-study heterogeneity. The power to detect as statistically significant a moderate mean effect size of 0.50 was 0.98 for a small heterogeneity of $I^2 = 25\%$, 0.90 for a moderate heterogeneity of $I^2 = 50\%$, and 0.63 for a large heterogeneity of $I^2 = 75\%$. Similarly, the power to detect a small mean effect size of 0.30 was 0.69 for a small heterogeneity of $I^2 = 25\%$, 0.52 for a moderate heterogeneity of $I^2 = 50\%$, and 0.29 for a large heterogeneity of $I^2 = 75\%$. Thus, combining five studies with a typical sample size of 30, we anticipated reasonable power for a moderate effect size with any amount of heterogeneity and for a small effect size with small heterogeneity, but little power for a small effect size with moderate or large heterogeneity (see full report of the power analysis in the online supplementary material, Appendix C). A p value of 0.05 was used throughout as a threshold for statistical significance. All statistical analyses were conducted with Stata v. 11 (Stata Corporation, College Station, Texas).

3. Results

3.1. Study characteristics

From the 220 distinct studies originally identified through our search strategy, 97 were not peer-reviewed empirical studies. An additional 103 studies were excluded due to not reporting a TEACCH intervention trial, or having an intervention group composed of less than five individuals. Of the remaining 22 trials potentially appropriate for meta-analysis, eight did not provide sufficient data to compute effect sizes (Deprez, 2004; Keel, Mesibov, & Woods, 1997; Ono, 1994; Panerai, Ferrante, & Caputo, 1997; Rúa, 2004; Schopler, Mesibov, & Baker, 1982; Schultheis, Boswell, & Decker, 2000; Short, 1984), and one was an earlier version of a study that was included in the meta-analysis (Leppert & Probst, 2005). Thirteen studies were included in the meta-analysis (between-group studies: Braiden et al., 2012; Persson, 2000; Probst & Leppert, 2008; Siaperas & Beadle-Brown, 2006; Siaperas et al., 2007; Van Bourgondien, Reichle, & Schopler, 2003; pre–post studies: Aoyama, 1995; McConkey et al., 2010; Ozonoff & Cathcart, 1998; Panerai et al., 2002, 2009; Tsang, Shek, Lam, Tang, & Cheung, 2007; Welterlin, Turner-Brown, Harris, Mesibov, & Delmolino, 2012). The study by Panerai et al. (2009) was analyzed as two separate between-group studies with different treatment settings (home vs. center). Van Bourgondien et al. (2003) was included as a pre–post study because the pre-test measurements of the control groups were not available. Fig. 1 summarizes the search and selection processes.

None of the individuals in the control groups received TEACCH for the duration of each of the studies. Control groups attended mainstream schools with special education support (Panerai et al., 2002, 2009), or without it (Tsang et al., 2007), received some form of specialized eclectic treatment for autism (Ozonoff & Cathcart, 1998), or underwent some form of specialized treatment including physical or speech therapy (McConkey et al., 2010; Panerai et al., 2002, 2009; Tsang et al., 2007). The control group reported by McConkey et al. (2010) received a placebo intervention consisting of home visits (unstructured presentation of toys and instructions).

The specific outcomes that were part of the meta-analysis were (number of studies in parenthesis): activities of daily living (ADL) (6), cognitive functioning (5), communication skills (5), developmental/mental age (5), language/verbal skills (9), eye-hand coordination (6), motor functioning (4), fine motor skills (6), gross motor skills (6), imitation (6), social repertoire (7), perception (6), maladaptive behavior (4), PEP-R total (6), and VABS adaptation composite (4). In addition, 31 additional isolated outcomes reported only in one or two studies, were included in the mean effect size meta-analysis. Aoyama (1995) and Van Bourgondien et al. (2003) reported only isolated outcomes that were not matched by the outcomes reported in any of the studies included in the meta-analysis. In addition, five studies reported both isolated outcomes and outcomes that were present in at least two other studies (McConkey et al., 2010; Persson, 2000; Siaperas & Beadle-Brown, 2006;

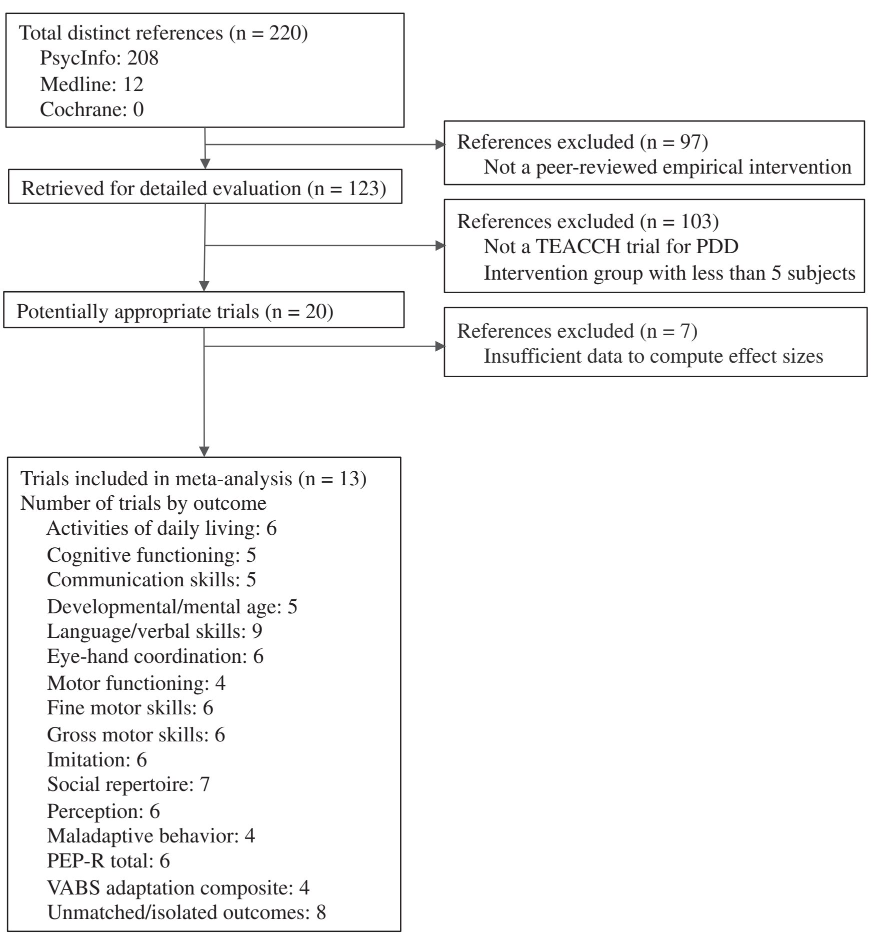

Fig. 1. Search and selection of references. ASD = autism spectrum disorders; PEP-R = Psychoeducational Profile—Revised; VABS = Vineland Adaptive Behavioral Scales.

Siaperas et al., 2007; Tsang et al., 2007; Welterlin et al., 2012; see a complete listing of isolated outcomes in the online supplementary material, Appendix D). The outcomes included in the meta-analysis allow for an ample evaluation of the perceptual, motor, social, and cognitive effects of TEACCH as discussed in the Introduction section.

The standardized assessments included the Vineland Adaptive Behavior Scale (VABS, Sparrow, Cicchetti, & Balla, 2005) for the outcomes: activities of daily living, motor functioning, social repertoire, maladaptive behavior, and VABS summary index. The Psychoeducational Profile scales (AA-PEP, Mesibov, Schopler, Schaffer, & Landrus, 1988; PEP-R, Schopler, 1990; PEP3, Fu, Chen, Tseng, Chiang, & Hsieh, 2012) informed perception, cognitive functioning, communication skills, language skills, eye–hand coordination, fine motor skills, gross motor skills, imitation, developmental age, and PEP summary index. In addition, developmental age was estimated by the mental age of the Merrill-Palmer—Revised Scales of Development (M-P-R; Roid & Sampers, 2004) and the developmental quotient of the Mullen Scales of Early Learning (MSEL; Mullen, 1995). Problem behavior was also assessed through the Classroom Child Behavioral Symptom Scale (CCBSS; Probst & Leppert, 2008) and through direct behavioral observation. The assessment methods used in each study are indicated in Table 1. Appendix D (online supplementary material) presents further details about the assessment methods used in studies with isolated outcomes.

The VABS is a measure of personal and social skills for individuals of all ages. It is frequently used with individuals with intellectual disability and ASD. The test examines four major domains: communication (receptive, expressive, written); daily living skills (personal, domestic, community); socialization (interpersonal relationships, play and leisure time, coping skills); and motor skills. The VABS includes an optional index of maladaptive behavior. On the other hand, the PEP-R is an inventory of behaviors and skills designed to identify idiosyncratic learning patterns among children with ASD. It is designed primarily for planning individualized intervention and curricula. It provides information on developmental functioning in imitation, perception, fine motor, gross motor, eye–hand coordination, cognitive performance, and cognitive verbal areas. The PEP-R assesses children aged six months to seven years functioning at or below the preschool range. It can also be used for children of up to twelve years if they are still functioning at preschool levels. Persson (2000) used a variation of PEP for adolescents and adults. Both VABS and PEP-R have well-established reliability and validity (Fu et al., 2012; Perry & Factor, 1989; Villa et al., 2010). The PEP assessments include items on a range of behavioral repertoires that are highly relevant to the core features of autism (imitation, social behavior, motor functioning, communication, and language), but it is not considered a measurement of the core symptoms of ASD and tends to be used more often within TEACCH circles. We refer the reader to the original psychometric literature on these assessment tools for a detailed operational definition of the constructs used as outcomes.

The average methodological quality index across studies (Downs & Black, 1998) was 60% (range 44–77). Results by methodological quality domains indicated that major concerns had to do with external validity (46%, range 30–100), and statistical power (15%, range 0–60), while reporting and internal validity, both in terms of bias and confusion, obtained average quality indices above 65%. None of the studies selected for meta-analysis concealed intervention assignment from staff (item 24), and only one reported treatment fidelity measures (item 19) (Welterlin et al., 2012). However, the individuals delivering the intervention were in most cases trained by the original TEACCH center in North Carolina (Braiden et al., 2012; McConkey et al., 2010; Ozonoff & Cathcart, 1998; Persson, 2000; Probst & Leppert, 2008; Tsang et al., 2007; Van Bourgondien et al., 2003; Welterlin et al., 2012) or were trained indirectly through individuals trained by the original developers (Panerai et al., 2002, 2009; Siaperas & Beadle-Brown, 2006; Siaperas et al., 2007) (correspondence with Dr. Panerai, Dr. Siaperas, Dr. McConkey, and Dr. Tsang is available upon request). Only two studies concealed intervention assignment from the individuals measuring the intervention outcomes (item 15) (Panerai et al., 2002; Welterlin et al., 2012). Three studies made an effort to report potential adverse events associated with the intervention (item 8) (McConkey et al., 2010; Panerai et al., 2002; Welterlin et al., 2012). Two studies used randomization (item 23) and intention to treat analysis (item 25) (Panerai et al., 2002; Tsang et al., 2007). Finally, only three of the studies provided sufficient information on the recruitment process to inform the representativeness of their study samples (items 11, 12) (Siaperas & Beadle-Brown, 2006; Siaperas et al., 2007; Tsang et al., 2007). The quality assessment is fully reported in the online supplementary material (Appendix B). While the quality of the studies selected was suboptimal, quality levels were, on average, superior to those of other intervention studies in ASD (see for instance Virués-Ortega, 2010).

The pooled sample of individuals receiving TEACCH across studies was 172. Two studies included partially overlapping samples but the outcomes reported in these studies did not overlap (Siaperas & Beadle-Brown, 2006; Siaperas et al., 2007). During the meta-analysis and meta-regression of mean effect sizes by study, the two studies with partially overlapping samples were combined into a single mean effect size. Therefore, the same individuals were never included more than once in a single meta-analysis (correspondence with the original authors is available upon request). Five studies evaluated the intervention effects among young children (average age range: 2.5 to 4.6), five studies assessed school-age children (average age range: 8.7 to 10), and four studies evaluated the intervention effects among young adults (average age range: 21.3 to 32.3). As it is common in ASD literature, the majority of participants were males (range of male participants: 67% to 100%). In six studies the intervention was delivered through a center-based TEACCH program (Panerai et al., 2002, 2009; Siaperas & Beadle-Brown, 2006; Siaperas et al., 2007; Tsang et al., 2007; Van Bourgondien et al., 2003). In three studies the intervention was delivered by trained parents in their homes (Braiden et al., 2012; Ozonoff & Cathcart, 1998; Welterlin et al., 2012). In two additional cases, parental intervention was supplemented with either specialized staff working at home (McConkey et al., 2010) or with a support teacher at school (Panerai et al., 2009). Probst and Leppert (2008) relied solely on trained teachers at school. Persson (2000) relied on trained staff at a group home for adults with autism. Finally, Aoyama (1995) used a short-duration intervention in a room for pre-vocational and work-related activities. The hours of intervention varied from 1.5 to 30 every week. However, weekly hours of intervention were not always reported. Intervention duration varied from 1 to 36 weeks. A systematic summary of the characteristics of the selected studies is available in Table 1.

Fig. 2 presents the mean effect size meta-analysis of the TEACCH program across all studies and outcomes, including the outcomes that were included only in one or two studies (unmatched outcomes). The pooled effect size across studies was 0.47 (95% confidence interval [CI] 0.30 to 0.70, $p < .001$). The effect was somewhat lower when the analysis was limited to the controlled studies (effect size [ES] = 0.38, 95% CI 0.12 to 0.65, $p = .005$). Heterogeneity was moderate ($I^2 = 64\%$). The Egger's test provided no evidence of publication or small-study bias ($p > .1$). According to the sensitivity analysis by age group, effects among young children were small and non-significant ($ES = 0.29$), while school-age children and adults demonstrated moderate ($ES = 0.53$) and large gains ($ES = 0.81$), respectively (Table 3).

3.2. Perceptual and motor skills

Six studies incorporating a total of 93 individuals receiving TEACCH reported the following perceptual and motor outcomes: eye–hand coordination, fine and gross motor functioning, imitation, and perception. These outcomes were assessed by means of the PEP-R or a variation of this instrument. Five of these studies used between-group designs, which included 61 participants receiving the intervention. Fig. 3 presents the forest plots for these outcomes.

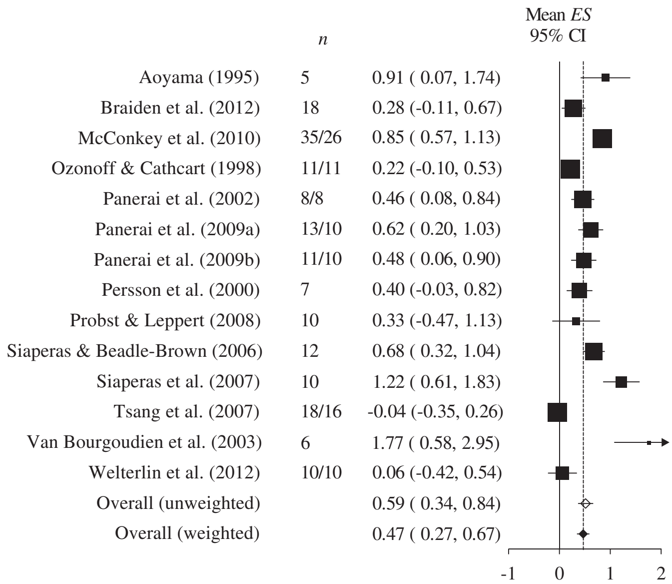

Fig. 2. Forest plots of mean effect sizes and 95% confidence intervals across all outcomes and studies. Square size is proportional to the weight of each study in the pooled analysis. The sample size (n) is indicated as the total number of subjects for pre–post studies and number of subjects in the intervention/control groups for between-group studies. The open diamond portrays the overall estimate of the unweighted meta-analysis. The solid diamond portrays the overall estimate of the inverse-variance weighted random-effects meta-analyses.

The pooled effect size across studies for eye–hand coordination was 0.26 (95% confidence interval [CI] −0.03 to 0.55, $p = .084$). The effect did not change when the analysis was limited to the controlled studies ($ES = 0.23$, 95% CI −0.15 to 0.60, $p > .1$). Heterogeneity was moderate ($I^2 = 74\%$). The Egger's test provided no evidence of publication or small-study bias ($p > .1$).

Similarly, the effect size for fine motor functioning was 0.36 (95% CI 0.08 to 0.65, $p = .012$). A somewhat smaller effect was established in the sensitivity analysis ($ES = 0.23$, 95% CI −0.06 to 0.40, $p = .007$). Heterogeneity was high ($I^2 = 85\%$). No evidence of publication bias was established ($p > .1$).

The pooled effect size for gross motor function was 0.58 (95% CI 0.25 to 0.91, $p = .001$) and 0.41 (95% CI 0.21 to 0.61, $p < .001$) when limiting the analysis to studies that included a control group. Heterogeneity across studies was moderate ($I^2 = 71\%$) and no publication bias or small-study effects were established ($p > .1$).

The pooled effect size for imitation was 0.41 (95% CI 0.12 to 0.70, $p = .005$) and 0.29 (95% CI 0.06 to 0.52, $p = .015$) according to the sensitivity analysis. Heterogeneity across studies was acceptable ($I^2 = 65\%$) and there was no evidence of publication bias ($p = .192$).

Perception showed a pooled effect size of 0.40 (95% CI 0.17 to 0.63, $p = 0.01$). Results were comparable when meta-analysis was restricted to the studies with a control group ($ES = 0.30$, 95% CI 0.12 to 0.49, $p = .002$). Heterogeneity was moderate ($I^2 = 53\%$,). Egger's test suggested the potential for publication bias or small-study effects ($p = .035$).

The meta-regression of the outcomes within this domain using setting, developmental age, study quality, intervention intensity and duration as potential moderators was all non-significant with one exception: higher developmental age was associated with more modest gains in eye–hand coordination (coefficient $\phi = -0.09$, 95% CI −0.15 to −0.28, $p = .016$). Sensitivity analyses by age group resulted in effect sizes of similar magnitude across young and school-age children, with the exception of eye–hand coordination, for which there was no effect among the young children (Table 3).

3.3. Adaptive behaviors

Communication assessed by means of the VABS was reported in five studies ($n = 74$) resulting in a pooled effect size of 0.34 (95% CI −0.13 to 0.82, $p > .1$) (Fig. 4). The sensitivity analysis limited to four controlled studies ($n = 64$) offered an almost identical result ($ES = 0.35$, 95% CI −0.29 to 0.99, $p > .1$). Heterogeneity across studies was moderate ($I^2 = 73\%$, 95%) and there was no evidence of publication bias or small-study effects ($p > .1$).

Six studies ($n = 81$) assessed activities of daily living by means of the VABS and the Adolescent and Adult Psychoeducational Profile (AA-PEP). The pooled effect size was 0.32 (95% CI −0.09 to 0.73, $p > .1$) (Fig. 4). We replicated the meta-analysis including only the four controlled studies available ($n = 64$) and the results did not change ($ES = 0.32$, 95% CI −0.33 to 0.98, $p > .1$). There was evidence of a relatively high heterogeneity across studies ($I^2 = 78\%$). The Egger's test did not suggest publication bias or small-study effects ($p > .1$).

The pooled effect size of motor functioning as established by the VABS was 0.34 (95% CI 0.001 to 0.65, $p < .05$) (Fig. 4). Meta-analysis limited to the four studies that included a control group ($n = 76$) resulted in a similar effect size ($ES = 0.33$, 95% CI 0.001 to 0.66, $p < .05$). There was no evidence of heterogeneity across studies ($I^2 = 0$) or publication bias ($p > .1$).

Social functioning assessed by VABS, AA-PEP, and MSEL was reported in seven studies ($n = 91$). The pooled effect size was 0.65 (95% CI 0.15 to 1.15, $p = 0.011$) (Fig. 4). When the analysis was limited to the five controlled studies available ($n = 74$), the pooled effect size was 0.64 (95% CI −0.06 to 1.33, $p = 0.072$). Heterogeneity was high ($I^2 = 78\%$ CI 55 to 89%), while the Egger's test was not significant ($p > .1$). The meta-regression analysis revealed that social functioning was moderated by study quality (coefficient $\phi = -0.13 \pm 0.44$, $p = 0.032$). Namely, the stronger the methodological quality of the study the smaller the effect of TEACCH over social functioning (Table 4).

Problem and maladaptive behaviors were reported in four studies ($n = 44$) implementing various assessment methods (Table 1). Pooled effect size was −0.92 (95% CI −1.51 to −0.33, $p < .001$) (Fig. 4). Similar results were obtained when the analysis was restricted to the two studies that included a control group ($ES = 1.04$, 95% CI −1.81 to −0.27,

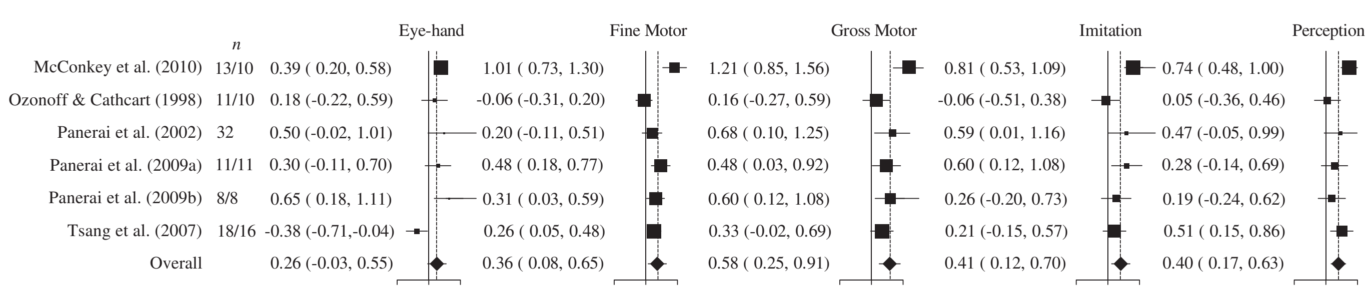

Fig. 3. Forest plots of effect sizes and 95% confidence intervals of perceptual and motor outcomes of the TEACCH program in individuals with autism.

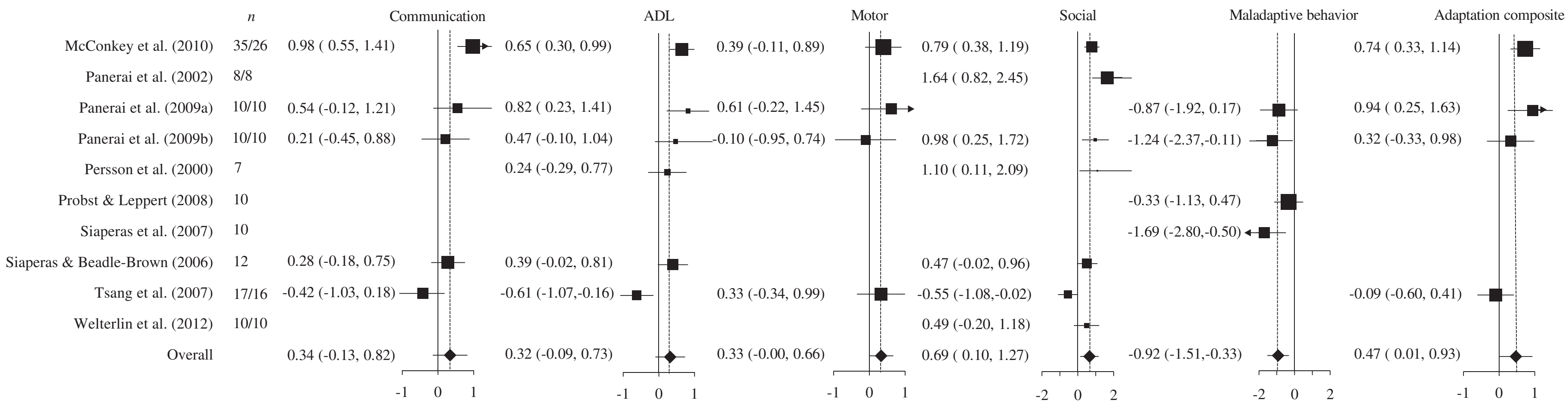

Fig. 4. Forest plots of effect sizes and 95% confidence intervals of adaptive outcomes of the TEACCH program in individuals with autism including communication skills, activities of daily living (ADL), motor and social functioning, presence of maladaptive or problem behavior, and a composite adaptation score.

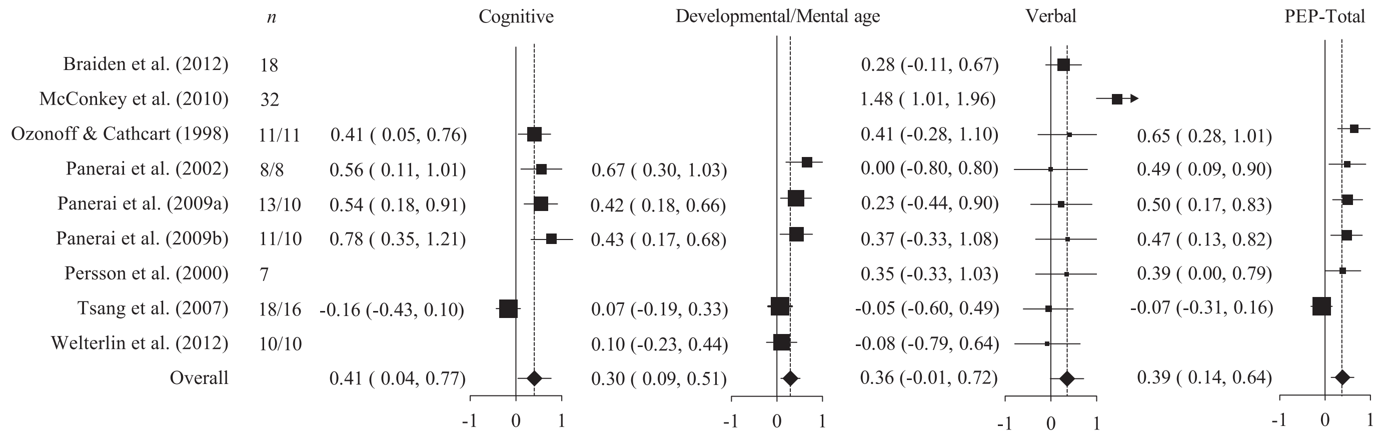

Fig. 5. Forest plots of effect sizes and 95% confidence intervals of verbal and cognitive outcomes of the TEACCH program in individuals with autism including the cognitive, verbal and total scores of the Psychoeducational Profile, Revised (PEP-R).

$p = .008$, $n = 19$). Heterogeneity was low ($I^2 = 25\%$), while the Egger's test showed evidence of potential publication bias or small-study effects ($p = .017$).

The pooled effect size of VABS total score based on the four controlled studies reporting the outcome was 0.47 (95% CI 0.01 to 0.93, $p = .047$, $n = 72$). Heterogeneity was moderate ($I^2 = 65\%$) and no evidence of publication bias was established ($p > .1$).

Sensitivity analyses showed that gains in adaptive behaviors were not always consistent across all age groups. Specifically, gains in social functioning and ADL were more pronounced among school-age children and adults, while gains in these outcomes were negligible among young children (Table 3).

3.4. Language and cognition

Five controlled studies assessed cognitive functioning using the PEP-R ($n = 43$) resulting in a pooled effect size of 0.41 (95% CI 0.04 to 0.77, $p = 0.028$) (Fig. 5). This outcome demonstrated a relatively high heterogeneity across studies ($I^2 = 80\%$). The Egger's test was significant ($p = .016$).

Five controlled studies assessed developmental/mental age in a total number of 53 individuals before and after being exposed to the TEACCH program. Mental/developmental age was established by means of the MSEL, M-P-R, and PEP-R. The pooled effect size was 0.30 (95% CI 0.09 to 0.51, $p = .006$) (Fig. 5). Heterogeneity was within the acceptable range ($I^2 = 48\%$). The Egger's test was not significant ($p > .05$).

Verbal skills assessed by means of the PEP scales and the MSEL were reported in nine studies totaling 121 participants exposed to TEACCH. The pooled effect size was 0.36 (95% CI, −0.01 to 0.72, $p = .052$) (Fig. 5). The effect magnitude obliterated when the meta-analysis was restricted to the six controlled studies available ($ES = 0.14$, 95% CI −0.14 to 0.96, $p > .1$, $n = 63$). Heterogeneity across studies was moderate ($I^2 = 69\%$), and no publication bias or small-study effects were established ($p > .1$).

PEP-R total score was reported in six studies ($n = 78$) resulting in a pooled effect size of 0.39 (95% CI 0.14 to 0.64, $p = .002$) (Fig. 5). The effect did not change in the sensitivity analysis conducted with

Table 2
Summary of effect sizes by outcome and study.

	Aoyama (1995)	Braiden et al. (2012)	McConkey et al. (2010)	Ozonoff and Cathcart (1998)	Panerai et al. (2002)	Panerai et al. (2009)	Panerai et al. (2009)	Persson (2000)	Probst and Leppert (2008)	Siaperas & Beadle-Brown (2006)	Siaperas et al. (2007)	Tsang et al. (2007)	Van Bourgondien et al. (2003)	Welterlin et al. (2012)
Perceptual and motor														
Eye–hand	–	–	0.39	0.18	0.50	0.30	0.65	–	–	–	–	−0.38	–	–
Fine motor	–	–	1.01	−0.06	0.20	0.48	0.31	–	–	–	–	0.26	–	–
Gross motor	–	–	1.21	0.16	0.68	0.48	0.60	–	–	–	–	0.33	–	–
Imitation	–	–	0.81	−0.06	0.59	0.60	0.26	–	–	–	–	0.21	–	–
Perception	–	–	0.74	0.05	0.47	0.28	0.19	–	–	–	–	0.51	–	–
Adaptive outcomes														
Communication	–	–	0.98	–	–	0.54	0.21	–	–	0.28	–	−0.42	–	–
ADL	–	–	0.65	–	–	0.82	0.47	0.24	–	0.39	–	−0.61	–	–
Motor	–	–	0.39	–	–	0.61	−0.10	–	–	–	–	0.33	–	–
Social	–	–	0.79	–	–	1.64	0.98	1.10	–	0.47	–	−0.55	–	0.49
Maladaptive behavior	–	–	–	–	–	−0.87	−1.24	–	−0.33	–	−1.69	–	–	–
Adaptation composite	–	–	0.74	–	–	0.94	0.32	–	–	–	–	−0.09	–	–
Verbal and cognitive														
Cognitive	–	–	–	0.41	0.56	0.54	0.78	–	–	–	–	−0.16	–	–
Developmental age	–	–	–	–	0.67	0.42	0.43	–	–	–	–	0.07	–	0.10
Verbal	–	0.28	1.48	0.41	0.00	0.23	0.37	0.35	–	–	–	−0.05	–	−0.08
PEP-total	–	–	–	0.65	0.49	0.50	0.47	0.39	–	–	–	−0.07	–	–
Mean *ES* (unmatched)	0.91	–	0.90	–	–	–	–	0.23	–	0.77	1.14	−0.01	–	−0.11
Mean *ES* (all)	0.91	0.28	0.85	0.22	0.46	0.62	0.48	0.40	0.33	0.68	1.22	−0.04	1.77	0.06

Notes. Unmatched outcomes were reported only in one or two studies. *ES* = effect size.

Table 3
Sensitivity analysis by age group.

	Age group 1 (0–5)		Age group 2 (6–17)		Age group 3 (≥18)		Overall	
	n	*ES* [95% CI]	*n*	*ES* [95% CI]	*n*	*ES* [95% CI]	*n*	*ES* [95% CI]
Perceptual and motor								
Eye–hand	3	.08 [−.41, .56]	3	.46** [.20, .72]		–	6	.26 [−.03, .55]
Fine motor	3	.40 [−.17, .97]	3	.33** [.16, .50]		–	6	.36* [.08, .65]
Gross motor	3	.57 [−.07, 1.21]	3	.57** [.28, .85]		–	6	.58** [.25, .91]
Imitation	3	.34 [−.19, .87]	3	.47** [.18, .76]		–	6	.41** [.12, .70]
Perception	3	.46* [.07, .84]	3	.29* [.03, .55]		–	6	.40** [.17, .63]
Adaptive								
Communication	2	.30 [−1.08, 1.68]	2	.38 [.38, −.09]	1	.29 [−.18, .75]	5	.34 [−.13, .82]
ADL	2	.03 [−1.21, 1.26]	2	.64** [.23, 1.05]	2	.34* [.01, .66]	6	.32 [−.09, .73]
Motor	2	.37 [−.03, .77]	2	.26 [−.44, .96]		–	4	.33* [.00, .66]
Social	3	.25 [−.64, 1.11]	2	1.28** [.65, 1.92]	2	.63* [.09,1.18]	7	.65* [.15, 1.15]
Problem behavior		–	3	−.70* [−1.25, −.15]	1	−1.69 [−2.89, −.50]	4	−.92** [−1.51, −.33]
Adaptation composite	2	.34 [−.48, 1.15]	2	.62* [.02, 1.23]		–	4	.47* [.01, .93]
Verbal and cognitive								
Cognitive	2	.11 [−.45, .67]	3	.62** [.38, .86]		–	5	.41* [.04, .77]
Developmental/mental age	2	.08 [−.13, .29]	3	.48** [.26, .70]		–	5	.30** [.09, .51]
Verbal	5	.43 [−.17,1.02]	3	.22 [−.20, .63]	1	.35 [−.33,1.03]	9	.36 [.00, .72]
PEP total	2	.27 [−.43, .98]	3	.49** [.28, .70]	1	.39 [.00, .79]	6	.39** [.14, .64]
Mean effect size	5	.29 [−.06, .63]	5	.53** [.31, .75]	4	.81*[.25,1.38]	6	.47** [.27, .67]

Notes. Mean effect sizes included all outcomes within a study. *ES* = effect size; CI = confidence interval; *n* = studies included.
* $p < .05$.
** $p < .01$.

the five controlled studies available ($ES = 0.39$, 95% CI 0.09 to 0.687, $p = 0.010$, $n = 71$) (Fig. 5). Heterogeneity across studies was within the moderate range ($I^2 = 69\%$). There was some potential for publication bias and small-study effects according to the Egger's test ($p = .049$).

Meta-regression analyses did not establish any conclusive factors as significantly associated with a reduced level of heterogeneity for the cognitive and verbal outcomes under analysis. There was a trend toward smaller intervention effects for cognitive functioning among those with higher developmental ages; and larger intervention effects in the same outcome for older individuals (Table 4). Individuals with higher developmental ages and exposed to higher intervention intensities (above 12 weekly hours) showed a trend toward lower intervention effects in the PEP-R total score (Table 4). Sensitivity analyses across age groups showed clear disparities of treatment effects for cognitive functioning and developmental age. Specifically, school-age participants showed moderate gains in both outcomes, while young children did not show those effects (Table 3).

4. Discussion

The present study represents a thorough attempt to summarize the evidence in support of the TEACCH program based on its effect on a variety of standardized outcomes. While TEACCH is a relatively common treatment option for individuals with developmental disabilities and autism (Green et al., 2006), no meta-analyses have been reported. By contrast, pharmacological approaches and interventions based on applied behavior analysis, to mention only two examples, have been tested on a number of occasions using pre–post and controlled studies, and their effectiveness have been documented in several meta-analyses (e.g., Chavez, Chavez-Brown, Sopki, & Rey, 2007; Virués-Ortega, 2010). We were able to identify 13 studies reporting six between-group and seven pre-post trials. While the amount of data available is limited, the studies allowed for a relatively wide evaluation of the impact of TEACCH over perceptual and motor skills, social behavior, activities of daily living, maladaptive behavior, cognition, and language. Our results indicated that TEACCH effects on perceptual, motor, verbal and cognitive skills were of small magnitude in the meta-analyzed studies. Effects over adaptive behavioral repertoires including communication, activities of daily living, and motor functioning were within the negligible to small range. There were moderate to large gains in social behavior and maladaptive behavior (see a summary of individual effect sizes in Table 2). The overall effect of the intervention across all outcomes was moderate (Fig. 2) and effects seemed to increase with age (Table 3). The adult population experienced the greatest overall benefit. However, due to the limited number of appropriately designed studies, we are still far from establishing an evidence base for the TEACCH program as effective or ineffective across children and adults. Therefore, our analysis should be considered a preliminary evaluation pointing to the outcomes that demonstrate greater promise.

In terms of large areas of functioning, our results suggested that TEACCH had a moderate positive effect over perceptual and motor skills ($ES = 0.37$–0.63). Nonetheless, the effect size for fine motor functioning lost statistical significance when the analysis was limited to controlled studies (sensitivity analysis). Interestingly, when the meta-analysis of perceptual and motor skills was restricted to control studies, the magnitude of the effects fell across all outcomes ($ES = 0.22$–0.45), suggesting that gains at post-test were to some extent driven by factors other than the intervention.

Adaptive skills, communication, activities of daily living, and motor functioning obtained effect magnitudes within the small range. Moreover, the effect sizes of communication and activities of daily living were not statistically significant. The impact of TEACCH in social adaptive functioning was moderate ($ES = 0.65$). This finding endured the sensitivity analysis and no evidence of publication bias or small study effects was established. However, the meta-regression analysis revealed that the effects over social functioning were negligible among higher quality studies. Also, heterogeneity levels were high for this outcome. The excessive heterogeneity in social functioning may have been driven by the various assessment strategies used to quantify the outcome. Namely, Siaperas and Beadle-Brown (2006), Persson (2000), McConkey et al. (2010), and Welterlin et al. (2012) used different assessment methods to establish social functioning.

The favorable effect of TEACCH on maladaptive behavior ($ES = -0.91$) remained unchanged when the analysis was replicated solely with controlled studies (Panerai et al., 2009). This is of some interest, as pre–post studies implemented non-standardized assessment methods (Probst & Leppert, 2008; Siaperas et al., 2007), which suggests

Table 4
Meta-regression and sensitivity analysis by specific predictors over selected TEACCH intervention outcomes.

	Cognitive *n*, *ES* [95% CI]	Verbal *n*, *ES* [95% CI]	Social *n*, *ES* [95% CI]	PEP-R total *n*, *ES* [95% CI]	*n*, mean *ES* [95% CI]
Developmental age	$p = 0.061$	*ns*	*ns*	$p = 0.067$	*ns*
≤23	3, 0.56 [0.33, 0.79]	5, 0.54 [−0.02, 1.10]	2, 0.83 [0.48, 1.19]	3, 0.54 [0.33, 0.75]	4, 0.51 [0.21, 0.81]
>23	2, 0.18 [−0.51, 0.87]	3, 0.02 [−0.34, 0.39]	3, 0.50 [−0.74, 1.73]	2, 0.20 [−0.36, 0.76]	4, 0.31 [−0.10, 0.72]
Intensity, h/week	*ns*	*ns*	*ns*	$p = 0.067$	*ns*
≤12.5	–	3, 0.73 [−0.08, 1.54]	–	–	4, 0.46 [0.10, 0.81]
>12.5	3, 0.37 [−0.25, 1.00]	3, 0.81 [−0.30, 0.46]	2, 0.19 [−1.31, 1.70]	3, 0.28 [−1.33, 0.68]	2, 0.21 [−0.29, 0.70]
Duration, weeks	*ns*	*ns*	*ns*	*ns*	*ns*
≤12.5	3, 0.24 [−0.22, 0.71]	5, 0.45 [−0.15, 1.04]	3, 0.25 [−0.52, 1.02]	3, 0.34 [−1.15, 0.82]	8, 0.48 [0.19, 0.77]
>12.5	2, 0.64 [0.37, 0.92]	4, 0.22 [−0.12, 0.57]	4, 1.02 [0.53, 1.50]	3, 0.46 [0.26, 0.67]	5, 0.47 [0.27, 0.71]
Main setting	*ns*	*ns*	*ns*	*ns*	*ns*
Center-based	–	2, 0.10 [−0.32, 0.53]	3, 0.27 [−0.61, 1.16]	2, 0.13 [−0.32, 0.59]	10, 0.57 [0.32, 0.83]
Home-based	4, 0.56 [0.36, 0.75]	7, 0.42 [−0.02, 0.86]	4, 0.91 [0.51, 1.31]	4, 0.53 [0.35, 0.71]	4, 0.35 [−0.04, 0.71]
Study quality	*ns*	*ns*	$p = 0.032$	*ns*	*ns*
≤65	3, 0.55 [0.33, 0.77]	7, 0.46 [−0.04, 0.88]	5, 0.92 [0.59, 1.25]	4, 0.51 [0.33, 0.69]	9, 0.55 [0.38, 0.73]
>65	2, 0.18 [−0.53, 0.88]	2, 0.04 [−0.49, 0.41]	2, −0.04 [−1.04, 0.96]	2, 0.19 [−0.37, 0.74]	5, 0.33 [−0.17, 0.83]

Notes. Mean effect sizes included all outcomes within a study. *p* values of random-effect meta-regression (only values below .1 reported). All continuous predictors except 'main setting'. Center-based programs were delivered partially or totally at a school, center, residence, or group home. Panerai et al. (2009) was considered a home-based program. Predictor medians were used as cut-off points in the sensitivity analysis by predictor level. 95% CI = 95% confidence interval; *ES* = effect size; *n* = number of studies; *ns* = non-significant.

that the effect is not the result of the study design, but a function of a true change of the outcome. Finally, while there was no evidence of heterogeneity, publication bias and small study effects were significant. In summary, according to the limited pool of studies currently available, TEACCH effects on adaptive skills were non-significant (communication, activities of daily living), small (e.g., motor), or were subject to quality and publication bias (social functioning, maladaptive behavior). Nonetheless, social functioning and maladaptive behavior showed significant promise. Future evaluations incorporating these outcomes would help to mitigate the quality and publication biases reported in the present analysis.

Treatment gains in language and cognition were within the small to moderate range ($ES = 0.36–0.42$). In terms of specific outcomes, TEACCH had a small favorable impact on cognitive functioning ($ES = 0.41$) as established by the PEP-R. However, the extent to which this result represented a true treatment gain was undermined by excessive heterogeneity and publication bias. Similar comments apply to mental/developmental age estimates, for which we identified a small effect ($ES = 0.36$). Finally, the effect of TEACCH over verbal skills ($ES = 0.42$) did not reach statistical significance. Furthermore, when the meta-analysis was restricted to controlled studies the effect magnitude became negligible ($ES = 0.17$). While these findings are compatible with the view that the intervention may have small or negligible effects on cognition and language, we should be extremely cautious while interpreting these findings. As noted above, effects not reaching statistical significance or subject to publication bias and small study effects may become firmly established upon further evaluation. Interestingly, TEACCH and early intensive behavioral intervention seem to diverge in their respective effects over language and cognition. Recent meta-analyses have pointed out that behavioral intervention has large effects on language and cognition with effect sizes around 1.5 or higher when accounting for dose-dependent effects (Eldevik et al., 2009; Virués-Ortega, 2010). By contrast, no significant moderators were established for this cluster of outcomes in the current meta-analysis.

While results were generally consistent across ages, the sensitivity analysis by age group showed that effects were, in general, of less magnitude and lower statistical significance among young children relative to school-age children. Discrepancies between age groups were found mainly in ADL, cognitive functioning, eye–hand coordination, social behavior, and developmental age. Also, the effects found among adults were consistent with those found among school-age children. These findings suggest that intervention effects are more variable at younger ages opposing the view that earlier interventions induce greater treatment effects (Ozonoff & Cathcart, 1998). Due to the limited pool of studies in each age group, our sensitivity analysis should be considered preliminary. Further studies are needed to substantiate these exploratory findings.

Our analysis suggests that TEACCH has modest effects on specific domains, while the outcomes that show greater promise (social functioning, maladaptive behavior) require further replication. Unfortunately, the meta-regression did not help to identify common factors driving the magnitude of the intervention effects. In particular, treatment effects could not be correlated reliably with prominent aspects of the intervention. The meta-regression analysis was conducted upon the premise of appropriate variability of the moderators (Table 1). Data dredging was avoided by examining the potential impact of all possible combinations of predictors and outcomes (Thompson & Higgins, 2002). However, the results of meta-regression analyses based upon a limited pool of studies may be variable. It remains to be seen if our general findings will be consistent with future meta-analyses based on a larger pool of studies.

In summary, the present meta-analysis suggests that (a) TEACCH effects over perceptual, motor, verbal and cognitive skills may be of small magnitude; (b) effects over adaptive behavioral repertoires including communication, and activities of daily living may be within the negligible to small range; (c) effects over social behavior and maladaptive behavior may be moderate to large; (d) the evidence base currently available does not allow to identify specific characteristics of the intervention (duration, intensity, and setting) and the target population (developmental age) that could be driving the magnitude of effects; and (e) effects are, in general, replicated across age groups, although the magnitude and consistency of intervention effects are greater among school-age children and adults. Again, it is important to acknowledge that these preliminary conclusions are grounded in very limited data. Namely, only two of the meta-analyzed studies were randomized controlled trials (RCT), all studies had small samples, only one study monitored treatment fidelity, and only two studies conducted blinded assessments. Moreover, several outcomes showed evidence of excessive heterogeneity and potential for publication bias. Therefore, our conclusions should be considered preliminary.

There seemed to be a general correspondence between the outcomes identified as specific targets of the TEACCH program and those for which significant, albeit generally small effects were established. For instance, verbal, cognitive, and social skills are acknowledged as important components of the *skill acquisition* and *structured teaching* principles that inspire the TEACCH program (Schopler, 1997, 2005).

Also, *improved adaptation*, a guiding principle generally accomplished by both skill acquisition and environmental accommodation, should have a positive impact on maladaptive behavior. These behaviors are known to serve psychological functions (e.g., escape from demands), which could be offset by environmental adaptations (e.g., modified demands), or the acquisition of new skills (e.g., those needed for compliance) (Hanley, Iwata, & McCord, 2003).

The scope of our findings is limited by the inherent methodological quality of the original studies composing our analysis. Lack of randomization, blinded assessment, and treatment fidelity measures were concerns among some of the studies included in the meta-analysis. In particular, the dearth of details on treatment fidelity allows for a potentially inconsistent implementation of the TEACCH program across studies beyond the disparities on intensity, duration, and setting analyzed here. Nonetheless, the general methodological quality was commensurable to what is typically encountered in the ASD outcome literature. For example, Virués-Ortega (2010), also using the Downs and Black questionnaire, reported an average quality index among interventions based on applied behavior analysis that was lower by comparison (50 vs. 60). We incorporated three strategies into our analysis to evaluate the potential impact of methodological quality in the validity of our findings. First, we conducted a thorough evaluation of methodological quality and reported it in detail. Second, we conducted sensitivity analyses to determine if pre–post studies biased treatment effects systematically. Finally, we reported meta-regression analyses using study quality as a continuous moderator. Meta-regression showed that study quality tended to be negatively correlated with treatment effects. For instance, quality was a significant moderator for social functioning, i.e., studies with higher methodological quality showed lower gains in social functioning. Similar non-significant trends were observed for all other outcomes in Table 4. In summary, study quality had a marginal potential for false-positive bias. Therefore, studies with limited quality were unlikely to overreport treatment efficacy within the small pool of studies available.

Few comprehensive psychosocial interventions for autism have accrued significant empirical support. A review by Rogers and Vismara (2008) determined that only the early intensive behavioral intervention based on applied behavior analysis (UCLA Young Autism Project model, Lovaas, 1987) has achieved the category of "well-established" according to the treatment classification criteria by Chambless and Hollon (1998). In addition, three parent-training models, two of them based on applied behavior analysis, were labeled as "possibly efficacious". The Early Start Denver Model may be added to this list due to a recently published RCT (Dawson et al., 2010). The evidence reviewed here suggests that the TEACCH approach meets Chambless and Hollon (1998) criteria for a "well-established" intervention, i.e., two independent RCT showing the treatment to be better than placebo or equivalent. Nonetheless, further inquiries on the methodological quality and the clinical significance of TEACCH relative to other evidence-based interventions, are warranted. We highlight three points of discussion that could be addressed by future research.

4.1. Methodological standards of treatment evaluation

A common concern in the treatment evaluation literature is the methodological standard of clinical trials. As indicated by Rogers and Vismara (2008), "there is a low number of RCT studies, and these use small samples and examine different treatments with radically different delivery approaches and intensities, delivered over different time spans, and using different measurement approaches" (p. 19). Varying measurement approaches is a smaller concern in the TEACCH literature as the studies available use almost invariably the same assessments (Table 1). However, several methodological considerations should be addressed in future evaluations of TEACCH. First, while randomization is important, small sample sizes prevent randomization effectiveness. Second, treatment fidelity is rarely reported (see an exception in Welterlin et al., 2012). In the present study we have characterized the training background of those delivering the intervention as a proxy to treatment fidelity. Future studies should provide more direct evidence in this respect. Finally, the range of assessments used to evaluate treatment effects is limited. Using wide spread measures of intellectual assessment and development will allow a more comprehensive view of treatment effects and will provide the opportunity to draw more direct comparisons between TEACCH and other approaches to treatment. It is important to note that the range of sample sizes in the RCT examining TEACCH (23–34) is equivalent to the range of sample sizes of RCT evaluating other comprehensive psychosocial interventions for autism. For instance, Rogers and Vismara (2008) described five RCT with a sample size range of 15 to 35. Moreover, the RCT that provided the highest level of evidence, according to these authors, rarely reported treatment fidelity. In summary, the need for higher methodological standards is relevant to all comprehensive psychosocial interventions for autism.

4.2. Effectiveness relative to other interventions

Currently, there are no published studies directly comparing the effects of TEACCH relative to other models of intervention. However, three studies have described TEACCH as a background (control) intervention during the evaluation of other target treatments (Aldred, Green, & Adams, 2004; Eikeseth, Smith, Jahr, & Eldevik, 2002; Howard, Sparkman, Cohen, Green, & Stanislaw, 2005). These studies provide a preliminary comparison of TEACCH relative to other interventions. Also, published meta-analyses incorporating similar or identical outcomes provide a second means for drawing a preliminary comparison across treatments. For instance, Eldevik et al. (2009) and Virués-Ortega (2010) reported several treatment outcomes also included here. Both means of indirect comparison suggest that TEACCH does not compare favorably to other approaches. Nonetheless, direct comparisons are needed to substantiate these initial observations.

4.3. Potential downsides of a name-brand

A name-branded intervention frequently combines a variety of elements that are not assessed separately. Namely, the brand frequently advocates for a "treatment package" preventing separate analyses of specific treatment components. Unlike TEACCH, approaches to treatment based on applied behavior analysis are supported by a large corpus of molecular analyses in the form of single-subject experimental studies. Treatment evaluations of TEACCH focusing on specific treatment components and specific treatment outcomes are sorely needed. Finally, branded interventions tend to centralize services and training, which may harm the scientific integrity of the intervention by limiting the opportunities of widespread dissemination and evaluation.

In summary, the following considerations may be relevant for future research: (a) adherence to the general quality standards of controlled trials including randomization, intention to treat analysis, and treatment fidelity; (b) prioritization of controlled studies as pre–post designs may be subjected to developmental confounders; (c) implementation of standardized assessments developed outside the TEACCH tradition including standardized measures of both intellectual functioning and actual achievement; (d) improvement of the comparability of the intervention and control groups in terms of hours of intervention on occasions when a no-intervention control group is not feasible, and (e) prioritization of studies comparing more than one form of intervention in order to establish the relative advantages and disadvantages of TEACCH.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.cpr.2013.07.005.

References[1]

Aldred, C., Green, J., & Adams, C. (2004). A new social communication intervention for children with autism: Pilot randomized controlled treatment study suggesting effectiveness. *Journal of Child Psychology and Psychiatry*, *45*, 1420–1430. http://dx.doi.org/10.1111/j.1469-7610.2004.00338.x.

American Psychological Association (2009). *Publication Manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

*Aoyama, S. (1995). The efficacy of structuring the work system: Individualization of the work format and the use of a 3-level paper rack in a special education class. *Japanese Journal of Special Education*, *32*(5), 1–5.

Autism Speaks (2012). Treatments & therapies. Retrieved from. http://www.autismspeaks.org/family-services/tool-kits/100-day-kit/treatments-therapies

Baio, J., & Baio, J. (2012). *Prevalence of autism spectrum disorders: Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. MMWR Surveillance Summaries*, *61(SS03)*. (pp. 1–19), 1–19 (Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss6103a1.htm?s_cid=ss6103a1_w).

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, *41*, 257–278. http://dx.doi.org/10.1111/j.2044-8317.1988.tb00901.x.

*Braiden, H. J., McDaniel, B., McCrudden, E., Hanes, M., & Crozier, B. (2012). A practice-based evaluation of Barnardo's forward steps early intervention programme for children diagnosed with autism. *Child Care in Practice*, *18*, 227–242. http://dx.doi.org/10.1080/13575279.2012.683773.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, *66*, 7–18. http://dx.doi.org/10.1037//0022-006X.66.1.7.

Chavez, B., Chavez-Brown, M., Sopki, M.A., & Rey, J. A. (2007). Atypical antipsychotics in children with pervasive developmental disorders. *Pediatric Drugs*, *9*, 249–266. http://dx.doi.org/10.2165/00148581-200709040-00006.

Chootrakool, H., Shi, J. Q., & Yue, R. (2011). *Meta-analysis and sensitivity analysis for multi-arm trials with selection bias. Statistics in Medicine*. http://dx.doi.org/10.1002/sim.4143P2011 (Advance online publication).

Cochrane Library (2012). Cochrane Central Register of Controlled Trials. Retrieved from. http://onlinelibrary.wiley.com/o/cochrane/cochrane_search_fs.html?newSearch=true

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.)Mahwah, NJ: Lawrence Erlbaum Associates.

Cottrell, C., Drew, J., Gibson, J., Holroyd, K., & O'Donnell, F. (2007). Feasibility assessment of telephone-administered behavioral treatment for adolescent migraine. *Headache*, *47*, 1293–1302. http://dx.doi.org/10.1111/j.1526-4610.2007.00804.x.

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., et al. (2010). Randomized, controlled trial of an intervention for toddlers with autism: The Early Start Denver model. *Pediatrics*, *125*, e17–e23. http://dx.doi.org/10.1542/peds.2009-0958.

Deprez, M. (2004). Avant et après la formation TEACCH: Les perceptions des participants [Perceptions of participants before and after the TEACCH program]. *Revue Francophone de la Déficience Intellectuelle*, *15*(1), 75–91.

Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality of both randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, *53*, 377–384. http://dx.doi.org/10.1136/jech.52.6.377.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629–634. http://dx.doi.org/10.1136/bmj.315.7109.629.

Eikeseth, S. (2009). Outcome of comprehensive psychoeducational interventions for young children with autism. *Research in Developmental Disabilities*, *30*, 158–178. http://dx.doi.org/10.1016/j.ridd.2008.02.003.

Eikeseth, S., Smith, T., Jahr, E., & Eldevik, S. (2002). Intensive behavioral treatment at school for 4- to 7-year-old children with autism: A 1-year comparison controlled study. *Behavior Modification*, *26*, 49–68. http://dx.doi.org/10.1177/0145445502026001004.

Eldevik, S., Hastings, R. P., Hughes, J. C., Jahr, E., Eikeseth, S., & Cross, S. (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child and Adolescent Psychology*, *38*, 439–450. http://dx.doi.org/10.1080/15374410902851739.

Fillipin, M., Reska, S., & Watson, L. (2010). Effectiveness of the Picture Exchange Communication System (PECS) on communication and speech for children with autism spectrum disorders: A meta-analysis. *American Journal of Speech-Language Pathology*, *19*, 178–195. http://dx.doi.org/10.1044/1058-0360(2010/09-0022).

Fu, C., Chen, K., Tseng, M., Chiang, F., & Hsieh, C. (2012). Reliability and validity of the Psychoeducational Profile — Third edition caregiver report in children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *6*(1), 115–122. http://dx.doi.org/10.1016/j.rasd.2011.03.011.

Green, V. A., Pituch, K. A., Itchon, J., Choi, A., O'Reilly, M., & Sigafoos, J. (2006). Internet survey of treatments used by parents of children with autism. *Research in Developmental Disabilities*, *27*, 70–84. http://dx.doi.org/10.1016/j.ridd.2004.12.002.

Hanley, G. P., Iwata, B.A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis*, *36*(2), 147–185. http://dx.doi.org/10.1901/jaba.2003.36-147.

Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, *6*, 203–217. http://dx.doi.org/10.1037//1082-989X.6.3.203.

Hess, K. L., Morrier, M. J., Heflin, L. J., & Ivey, M. L. (2008). Autism treatment survey: Services received by children with autism spectrum disorders in public school classrooms. *Journal of Autism and Developmental Disorders*, *38*, 961–971. http://dx.doi.org/10.1007/s10803-007-0470-5.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558. http://dx.doi.org/10.1002/sim.1186.

Howard, J. S., Sparkman, C. R., Cohen, H. G., Green, G., & Stanislaw, H. (2005). A comparison of intensive behavior analytic and eclectic treatments for young children with autism. *Research in Developmental Disabilities*, *26*, 359–383. http://dx.doi.org/10.1016/j.ridd.2004.09.005.

Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q test or I^2 index? *Psychological Methods*, *11*, 193–206. http://dx.doi.org/10.1037/1082-989X.11.2.193.

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, *63*, 146–159. http://dx.doi.org/10.1037/0003-066X.63.3.146.

Keel, J. H., Mesibov, G. B., & Woods, A. V. (1997). TEACCH-supported employment program. *Journal of Autism and Developmental Disorders*, *27*, 3–9. http://dx.doi.org/10.1023/A:1025813020229.

Kielinen, M., Linna, S. L., & Moilanen, I. (2002). Some aspects of treatment and habilitation of children and adolescents with autistic disorder in Northern-Finland. *International Journal of Circumpolar Health*, *61*(Suppl. 2), 69–79.

Kim, Y. S., Leventhal, B.L., Koh, Y. -J., Fombonne, E., Laska, E., Lim, E. -C., et al. (2011). Prevalence of autism spectrum disorders in a total population sample. *The American Journal of Psychiatry*, *168*, 904–912. http://dx.doi.org/10.1176/appi.ajp.2011.10101532.

Leppert, T., & Probst, P. (2005). Entwicklung und Evaluation eines psychoedukativen Gruppentrainings für Lehrer von Schülern mit einer autistischen Entwicklungsstörung und Intelligenzminderung [Development and evaluation of a psychoeducational training group for teachers of students with autism and mental retardation]. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, *33*, 49–85. http://dx.doi.org/10.1024/1422-4917.33.1.49.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, *55*, 3–9. http://dx.doi.org/10.1037//0022-006X.55.1.3.

*McConkey, R., Truesdale-Kennedy, M., Crawford, H., McGreevy, E., Reavey, M., & Cassidy, A. (2010). Preschoolers with autism spectrum disorders: Evaluating the impact of a home-based intervention to promote their communication. *Early Child Development & Care*, *180*, 299–315. http://dx.doi.org/10.1080/03004430801899187.

Mesibov, G. B. (1997). Formal and informal measures on the effectiveness of the TEACCH programme. *Autism*, *1*, 25–35. http://dx.doi.org/10.1177/1362361397011005.

Mesibov, G. B., Schopler, E., Schaffer, B., & Landrus, R. (1988). *Adolescent and Adult Psychoeducational Profile (AAPEP)*. Austin, TX: Pro-Ed.

Mesibov, G. B., & Shea, V. (2010). The TEACCH program in the era of evidence-based practice. *Journal of Autism and Developmental Disorders*, *40*, 570–579. http://dx.doi.org/10.1007/s10803-009-0901-6.

Mesibov, G. B., Shea, V., & Schopler, E. (2005). *The TEACCH approach to autism spectrum disorders*. New York, US: Springer.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*, 264–269. http://dx.doi.org/10.1371/journal.pmed.1000097.

Morris, S. B. (2008). Estimating effect sizes from pre-test–post-test-control group designs. *Organizational Research Methods*, *11*, 364–386. http://dx.doi.org/10.1177/1094428106291059.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105–125. http://dx.doi.org/10.1037//1082-989X.7.1.105.

Mullen, E. M. (1995). *Mullen scales of early learning* (AGS ed.)Circle Pines, MN: American Guidance Service.

National Institutes of Health (2011). Research portfolio online reporting tools (RePORT). Retrieved from. http://projectreporter.nih.gov/reporter.cfm

Ono, M. (1994). A trial in applying the TEACCH program in a children's counseling service. *Japanese Journal of Special Education*, *31*(5), 15–22.

Ospina, K., Krebs, S. J., Clark, B., Karkhaneh, M., Hartling, L., Tjosvold, L., et al. (2008). Behavioural and developmental interventions for autism spectrum disorder: A clinical systematic review. *PLoS One*, *3*. http://dx.doi.org/10.1371/journal.pone.0003755.

*Ozonoff, S., & Cathcart, K. (1998). Effectiveness of a home program intervention for young children with autism. *Journal of Autism and Developmental Disorders*, *28*, 25–32. http://dx.doi.org/10.1023/A:1026006818310.

Panerai, S., Ferrante, L., & Caputo, V. (1997). The TEACCH strategy in mentally retarded children with autism: A multidimensional assessment: Pilot study. *Journal of Autism and Developmental Disorders*, *27*(3), 345–346.

*Panerai, S., Ferrante, L., & Zingale, M. (2002). Benefits of the Treatment and Education of Autistic and Communication Handicapped Children (TEACCH) programme as compared with a non-specific approach. *Journal of Intellectual Disability Research*, *46*, 318–327. http://dx.doi.org/10.1046/j.1365-2788.2002.00388.x.

*Panerai, S., Zingale, M., Trubia, G., Finocchiaro, M., Zuccarello, R., Ferrei, R., et al. (2009). Special education versus inclusive education: The TEACCH program. *Journal of Autism and Developmental Disorders*, *39*, 874–882. http://dx.doi.org/10.1007/s10803-009-0696-5.

Perry, A., & Factor, D. C. (1989). Psychometric validity and clinical usefulness of the Vineland Adaptive Behavior Scales and the AAMD Adaptive Behavior Scale for an autistic sample. *Journal of Autism and Developmental Disorders*, *19*(1), 41–55. http://dx.doi.org/10.1007/BF02212717.

*Persson, B. (2000). Brief report: A longitudinal study of quality of life and independence among adult men with autism. *Journal of Autism and Developmental Disorders*, *30*, 61–66. http://dx.doi.org/10.1023/A:1005464128544.

Preston, D., & Carter, M. (2009). A review of the efficacy of the picture exchange communication system intervention. *Journal of Autism and Developmental Disorders*, *39*, 1471–1486. http://dx.doi.org/10.1007/s10803-009-0763-y.

[1] References marked with an asterisk were included in the meta-analysis.

Probst, P., Jung, F., Micheel, J., & Glen, I. (2010). Tertiary-preventive interventions for autism spectrum disorders (ASD) in children and adults: An evaluative synthesis of two TEACCH based outcome studies. *Life Span and Disability*, *13*(2), 129–167.

*Probst, P., & Leppert, T. (2008). Brief report: Outcomes of a teacher training program for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*, 1791–1796. http://dx.doi.org/10.1007/s10803-008-0561-y.

Reichow, B. (2011). Overview of meta-analyses on early intensive behavioral intervention for young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*. http://dx.doi.org/10.1007/s10803-011-1218-9.

Reichow, B., & Wolery, M. (2009). Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA Young Autism Project model. *Journal of Autism and Developmental Disorders*, *39*, 23–41. http://dx.doi.org/10.1007/s10803-008-0596-0.

Rogers, S. J., & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology*, *37*, 8–38. http://dx.doi.org/10.1080/15374410701817808.

Roid, G. H., & Sampers, J. L. (2004). *Merrill-Palmer—Revised Scales of Development*. Wood Dale, IL: Stoelting.

Rúa, M.O. (2004). Intercención psicoeducativa en estudiantes con autismo [Psychoeducational intervention in students with autism]. *Revista de Psicología General y Aplicada*, *57*(1), 69–82.

Schopler, E. (1990). *Psychoeducational Profile, Revised.* Austin, TX: Pro-Ed.

Schopler, E. (1997). Implementation of TEACCH philosophy. In D. J. Cohen, & F. R. Volkmar (Eds.), *Handbook of autism and developmental disorders* (pp. 767–798) (2nd ed.). New York: Wiley.

Schopler, E. (2005). *Cross-cultural program priorities and reclassification of outcome research methods.* Hoboken, NJ, US: John Wiley & Sons, 1174–1189.

Schopler, E., Mesibov, G., & Baker, A. (1982). Evaluation of treatment for autistic children and their parents. *Journal of the American Academy of Child Psychiatry*, *21*, 262–267. http://dx.doi.org/10.1016/S0002-7138(09)60881-5.

Schultheis, S. F., Boswell, B. B., & Decker, J. (2000). Successful physical activity programming for students with autism. *Focus on Autism and Other Developmental Disabilities*, *15*, 159–162. http://dx.doi.org/10.1177/108835760001500306.

Shadish, W. R., & Haddock, C. K. (2009). Combining estimates of effect size. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 257–278) (2nd ed.). New York: Russell Sage Foundation.

Short, A.B. (1984). Short-term treatment outcome using parents as co-therapists for their own autistic children. *Journal of Child Psychology and Psychiatry*, *25*, 443–458. http://dx.doi.org/10.1111/j.1469-7610.1984.tb00163.x.

*Siaperas, P., & Beadle-Brown, J. (2006). A case study of the use of a structured teaching approach in adults with autism in a residential home in Greece. *Autism*, *10*, 330–343. http://dx.doi.org/10.1177/1362361306064433.

*Siaperas, P., Higgins, S., & Proios, P. (2007). Challenging behaviours on people with autism: A case study on the effect of a residential training programme based on structured teaching and TEACCH method. *Psychiatriki*, *18*(4), 343–350.

Simpson, R. L., de Boer-Ott, S. R., Griswold, D. E., Myles, B.S., Byrd, S. E., Ganz, J. B., et al. (2005). *Autism spectrum disorders: Interventions and treatments for children and youth.* Thousand Oaks, CA: Corwin Press.

Sinha, Y., Silove, N., Williams, K., & Hayen, A. (2004). Auditory integration training and other sound therapies for autism spectrum disorders. *Cochrane Database of Systematic Reviews*(1). http://dx.doi.org/10.1002/14651858.CD003681.pub2 (Art. No.: CD003681).

Sparrow, S. S., Cicchetti, D.V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales* (2nd ed.)Circle Pines, MN: AGS Publishing.

Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, *21*, 1559–1573. http://dx.doi.org/10.1002/sim.1187.

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in Medicine*, *18*, 2693–2708. http://dx.doi.org/10.1002/(SICI)1097-0258(19991030)18:20<2693::AID-SIM235>3.3.CO;2-M.

Tien, K. (2008). Effectiveness of the Picture Exchange Communication System as a functional communication intervention for individuals with autism spectrum disorders. *Education and Training in Developmental Disabilities*, *43*, 61–76. http://dx.doi.org/10.1044/1058-0360(2010/09-0022).

*Tsang, S., Shek, D., Lam, L., Tang, F., & Cheung, P. (2007). Brief report: Application of the TEACCH program on Chinese pre-school children with autism — Does culture make a difference? *Journal of Autism and Developmental Disorders*, *37*, 390–396. http://dx.doi.org/10.1007/s10803-006-0199-6.

*Van Bourgondien, M. E., Reichle, N. C., & Schopler, E. (2003). Effects of a model treatment approach on adults with autism. *Journal of Autism and Developmental Disorders*, *33*, 131–140. http://dx.doi.org/10.1023/A:1022931224934.

Van Bourgondien, M. E., & Schopler, E. (1996). Intervention for adults with autism. *Journal of Rehabilitation*, *62*, 65–71.

Villa, S., Micheli, E., Villa, L., Pastore, V., Crippa, A., & Molteni, M. (2010). Further empirical data on the Psychoeducational Profile—Revised (PEP-R): Reliability and validation with the Vineland Adaptive Behavior Scales. *Journal of Autism and Developmental Disorders*, *40*, 334–341. http://dx.doi.org/10.1007/s10803-009-0877-2.

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood. *Clinical Psychology Review*, *30*, 387–399. http://dx.doi.org/10.1016/j.cpr.2010.01.008.

*Welterlin, A., Turner-Brown, L. M., Harris, S., Mesibov, G., & Delmolino, L. (2012). The home TEACCHing program for toddlers with autism. *Journal of Autism and Developmental Disorders*, *42*, 1827–1835. http://dx.doi.org/10.1007/s10803-011-1419-2.