

Challenges in Evaluating Psychosocial Interventions for Autistic Spectrum Disorders

Catherine Lord,^{1,24} Ann Wagner,² Sally Rogers,³ Peter Szatmari,⁴ Michael Aman,⁵ Tony Charman,⁶ Geraldine Dawson,⁷ V. Mark Durand,⁸ Lee Grossman,⁹ Donald Guthrie,¹⁰ Sandra Harris,¹¹ Connie Kasari,¹² Lee Marcus,¹³ Susan Murphy,¹⁴ Samuel Odom,¹⁵ Andrew Pickles,¹⁶ Lawrence Scahill,¹⁷ Evelyn Shaw,¹⁸ Bryna Siegel,¹⁹ Marian Sigman,²⁰ Wendy Stone,²¹ Tristram Smith,²² and Paul Yoder²³

In 2002, the National Institutes of Health sponsored a meeting concerning methodological challenges of research in psychosocial interventions in Autism Spectrum Disorders. This paper provides a summary of the presentations and the discussions that occurred during this meeting. Recommendations to federal and private agencies included the need for randomized clinical trials of comprehensive interventions for autism as the highest, but not the sole priority. Ongoing working groups were proposed to address psychosocial interventions with a focus on relevant statistics, standardized documentation and methods of diagnosis, development of outcome measures, establishment of standards in research; and the need for innovative treatment designs, including application of designs from other research areas to the study of interventions in ASD.

In September 2002, a National Institutes of Health-sponsored meeting concerning methodological challenges of research in psychosocial interventions in autism took place over a day and a half. The purpose of the meeting was to convene a group

of investigators and stakeholders to inform the National Institutes of Health and other relevant government agencies about the current state of the science with regard to (a) psychosocial, behavioral and educational interventions for children with

¹ University of Michigan, Ann Arbor, USA.

² Child & Adolescent Psychosocial Intervention Program – NIMH, Bethesda, MD, USA.

³ University of California – Davis, Tulare, USA.

⁴ McMaster University, Hamilton, Canada.

⁵ Ohio State University, Columbus, USA.

⁶ Institute of Child Health, University College London, London, UK.

⁷ University of Washington, Seattle, USA.

⁸ University of South Florida, St. Petersburg, USA.

⁹ Autism Society of America, Bethesda, MD, USA.

¹⁰ University of California, Los Angeles, USA.

¹¹ University of New Jersey Rutgers, New Brunswick, USA.

¹² University of California, Los Angeles, USA.

¹³ University of North Carolina, Chapel Hill, USA.

¹⁴ University of Michigan, Ann Arbor, USA.

¹⁵ Indiana University, Indianapolis, USA.

¹⁶ University of Manchester, Manchester, UK.

¹⁷ Yale University, New Haven, USA.

¹⁸ NECTAS, University of North Carolina, Chapel Hill, USA.

¹⁹ University of California, San Francisco, USA.

²⁰ University of California, Los Angeles, USA.

²¹ Vanderbilt University, Nashville, USA.

²² University of Rochester, Rochester, USA.

²³ Kennedy Center, Vanderbilt University, Nashville, USA.

²⁴ Correspondence should be addressed to: Catherine Lord, University of Michigan, Autism and Communication Disorders Center 1111, East Catherine Street, Ann Arbor, MI, 48109, USA; e-mail: celord@umich.edu

autistic spectrum disorders (ASD), (b) barriers to progress in the field, and (c) potential strategies for overcoming those barriers. Presentations and discussion by parent advocates, professional consumers, and investigators addressed these issues from various perspectives.

The conference began with participants providing brief presentations concerning the state of the science in five areas. Each presentation was followed by comments by one or two discussion facilitators who then led the discussion in that area. The five areas consisted of (1) a general update on how far we have come and where we need to go; (2) recruitment challenges and solutions; (3) measurement of outcome; (4) challenges in methodology; and (5) strategies about how to move from the lab to the community and from efficacy to effectiveness. On the second day of the meeting, discussion focused on strategies for facilitating collaboration, increasing the quality of research, partnering with families and communities, and identifying ways in which governmental support could be used to facilitate progress.

WHAT ARE UNIQUE ISSUES IN CONSIDERING THE PSYCHOSOCIAL TREATMENTS OF ASD?

Although the purpose of the meeting was to address issues specifically related to psychosocial interventions for individuals with Autism Spectrum Disorders (ASD), one of the secondary goals was to apply knowledge from other types of treatment research to our field. Thus, many of the ideas discussed here are relevant to other disorders. Nevertheless, there are a number of features of ASD that are unique or at least, that differ from many other psychiatric disorders beginning in childhood. There is tremendous variability in ASD, in part because the disorders are defined by a pattern of behaviors that are quite different from each other (e.g., very basic social behaviors, unusual motor patterns, specific language characteristics) and in part, because of the frequent, but not universal association of ASD with expressive and receptive language delays and mental retardation. This variability means that a comprehensive assessment is a critical part of treatment in ASD in order to define the characteristics and ability levels, as well as the needs of individuals (Howlin, 1998).

The deficits in very basic skills that are usually present in typical infants and toddlers that define ASD (Klin, Volkmar, & Sparrow, 1992), the

pervasiveness of these deficits across contexts, and the very early onset noted in most children also affect the design and selection of treatments, in that interventions are needed that are as comprehensive as the disorders are pervasive, and that begin as early as the disorders are recognized. Thus, the age suggested for optimum intervention to begin for autism is far younger than for children with Attention Deficit Hyperactivity Disorder, anxiety or depression. Moreover, research suggests that early intervention can result in significant improvements for some children with autism, though for how many children and for how much gain has varied considerably across studies. (Anderson *et al.*, 1987; Birnbrauer, 1997; Dawson & Osterling, 1997; Fenska, Krantz, & McClannahan, 1985; Green *et al.*, 2002; Harris *et al.*, 1991; Lovaas, 1987; McEachin, Smith, & Lovaas, 1993; Rogers, 1998; Sheinkopf & Siegal, 1998). Traditional 1 h a week treatments for language or social skills or behavior used in the US mental health system are rarely sufficient to produce generalizable improvements in core areas of ASD (Howlin, 1998; Wetherby, Schuler, & Prizant, 1997). On the other hand, comprehensive treatments are extremely expensive to carry out and require large numbers of highly trained staff and infrastructures that greatly limit who can study the effects of intervention on groups of children. Although other disorders also affect learning and social opportunities, the degree to which the deficits associated with autism affect and are affected by other areas of development is generally greater, which also contributes to the need for comprehensive treatments (Lord & Bailey, 2002).

There is currently no medical or biological treatment of the core features of ASD, though there are medications that ameliorate certain aspects of behavior or mood, such as activity level (Aman, 1996) or irritability (Arnold *et al.*, 2003). Thus, the primary source of intervention for most children is through their families and through the educational system. The majority of research papers about treatment of individuals with ASD is from a behavioral perspective, but there are a number of widely-used programs with alternative models, including the TEACCH program (Schopler, Mesibov, Shigley, & Bashford, 1984) that often include both developmental and behavioral components (Harris & Weiss, 1998). In general, there is a wide gap both between studies of very specific techniques (such as timing in prompting conversational skills or replacement of a repetitive behavior with a more appropriate action (Durand & Carr, 1987) and outcomes that range

Challenges in Evaluating Psychosocial Interventions

from very specific (e.g., asking a certain kind of question) to very general (e.g., full inclusion in a regular school program; looking indistinguishable from normal). None of these issues is completely unique to ASD, but, taken together, they mean that research designs about treatments must be approached with care and creativity.

Below, the presentations and discussions are summarized briefly within the headings of the meeting.

State of the Art Evidence-based Treatment: How Far have We Come and How Far do We Need to Go?

A great deal of knowledge about psychosocial treatments in ASD exists. Nevertheless, this knowledge represents only the beginning of understanding why and how best to design interventions for individuals with ASD. Providing intervention for individuals with autism, most often across the life-span, is an expensive proposition for most children and adults. Recent reports, based on Lovaas' original study (1987), suggested that improvements produced by early intervention in Autistic Spectrum Disorder may result in cost-saving to society of nearly \$200,000 by the time an individual reaches age 22 years and a million dollars by the time an individual was age 55 years (Columbia Pacific Consulting, 1999; Jacobson, Mulick & Green, 1998). Cost effectiveness of treatment has not yet been broken down by individual differences in response to various interventions.

Most group-design research in autism has concerned comprehensive curricula. Consequently, it has been difficult to identify the effects of components of these programs (see Dawson & Osterling, 1997; Kasari, 2002; National Research Council, 2001; Rogers, 1999) or the underlying mechanisms for change. Research in developmental psychology suggests that sensitivity to sequences of typical development and to individual differences, including transactions (e.g., between the individual with ASD and parents or family members or teachers, or between higher level organizational factors such as school systems and families) is important (Dunst, 1999; Rogers, Hall, Osaki, Reaven, & Herbison, 2001; Schuler, Prizant & Wetherby, 1997).

Many single-subject designs have already indicated that specific interventions are associated with learning in ASD. A number of different models of interventions have shown significant improvements in a subset of children, but rarely have they resulted in complete amelioration of social difficulties (see Dawson & Osterling, 1997; National Research Council,

2001; Rogers, 1999). The question then becomes whether a particular treatment results in more improvement than another treatment. Once this is demonstrated, one might ask (a) why this is the case, (b) how much more improvement does it offer, (c) what are the consequences of this difference, (d) with whom is it evident, and (e) how long does this effect last and how generalizable is it? A lack of well-developed theories about how treatment effects are mediated has left the field with some sense of predictors of outcome, but little knowledge about how and why improvements come about.

For example, in most longitudinal studies, early IQ scores have played a strong role in predicting independence and achievement (Howlin, Mawhood, & Rutter, 2000; Venter, Lord, & Schopler, 1992). Early social-cognitive functioning has also predicted language growth in autism (Sigman *et al.*, 1999). Longitudinal and intervention studies have shown that individuals who are more mildly affected by autism, defined by severity of symptoms (e.g., those diagnosed with PDD-NOS versus autism) and higher intellectual level and more typical early social-cognitive functioning (Bartak & Rutter, 1973; Sheinkopf & Siegel, 1998; Smith, Groen, & Wynn, 2000b; Yirmiya & Sigman, 1991) show better outcomes.

However, such conclusions, for example, that initial IQ score and severity of autism are predictors of response to treatment, are based on studies of development, often within a treatment group. These studies do not necessarily inform us about response to treatment, but tell us about the relationship between child characteristics and developmental trajectories. To understand response to treatment, one needs to know if a difference between treatment groups varies as a function of pre-treatment characteristics, that is, is there a statistical interaction between the characteristics and treatment? (i.e., do children with ASD and better verbal skills show greater improvements than we would expect, given their stronger language skills, when they receive a particular treatment than if they had no treatment, compared to children with weaker verbal skills?)

In psychopharmacological studies of children with autism (Owley *et al.*, 2001), placebo effects were quite strong. They demonstrate the important effects of study enrollment and the perception of receiving a potentially effective treatment. Yet there are only three published studies that have used a randomized control design to determine intervention outcome in response to psychosocial treatments (Drew *et al.*, 2002; Jocelyn, Casiro, Beattie, Bow, & Kneisz, 1998;

Smith *et al.*, 2000a, b). In the Smith and colleagues' study of a comprehensive program of the Lovaas treatment model, at follow-up, the intensive treatment group had a statistically significant advantage over a parent-training group in IQ and visual-spatial skills, but not adaptive behavior or language. Within both groups, there were large individual differences in response. Though the study was a randomized group experiment, it also looked for differences between "best-outcome" children and "non-best-outcome" children. Interpreting these differences as "predictors of treatment response" requires the assumption that the only reason for the best-outcome results is due to the treatment, which may not have been the case.

As discussed later, this lack of information has led some investigators to call for Randomized Controlled Trials (RCTs) with sufficient numbers of participants that moderator variables, such as environmental features, child variables, and family factors, can be studied. Such trials would deliberately manipulate intensity and address a broader range of outcomes than has typically been studied (see below). To adequately address mediator and/or moderator variables, there need to be at least 10 participants per variable, which would require trials coordinated across a number of sites, as have occurred in other national cross-site studies, such as treatment trials for cancer, depression, and ADHD.

In much of medicine, the view of a single RCT as providing a completely decisive answer is now being replaced by one in which a study is seen as contributing to a larger body of evidence, by meeting standards that allow inclusion in a systematic review (McLeod & Weisz, 2004). All treatment studies should be designed in such a way that they can contribute to a greater program of knowledge accumulation beyond whether a particular treatment is more effective than another treatment in a particular sample. This expectation implies attention to eventual replicability, including treatment manuals, standard subject selection and measures, and some links between theory and predictions in all studies. Methodologies such as partial factorial designs may be particularly relevant to ASD.

There was much discussion regarding the need for well-monitored control groups and the possibility that "no treatment" controls may be less appropriate than an alternative treatment group because of the known effects of positive expectations and hope on the examiners, parents, and the children. Nevertheless, very little is known about the effects of intervention on parents, or about the accessibility and feasibility,

particularly the affordability, of the most well documented models in public schools or public health settings. A minority considered the continued use of no treatment control groups justified because such a design could detect whether new treatments slow down or had negative effects on behavior or development. A further complexity, however, is that a "no treatment" control raises ethical concerns, particularly if the study period is lengthy, as it typically has been in early intervention studies. In addition, a considerable literature already exists in autism (though primarily not formal RCTs), showing that some treatment is better than treatment with no evidence of negative effects of treatment (see Dawson & Osterling, 1997; National Research Council, 2001; Rogers, 1999), at least in early childhood interventions.

Many single-subject multiple baseline designs in older children and adults have also shown positive effects of treatment, though single subject designs also do not necessarily indicate if development was "slowed" by treatment. In small studies, there is the need to consider case-wide influence statistics to ensure that effects were not due to a small number of unusual individuals. Smaller studies including single subject designs and detailed qualitative analyses most appropriately introduce innovative treatments and may raise important questions about factors that maintain gains made in response to the more comprehensive interventions.

Reviews of 48 different journals in research carried out in the 1990s revealed 835 articles about psychosocial and educational intervention, which included 185 single-subject designs. However, studies that included children with autism were relatively rare, accounting for only about 10% of the participants. As of 2002, a total of only 30 papers concerned children under age 6 with ASD. Many of the practices most strongly supported by published papers, for example, positive behavioral supports, have not been studied specifically for children with autism, though there is reason to believe that these strategies may be effective (Horner, Carr, Strain, Todd & Reed, 2002). The most common designs are multiple baseline studies in which a treatment is introduced after varying lengths of baseline assessment. Topics addressed in these studies included individualized interventions, manipulation of consequences, use of naturalistic teaching approaches, peer-mediated intervention, use of prompting procedures, functional assessment, a range of measures involving changes in environment (including physical space), and the development of many different skills.

Challenges in Evaluating Psychosocial Interventions

The largest collections of studies about children with autism to date concern reduction of behavior problems as communication increases, peer-mediation and the use of prompting procedures.

One single-subject methodology, the reversal design, involves systematically introducing and withdrawing a treatment two or more times to test the effects of the treatment on the occurrence of a particular behavior. Reversal designs are useful for examining how to modify already acquired behaviors that vary primarily in frequency. However, because they require repeatedly withdrawing treatment, they are usually inappropriate for studying how to develop new skills. They are also inappropriate when teaching developmentally important skills in part because, if the skill is well-chosen, then typical environmental reinforcers will support the skill and withdrawal of the study intervention would not necessarily withdraw all the reinforcement. For this reason, although reversal designs are easily interpreted, they have had limited value for many of the developmental outcome goals addressed in autism.

Single subject reversal trials are often treated as multiple baseline designs when only the first withdrawal of treatment is interpreted. However, such designs demonstrate obvious effects, only when targeted behaviors change abruptly with the onset of the treatment, something that is rare in children with autism when developmentally important dependent variables are used. On the other hand, single-subject design approaches have paid more attention to generalization, maintenance, and social validity than other approaches in the field of autism (see National Research Council, 2001). Single-subject designs continue to offer strengths as initial tests of innovative treatment when individual differences and general effects or major changes in development are not an issue. We know relatively little about generalization and maintenance of the effects of comprehensive programs, though single-subject designs have increasingly included measurement of these aspects of change (Goldstein, 2002; McConnell, 2002). More recently, some studies of comprehensive behavioral interventions have discussed short-term changes as well as more general outcomes (Eikeseth, Smith, Jahr, & Eldevik, 2002).

A number of discussion points were raised. Because almost all studies of ASD treatments used behavioral techniques (Lovaas, 1987) with some attention to developmental sequences and communicative intent (Ozonoff & Cathcart, 1998; Wetherby *et al.*, 1997), the content of the interventions and the

fidelity and intensity with which they were carried out often varied more than the methods used. Differences in effectiveness, therefore, may have been due to many factors in addition to or besides the targeted treatment comparison. Although clinical observation suggests that several approaches (e.g., developmental, behavioral, structured teaching) are often used together in practice in ASD, the most effective ways to integrate developmental and other approaches with behavioral methods has not yet been tested systematically. Treatments vary as to whether they are addressing a specific behavior or symptom, or whether they address the “syndrome” of autism. In recent collaborative psychopharmacological clinical trials, investigators agreed to focus interventions on symptoms such as irritability or aggression and not to “oversell” treatments as cures for the syndrome of autism (Arnold *et al.*, 2000; Scahill & Lord, 2004). Intervention studies using the Applied Behavior Analysis (ABA) or Early Intensive Behavior Intervention (EIBI) have not generally specified what measures are expected to change (e.g., core symptoms of autism, degree of cognitive impairment, functional use of language) nor have they followed standard practices in clinical trials, such as indicating the effect size that would be considered a significant change.

Recruitment

Discussion of recruitment dealt generally with how to insure that results from individual studies could be interpreted in terms of more general populations. Three levels of challenges were outlined: (1) How to identify children and adults to participate in autism studies, (2) How to arrange for participation in a way that affects families positively and (3) How to maintain good community relations. Different motivations for families to participate in research were identified including (a) access to direct, better, or free services; (b) helping the “greater good,” (c) financial incentives, (d) and desperation (in the case of families seeking any new treatment when little else has been helpful).

Various recruitment strategies were discussed. A number of recruitment options were proposed, including (1) recruiting only children with existing diagnoses from established diagnostic centers, (2) recruiting children at risk or suspected of ASD and providing diagnoses through the research—a strategy that can become very expensive, but results in more standardized diagnoses or (3) establishing a feeder diagnostic clinic from which to recruit. The pros and

cons to the alternatives were discussed. Large studies demand a focused, dedicated recruitment effort. This cannot be done by mail; personal approaches are necessary. Recruitment aspects of intervention research must be a priority in autism research, meriting a dedicated financial investment. Project coordinators/recruiters and principal investigators must be willing to take initiative in finding participants, with one person or one arm of a project “responsible” for recruitment.

Specific recommendations for responsible inclusion of participants with autism were made, including offering written reports of assessments, flexibility in scheduling, providing clinical interpretation of diagnostic and developmental results, and building into research planning family support, access to expert consultation, or help in identifying quality resources. Families must perceive all treatment options as approaches that are strongly recommended by the investigators. It is difficult to keep families in a treatment-as-usual group and impossible to separate the effects of hope, enthusiasm, confidence and often training of the interveners from specific aspects of the treatment (see Eikeseth, Smith, Jahr & Eldevik, 2002), if deliberate plans are not made to take these factors into account.

Although in recent years, the bulk of parent advocacy for research has supported biomedical studies, there are groups of parents who are strongly committed to addressing the need for psychosocial and behavioral research. From a parent’s perspective, the possibility of being assigned to a placebo control group rather than treatment is a concern. Time commitments, needs for excessive documentation, and questions about the family’s role in experimental research (“being a guinea pig”) are also of importance. Researchers were urged to recognize that most individuals with autism will be receiving multiple interventions and this must be taken into account within research designs.

Recommendations beyond having a focused recruitment plan included having experienced clinicians on research teams and promoting community awareness of early signs of autism, both through personal contact with referral sources such as pediatricians, general practitioners, and local practitioners including speech and language pathologists, audiologists or occupational therapists, and through collaboration with existing early intervention settings. The ethical aspects of families’ needs for access to services and clinical information must also be considered. Another practical recommendation was that studies

keep standard data from screening contacts so that records are kept about which individuals participated and did not participate in the research. It was noted that, in RCTs for a complex disorder such as autism, there should be as much attention paid to the control group as to the experimental treatment group. Furthermore, it is clear that most children and families who participate in treatment research do not represent the general population. Information about which families took part in different studies and which did not (and why or why not) could provide invaluable information concerning the eventual potential effectiveness of a treatment for ASD.

Outcome Measures

Outcome measures can be grouped according to whether they are measured close in time to the intervention or in longer-term follow-up, whether they are near or distant to the treatment in terms of the conceptualization, and to what degree, inferences about specific effects of a treatment can be drawn (Charman & Howlin, 2003; Drew *et al.*, 2002). A distinction can be made between skills directly taught by the treatment and general measures of development likely to be affected by these skills. One of the inherent “trade-offs,” is that outcomes that are close to interventions in theoretical conceptualization may be less practically meaningful for families and clinicians. For example, in autism, improvements in spontaneous imitation or emotional responsiveness may be easier to link to theory, but less appealing to parents as measurements of treatment success than full-inclusion or “normal appearance”.

Ideally, the anticipated relationships between general areas of development and specific behaviors can be identified. For example, for an intervention aimed at improving children’s joint attention, one might predict, as outcomes, that a child would (1) initiate more joint attention and respond more consistently to others during attentional acts, (2) produce more language in coordination with joint attention, and (3) show general increases in social interest and initiation. Similarly, in a study examining improvement in communication through use of a visual system (Picture Exchange Communication System or PECS: Bondy & Frost, 1994), one might first look at (1) a child’s ability to learn to understand and use the system in a teaching situation, (2) then the extent to which the child uses the system in comprehension throughout a number of circumstances and whether the child initiates using the

Challenges in Evaluating Psychosocial Interventions

system across contexts and (3) then whether there were general changes in the child's social interaction or language level.

A recommendation was made that research designs should initially focus on specific target behaviors and then study the relationship between specific targets and theoretically based predictions of changes in more general development. Studies targeting general changes (e.g. ability to learn in a regular classroom) are likely to be uninterpretable without evidence of the links between changes in specific targets and more general outcome. Group differences in general outcomes in a randomized group experiment can be interpreted as a treatment effect, but the mechanism by which the treatment had an effect on these outcomes is uninterpretable without information about more specific changes. Longitudinal studies may also help us understand the relationship between different levels of change that are targeted. For example, is it necessary that a child with ASD improve in joint attention in order to show improvements in language in response to a particular intervention?

Given variation in baseline levels of earlier-emerging skills, it seems unlikely that any treatment will work comparably for all children or adults with ASD. Effects of both chronological age and developmental levels in various areas are probably not linear, and the magnitude of effects may vary according to where in a developmental trajectory teaching of a skill begins. For example, teaching a pre-verbal child 5 words may be different than teaching 5 more words to a child with 5 words already. Eventually, when studies of psychosocial interventions in autism address broader populations, the need to control sources of variation such as family variables, level of parents' IQ scores, and ethnicity will be crucial. As noted earlier, specific relations between treatment characteristics and individual differences (of both children and families) need to be considered. This will require complex studies with large samples.

Discussants raised the importance of attending to child behaviors in ASD that elicit or enable adult use of facilitating actions. These child behaviors include "pivotal responses" that provide a child with important ways to learn from others (Koegel, Koegel, & McNERney, 2001) and those eliciting "boot-straping operations," that is, that result in positive responses from other people that may lead to learning of other behaviors (Shatz & Ebeling, 1991). The need to attend to changes in developmental level, such as age equivalents, or acquisition of specific skills (as

opposed to changes in level of impairment such as IQ scores) was raised. However, when age equivalents are compared across a variety of measures, the limitations of this strategy must be acknowledged *a priori* (Mervis & Robinson, 1999).

In addition, not all children with ASD benefit from treatments addressing the same goals. One research design has been to match individuals with ASD in pairs on certain variables and then randomize them into treatment groups (Van Bourgondien, Reichle, Campbell, & Mesibov, 1998) from the pairs. Another alternative is to compare actual to expected progress within different domains (Rogers, 1999). Eventually, designs that compare different kinds of treatment will provide important options in this regard. An example was presented from a study by Connie Kasari (in preparation), in which treatments emphasizing symbolic play versus joint attention were compared in terms of their effects on joint attention behaviors and a number of different types of play in both structured and less-structured settings.

Several important factors in selecting outcome measures in ASD were identified. As in standard protocols for clinical trials, outcome measures must be decided in advance; investigators should prioritize which measures will and will not respond to treatment (Aman *et al.*, 2004). This has not typically occurred in comprehensive autism treatments.

Foremost was attention to the need for measures that assess generalization and measures that can be used across activities, interaction styles, and task demands. Thus, outcomes need to match the requirements of research design without sacrificing the social importance of variables. Measures of social validity relevant to autism from both parent and teacher measures were discussed. One example was the recent use by Research Units on Pediatric Psychopharmacology Autism Network (Arnold *et al.*, 2003) of changes in parent-nominated symptoms. Another example was the use of teacher-rating scales. The importance of using measures that are responsive to change was reiterated. There has recently been much interest in using the Autism Diagnostic Observation Schedule (ADOS – Lord, Rutter, DiLavore & Risi, 1999) and the Childhood Autism Rating Scale (CARS – Schopler, Reichler & Renner, 1988) as treatment outcome measures. Broadening the ranges of scores in the instruments might make both more useful in this regard by extending the range of codes beyond 0–2 or 1–5 (for the ADOS and CARS, respectively). However, in at least one recent study, placebo effects were of more concern than sensitivity

to change (Owley *et al.*, 2001). Acknowledgment of practice effects, which have been measured on small samples, is also important (Lord *et al.*, 1999). Integrating the use of multiple instruments, both general and specific, within a domain, such as designing communication samples collected during the ADOS or other observations, and employing broader measures such as the Children's Global Index (CGI) were all suggestions (Scahill & Lord, 2004).

Methodological Challenges

The discussion of methodological challenges included a number of different perspectives in determining appropriate methodology in treatment trials. It was recognized that Randomized Clinical Trials (RCTs) have made a critical contribution to medicine and clinical practice (Medical Research Council Health Services and Public Health Research Board, 2000). RCTs protect the public from deliberate false claims and exploitation and from well-intentioned and sincerely-believed views that may nonetheless be false. The use of RCTs enable more efficacious treatments to be shown as better than less efficacious treatments. On the other hand, RCTs, while crucial, do not by themselves form a complete and coherent program of treatment development and evaluation. Without sufficient sample size and attention to possible mediating factors, treatment by child interactions cannot be assessed. RCTs will not necessarily indicate the trajectory of change or time. Single-blind studies that follow conventional guidelines may be nearly impossible, even though historically, the less rigorous that randomized control trials are, the stronger the findings, so awareness of these trends is critical.

In general, neither the "autism community" nor most educational researchers, have emphasized RCT standards for design, implementation and reporting of studies. In addition, because many parents of children with autism are well-informed about different treatments, conducting a highly controlled randomized clinical trial is difficult due to treatment switching, and "contamination," as parents seek out treatments that they perceive as most useful for their children (and as clinicians and educators include aspects of these approaches in their interventions). When a new treatment is designed and assessed and then is changed after alternative treatments have been compared, assessing the effect of the original treatment becomes complicated. Yet, clinical and educational practice is often changed as interventions are

put into play. Comparisons of a particular intervention to treatment-as-usual at one point in time are not necessarily the same as comparisons of that same treatment a year later to treatment-as-usual (i.e., it is likely both the experimental treatment and the treatment-as-usual have changed).

Ideally, RCTs would allow comparisons of treatments with minimal differences so that the active elements linked to effectiveness could be identified. However, in many cases, such an approach would be viewed as not sufficiently informative to merit large-scale support. Currently, most treatment studies in ASD are carried out with very limited funding compared to what would be needed to conduct an informative RCT. Only a few treatments can be tested at any one time in a standard, large RCT. While there was general, though not universal, opinion that a standard "control" treatment ought to be selected and new treatments be compared to it, there was much disagreement as to what this standard treatment ought to be. Interventions based on Applied Behavior Analysis (ABA), such as the Lovaas model (Lovaas, 1987; Smith *et al.*, 2000a), have received the most empirical study using group designs and so might be considered the appropriate treatment for comparison. An alternative would be to use a less expensive treatment that is feasible in public school programs, such as TEACCH (Ozonoff & Cathcart, 1998; Schopler *et al.*, 1984). Consensus about such an approach was not achieved in this meeting.

Across-site replication by independent research teams was considered the most useful source of information about the efficacy of treatment in autism. Study designs should allow pooled data, whether the initial methodology is single-subject or RCT. The need for researchers to use standard evaluation tools, as well as to develop unique assessments for their own projects, and the need for evaluators to have had similar training and to pool research findings were stressed. With children with autism, experimenter and test differences at baseline and outcome assessments (e.g., Koegel *et al.*, 1997) may yield differences comparable in size to likely intervention effects. For example, several studies have shown marked increases in IQ moving from the Bayley Scales (Bayley, 1993) to performance tests, especially the Merrill-Palmer Scales (Stutsman, 1931), (Lord, 1991; Schopler, Reichler, Devellis, & Daly, 1980). In addition, developmental change over time in ASD may often exceed the effects of interventions. A number of implications of these complexities were noted. Maintaining the

Challenges in Evaluating Psychosocial Interventions

blindness and skill of testers becomes important, and is quite difficult to achieve without extensive planning.

Finally, the need for further information regarding subgroups of children, such as high-functioning children who are speaking spontaneously by age 2 or 3 years, children who use only minimal spoken language by later preschool, and older children and adults with behavior problems, should not be neglected in a focus on broad-based early intervention. For example, creating interventions that differ, not necessarily on the first treatment, but on what happens if the first treatment is effective or ineffective may be one approach. Such strategies might be able to incorporate family preference and clinician preference within a model that could be replicated across individuals.

From the Lab to the Community

Despite complexities, many recent reviews (Dawson & Osterling, 1997; Rogers, 1999) have emphasized how much useful information is available about how to teach and work with children and adults with autism. There is a great deal of information that has been generated by researchers that can be helpful to classroom teachers about the technology needed to help students with autism master many discrete skills from tying their shoes to toilet training, reading words, making initial social approaches, and carrying out basic vocational tasks (National Research Council, 2001). There is a body of knowledge on assessing the functions of maladaptive behavior and helping students learn more adaptive ways to solve problems, and much information regarding the role of communication in behavior management (Horner *et al.*, 2002). However, there are also significant areas of “uncharted territory,” including very basic aspects of treatment such as the number of hours of intervention needed to accomplish major changes (see Sheinkopf & Siegel, 1998) and whether common treatments such as the Developmental, Individual-Difference, Relationship-Based Model (DIR) (Wieder & Greenspan, 2003) or TEACCH (Schopler *et al.*, 1984) or the Denver model (Rogers *et al.*, 2001) are equally strong as various models based on Applied Behavior Analysis in supporting children’s social and cognitive development. Many of the comprehensive programs described in the Educating Children with Autism book (National Research Council, 2001) have detailed curricula that are not yet generally available to the public. The need for manualization of treatment is crucial to dissemination.

The need to recruit underrepresented populations, such as those living in rural areas, ethnic minorities, and non-English speaking children, the potential of internet capability to reach a broad audience, and the force of advocacy organizations to support research about psychosocial interventions was also highlighted. Almost all of the current psychosocial interventions in autism have been built on studies with white middle class populations. There has been little attempt to test the relevancy of different treatments to other ethnic or socioeconomic groups. Almost all of the intervention programs are university-based.

Another issue is how to measure the effects of the multiple other treatments in which families are involved, because of the potential statistical interactions with the within-project treatment. There have been relatively few attempts to investigate how these strategies might apply in public school programs for children whose families have fewer resources or less education (Rogers, Lewis, & Reis, 1987). Many rural and inner city children in the United States receive minimal interventions of any type. Consequently, these children may be the most likely to change as the result of a particular intervention, though they add other complexities. The advantages of parental involvement in research planning in terms of communication, developing measures of meaningful outcomes, and helping to identify some immediate benefits for all who participate were emphasized.

The Canadian Autism Intervention Research Network was described (CAIRN: Canadian Autism Intervention Research Network, 2001; see also www.cairn-site.com). This national multi-disciplinary network involves scientists, clinicians, parents, and policy makers. It was proposed that one of the reasons that evidence-based treatments are often not implemented in community settings is that parents and policy makers do not see the relevance of the research. If families and politicians and administrators are involved in the development of a research agenda from the beginning, more effective treatments may become available as a matter of policy. Focus groups including parents, teachers, therapists, and scientists emphasized the need for evaluation of intensive multi-disciplinary interventions that use existing resources and focus on clinically-relevant outcomes for individuals with ASD. Findings from research need to be tied explicitly to training and advocacy. The importance of transportability and dissemination of techniques was emphasized. A working group in the UK also produced a report

about the design of autism treatment studies. The earlier discussion in this paper on “levels of inference” was based on this document (Charman & Howlin, 2003; Drew *et al.*, 2002).

In a recent program run by the National Early Childhood Teacher Assistance System (NECTAS) in the United States, state representatives identified several of their greatest concerns as follows: (a) the increasing number of children diagnosed as having Autistic Spectrum Disorder (ASD), (b) the lack of availability of diagnostic centers or experts, and (c) the increasing number of children identified at younger ages (with some local service providers having little experience working with young children with ASD). Inequities in access to expertise, resources, and appropriate services occur across and within states. Questions continue to emerge as to whether individual communities should endorse and fund one model, an array of models, or work primarily to embed effective practices into existing programs. While selecting from many resources implies a wealth of opportunities to individualize, one concern is that many providers engage in eclectic practices for which there has not been systematic planning or research.

Few states require special certification or endorsement to work with children with autism. Teachers’ and professionals’ knowledge of intervention research in autism is often minimal. Though some have begun to use the Educating Children with Autism report from the National Research Council (National Research Council, 2001), the book has not yet been disseminated as widely as possible in a form useful to practitioners or parents.

Equitable allocation of resources has not been based on treatment research; instead resources have often been determined by parents and professionals’ hope for the most dramatic gains. There are vast inequities in per pupil expenditures based on the success or failure of litigation by small numbers of families. Challenges for the states include the need (a) to maximize use of the existing public resources (federal, state, and local); (b) to access and understand multiple approaches across the state; (c) to extend the use of Medicaid Waivers to fund services such as in-home treatment and personal care services; and (d) to define medical necessity in a way that includes symptoms associated with ASD. Suggested ways to maximize private resources include working proactively with insurance companies, foundations, community organizations and local businesses. Payment structures need to be clarified and used crea-

tively, and lawsuits need to be avoided due to the tremendous drain of money and energy that they cause.

A final issue raised was how the field should move from efficacy trials to effectiveness. Efficacy trials (also called explanatory trials) ask the question of whether the treatment works in ideal circumstances; effectiveness research asks whether it works in the real world. That is, do the essential elements of a therapy lead to positive outcomes in the usual practice setting, and how can these be modified without losing the benefit found in the efficacy trials? Efficacy trials generally use small sample sizes consisting of volunteers who are defined by strict inclusion and exclusion criteria. Treatment is carried out by trained therapists who are supervised and monitored using manualized programs. Outcomes are predicted in advance and are specifically described and limited. While attrition should be recorded in efficacy trials, dropouts are sometimes eliminated from analysis because the investigators’ priority is to know if the intervention works for those who received it.

In contrast, effectiveness studies (also called management trials or large sample trials) are large clinical trials with relatively few inclusion and exclusion criteria. Treatment is allowed to occur as it does in the natural world. A child might receive multiple interventions; the usual clinic therapist would be employed with standard clinic supervision. Outcomes measured are very general and analyses would be conducted on the intention to treat “last observation carried forward” model. In other words, the last observation of those who dropped out is included in the analysis.

There are many examples in mental health of interventions that were successful in efficacy trials but were not useful when evaluated in the community. Indeed, few effectiveness trials have been carried out in children’s mental health and in autism, even though this is seen as an essential step by some researchers. The appropriateness of effectiveness studies in autism at this point was a source of discussion, without consensus, in part because there have been only a few well-controlled efficacy trials, and some, but not all participants felt that scientifically, understanding what treatments are most beneficial, when done well, must precede the question of their effectiveness in the real world. Other participants disagreed and pointed out that, in other areas of medicine, many treatments have gone straight from development to dissemination.

Challenges in Evaluating Psychosocial Interventions

Recommendations

Investigators and stakeholders, including representatives of parents' groups, felt strongly that recommendations from this meeting should be shared with the public. These recommendations were addressed to the National Institutes of Health, other federal agencies, and private organizations involved in funding and/or supporting (e.g. through parent participation) research and training in autism, and are as follows:

- (1) Ongoing working groups that include investigators and federal staff should continue to define critical gaps in the evidence base, opportunities for moving the field forward, and creative use of existing or new mechanisms to support and facilitate advances in the field of psychosocial intervention research in autism. This information should be used in considering initiatives such as requests for applications through the National Institute of Health and program announcements.

Ideas for ways to maximize the impact of the information included (a) informing federal staff who provide technical assistance for grant applications and review administration and (b) providing background information to peer reviewers.

- (2) Inclusion of a diversity of study participants in autism treatment research should be a priority.
- (3) (a) Study designs that would yield data that could be pooled should be given priority. Standard documentation including forms for screening developed through the NIMH Autism RUPP, records of current and past treatments developed through the CPEA centers and standard databases through ISAAC (The Internet System for Assessing Autistic Children) or individual investigators are all available. Dissemination of practical information concerning management of data and research methods, as well as manualization of interventions of wide interest should be considered a priority for federal agencies and foundations supporting autism treatment research. (b) Minimal standards for autism research in psychosocial treatments were identified in terms of examiner blindness, records of attrition, diagnosis of subjects, manualization of treatment, measures of maintenance and generalization, and the

independence of pre- and post-treatment assessments from the provision of treatment.

- (4) There is a need for training workshops for autism intervention researchers in areas such as relevant statistics, standardized methods of diagnosis, and the development of better outcome measures.
- (5) Modification of existing measures and creation of new measures of outcome in autism should be a high priority in order better to document the goals for and effects of treatment. Areas of special interest include measures of parent-child interaction, family satisfaction, negative effects of treatment, and semi-structured observations or probes for targeted areas of change identified by parents or teachers of individuals with autism.
- (6) More controversial, particularly given the recognition of the need for relatively large samples to provide sufficient power to measure children by treatment interactions, was the question of whether a very large-scale collaborative intervention should be organized or whether the emphasis should be on traditional investigator-initiated NIH grants such as R01 and R21 grants through NIH. Given evidence for strong placebo effects, comparisons of a single treatment versus no treatment or treatment as usual were felt to be much less informative than comparisons of treatments that were equally weighted in terms of the training, belief systems, and enthusiasm generated in investigators and in parents. However, in developing new interventions or to address particular questions for which they are specifically suited, treatment versus no treatment designs might be considered.
- (7) Overall, it was felt that randomized clinical trials measuring the efficacy of early and then later interventions should be the highest, but not the sole, priority of research evaluating psychosocial interventions. Room for innovation in treatment development and in designs and statistical methods of assessing outcomes should also be a part of any national agenda. Nevertheless, there should be clear expectations for meeting consensually defined standards relevant to each project.
- (8) Innovative treatment designs and application of statistical models from other areas should be part of any national agenda. These could

include fractional factorials, comparisons of short-term treatments that differ from each other on one element only, methods and statistics to determine interactions between psychosocial and other treatments (including psychopharmacological), and the use of multiple baseline techniques with random assignment, as well as other approaches that directly address major questions in the effects of psycho-social interventions in ASD.

ACKNOWLEDGMENTS

The meeting during which this manuscript was initiated was sponsored by National Institute of Mental Health, National Institute of Neurological Disorders and Stroke, National Institute on Deafness and Other Communication Disorders, National Institute of Child Health and Human Development, and the NIH Office of Rare Diseases.

The views expressed in this article are those of the authors and do not necessarily reflect the official position of the National Institute of Mental Health, the National Institutes of Health, or any other part of the U.S. Department of Health and Human Services.

REFERENCES

- Aman, M. G. (1996). Stimulant drugs in the developmental disabilities revisited. *Journal of Developmental and Physical Disabilities, 8*(4), 347–365.
- Aman, M. G., Novotny, S., Samango-Sprouse, C., Lecavalier, L., Leonard, E., Gadow, K. D., King, B. H., Pearson, D. A., Gernsbacher, M. A., & Chez, M. (2004). Outcome measures for clinical drug trials in autism. *CNS Spectrums, 9*(1), 36–47.
- Anderson, S. R. et al. (1987). Intensive home-based early intervention with autistic children. *Education & Treatment of Children: Special Issue: New developments in the treatment of persons exhibiting autism and severe behavior disorders, 10*(4), 352–366.
- Arnold, L. E., Aman, M. G., Martin, A., Collier-Crespin, A., Vitiello, B., Tierney, E. et al. (2000). Assessment in multisite randomized clinical trials of patients with autistic disorder: The Autism RUPP Network. *Journal of Autism and Developmental Disorders, 30*(2), 99–111.
- Arnold, L. E., Vitiello, B., McDougle, C., Scahill, L., Shah, B., Gonzalez, N. M. et al. (2003). Parent-defined target symptoms respond to risperidone in RUPP autism study: Customer approach to clinical trials. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(12), 1443–1450.
- Bartak, L., & Rutter, M. (1973). Special educational treatment of autistic children: A comparative study: I. Design of study and characteristics of units. *Journal of Child Psychology and Psychiatry, 14*, 161–179.
- Bayley, N. (1993). *Bayley scales of infant development* (2nd ed.). San Antonio: The Psychological Corporation.
- Birnbrauer, J. S. (1997). Empirically validated learning environments. *Behaviour Change, 14*(1), 21–24.
- Bondy, A. S., & Frost, L. A. (1994). The picture exchange communication system. *Focus on Autistic Behavior, 9*(3), 1–19.
- Canadian Autism Intervention Research Network. (2001). *Developing a Canadian research agenda in early intervention of autism*. Canada: Report of a workshop held in Toronto.
- Charman, T., & Howlin, P. (2003). Research into early intervention for children with autism and related disorders: Methodological and design issues – Report on a workshop funded by the Wellcome Trust, Institute of Child Health, London, UK, November 2001. *Autism, 7*(2), 217–225.
- Columbia Pacific Consulting. (1999). Preliminary Report: Cost-benefit analysis of Lovaas Treatment for autism and Autism Spectrum Disorder (ASD), Prepared for: Harper Grey Easton Barrister and Solicitors.
- Dawson, G., & Osterling, J. (1997). Early intervention in autism. In M. J. Guralnick (Ed.), *The effectiveness of early intervention*. Baltimore. (pp. 307–326). MD: Brookes.
- Drew, A., Baird, G., Baron-Cohen, S., Cox, A., Slonims, V., Wheelwright, S. et al. (2002). A pilot randomised control trial of a parent training intervention for pre-school children with autism – Preliminary findings and methodological challenges. *European Child & Adolescent Psychiatry, 11*(6), 266–272.
- Dunst, C. J. (1999). Placing parent education in conceptual and empirical context. *Topics in Early Childhood Special Education, 19*(3), 141–147.
- Durand, V. M., & Carr, E. G. (1987). Social influences on “self-stimulatory” behavior: Analysis and treatment application. *Journal of Applied Behavior Analysis, 20*(2), 119–132.
- 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*(1), 49–68.
- Fenska, E. C., Krantz, P. J., & McClannahan, L. E. (1985). Age of intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis and Intervention in Developmental Disabilities, 5*, 49–58.
- Goldstein, H. (2002). Communication intervention for children with autism: a review of treatment efficacy. *Journal of Autism and Developmental Disorders, 32*(5), 373–396.
- Green, D., Baird, G., Barnett, A. L., Henderson, L., Huber, J., & Henderson, S. E. (2002). The severity and nature of motor impairment in Asperger’s syndrome: a comparison with Specific Developmental Disorder of Motor Function. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 43*(5), 655–668.
- Harris, S. L., Handleman, J. S., Gordon, R., Kristoff, B. et al. (1991). Changes in cognitive and language functioning of preschool children with autism. *Journal of Autism & Developmental Disorders, 21*(3), 281–290.
- Horner, R. H., Carr, E. G., Strain, P. S., Todd, A. W., & Reed, H. K. (2002). Problem behavior interventions for young children with autism: A research synthesis. *Journal of Autism and Developmental Disorders, 32*(5), 423–446.
- Howlin, P. (1998). Practitioner review: Psychological and educational treatments for autism. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 39*(3), 307–322.
- Howlin, P., Mawhood, L., & Rutter, M. (2000). Autism and developmental receptive language disorder – a follow-up comparison in early adult life. II: Social, behavioural, and psychiatric outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(5), 561–578.
- Jacobson, J. W., Mulick, J. A., & Green, G. (1998). Cost-benefit estimates for early intensive behavioral intervention for young children with autism-general model and single state case. *Behavioral Interventions, 13*, 201–226.
- Jocelyn, L. J., Casiro, O. G., Beattie, D., Bow, J., & Kneisz, J. (1998). Treatment of children with autism: A randomized controlled trial to evaluate a caregiver-based intervention

Challenges in Evaluating Psychosocial Interventions

- program in community day-care centers. *Journal of Developmental & Behavioral Pediatrics*, 19(5), 326–334.
- Kasari, C. (2002). Assessing change in early intervention programs for children with autism. *Journal of Autism and Developmental Disorders*, 32(5), 447–461.
- Kasari, C. (in preparation). Joint attention and symbolic play in young children with autism: A randomized controlled intervention study.
- Klin, A., Volkmar, F. R., & Sparrow, S. S. (1992). Autistic social dysfunction: Some limitations of the theory of mind hypothesis. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 33(5), 861–876.
- Koegel, L. K., Koegel, R. L., & Smith, A. (1997). Variables related to differences in standardized test outcomes for children with autism. *Journal of Autism and Developmental Disorders*, 27(3), 233–243.
- Koegel, R. L., Koegel, L. K., & McNeerney, E. K. (2001). Pivotal areas in intervention for autism. *Journal of Clinical Child Psychology: Special Issue*, 30(1), 19–32.
- Lord, C., & Bailey, A. (2002). Autism spectrum disorders. In M. Rutter, & E. Taylor (Eds.), *Child and adolescent psychiatry: Fourth edition*. Oxford: Blackwell Publications.
- Lord, C., & McGee, J. P. (2001). *Educating children with autism*. Washington, DC: National Academy Press.
- Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism diagnostic observation schedule*. Los Angeles: Western Psychological Services.
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting & Clinical Psychology*, 55(1), 3–9.
- McConnell, S. R. (2002). Interventions to facilitate social interaction for young children with autism: review of available research and recommendations for educational intervention and future research. *Journal of Autism and Developmental Disorders*, 32(5), 351–372.
- McEachin, J. J., Smith, T., & Lovaas, O. I. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation*, 97(4), 359–372.
- McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias in child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology*, 72(2), 235–251.
- Medical Research Council Health Services and Public Health Research Board. (2000). *A framework for development and evaluation of RCTs for complex interventions to improve health*. Unpublished manuscript, London, MRC.
- Mervis, C. B., & Robinson, B. F. (1999). Methodological issues in cross-syndrome comparisons: Matching procedures, sensitivity and specificity. *Monographs of the Society for Research in Child Development*, 64(1), 115–130.
- National Research Council (2001). *Educating Children with Autism*. Lord, C., & McGee, J.P., Washington DC: National Academy Press.
- Owley, T., McMahon, W., Cook, E. H., Laulhere, T., South, M., Mays, L. Z. et al. (2001). Multisite, double-blind, placebo-controlled trial of porcine secretin in autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(11), 1293–1299.
- Ozonoff, S., & Cathcart, K. (1998). Effectiveness of a home program intervention for young children with autism. *Journal of Autism and Developmental Disorders*, 28(1), 25–32.
- Research Units on Pediatric Psychopharmacology Autism Network. (2002). Risperidone in children with autism and serious behavioral problems. *New England Journal of Medicine*, 347, 314–321.
- Rogers, S. J. (1999). Intervention for young children with autism: From research to practice. *Infants and Young Children*, 12(2), 1–16.
- Rogers, S. J. (1998). Neuropsychology of autism in young children and its implications for early intervention. *Mental Retardation & Developmental Disabilities Research Reviews: Special Issue: Autism*, 4(2), 104–112.
- Rogers, S. J., Hall, T., Osaki, D., Reaven, J., & Herbison, J. (2001). The Denver model: A comprehensive, integrated educational approach to young children with autism and their families. In J. S. Handelman, & S. L. Harris (Eds.), *Preschool education programs for children with autism*. (2nd ed) , pp. 95–134. Austin: pro-ed.
- Rogers, S. J., Lewis, H. C., & Reis, K. (1987). An effective procedure for training early special education teams to implement a model program. *Journal of the Division for Early Childhood*, 11, 180–188.
- Scahill L., & Lord C. (2004). Subject selection in clinical trials in children with autism. *CNS Spectrums*, 9(1).
- Schopler, E., Mesibov, G., Shigley, R., & Bashford, A. (1984). Helping autistic children through their parents: The TEACCH model. In C. R. Reynolds, & T. R. Gutkin (Eds.), *The handbook of school psychology*. (pp. 629–643). New York: Wiley.
- Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood autism rating scale (CARS). *Journal of Autism & Developmental Disorders*, 10(1), 91–103.
- Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The childhood autism rating scale (CARS)*. Los Angeles: Western Psychological Services.
- Schuler, A. L., Prizant, B. M., & Wetherby, A. M. (1997). Enhancing language and communication development: Pre-linguistic approaches. In F. R. Volkmar, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders*. (2nd ed) New York: John Wiley & Sons, Inc.
- Shatz, M., & Ebeling, K. (1991). Patterns of language learning-related behaviours: evidence for self-help in acquiring grammar. *Journal of Child Language*, 18(2), 295–313.
- Sheinkopf, S. J., & Siegel, B. (1998). Home based behavioral treatment of young children with autism. *Journal of Autism and Developmental Disorders*, 28(1), 15–23.
- Sigman, M., Ruskin, E., Arbeile, S., Corona, R., Dissanayake, C., Espinosa, M., Kim, N., Lopez, A., & Zierhut, C. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development*, 64(1), 1–114.
- Smith, T., Groen, A. D., & Wynn, J. W. (2000a). Randomized trial of intensive early intervention for children with pervasive developmental disorder. *American Journal on Mental Retardation*, 105(4), 269–285.
- Smith, T., Groen, A. D., & Wynn, J. W. (2000b). Randomized trial of intensive early intervention for children with pervasive developmental disorder (vol 105, p. 269, 2000). *American Journal on Mental Retardation*, 105(6), 508–508.
- Stutsman, R. (1931). Guide for administering the Merrill–Palmer scale of mental tests, mental measurement of preschool children. In L. M. Terman (Ed.), . (pp. 139–262). New York, NY: Harcourt, Brace & World.
- Van Bourgondien, M. E., Reichle, N. C., Campbell, D. G., & Mesibov, G. B. (1998). The environmental rating scale (ERS): A measure of the quality of the residential environment for adults with autism. *Research in Developmental Disabilities*, 19(5), 381–394.
- Venter, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 33(3), 489–507.
- Wetherby, A. M., Schuler, A. L., & Prizant, B. M. (1997). Enhancing language and communication: Theoretical foundations. In D. Cohen, & F. Volkmar, (Eds.), *Handbook of*

- Autism and Pervasive Developmental Disorders* (vol. 2, pp. 513–538). New York: Wiley.
- Wieder, S., & Greenspan, S. I. (2003). Climbing the symbolic ladder in the DIR model through floor time/interactive play. *Autism*, 7(4), 425–435.
- Yirmiya, N., & Sigman, M. (1991). High functioning individuals with autism – diagnosis, empirical-findings, and theoretical issues. *Clinical Psychology Review*, 11(6), 669–683.