

Evaluation of Comprehensive Treatment Models for Individuals with Autism Spectrum Disorders

Samuel L. Odom · Brian A. Boyd · Laura J. Hall · Kara Hume

Published online: 25 July 2009
© Springer Science+Business Media, LLC 2009

Abstract Multiple dimensions of comprehensive treatment models (CTMs) for learners with autism were evaluated in this study. The purpose of the study was to provide evaluative information upon which service providers, family members, and researchers could make decisions about model adoption, selection for a family member, or future research. Thirty CTMs were identified, with the majority based on an applied behavior analysis framework, although a substantial minority followed a developmental or relationship-based model. As a group, CTMs were strongest in the operationalization of their models, although relatively weaker in measurement of implementation, and with notable exceptions, weak in evidence of efficacy.

Keywords Comprehensive · Treatment · Model · Autism · Program · Evaluation

The increase in the reported prevalence of children with autism spectrum disorders (ASD), estimated at 1 in 150 (Stevens et al. 2007), means that personnel from early intervention agencies, school districts, and educational programs serving all ages will experience greater demand for providing high quality educational and clinical services. Current legislation, such as the No Child Left Behind Act,

requires that service providers use scientifically proven practices in their instruction (Yell et al. 2005). That is, models for treatment need to have evidence of efficacy demonstrated by researchers and program developers through the use of rigorous research designs (Odom et al. 2003; Reichow et al. 2008). Stakeholders are looking to adopt those models and practices that yield positive outcomes for individuals with ASD and their families. Yet such adoption requires focused intervention practices and comprehensive treatment models (CTMs) that have procedures developed well enough to be feasibly implemented by stakeholders. A comprehensive, consumer-oriented evaluation of practices and models requires a multi-dimensional approach to evaluation that extends beyond only an analysis of the efficacy of the intervention, although efficacy is a central feature of such an evaluation.

In the professional literature, there are two classifications of intervention. The first, focused intervention practice, are designed to produce specific behavioral or developmental outcomes for individual children with ASD (Odom et al. *in press*). Examples of focused intervention practices include prompting, reinforcement, discrete trial teaching, social stories, or peer-mediated interventions. In the applied research literature, these focused interventions are used with individuals with ASD for a limited time period with the intent of demonstrating change in the targeted behavior(s). Focused interventions occur over a relatively short period of time (e.g., 3 months), although probes for maintenance are often included in the studies. Past and current efforts have identified the focused intervention practices having evidence of efficacy (Odom et al. 2003). Practitioners may adopt focused intervention approaches based on the evidence of efficacy, the type of outcome produced by the intervention, and individual needs of the learner with ASD. We refer the reader to Hall (2009) for a detailed description

S. L. Odom (✉) · B. A. Boyd · K. Hume
Frank Porter Graham Child Development Institute, University of
North Carolina at Chapel Hill, CB 8180, 105 Smith Level Road,
Chapel Hill, NC 27599-8180, USA
e-mail: slodom@mail.fpg.unc.edu; slodom@unc.edu

L. J. Hall
Department of Special Education, College of Education,
San Diego State University, 5500 Campanile Dr.,
San Diego, CA 92182-1170, USA

of focused intervention practices. Comprehensive treatment models are a second classification of intervention approaches that appear in the professional literature. CTMs consist of a set of practices designed to achieve a broader learning or developmental impact on the core deficits of ASD (National Research Council 2001). Rogers and Vismara (2008) described these as “branded” interventions in that they are often identifiable by a consistently used name (e.g., Denver Model, Lovaas Model, Walden Model). They occur over an extended period of time (e.g., a year or years), are intense in their application (e.g., 25 h per week), and usually have multiple components. CTMs have been in existence for over 30 years and new models continue to be created along with the advancement of the scientific knowledge about ASD. Examples of historic CTMs are the UCLA Young Autism Project (now the Lovaas Institute), Treatment and Education of Autistic and Communication Handicapped Children (TEACCH), the Denver Model, and the Princeton Child Development Institute (PCDI).

Although some models have been disseminated widely, there remains a scarcity of comparative evaluation information about CTMs. A committee convened by the National Academy of Sciences to review research on effective practices for children with ASD identified a set of CTMs that had been developed, as of the late 1990s (National Research Council 2001). Their review primarily described the CTMs and, at that time, the limited evidence of efficacy. In a recent critical review of CTMs for young children with ASD and their families, Rogers and Vismara (2008) evaluated the current research on comprehensive treatments for young children with ASD, still finding limited evidence of efficacy for all but the Lovaas model, with some limited support for Pivotal Response Treatment (PRT) (Koegel et al. 1999). Focusing only on a single model, Reichow and Wolery (2009) documented substantial support for the UCLA Young Autism Project Model (the Lovaas model), especially when the researchers were trained by model developers, hours of therapy were high, and the duration of the intervention was long. A second independent review by Howlin et al. (2009) drew similar conclusions, but also found that documentation of the fidelity of implementation of the CTMs in the studies was often limited. These summaries of the empirical evidence reveal that (a) stronger empirical demonstrations of efficacy are being published (Rogers and Vismara 2008); (b) efficacy studies are beginning to be conducted by investigators independent from the original model developers (Reichow and Wolery 2009); and (c) one may draw practical recommendations by examining features of CTMs that appear to be efficacious (National Research Council 2001; Rogers and Vismara 2008).

To realize the benefits of CTMs for learners with ASD and their families, however, one must look to a broader set

of information than usually found in research studies. To enhance the research to practice process, practitioners' implementation of the CTM is a necessary feature (Fixsen et al. 2005). To foster implementation, a CTM has to be developed and documented well enough to be exported to applied settings (Durlak and DuPre 2008). A multi-dimensional evaluation that systematically incorporates data from the published literature, procedural information (e.g., curriculum used by model) and data retrieved from CTM program developers may generate a cohesive set of information upon which stakeholders can base their decisions about adoption of CTMs.

The goal of the current study was to provide information about multiple dimensions of CTMs upon which stakeholders (i.e., service providers, families, researchers) may base decisions about adoption, selection of a model for an individual, and/or future research. To provide this information, we followed a multidimensional evaluation process. Evaluation differs from research in that its purpose is to provide information that informs decision making (Cronbach 1982). In this case, the information provided may assist service providers in selecting a CTM for adoption by their agency, a family member in selecting a CTM for the individual with ASD within their family, or a researcher in determining questions for future research.

Method

For the purpose of this evaluation study, CTMs are operationally defined by six criteria, which represent the inclusion/exclusion criteria. First, a description of the model and its components had to be published in a refereed journal article, book chapter, or book. Second, at least a single procedural guide, manual, curriculum, or description should exist to define the model. Third, the model must have a clear theoretical or conceptual framework. This framework must be published in one of the formats noted previously. Fourth, the model must address multiple developmental or behavioral domains that, at a minimum, represent the core features of autism spectrum disorder (i.e., social competence, communication, repetitive behaviors). Fifth, the model must be intensive. Intensity is defined by the number of hours the model is implemented per week (e.g., National Academy of Sciences Committee recommended 25 h or more), longevity (e.g., model implementation extends across a period equal to or greater than a typical school year, which is usually 9–10 months), and/or engagement (i.e., a planned set of activities or procedures actively engage the child/person with autism in learning experiences consistent with the model). Finally, the sixth criterion is that the CTM must have been implemented in at least one site in the United States.

Identification of CTMs

CTMs were located through several sources. First, a systematic and thorough review of the literature was conducted. A search in PsychINFO and EBSCO was initiated using the keywords *autism*, *comprehensive*, *treatment*, *program description*, and *intervention*. CTMs also tend to be described in book chapters and literature reviews, so the evaluators surveyed well known books that include chapter identifications of such models (e.g., Handleman and Harris 2006, 2008), books that have reviewed the evidence from the field (e.g., National Research Council 2001), and highly visible handbooks (e.g., Volkmar et al. 2005). Second, the National Standards Project (NSP)¹ has conducted an exhaustive search of the literature and identified CTMs from that search, which were incorporated into this review. Third, evaluators presented the information from this project in public forums and at conference presentations. Audience members were asked to review the list of CTMs identified, and to recommend models that were missed. When CTMs were identified, they were first examined to determine if they met the inclusion criteria noted previously.

Evaluators

Four evaluators, who are the four authors of this paper, participated in this project. They all hold Ph.Ds, have extensive training and experience related to ASD, have published articles about autism treatments in refereed journals, and are currently engaged in research and scholarship about instructional and treatment programs for children and youth with ASD and their families.

Assembly and Review of the Evidence

When a CTM was identified as having met the inclusion criteria, evaluators assembled a portfolio of evidence needed for evaluating the model. Each evaluator had primary responsibility for seven or eight CTMs. The evidence included published journal articles and/or chapters, web pages, books, curricula, and/or procedural manuals. Evaluators contacted each CTM developer and conducted a

standard telephone interview. A letter or email describing the components of the interview was sent to model developers before the call; the calls lasted from 30 to 90 min; and the evaluators summarized in writing information from the call. Model developers initially were asked to confirm that the model met the criteria for a CTM and, if so, to provide copies of or samples from their procedural manuals and/or their curricula. Also, evaluators asked for copies of fidelity of implementation measures and any information about inter-rater agreement, reliability, or other psychometric evidence. Model developers were asked to identify empirical studies that documented the efficacy for their model, as well as studies of focused interventions that had been conducted to examine efficacy of key procedures and strategies. The four evaluators also conducted a separate literature search to identify published empirical studies on the efficacy of identified CTMs.

Conceptual Overview of Evaluation System

We followed the guiding principles established by the American Evaluation Association (2008) in which we combined methods employed in meta-evaluation (Hassen et al. 2008), as well as guidelines from leaders in the ASD treatment research field (e.g., Lord et al. 2005; Smith et al. 2007). This multidimensional evaluation consisted of six evaluation features:

Operationalization. Interventions must be documented in manuals or procedural guides (i.e., manualized) well enough for individuals from outside the project to be able to use the intervention (Lord et al. 2005; Smith et al. 2007). Manuals or guides should describe content (i.e., what to implement) and/or procedures (i.e., how to implement).

Implementation Measures. To document quality of replication, a fidelity of implementation measure should be developed and ideally have some evidence of reliability and validity (Smith et al. 2007).

Replication. The adoption of CTMs by individuals independent of the original developers and programs provide confidence that the model is replicable. Replication is defined by the adopter not being employed by the developer; the adoptee is not supervised directly by the adopter, although the developer may have provided initial training and feedback on model implementation; and the site of implementation is located separately from the developer's site(s).

Type of empirical evidence. Evidence of efficacy may appear in different venues, with the strongest venue being peer-reviewed journals (Rogers and Vismara 2008), because the judgment about the quality of the evidence comes from outside the developers program. Less convincing evidence may appear in book chapters or reports from the developer, but it still may provide information

¹ The National Standards Project is located at the National Autism Center in MA. Its purpose is to identify research that documents efficacy of procedures and programs for learners with ASD. Investigators with this project conducted a comprehensive review of the literature and assembled a national review team to evaluate all articles identified. NSP staff is now aggregating the analyses by focused intervention practices and also CTMs. The four authors were reviewers with the NSP. The website for the NSP is: <http://www.nationalautismcenter.org/about/national.php>.

about outcomes that stakeholders may consider in decisions about CTMs. This latter and weaker (than peer-reviewed journal articles) form of evidence was included to allow the readers to make their own judgments about efficacy support for the CTM.

Quality of the research methodology. A critical evaluation of the experimental methodology utilized in efficacy research provides a necessary basis upon which consumers could make judgments about the study findings and the potential impact of a CTM (Rogers and Vismara 2008). This evaluation could only be made if at least one efficacy study for a CTM has been published in a peer-reviewed journal.

Complementary evidence from studies of focused interventions. Developers may provide evidence (from peer reviewed studies) of focused interventions that are used as part of the models. For example, structured work systems are a type of focused intervention practice about which individuals studies could be conducted (e.g., Hume and Odom 2007) and also are a component of the TEACCH CTM. Focused intervention research does not document efficacy for the entire model, but such evidence may provide complementary information about components of the CTM.

Evaluation Rating Scale

Following this conceptual framework an evaluation instrument, the summative evaluation of evidence for comprehensive treatment models, was developed to guide the evaluation of portfolios of evidence. The rating form was organized into six dimensions that reflected the six areas of evaluation described in the previous section. A six-point (0–5) rating system was developed for each of dimensions. Each point on the scale was anchored with a description of evidence necessary for meeting the criteria for the respective item. The rating measure appears in the Appendix. A coding manual was created to provide definitions and detailed guidelines for completing the scale. This manual can be obtained from the first author. Evaluators developed the rating measure and manual together, pilot tested the evaluation system on one CTM that was coded by all evaluators, revised the rating measure, and then piloted it again on another individual CTM. Evaluators engaged in weekly conference calls, and when questions about specific criteria occurred, they were addressed on the call. Any modifications of the criteria resulted in the evaluators recoding (for all previously coded CTMs) the particular dimension for which the rating criteria were revised.

A more detailed explanation is needed for the rating of Quality of Research as these data were drawn from a parallel review that is currently being conducted by

the NSP, located at the National Autism Center (www.nationalautismcenter.org/about/national.php). Our evaluators used the Scientific Merit Rating System (SMRS) developed by the NSP to generate a “score” for each CTM efficacy study published in a refereed journal. The NSP trained our evaluators to use the SMRS and each rater established inter-rater agreement with their system. The SMRS protocol included evaluations of empirical studies based on subject description, setting, design, and dependent variables. The NSP aggregated this information into a scoring rubric that generated an overall numerical score, which could range from 0 to 5. For the current study, we used the NSP summary score (or average of summary scores if there were more than one article for a CTM) to generate a rating score for the Quality of Research dimension.

To establish inter-rater agreement, two evaluators coded each portfolio. Agreement was defined as each evaluator scoring the same rating on a dimension. Inter-rater agreement was calculated (i.e., total exact agreements for a rating/agreements plus disagreements X 100). The mean inter-rater agreement for the evaluations was 83%. When disagreements occurred, both evaluators re-reviewed the portfolio and reached a consensus on the rating, which appeared in the final rating summaries.

Results

CTMs

Evaluators identified a total of 30 CTMs using the inclusion criteria previously described. A list of models and their contact information is included in Table 1. Additional descriptions of the models can be obtained from the first author. The CTMs were based on different conceptual and theoretical frameworks, as can be seen from the list in Table 1. The highest proportion of models was based on an applied behavior analysis theoretical framework, but procedural features and contexts of implementation differed. Some applied behavior analysis CTMs operated in clinic settings, in homes, or in a combination of clinic and home settings. The CTM from the Lovaas Institute and PRT are examples of two clinic/home models based on applied behavior analysis but following quite different procedures. The Lovaas Institute CTM emphasized, at least initially, discrete trial training, while PRT emphasizes self-management, child initiation, and attention to multiple stimuli and the motivational quality of the learning context.

The largest and perhaps most similar set of CTMs also followed an applied behavior analysis theoretical framework but generally operate in classroom settings. These CTMs often incorporated a variety of applied behavior

Table 1 Comprehensive treatment model summaries organized by conceptual classification

Model name & Site of origin	Citation	Contact information	Ages served
Applied behavior analysis—clinic or home based			
Autism partnership seal beach, CA	Leaf and McEachin (1999)	www.autismpartnership.com	Preschool-adult
Center for autism and related disorders (CARD) Tarzana, CA	Keenan et al. (2006)	www.centerforautism.com	2–21
Lovaas Institute Los Angeles, CA	Cohen et al. (2006)	www.lovaas.com	2–8
Pivotal response treatment santa Barbara, CA	Koegel and Koegel (2006)	education.ucsb.edu/autism	3–21
Applied behavior analysis—classroom based			
Alpine Learning Group Paramus, NJ	Meyer et al. (2006)	www.alpinelearninggroup.org	3–21
Eden Institute Princeton, NJ	Holmes (1998)	www.edenservices.org/index.jsp	Preschool-adult
Douglass Developmental Disabilities Center New Brunswick, NJ	Harris et al. (1991)	dccc.rutgers.edu	Preschool-adult
Institute for Child Development-SUNY Binghamton Vestal, NY	Romanczyk et al. (2006)	http://icd.binghamton.edu	1–11
Lancaster-Lebanon IU 13 Lancaster County, PA	Bruey and Vorhis (2006)	www.iu13.org	3–21
May Institute Randolph, MA	Campbell et al. (1998)	www.mayinstitute.org	Infant-adult
Princeton Child Development Institute Princeton, NJ	Fenske et al. (1985)	www.pcdi.org	Infant-adult
Pyramid approach to education Newark, DE	Bondy and Battaglini (2006)	www.pecs.com	2–21
Strategies for teaching based on autism research (STAR) Portland, OR	Arick et al. (2003)	www.starautismprogram.com	3–12
Summit Academy Getzville, NY	Anderson et al. (2006)	www.summited.org	0–21
Therapeutic pathways/Kendall School Modesto, CA	Howard et al. (2005)	www.tpathways.org	1–7
Valley Program Bergen County, NJ	McKeon et al. (2006)	www2.nvnet.org/nvhs/region3/valley/valley.html	3–14
Applied behavior analysis—inclusive			
Children’s Toddler School San Diego, CA	Stahmer and Ingersoll (2004)	http://www.chsd.org/body.cfm?id=35&action=detail&ref=95	18 months-3
Learning experiences: an alternative program for preschoolers and parents (LEAP) Denver, CO	Hoyson et al. (1984)	phil.strain@cudenver.edu	3–6
Project DATA (Developmentally Appropriate Treatment for Autism)	Schwartz et al. (2004)	depts.washington.edu/dataproj/index	0–7
Walden model Atlanta, GA	McGee et al. (1999)	www.psychiatry.emory.edu/PROGRAMS/autism/Walden.html	0–5
Developmental and relationship-based			
Denver model Denver, CO	Rogers et al. (2006)	Sally.rogers@ucdmc.ucdavis.edu	2–5
DIR/Floortime (Developmental, Individual difference, relationship-based model) Bethesda, MD	Solomon et al. (2007)	www.floor-time.org	0–5
Hanan model Toronto, Ontario, Canada	McConachie et al. (2005)	www.hanan.org	0–5
Relationship development intervention (RDI) Houston, TX	Gutstein et al. (2007)	www.rdi-connect.com	Infant-adolescence
Responsive teaching Cleveland, OH	Mahoney and Perales (2005)	www.responsiveteaching.org	0–6

Table 1 continued

Model name & Site of origin	Citation	Contact information	Ages served
SCERTS model (social communication, emotional regulation, transactional supports) Cranston, RI	Prizant et al. (2006)	www.scerts.com	0–10
Son-rise program	Kaufman (1981)	www.autismtreatmentcenter.org	Infant-adult
Idiosyncratic			
Higashi School Boston, MA	Larkin and Gurry (1998)	www.bostonhigashi.com	3–22
Miller Method Newton, MA	Miller and Miller (1973)	www.millermethod.org	2–14
Division TEACCH (Treatment and education of autistic and related communication handicapped children)	Panerai et al. (2002)	www.teacch.com	Infant-adult

analysis procedures, such as prompting, reinforcement, and/or discrete trial training. Yet within these CTMs there were also some emphasized focused intervention strategies, such as the Pyramid Program's use of the Picture Exchange Communication System (PECS) (Bondy and Frost 2002) and photographic activity schedules used by the Princeton Child Development Institute. A third variety of applied behavior analysis CTMs occurred also in classroom settings, but we identified them as inclusive because they actively incorporated typically developing children into their teaching procedures. For example, both the LEAP and Walden models included interactions between young children with ASD and typical peers as well as a naturalistically oriented, behavioral framework. It should be noted that some of the models in the other classifications also had phases of the model that involved typically developing children, but their presence was not an integral feature of the entire model.

A fourth type of CTM was based on developmental or relationship-based conceptual frameworks. Often following a constructivist approach, these models drew from the developmental science literature to establish modes of interaction between adults/parents and (primarily young) children with ASD to promote the essential processes necessary for development and formation of social relationships. The Denver Model is an example of a model that has evolved over the years and takes place in an early childhood context, while the DIR (i.e., Floortime) model was initially applied in clinic and home settings, although is now expanding to school settings.

A fifth set of CTMs were so varied in their approach that they were difficult to categorize and are termed idiosyncratic models. For example, TEACCH, a historical CTM, is based on several theoretical frameworks, (i.e., social learning theory, developmental, and behavioral), while the Higashi model focuses on Daily Living Therapy, which emphasized physical exercise, emotional stability, and intellectual stimulation.

Evaluation Ratings

The ratings for each model are presented in Table 2. As noted, the ratings range from 0 to 5, with 5 being the highest. The descriptions of features of the CTM necessary for rating each domain may be found in the appendix.

Operationalization. As noted, operationalization refers to the degree to which the procedures and curriculum content of a model is specified well enough so that, with training, others could replicate it in their school system, classroom, home, or community. There was considerable variation among CTMs. About a third (12 out of 30) of the CTMs had specified both the procedures they follow and the content of the curriculum in clear details (score of 5). Thirteen other

Table 2 Comprehensive treatment models and evaluation ratings

Program	Operation- alization	Fidelity	Replication	Outcome data	Quality	Additional studies
Alpine	3	3	5	3	N/A	2
Autism partnerships	5	3	5	0	N/A	1
CARD	5	4	4	3	N/A	2
Children's toddler	2	3	1	5	3	2
DATA	3	1	5	3	N/A	2
Denver	5	4	5	5	2	0
DIR	5	3	5	4	2	0
Douglass	5	3	0	5	3	5
Eden	3	2	0	0	N/A	0
Hanen	2	0	1	3	N/A	2
Higashi	2	0	2	3	N/A	0
Institute for child. dev.	3	2	0	3	N/A	0
Lancaster	2	0	0	0	N/A	0
LEAP	4	5	5	4	2	5
Lovaas institute	5	4	5	5	3	5
May institute	5	4	5	4	2	5
Miller	3	1	5	4	0	1
PCDI	5	4	5	4	2	5
PRT	4	3	5	2	N/A	5
Pyramid	2	3	4	3	N/A	5
Responsive teaching	3	3	0	5	3	0
RDI	5	3	0	4	2	0
SCERTS	5	0	0	0	N/A	4
Son Rise	3	0	0	2	N/A	0
STAR	5	3	5	4	2	0
Summit	3	4	0	0	N/A	0
TEACCH	3	3	5	5	2	2
Therapeutic pathways	5	4	3	4	3	0
Valley	3	3	5	0	N/A	0
Walden	4	3	4	3	N/A	2

CTMs provide either specifically documented procedures or well described content, but not both (ratings 3–4). Only five CTMs received ratings of 2 and none scored lower.

Implementation. High fidelity of implementation is necessary if a CTM is to be successfully replicated. Quality fidelity measures should have some psychometric evidence of reliability and/or validity. Based on this criterion, only one CTM, the LEAP model, received the highest rating for this domain. Only seven of the CTMs were able to provide evidence of inter-rater agreement data from more than one supervisor who evaluated the intervention procedures by staff. Fifteen CTMs had at least some method of observing implementation practices that staff collected on a regular basis (e.g., completion of checklists). A minority of the programs ($N = 2$) indicated only informal assessment of fidelity (rating of 1) and several programs ($N = 5$; denoted

by 0) had no formal or informal methods of collecting fidelity data.

Model Replication. Nearly half the CTMs (14 out of 30) had three or more independent replications (denoted by a 5). Striking in its contrast, a substantial minority (9 out of 30) of the CTMs had no replications at all as denoted by a 0. It should be noted that if individuals using a model were trained by a central site and then implemented the model in their local program under the supervision of the original CTM developer, this was not counted as an independent replication. For example, RDI, a widely used model, does not have a replication listed, although there are practitioners all over the country practicing RDI, under the supervision of the model developer (i.e., practitioners video tape their performance, send it to RDI developers to evaluate, and receive yearly certification).

Outcome Data. This strand of the evaluation assessed the type of evidence on child outcomes provided and/or published by the model developers. To achieve the highest ratings, developers had to have published at least two articles documenting efficacy in refereed journals, while to achieve the next lower rating (a 4), developers had to publish one article. Nearly half of the CTM developers (14 out of 30) had published at least one article documenting efficacy of their model through child outcomes. Many other developers had published child outcome data in book chapters or books ($N = 10$). We were unable to find any published outcome data for several programs (6 out of 30), as denoted by a 0.

Quality. The SMRS, from the National Standards Project, provided a single rating score for a peer-reviewed journal article rounded to the nearest whole number, which ranged from 0 to 5. When there was more than one article we averaged scores across articles. The number of studies supporting CTMs ranged from eight for the Lovaas Institute to one study for seven of the programs. The Lovaas Institute, Douglass Developmental Disabilities Center, Responsive Teaching, Children's Toddler Program, and Therapeutic Pathways had the highest ratings, although the quality score for these models never rose above the 3 rating. Also, it is important to note that the quality evaluation only applied to peer-reviewed journal articles, and when no peer-review articles had been published, a rating of Not Applicable (N/A) was applied.

Additional Studies. Some CTM developers had published efficacy studies of focused interventions that make up components of their model. Seven of the CTMs had over 20 journal publications to document the effects of focused interventions, as indicated by a 5 rating. Also, notably, many CTMs (15 out of 30), had no or very few focused intervention studies (less than 6) associated with their model, indicated by 1 or 0 ratings.

Discussion

The purpose of this evaluation study was to determine the quality of multiple dimensions of CTMs that have been developed for learners with ASD and their families. Previous authors have conducted authoritative reviews of the research methodology and efficacy results for CTMs (Howlin et al. 2009; Reichow and Wolery 2009; Rogers and Vismara 2008). The current review compliments those rigorous efforts by providing information about the quality of the procedures and implementation, the number of replications of the model, and the associated evidence generated by efficacy studies of focused interventions that are components of the models.

Rather than generating a single summary score, this review established a profile across dimensions of the

evaluation. This profile depicts the CTMs that have the stronger and weaker evidence for each dimension. Models that have ratings of four or five across at least four dimensions of the evaluation have stronger evidence of what we will call "model development." That is, they are procedurally well documented, they have been replicated, and there is some evidence of efficacy, even if limited. These models include: Denver, LEAP, Lovaas Institute, May Institute, and PCDI. There are other models that have ratings of five or four on less than four domains, but there remain features of the model that are relative strengths. These include Autism Partnerships, CARD, Children's Toddler Program, DIR, Douglass, PRT, Responsive Teaching, SCERTS, and TEACCH. Also, there are some models that received very low evaluation ratings across the board. These included Hanen, Higashi, Eden, Summit, Lancaster, and Son Rise. Again, potential consumers may use this as a general guide when determining the degree of model development and level of efficacy evidence that exists for individual CTMs.

As a group, CTMs were strongest in the operationalization of their models. From an implementation science perspective, clearly stated procedures and materials are a requisite for high implementation (Durlak and DuPre 2008; Fixsen et al. 2005; Greenberg et al. 2005). Both Lord et al. (2005) and Smith et al. (2007) emphasized the necessity of manualization of treatment procedures for learners with ASD.

Given the strength of the operationalization dimension, it was surprising that measurement of implementation was relatively weak in comparison. Only a few CTMs had a tool to collect implementation data, and only one project had any evidence to share about the psychometric qualities of their implementation instrument. When CTMs are designed as models that may be adopted by service providers, the availability of implementation instruments would seem to be a critical feature of all CTMs because it would allow (a) the developer to observe and document the use of their model and (b) the adoptee to evaluate their own implementation. This is a much needed direction for future developers of CTMs.

As a group, CTM projects reported frequent replications by individuals outside of the central project. Nearly half the CTM developers (14 out of 30) reported that two or more independent sites had replicated their model. Some sites reported replication by dozens of programs (e.g., TEACCH, Lovaas, DIR, LEAP). With the increased demand for efficacious treatment models, a process for training and supporting implementation at replication sites is a direction needed for the future. Again, the emerging area of implementation science (Fixsen et al. 2005) may well assist in this process. Current models for implementation (Domitrovich et al. 2008; Fixsen et al. 2005)

are built on ecological systems conceptual frameworks that require not only demonstration of efficacy but supports provided at different levels of the system (e.g., the local administrative and larger program levels). Scaling up of CTMs however, may well be limited by the development of implementation assessment. Some CTMs did report replication of their models by service providers using brief or no instrumentation for assessing implementation, which should be a concern for the field.

Across CTMs, the published evidence of efficacy was not strong. Over half (16 out of 30) of the CTMs had no publication of efficacy in peer reviewed journals, although often they did report efficacy data in book chapters or from their program reports. Some models (Lovaas Institute, Responsive Teaching, Children's Toddler School, TEACCH) had at least two publications of efficacy in peer-reviewed journals, and among these the Lovaas Institute has the most extensive record of publications, which has also been documented by Howlin et al. (2009) and Reichow and Wolery (2009). Yet, the quality of the experimental methodology even for these aforementioned CTMs never rose into the upper levels of our rating range. It is important to note, however, that because this evidence does not exist does not mean that the CTM is ineffective. Often, the lower quality rating was the result of CTM researchers not employing RCT methodology. Although RCT methodology has been identified as the gold standard in efficacy research (Lord et al. 2005; Smith et al. 2007), Rogers and Vismara (2008) noted the difficulty of conducting RCTs with this population. This is particularly true for classroom-based interventions, in which classrooms or even schools are the units of analysis that are randomly assigned to conditions, rather than individual children (i.e., to account for nesting within groups). These are expensive studies to conduct and only recently (i.e., the 2000s) have governmental agencies been willing to provide funds at the magnitude needed to conduct such rigorous efficacy research. As a result, a number of the CTMs are now being evaluated through RCTs or rigorous quasi-experimental designs (e.g., Denver, LEAP, TEACCH, SCERTS, STAR). We anticipate that data documenting the efficacy of CTMs will be forthcoming in the near future. But, even with the anticipation of new data becoming available, there may be concern that wide-spread replication and use of some CTMs already is occurring when there is little evidence that effective implementation can be documented on a reliable basis.

This evaluation of CTMs was limited by several factors. First, although the authors made extensive efforts to locate CTMs meeting their criteria, it is possible that some models were missed. Second, one dimension of the evaluation system is labeled replication, which may lead to

some confusion because of the slightly different meaning of the term in experimental research. "Adoption" may have been a better term, but in fact when service providers adopt a CTM, their intent is to replicate the effects of the original model, so we decided to stay with the original descriptor. Third, in this study, the evaluators participated in the development of the evaluation instrument and criteria; involving additional evaluators who had not been involved in the instrument development process may have strengthened the study. Replicating this evaluation process with a larger set of evaluators could be a useful direction for the future. Fourth, this is a very active literature. Even when using the most current information, it is possible new information will become available in the near future. It would be important to view the findings of this evaluation with the perspective that new information may supplement or extend these findings.

Last, it is only fair to note that some of the CTMs operate in public school districts, (e.g., Lancaster, Valley), and we included them in our evaluation because a description of their model appeared in a chapter or other publication. It may well be that their primary intention is to implement evidence-based practices and provide high quality services to the children they serve, and they might not have the funds necessary for the expensive work of conducting efficacy research or validating implementation measures. A set of low ratings does not mean they, or any of these CTMs, necessarily are ineffective models; they just have not published their procedures or models for external review.

In conclusion, this evaluation reveals that there are a set of CTMs that have fairly well established evidence for their model, others with mixed evidence, while still others with very weak evidence. This information may be used by researchers, service providers, and families to make decisions about selection of a CTM for use with individual children or even for use in future research. The next steps for the field will be in the development of measures that can document the appropriate implementation of CTM procedures, the further publication of research to document the efficacy of CTMs, and the use of both sets of information to "scale-up" the use of efficacious practices for learners with ASD and their families.

Acknowledgments This paper was based on a larger report funded by the California Department of Developmental Services and its Autism Spectrum Disorders: Guidelines for Effective Intervention project.

Appendix

See Table 3.

Table 3 Summative evaluation of evidence for comprehensive treatment models

Rating	Operationalization	Implementation measurement	Replication at remote sites	Outcome data presented	Average quality rating from outcome data	Additional studies on related focused interventions
5	Written document(s) that describe procedures for the treatment and content in enough detail to replicate the model	Robust psychometric data have been collected on treatment fidelity measures. This would include evaluating reliability (e.g., internal consistency, test-retest), inter-rater agreement and validity (e.g., construct/content, convergent) of the instrument	Model replicated at two or more remote locations (i.e., not the original model development site) by two or more investigators who were not part of the original model development team	Two or more <i>experimental</i> referred journal articles analyzing efficacy of the project	5.0–4.01	21+ studies conducted on focused interventions that are components of the comprehensive treatment model
4	Written document(s) in which procedures are well specified and content is described but poorly specified	Preliminary psychometric data have been collected on treatment fidelity measure. This would include the basic elements of reliability (i.e., inter-rater)	Model replicated at one remote location by one or more investigators who were not part of the model development team	One <i>experimental</i> referred journal article analyzing efficacy of the project	4.0–3.01	16–20 studies conducted on focused interventions that are components of the comprehensive treatment model
3	Written document(s) in which content is clearly specified and procedures are described but poorly specified	Formal or standardized treatment fidelity measure exists (e.g., checklists outlining model components) in a form that would allow the model implementer or other outside observer to evaluate implementation of essential model components, but no psychometric data have been collected on instrument	Model replicated by original model developers at two or more remote locations	A book chapter in a book edited by someone other than the model developer describing outcome data or a journal article providing descriptive data about child outcome for the project	3.0–2.01	11–15 studies conducted on focused interventions that are components of the comprehensive treatment model
2	Some written documentation provided that describes procedures and content but neither procedures nor content sufficiently specified	Formal measure only exists for some but not all essential model components	Model replicated by original model developers at one remote location	A book written by model developer(s) with section or chapter on outcome study	2.0–1.01	7–10 studies conducted on focused interventions that are components of the comprehensive treatment model
1	Some written documentation provided that describes EITHER procedures OR content but is not sufficiently specified	No formal or codified measure of treatment fidelity exists and only informal measures are used	Model currently implemented by original model developer at original site	A chapter in a book edited by model developer describing outcome data	1.00–0.01	3–6 studies conducted on focused interventions that are components of the comprehensive treatment model

Table 3 continued

Rating	Operationalization	Implementation measurement	Replication at remote sites	Outcome data presented	Average quality rating from outcome data	Additional studies on related focused interventions
0	Only general description of the model is available	No formal or informal measure of treatment fidelity exists	No information about model replication is provided	No outcome data published or outcome data published only in report form by model developer	0.00	Less than three studies conducted on focused interventions that are components of the comprehensive treatment model

References

- American Evaluation Association. (2008). Guiding principles for evaluators. *American Journal of Evaluation*, 29, 233–234.
- Anderson, S. R., Thomeer, M. L., & King, D. C. (2006). Summit Academy: Implementing a system-wide intervention. In S. Handleman & S. Harris (Eds.), *School-age programs for children with autism* (pp. 195–220). Austin, TX: PRO-ED.
- Arick, J., Young, H., Falco, R., Loos, L., Krug, D., Gense, M., et al. (2003). Designing and outcome study to monitor the progress of students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 18, 75–87.
- Bondy, A., & Battaglini, K. (2006). Application of the Pyramid approach to education model in a public school setting. In S. Handleman & S. Harris (Eds.), *School-age programs for children with autism* (pp. 163–194). Austin, TX: PRO-ED.
- Bondy, A., & Frost, L. (2002). *A picture's worth: PECS and other visual communication strategies in autism*. Bethesda, MD: Woodbine House.
- Bruce, C. T., & Vorhis, N. (2006). The Lancaster-Lebanon IU 13 autism support program. In S. Handleman & S. Harris (Eds.), *School-age programs for children with autism* (pp. 115–142). Austin, TX: PRO-ED.
- Campbell, S., Cannon, B., Ellis, J. T., Lifter, K., Luiselli, J. K., Navalta, C. P., et al. (1998). The May Center for early childhood education: Description of a continuum of services model for children with autism. *International Journal of Disability, Development and Education*, 45, 173–187.
- Cohen, H., Amerine-Dickens, M., & Smith, T. (2006). Early intensive behavioral treatment: Replication of the UCLA Model in a community setting. *Developmental and Behavioral Pediatrics*, 27(2), 145–155.
- Cronbach, L. J. (1982). *Designing evaluations of educational and social programs*. San Francisco, CA: Jossey-Bass.
- Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoogwood, K., Buckley, J. A., Olin, S., et al. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advance in School Mental Health Promotion*, 1, 6–28.
- Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327–350.
- Fenske, E. C., Zalenski, S., Krantz, P. J., & McClannahan, L. E. (1985). Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis and Intervention in Developmental Disabilities*, 5, 49–58.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).
- Greenberg, M. T., Domitrovich, C. E., Graczyk, P. A., & Zins, J. E. (2005). *The study of implementation in school-based preventive interventions: Theory, research, and practice. Promotion of mental health and prevention of mental and behavior disorders* (Vol. 3). Washington, DC: U.S. Department of Health and Human Services.
- Gutstein, S. E., Burgess, A. F., & Montfort, K. (2007). Evaluation of the Relationship Development Intervention. *Autism*, 11, 397–411.
- Hall, L. (2009). *Autism spectrum disorders: From theory to practice*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- Handleman, J. S., & Harris, S. L. (Eds.). (2006). *School-age education programs for children with autism*. Austin, TX: PRO-ED.
- Handleman, J. S., & Harris, S. L. (Eds.). (2008). *Preschool education programs for children with autism*. Austin, TX: PRO-ED.
- Harris, S., Handleman, J., Gordon, R., Kristoff, B., & Fuentes, F. (1991). Changes in cognitive and language functioning of preschool children with autism. *Journal of Autism and Developmental Disorders*, 21, 281–290.
- Hassen, C. E., Lawrenz, F., & Dunet, D. O. (2008). Concurrent meta-evaluation: A critique. *American Journal of Evaluation*, 29, 572–582.
- Holmes, D. (1998). *Autism through the lifespan: The Eden model*. Bethesda, MD: Woodbine House.
- 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.
- Howlin, P., Magiati, I., & Charman, T. (2009). Systematic review of early intensive behavioral interventions for children with autism. *American Journal of Intellectual and Developmental Disabilities*, 114, 23–41.
- Hoysen, M., Jamieson, B., & Strain, P. S. (1984). Individualized group instruction of normally developing and autistic-like children: The LEAP curriculum model. *Journal of the Division for Early Childhood*, 8, 157–172.
- Hume, K., & Odom, S. L. (2007). Effects of a work system on the work and play of children and individuals with autism. *Journal of Autism and Developmental Disorders*, 37, 1166–1180.
- Kaufman, B. (1981). *A miracle to believe in*. New York: Fawcett Crest.

- Keenan, M., Henderson, M., Kerr, K. P., & Dillenburger, K. (2006). *Applied behaviour analysis and autism: Building a future together*. London: Jessica Kingsley Publishers.
- Koegel, R., & Koegel, L. (2006). *Pivotal response treatments for autism*. Baltimore, MD: Paul H. Brookes.
- Koegel, L. K., Koegel, R. L., Harrower, J. K., & Carter, C. M. (1999). Pivotal response intervention I: Overview of the approach. *Journal of the Association for Persons with Severe Handicaps*, 24(24), 174–185.
- Larkin, A., & Gurry, S. (1998). Brief report: Progress reported in three children with autism using daily life therapy. *Journal of Autism and Developmental Disorders*, 4, 339–342.
- Leaf, R., & McEachin, J. (1999). *A work in progress: Behavior management strategies and a curriculum for intensive behavioral treatment of autism*. New York: DRL Books.
- Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., et al. (2005). Challenges in evaluating psychosocial interventions for autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, 35, 695–708.
- Mahoney, G., & Perales, F. (2005). Relationship-focused early intervention with children with pervasive developmental disorders and other disabilities: A comparative study. *Journal of Developmental and Behavioral Pediatrics*, 26(2), 77–85.
- McConachie, H., Randle, V., Hammal, D., & LeCouteur, A. (2005). A controlled trial of a training course for parents of children with suspected autism spectrum disorder. *The Journal of Pediatrics*, 33, 5–340.
- McGee, G. G., Morrier, M. J., & Daly, T. (1999). An incidental teaching approach to early intervention for toddlers with autism. *The Journal of the Association of Persons with Severe Handicaps*, 24, 133–146.
- McKeon, J., Vuoncino, K., Brenkert, R., Dinnell-LoPresti, E. D., Lampert, M., Madden-Perez, M., et al. (2006). The valley program. In S. Handleman & S. Harris (Eds.), *School-age programs for children with autism* (pp. 247–268). Austin, TX: PRO-ED.
- Meyer, L. S., Taylor, B. A., Cerino, K. E., Fisher, J. R., Moran, L., & Richard, E. (2006). Alpine learning group. In J. S. Handleman & S. L. Harris (Eds.), *School-age education programs for children with autism* (pp. 19–47). Austin, TX: Pro-Ed.
- Miller, A., & Miller, E. E. (1973). Cognitive-developmental training with elevated boards and sign language. *Journal of Autism and Childhood Schizophrenia*, 3, 65–85.
- National Research Council. (2001). *Educating children with autism*. Washington, DC: National Academy Press.
- Odom, S. L., Brown, W. H., Frey, T., Karasu, N., Smith-Carter, L., & Strain, P. (2003). Evidence-based practices for young children with autism: Evidence from single-subject research design. *Focus on Autism and Other Developmental Disabilities*, 18, 176–181.
- Odom, S., Collet-Klinenberg, L., Rogers, S., & Hatton, D. (in press). Evidence-based practices in interventions for children and youth with autism spectrum disorders. *Preventing School Failure*.
- Panerai, S., Ferrante, L., & Zingale, M. (2002). Benefits of the treatment and education of autistic and communication handicapped children (TEACCH) programme [sic] as compared with a non-specific approach. *Journal of Intellectual Disability Research*, 46, 318–327.
- Prizant, B., Wetherby, A., Rubin, E., Laurent, A., & Rydell, P. (2006). *The SCERTS model: A comprehensive educational approach for children with autism spectrum disorders*. Baltimore, MD: Paul H. Brookes Publishing.
- 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.
- Reichow, B., Volkmar, F., & Cicchetti, D. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders*, 38, 1311–1319.
- Rogers, S., & Vismara, L. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology*, 37, 8–38.
- Rogers, S. J., Hayden, D., Hepburn, S., Charlfue-Smith, R., Hall, T., & Hayes, A. (2006). Teaching young nonverbal children with autism useful speech: A pilot study of the Denver model and PROMPT interventions. *Journal of Autism and Developmental Disorders*, 36, 1007–1024.
- Romanczyk, R. G., Lockshin, S. B., Matey, L., & Gillis, J. M. (2006). The children's unit for treatment and evaluation. In S. Handleman & S. Harris (Eds.), *School-age programs for children with autism* (pp. 49–89). Austin, TX: PRO-ED.
- Schwartz, I., Sandall, S., McBride, B., & Boulware, G. (2004). Project DATA (Developmentally Appropriate Treatment for Autism): An inclusive school-based approach to educating young children with autism. *Topics in Early Childhood Special Education*, 24, 156–168.
- Smith, T., Scahill, L., Dawson, G., Guthrie, D., Lord, C., Odom, S., et al. (2007). Designing research studies on psychosocial interventions in autism. *Journal of Autism and Developmental Disorders*, 37, 354–366.
- Solomon, R., Necheles, J., Ferch, C., & Bruckman, D. (2007). Pilot study of a parent training program for young children with autism. *Autism*, 11, 205–224.
- Stahmer, A. C., & Ingersoll, B. (2004). Inclusive programming for toddlers with autism spectrum disorders: Outcomes from the Children's Toddler School. *Journal of Positive Behavior Interventions*, 6, 67–82.
- Stevens, M., Washington, A., Rice, C., Jenner, W., Ottolino, J., Clancy, K., et al. (2007). *Prevalence of the autism spectrum disorders (ASDs) in multiple areas of the United States, 2000 and 2002*. Atlanta, GA: Centers for Disease Control and Prevention.
- Volkmar, F., Paul, R., Klin, A., & Cohen, D. (2005). *Handbook of autism and pervasive developmental disorders*. Hoboken, NJ: Wiley.
- Yell, M. L., Drasgow, E., & Lowrey, A. K. (2005). No child left behind and students with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, 20(3), 130–139.