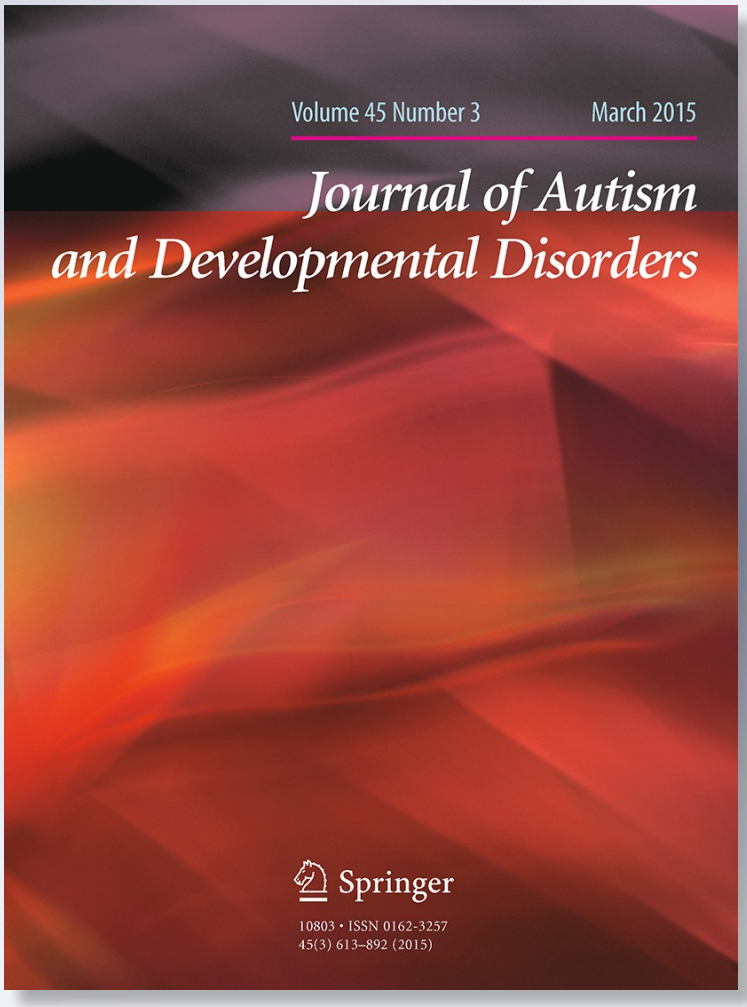
Measuring and Supporting Language Function for Children with Autism: Evidence from a Randomized Control Trial of a Social-Interaction-Based Therapy

Devin M. Casenhiser, Amanda Binns, Fay McGill, Olga Morderer & Stuart G. Shanker

Journal of Autism and Developmental Disorders

ISSN 0162-3257
Volume 45
Number 3

J Autism Dev Disord (2015) 45:846-857
DOI 10.1007/s10803-014-2242-3



Springer

J Autism Dev Disord (2015) 45:846–857
DOI 10.1007/s10803-014-2242-3

ORIGINAL PAPER

Measuring and Supporting Language Function for Children with Autism: Evidence from a Randomized Control Trial of a Social-Interaction-Based Therapy

Devin M. Casenhiser · Amanda Binns · Fay McGill · Olga Morderer · Stuart G. Shanker

Published online: 19 September 2014

Abstract In a report of the effectiveness of MEHRIT, a social-interaction-based intervention for autism, Casenhiser et al. (Autism 17(2):220–241, 2013) failed to find a significant advantage for language development in the treatment group using standardized language assessments. We present the results from a re-analysis of their results to illustrate the importance of measuring communicative language acts (formally called "speech acts"). Reanalysis confirmed that children in the MEHRIT group outperformed the community treatment group on measures of MLUm, number of utterances produced, and various speech act categories. The study underscores the importance of functional language measures in guiding and evaluating treatment for children with autism, and suggests that MEHRIT is effective in improving children's use of language during parent–child interactions.

Keywords Autism · Language development · MEHRIT · Speech acts · Communicative acts

Introduction

A deficit in communication ability is one of the hallmarks of the autism diagnosis. Individuals diagnosed with autism, for example, have been found to produce fewer communicative acts [traditionally called "speech acts" (Austin 1962; Bates 1974; Bruner 1975; Dore 1974, 1975; Searle 1969)] such as affirmations, initiations, requests, directions, and turn-taking vocalizations and fail to respond to parents' utterances significantly more often than typically developing children and language-matched children with developmental language delays (Loveland et al. 1988). Dennis et al. (2001) report that even high-functioning individuals with autism produced significantly fewer types of speech acts than matched controls.

More general impairments in communicative acts (henceforth CA) exist as well. Tager-Flusberg and Anderson (1991), for example, studied response contingency in a corpus of language samples collected from six children with autism over the course of a year. The results indicated that in comparison to a language- and age-matched group of children with Down syndrome, children with autism were significantly less likely to produce contingent utterances and to expand contingent responses by adding new information. Similarly, Capps and colleagues, reported that children with autism failed to respond to a comment or question directed to them significantly more often than typically developing children, produced significantly more verbatim repetitions of questions or comments directed at them, and extended a conversation significantly less often by offering new and relevant information (Capps et al. 1998).

It is therefore unfortunate for autism researchers and clinicians that standardized assessments (unlike language sample analysis) often fail to identify deficits in social communication and language pragmatics (Botting et al. 1997; Condouris et al. 2003; Dunn et al. 1996; Koegel et al. 1997; Tager-Flusberg 1994, 2000; and Wilkinson 1998), especially since deficits in communication are central to the

D. M. Casenhiser · A. Binns · F. McGill · O. Morderer · S. G. Shanker
Milton and Ethel Harris Research Initiative, York University, Toronto, Canada

D. M. Casenhiser (✉)
Department of Audiology and Speech Pathology, University of Tennessee Health Science Center, 548 South Stadium Hall, Knoxville, TN 37996, USA
e-mail: dcasenhi@uthsc.edu

autism diagnosis. Indeed, measuring formal aspects of language alone may constitute an incomplete measure of language ability in children diagnosed with autism. This is because the function of communication and self-expression is central to what makes language *language* (Baker and Hacker 1980; Wittgenstein 1953). The behavior of labeling or identifying an object, often a primary goal for autism treatment, for example, is not language unless the intent of the behavior involves the intent to communicate. This is what separates a child's use of the word *cracker* from that of a parrot's.

The distinction is worth contemplating when evaluating treatments for language deficits in autism; for foundational deficits in social interaction and communication, which are characteristic of autism, may indicate either a failure to understand the purpose of language as a communicative tool, or a restricted understanding of the range of language functions (e.g., language as a tool for obtaining help, sharing information, asking questions, and so on).

To illustrate this notion about the importance of measuring language function for children with autism, we point to a recently published study on a social-interaction-based treatment program called MEHRIT (Casenhiser et al. 2013). In this study, the authors reported that although the target treatment group made significant gains in language as measured by standardized language assessments over the course of 12 months, the gains made were not significantly greater than the gains made by the comparison group who were getting on average about 4 h of community-based treatment each week. Since treatment took place over the course of a year thereby giving participants ample opportunity to improve developmentally even without treatment, the study presents at best weak evidence for the effectiveness of the therapy in improving children's language abilities. At odds with this finding, however, was the apparent result that children in the target treatment group made significant gains over the community treatment (CT) group in a number of so-called proximal skills such as joint attention and social interaction that have been previously shown to correlate highly with language skills (for a review see Bruinsma et al. 2004).

We suggested that this apparent discrepancy could be accounted for by the fact that the standardized language tests used by Casenhiser and colleagues to measure language outcomes focused largely on formal properties of language (vocabulary, syntax, morphology) whereas the target treatment program emphasized improving social interaction and communication. MEHRIT facilitates language development strictly through functional social interaction (play during therapy sessions). There are no set teaching routines, games, or activities; no discrete trials or "table time". The primary focus is always on facilitating functional communicative interaction (including regulation, attention, reciprocation, social problem solving, and use of functional language during social interaction), and the topic of interaction is determined by the child's play preferences. Language form (syntax, morphology, and phonology) is not specifically targeted in activities, although it is of course present in the input if the caregiver is talking during interaction. Likewise, language content (semantics) is provided only through the language used to navigate social interaction. In this way, it is quite different from many standard approaches to treatment that may focus on teaching object naming through use of pictures and preselected items placed in the treatment room (e.g., types of fruit), or by teaching children to repeat, or produce particular sentence types to improve their mastery of language form. The fact that MEHRIT's focus is so different from standard form-focused approaches should logically lead to qualitative differences in outcomes. In particular, we might expect to find greater improvement in measures of language function when compared to treatments that focus solely on language form and content.[1]

This paper tests this notion through a reanalysis of the language outcomes of children previously reported in Casenhiser et al. (2013). We predict that an analysis which focuses on functional language use will be more sensitive to the sorts of gains made by children in the target treatment group, and that the MEHRIT group will be found to have made greater gains in comparison to the CT group in the amount of language that they used during interaction, in the number of different types of language functions they employed during interaction, in the rate with which they responded to a communicative partner, and in the contingency of responses (i.e., whether the responses are relevant to the previous communicative turn). In addition to these main analyses, we report group differences on several specific language functions for the sake of comparing possible differences in the types of CAs that each treatment group produced. Although there may be patterns of differences in the types of CAs facilitated by each treatment approach, this last analysis is largely heuristic in nature and we make no predictions about the kinds of differences we are likely to find.

[1] One might also suspect that treatments that focus on language content and form will show more improvement than those that focus on language function. Casenhiser et al. (2013) report some evidence contrary to this hypothesis since they found no difference on standard language measures between MEHRIT and the community treatment group (who received on average 3–4 h per week of treatment through mostly standard treatment programs). However, since the CT group's treatment was not controlled by the experimenters, this evidence should not be considered reliable.

Methods

Participants

For the sake of completeness, we review the methods used in Casenhiser et al. (2013). 51 children between the ages of 2;0 and 4;11 participated in the research. Families were recruited through diagnosing physicians, public service agencies and newspaper advertisements in the Greater Toronto Area. All children were previously diagnosed with a Pervasive Developmental Disorder (PDD) by either a developmental pediatrician or developmental psychologist using DSM-IV-TR (American Psychiatric Association 2000) criteria. All participants met the autism or autism spectrum category of the Autism Diagnostic Observation Schedule (ADOS) (Lord et al. 1999) and Autism Diagnostic Interview-Revised (ADI-R) (Lord et al. 1994). These assessments were administered by individuals who had completed the research training requirements of the test developers. Upon expressing interest in the study, families were pre-screened for eligibility requirements. Those with neurological or developmental diagnoses other than PDDs were excluded from the sample. Families who were not able to meet the time requirements of the study (2 h per week of therapy and approximately 3 h per day spent interacting with their child) were likewise excluded from the study prior to group assignment.

Treatment Groups

In each cohort, children were stratified by age and baseline level of language and cognitive function (Preschool Language Scale-IV/Comprehensive Assessment of Spoken Language; Bayley Scales of Infant Development), and were assigned to one of two groups using random.org's random number generator: (1) targeted treatment program (MEHRIT) or (2) CT. A sizeable number of participants were lost to attrition during the course of treatment (see Fig. 1). The majority were lost due to an unexpected increase in funding for government-funded intensive behavioral intervention. This therapy consisted of 20 h per week of therapy for the children and compared favorably to the 2 h per week being offered through the study. Time and scheduling did not permit parents to participate in both treatment programs simultaneously. It is not clear how the study outcomes would have changed if these participants were included. The primary concern with attrition is that there is a group bias in the rate of attrition. In consideration of this, we note that most parents who withdrew from both groups cited government-funded IBI as the reason for withdrawal, and that roughly equal numbers of participants withdrew from each group ($\chi^2 = .73$, $p = .39$), with more participants withdrawing for IBI in the CT group than in the MEHRIT group. There does not, therefore, appear to be a significant group bias in the attrition rates. We note, however, that one participant in the MEHRIT group did withdraw citing a mismatch in treatment expectations and the time commitment of the study.

The resulting dataset contained 25 children in the MEHRIT group and 26 in the CT group. Families in the CT group were encouraged to seek treatment for their child while awaiting treatment through the study. CT group treatment amount ranged from 30 min per week to 15 h, with an average of 3.9 h per week. Various treatments solely or in combination were reported by parents. By far the largest treatments included traditional speech therapy (69 % participating), some version of Applied Behavioral Analysis/Early Intensive Behavioral Intervention (62 %), and occupational therapy (46 %). Less than 12 % of the children also participated in social skills group, specialized part-time day care, and other alternative treatments (e.g., specialized diets). Although most parents in the treatment group reported no additional therapy, some received 3 months of weekly 20–30-min therapy sessions provided by the government. Groups did not differ significantly at intake on age, ADOS total scores, age equivalency scores derived from the Pre-School Language Scale (Zimmerman et al. 2006), age equivalency scores from the Bayley Scales of Infant Development (Bayley 2005) cognitive score or Wechsler Preschool and Primary Scale of Intelligence (Wechsler 2002) full scale IQ score (Table 1).[2]

Description of the Treatment Program

Treatment was a social-interaction based therapy and is described in greater detail in Casenhiser et al. (2013). Therapeutic goals included (1) the ability of a child to be regulated, (2) the ability to attend to social interaction, (3) the ability to engage in reciprocal interactions such as conversations or proto-conversations, (4) the ability to solve problems (as distinguished from learning solutions) in social interactions, and (5) the ability to use ideas and language functionally. A multidisciplinary team of specialists including a speech-language pathologist, an occupational therapist, and a clinical social worker determined a treatment plan for each child. Families spent a total of 2 h in therapy each week—1 h with one therapist, and 1 h with a second therapist. During therapy sessions, families were coached about how to best facilitate social interaction and social communication with their child. In addition to weekly therapy sessions, caregivers met approximately

[2] We were unable to obtain scores for two children in each group due to children's inability to complete the test even after multiple attempts. Children who were unable to complete the WPPSI (either due to age or developmental level), were tested using the Bayley Scales of Infant Development.

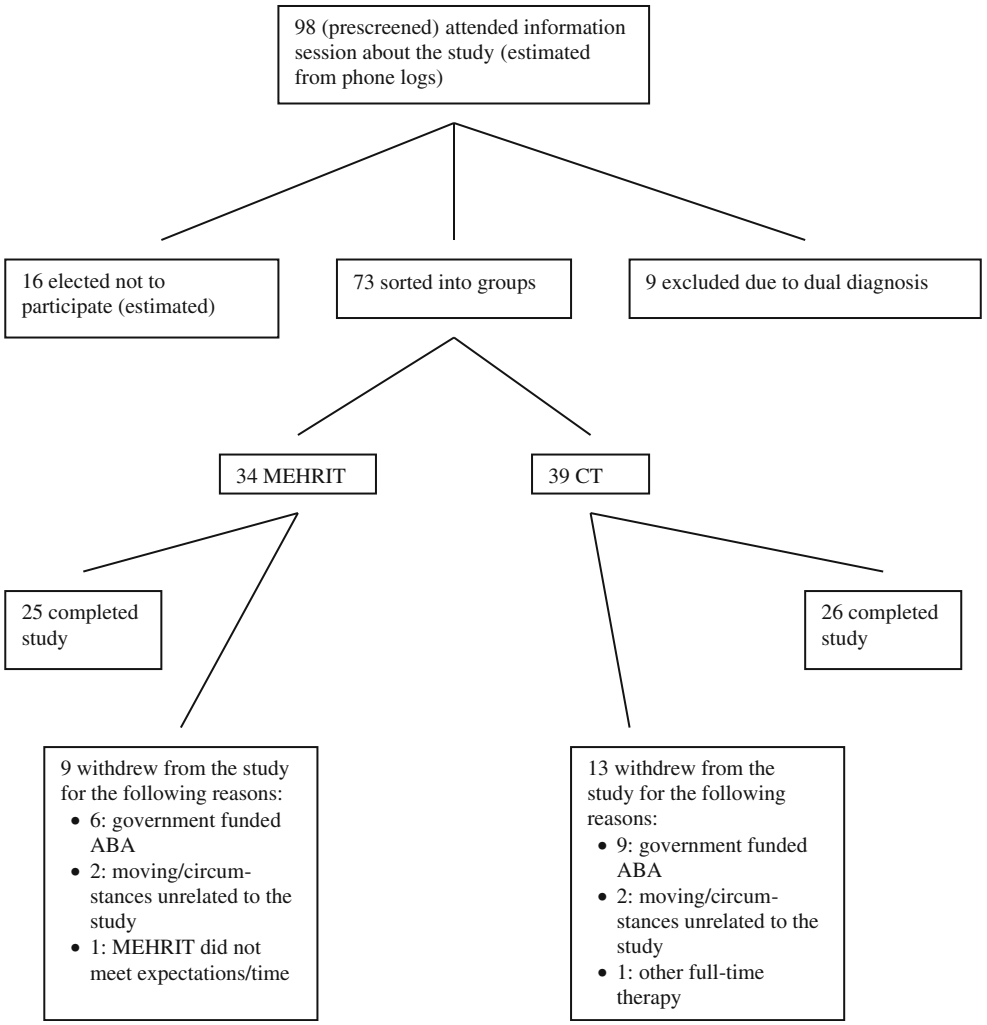


Fig. 1 Consort diagram

every 8 weeks with the therapy team to discuss progress and review videotaped sessions of themselves interacting with their child.

Measures and Coding Procedures

Children were videotaped during a 25-min play session with a parent both before beginning treatment and after 12 months of treatment. Play sessions consisted of 15 min of free play, 5 min of play with tactile objects, and 5 min of play with gross motor apparatus. Complete sessions were transcribed by C-units, and coded from videotape using the Child Language Data Exchange System (MacWhinney 2000). Children's responses were coded by pairs of coders from a team of six coders. Videos were numbered in such a way as to disguise group assignment and testing time. To reduce possible coding biases, all videos were coded for CAs by two independent coders. 35 % of the videos were also double coded for response rate and contingency. Coders discussed disagreements until they reached a consensus. When calculated across categories, average initial agreement for CAs was .841 and for response types, it was .91 (unweighted Cohen's Kappa). All disagreements were discussed until 100 % agreement was reached. Agreement of individual categories ranged from 73–100 %. Table 2 includes agreement values for individual categories and sub-categories.

Coding of Communication Acts

A summary of code descriptions and examples appears in Table 2. CA codes were adapted from Bloom and Lahey (1978), Lahey (1988), and Ninio and Wheeler (1986). Codes were selected that were developmentally appropriate for the age-range being examined, and/or pertained to the particular deficits common in children with PDDs (e.g., it is common for children with autism to lack sharing behaviors, so we included a CA for sharing). Coded acts fit into nine

Table 1 Summary of group characteristics at intake

		MEHRIT treatment mean (SD)	Community treatment mean (SD)
Child's age	t (49) = 1.619, $p = .112$	42.52 (8.76)	46.38 (8.29)
Language age equiv.	t (1,49) = .064, $p = .950$	24.47 (12.68)	24.69 (12.64)
ADOS total score	t (49) = .503, $p = .617$	14.76 (4.02)	14.23 (3.48)
Cognitive age equiv.	t (45) = .337, $p = .738$	34.69 (13.57)	35.87 (10.25)
Mother's education level	Advanced degree	2	6
	Bachelor's degree	15	11
	Associates degree	1	3
	Some university/ college	7	4
	High school	0	2
Income[a]	More than 100 K	12	11
	50–100 K	6	4
	Less than 50 K	4	8
Mother's native language	English	15	12
	Other	10	14
Language most often spoken at home	English	23	23
	Other	2	3
Marital status	Married/partnered	24	22
	Single/divorced/ separated	1	4

[a] Three families in each group elected not to provide information on their income. Statistics Canada reports the 2008 median gross income in Canada is approximately $76,000 (2010)

broad categories: Commenting, Labeling, Responding, Directing, Sharing, Obtaining Information (OI), Rejecting or Protesting (RP), Social Conventions and Routines, and Spontaneous Social Expressions. Several categories were further subdivided into subcategories to enable closer examination.

In addition to the CA categories above, we also coded response types. Responses to utterances requiring a response were coded as OR (obligatory response), and ORs were further coded as contingent if they reflected an understanding or acknowledgement of the previous speaker's turn. Since the number of ORs is affected by the number of utterances requiring an obligatory response, they were recorded as a proportion of the total number of utterances requiring a response. Contingent responses (CR) were coded as a proportion of total number of OR utterances. Contingent responses were further divided into categories reflecting the degree to which the responses were repetitions of the previous speaker's turn: Partial Repetition (PR), Complete Repetition (CR), Novel (N) and Expansive (E). Responses to comments (RtC), those statements that are directed to the child, but do not *require* a response, were also recorded.

And finally, because there has been some research suggesting that children with autism produce fewer filled pauses than do typically developing children (Lake et al. 2011) we included a category for productions of filled pauses.

Utterances were also coded for morphemes using the %mor tier in CHILDES. CLAN was used to calculate, mean length of utterance in morphemes (MLUm) and the number of utterances produced by each child in each video (MacWhinney 2000). Exact repetitions of the previous utterance were excluded from analyses for the purpose of calculating MLUm and total number of utterances produced.

Internal reliability calculated on the pretreatment measurements for the set of language function codes employed in this study was good: Cronbach's $\alpha = .830$ (post treatment measures were similar with Cronbach's $\alpha = .836$). In addition, some measure of validity for the set of spoken communication acts is provided in Table 3. This table shows correlations between each of the major category items and the child's developmental language quotient, joint attention, and involvement scores reported in Casenhiser et al. (2013). Significant, moderately strong correlations were obtained for all major categories of spoken language codes with the exception of filled pauses, obligatory and contingent response ratios, social conventions, and response to comments.

Results

All analyses were conducted using a mixed MANOVA analysis with Treatment Group (MEHRIT vs. CT) as a between-groups factor and Time (Pre vs. Post) as a within groups factor. Accordingly, only the significance of the Group × Time interactions are reported within the paper. To determine the location of individual effects, analyses with more than one dependent variable, as in the case of sub-category CAs listed in Table 5, were followed up with univariate mixed ANOVAs and adjusted using Holm's (1979) sequential Bonferroni adjustment for post hoc comparisons where appropriate. Results are further explained below and summarized in Tables 4 and 5.

Structural Analysis and Utterance Counts

Investigation of MLUm and total number of utterances shows a significant Time × Group interaction, $F(2,45) = 6.545$, $p = .003$, $\eta_p^2 = .225$. The MEHRIT group increased both the

Table 2 Inter-rater reliability, descriptions and examples for CA codes

Item	Inter-rater agreement	Description	Example (where appropriate)
Agree with statement	.96	General affirming response	MOT: let's play beans CHI: okay
Comment on an object in the immediate environment	.79	Talking about an object in the immediate environment	CHI: Tiger ate the banana
Comment on future or past event	.83	Comment about an object or event not in the here and now	CHI: This is like the round table at preschool
Comment on one's own action	.85	Child talks about what he or she is doing	CHI: I can do that!
Comment on other's action	.75	Child talks about what another person is doing	CHI: bouncing! (while parent bounces ball)
Commenting	.75	Summary item	
Complete repetition	.89	Child repeats utterance verbatim and with same intonation	MOT: what do you want to play? CHI: what do you want to play?
Contingent response ratio	.99	Contingent responses contain an indication that the child has heard and understood the previous speaker's statement.	MOT: what do you want to play? CHI: this a duck
Direct attention to other	.711	Direct the attention of another person to an object or 3rd person	CHI: look! (points to car)
Direct attention to self	.87	Direct attention oneself with no intention to get help	CHI: look at me jump!
Directing	.82	Summary item	
Expansive utterance	.80	A contingent response in which the child includes novel material	
Filled pauses	.73	Linguistic hedges employed to indicate that one is holding one's communicative turn	MOT: what do you want to do? CHI: um…
Labeling	.796	Child labels an object in the immediate environment	CHI: doggie (when looking at a stuffed animal)
MLU*m*	N/A	Mean length of utterance in morphemes	
Novel Utterance	.76	Utterances (generally initiations) not based upon a previous speaker's utterance	
Number of different CAs	N/A	Type count of language functions	
Number of utterances	N/A	Total number of utterances	
Obligatory response ratio	.97	Ratio of the number times a child responds to an utterance which requires a response. Comments like "It's hot in here!" do not *require* a response although they may invite one.	MOT: what's that? CHI: apple
Obtain help	.77	Child directs another person with the intent to get help	CHI: ba (while looking at ball that is out of reach)
Obtain information	.92	Child asks question to obtain information	CHI: what that?
Partial repetition	.81	Child's utterance contains a partial repetition of previous utterance	MOT: let's play ball CHI: ball
Protest other's action or communicative turn	.90	Child protests another's action or turn (must include clear negation)	CHI: don't do it!
Reject or protest	.91	Summary item	
Reject the suggested idea/object	.92	Child rejects an object or suggestion	MOT: let's play ball CHI: i don't want to
Responding	.95	Summary item	
Response to a choice question	1.0	Child responds when given a choice	MOT: do you want a duck or a cow? CHI: i want a duck
Response to a prompt	.91	Child gives an elicited (prompted) response. Includes prompted imitation, withholding for an expected response, routine response, cloze-type response	MOT: old McDonald had a… CHI: farm
Response to a WH question	1.0	Child responds to a "WH" question	MOT: where is the tomato? CHI: here it is!

Table 2 continued

Item	Inter-rater agreement	Description	Example (where appropriate)
Response to a yes/no question	.89	Child responds to a yes/no question	MOT: do you want to play with the beans?
			CHI: yes!
Response to Comments	.82	Child responds to an non-obligatory statement	MOT: it's sure hot in here!
			CHI: yeah! (sighs)
Scripted social convention	1.0	Conventional statements made in common situations	CHI: bye–bye. (when leaving)
Scripted social routine	.97	Child engages in a scripted social routine	CHI and MOT together: one…two…three! (when jumping off mini-tramp)
Share fact	.83	Child shares something factual with another person	CHI: two plus two is four
Share pretend play	.93	Child talks to or about a symbolic/miniature toy (such as a doll), pretends to be someone or something, or do something make-believe	CHI: the baby is sleeping (while covering a doll)
Share something about one's self	.73	Child shares information about him or herself, including feelings and like or dislikes	CHI: i like you!
Sharing	.83	Summary item	
Social conventions	.98	Summary item	
Spontaneous social expression	.92	Spontaneous use of an expression for purely social reasons (e.g., share happiness, excitement about an object)	CHI: wow, cool!

number and length of their utterances significantly more than the CT group: for utterances, $F(1,46) = 10.874$, $p = .002$, $\eta_p^2 = .191$; and for MLUm, $F(1,46) = 6.446$, $p = .015$, $\eta_p^2 = .123$. The number of different CAs also showed a significant Time × Group interaction with the MEHRIT group increasing the number of different CAs more than the CT group. We note, however, that since the number of utterances and the number of different CA types is correlated, we cannot determine whether the number of different CA types increases because the child is talking more or whether the child is talking more because he or she has a greater number of CAs in his or her language repertoire.

Response Type

Since response type data were found to evidence significant skew, the data were log10 transformed after adding .5 to each value to adjust for zeros. Results showed a significant Time × Group interaction for response type codes (Response to Comments, Obligatory Response, Contingent Response): $F(1, 48) = 3.847$, $p = .028$, $\eta_p^2 = .138$. The MEHRIT group showed a significantly greater increase in production of ORs, $F(1,49) = 5.731$, $p = .021$, $\eta_p^2 = .105$ and CRs, $F(1,49) = 9.67$, $p = .003$, $\eta_p^2 = .165$, but the groups did not perform differently on RtCs ($F(1,49) = .599$, $p = .443$, $\eta_p^2 = .012$). Of the CR sub-categories, it appears that the Expansive response type ($F(1,49) = 8.598$, $p = .005$, $\eta_p^2 = .188$) accounts for the effect since no other sub-categories of response type were significant. Finally, the number of filled pauses did not show a significant Time × Group interaction ($F(1,49) = .832$, $p = .366$, $\eta_p^2 = .017$.

Communicative Acts

Since CA data were found to evidence significant skew, the data were log10 transformed after adding .5 to each value to adjust for zeros. Each of the nine CA major categories were entered into a MANOVA for analysis. The MEHRIT group showed a significantly greater increase in the production of CAs as compared with the CT group as indicated by a significant Time × Group interaction: $F(1,41) = 2.516$, $p = .021$, $\eta_p^2 = .356$. Post-hoc analyses indicated significant Time × Group interactions at $\alpha = .05$ for commenting, labeling, sharing, OI, RP, and a marginally significant interaction for spontaneous social interaction. If, however, alpha is adjusted using Holm's (1979) sequential Bonferroni adjustment, only sharing and OI are significant (see Table 5). Complete results from post hoc analysis of variance on sub-category codes are included in Table 5 for the sake of completeness and data exploration. In addition, we note that the CT group produced significantly more comments at Time 1 than did the MEHRIT group (7.7 vs. 2.8). This may have affected the resulting trajectory of differences between Time 1 and Time 2.

Table 3 Correlations between major CA codes and scores from observational and standard measures reported in Casenhiser et al. (2013)

	Developmental language quotient	Joint attention	Involvement
MLUm	.838**	.705**	.649**
Total number of utterances	.814**	.630**	.642**
Directing	.806**	.650**	.649**
Number of CA types	.800**	.666**	.660**
Sharing	.672**	.485**	.589**
Commenting	.671**	.489**	.476**
Responding	.604**	.428**	.437**
Rejecting/protesting	.588**	.407**	.528**
Obtaining information	.582**	.533**	.374**
Labeling	.454**	.527**	.336*
Spontaneous social expression	.345*	.382**	.456**
Filled pauses	.181	.170	.237
Obligatory response ratio	.147	−.079	.144
Social conventions	.097	.074	.110
Contingent response ratio	.081	−.075	.167
Response to comments	.077	−.027	.244

* Significant at .05

** Significant at or below .01

Discussion

The purpose of this study is to illustrate the importance of evaluating functional aspects of language outcomes for children diagnosed with PDDs – especially for treatment programs that target functional aspects of speech. As we predicted, we found that the MEHRIT group made greater functional language gains in comparison to the CT group in the amount of language that they used during interaction, in the number of different types of language functions they employed during interaction, in the rate with which they respond to a communicative partner, and in the contingency of responses. Thus, Casenhiser et al. (2013) report that the MEHRIT group did not make greater gains in language than did the CT group should be amended to include the gains in spoken communication acts measured herein.

What then accounts for the difference in the original report and this re-analysis? We surmise that differences in what each approach to evaluation measures coupled with a general performance difficulty that children with autism often face during standardized tests contributed to discrepancy. If this is in fact the case, then the study underscores the importance of attending to CAs when making determinations about the effectiveness of autism intervention programs. Furthermore, the diagnostic information gained from transcript analysis is particularly useful for guiding and assessing programs aimed at improving children's ability to interact and communicate with others.

Implications

Finally, we return to the argument made in the introduction to motivate the study. That is, since the language of children with autism is especially constricted in the area of pragmatics and social communication, we argue that there is good reason for including therapeutic goals aimed at these particular challenges. This is not only for the sake of determining that a treatment program has reduced symptom severity, but also because skills in social interaction and communication facilitate language learning. Increased communicative competence better enables children to engage in interactions, thereby increasing the amount of language input (and intake). It also enables children to leverage these skills to broaden learning opportunities. Several studies, for example, suggest that children's understanding of the pragmatics of social interaction and communication play a pivotal role in language learning outside of ostensive contexts (e.g., saying "ball" while showing a ball), which as Jaswal and Markman (2001) point out, comprise less than 20 % of maternal utterances to young children.[3] Without the ability to learn through social interaction, a child is almost entirely dependent on the therapist/teacher to present and teach language through techniques involving ostensive context. But attempting to teach language in this manner is tantamount to programming a computer with all the language necessary to communicate well enough to pass as a human being (see Turing's (1950) test of artificial intelligence).

A few treatment studies have already investigated functional aspects of language in their programs outcomes assessments. Rogers et al. (2006) include a measure of the communicative function of novel words and phrases in their study of the Denver and PROMPT treatment programs, and Venker et al. (2011) and Yoder and Stone (2006) include measures of spontaneous (i.e., non-imitative, non-prompted) speech in their assessment of intervention programs. Likewise, in their assessment of the More than Words intervention, which focuses on children's communication, Carter et al. (2011) include measures such as initiation of joint attention from Mundy et al. (2003). Early Social Communication Scales.

Limitations

We feel it is important to reiterate some limitations resulting from the design of the original study (see

[3] For a discussion of many such studies, see Tomasello (2003).

Table 4 Summary of analysis of response and contingency codes

Major category (corrected item-total correlation)	Sub-category	Group	Mean (SD)		F	η_p^2
			Time 1	Time 2		
MLUm		MEHRIT	2.01 (.93)	2.42 (.99)	6.45	.123
		CT	2.10 (.92)	2.10 (.89)	$p = .015^{\dagger}$	
Number Of utterances		MEHRIT	84.22 (77.07)	156.65 (106.20)	10.87	.191
		CT	106.04 (97.31)	119.08 (94.66)	$p = .002^{\dagger}$	
Number of different spoken comm. acts		MEHRIT	10.40 (6.55)	14.16 (6.36)	12.90	.208
		CT	12.35 (5.85)	12.46 (6.00)	$p < .001^{\dagger}$	
Filled pauses		MEHRIT	.72 (1.72)	3.68 (6.58)	.832	.017
		CT	.77 (3.33)	1.27 (2.13)	$p = .366$*	
Response to comments		MEHRIT	12.84 (11.92)	27.80 (25.42)	.599	.012
		CT	15.23 (14.43)	17.58 (12.62)	$p = .443$*	
Obligatory response ratio		MEHRIT	.71 (.11)	.82 (.15)	5.73	.105
		CT	.73 (.12)	.77 (.12)	$p = .021^{\dagger}$	
Contingent response ratio		MEHRIT	.47 (.27)	.84 (.44)	9.67	.165
		CT	.51 (.31)	.60 (.30)	$p = .003^{\dagger}$	
	Complete repetition	MEHRIT	4.28 (6.34)	3.92 (3.87)	1.095	.022
		CT	4.19 (4.49)	3.92 (4.64)	$p = .300$*	
	Partial repetition	MEHRIT	6.00 (7.54)	5.40 (5.12)	.040	.001
		CT	8.04 (6.87)	7.65 (6.38)	$p = .842$*	
	Novel	MEHRIT	26.00 (28.60)	53.48 (44.16)	2.574	.050
		CT	31.62 (34.84)	46.50 (45.42)	$p = .115$*	
	Expansive	MEHRIT	7.28 (11.85)	21.72 (23.02)	8.598	.188
		CT	11.73 (15.71)	11.69 (12.41)	$p = .005^{*\dagger}$	

* Analyses performed on log10 transformed data to correct for skew

† Indicates a value that is significant. Sub-category CAs are adjusted for alpha using Holm's (1979) Bonferroni sequential adjustment for multiple comparisons

Casenhiser et al. 2013 for a more complete discussion). First, the CT group is not a no-treatment group, nor do the children in the CT group all receive one consistent treatment. This fact does not permit researchers and clinicians to determine that MEHRIT is better than another particular treatment. Secondly, families were expected to spend 2 h each week being coached on the method, and 3 h each day interacting with their child to implement what they had learned during coaching. Thus, there is an inherent bias in the study participants (both groups). Only individuals who were willing and able to attend 2-h weekly meetings at the university, as well as spend 3 h per day interacting with their child enrolled in the study. This means that MEHRIT may not work for families who do not have the time or inclination for such a large commitment. Indeed, one of the economic advantages of MEHRIT is the fact that it increases the number of treatment hours by training parents to provide the treatment rather than having parents pay a therapist for treatment.

In addition to limitations regarding the study's overall design, we note that while the present analysis may be better than standardized tests at capturing changes in children's functional language use during social interaction, unlike standardized assessments, children's performance on the measures reported in this paper may well have been influenced by the parents' skills at interacting with the children. This is especially likely given that parents in the MEHRIT group were coached to better interact with their children. Future work could address this limitation by including an additional interaction partner whose interaction skills are not improving over the course of the study (i.e., one who is already an expert, or one who is a novice such as a peer or caregiver who has not received coaching). This limitation also points to a need for those using parent-mediated interventions to be mindful of potential challenges that may be presented by situations where the child needs to interact with someone who has not been trained to interact with the child.

Table 5 Summary of results from analysis of CA categories and sub-categories

Major category	Sub-category	Group	Mean (sd)		F	η_p^2
			Time 1	Time 2		
Commenting		MEHRIT	2.80 (5.68)[a]	6.92 (7.69)	3.61	.239
		CT	7.73 (12.20)	5.54 (6.83)	$p = .012$	
	Comment on future or past event	MEHRIT	.08 (.28)	.64 (1.38)	3.736	.071
		CT	.35 (1.38)	.12 (.33)	$p = .0499^{\dagger}$	
	Comment on other's action	MEHRIT	.08 (.28)[a]	.32 (.69)	4.085	.077
		CT	.38 (.75)	.23 (.51)	$p = .048$	
	Comment on one's own action	MEHRIT	6.24 (9.03)	11.72 (12.48)	12.89	.208
		CT	11.19 (12.28)	8.00 (7.38)	$p < .001^{\dagger}$	
	Comment on an object in the immediate environment	MEHRIT	2.72 (5.60)[a]	6.28 (7.16)	5.31	.098
		CT	7.38 (11.21)	5.42 (6.80)	$p = .025$	
Labeling		MEHRIT	16.72 (18.99)	21.00 (19.98)	5.702	.104
		CT	22.73 (21.63)	18.69 (17.37)	$p = .021$	
Responding		MEHRIT	29.20 (28.35)	40.60 (35.87)	1.730	.161
		CT	31.46 (34.27)	42.00 (41.14)	$p = .147$	
	Agree with statement	MEHRIT	6.88 (6.78)	15.20 (13.99)	3.44	.066
		CT	12.81 (12.29)	15.77 (13.25)	$p = .069$	
	Response to a choice question	MEHRIT	.60 (1.00)	1.88 (2.85)	1.78	.035
		CT	1.27 (2.57)	1.15 (1.49)	$p = .188$	
	Response to a WH question	MEHRIT	6.52 (8.03)	14.04 (14.25)	3.29	.063
		CT	8.27 (9.44)	11.62 (11.90)	$p = .075$	
	Response to yes/no question	MEHRIT	15.60 (18.72)	21.80 (21.82)	.183	.004
		CT	15.54 (23.38)	22.65 (30.19)	$p = .671$	
	Response to a prompt	MEHRIT	6.48 (11.23)	2.88 (3.86)	.918	.018
		CT	6.38 (7.88)	6.58 (10.27)	$p = .343$	
Directing		MEHRIT	7.88 (7.47)	12.88 (12.10)	1.966	.111
		CT	10.81 (12.46)	10.23 (8.38)	$p = .132$	
	Obtain help	MEHRIT	3.52 (4.07)	3.92 (4.59)	.034	.001
		CT	4.54 (8.82)	5.23 (6.36)	$p = .854$	
	Direct attention to self	MEHRIT	1.60 (2.22)	4.56 (5.68)	5.38	.099
		CT	1.69 (4.38)	.77 (1.56)	$p = .025$	
	Direct attention to other	MEHRIT	2.76 (4.40)	4.40 (4.98)	3.18	.061
		CT	4.58 (5.26)	4.23 (5.20)	$p = .08$	
Sharing		MEHRIT	6.32 (12.87)	21.04 (25.04)	4.78	.234
		CT	5.92 (8.68)	5.73 (7.80)	$p = .005^{\dagger}$	
	Share fact	MEHRIT	.24 (.83)	.36 (1.32)	.193	.004
		CT	.27 (1.00)	.31 (.79)	$p = .663$	
	Share pretend play	MEHRIT	5.52 (11.36)	19.04 (24.18)	11.50	.190
		CT	5.04 (7.52)	4.81 (6.48)	$p = .001^{\dagger}$	
	Share something about one's self	MEHRIT	.56 (1.26)	1.64 (3.30)	4.88	.083
		CT	.62 (1.06)	.62 (1.50)	$p = .040$	
Obtaining information		MEHRIT	2.44 (6.13)	4.16 (5.44)	8.70	.151
		CT	5.00 (9.20)	3.04 (5.78)	$p = .005^{\dagger}$	
Rejecting or protesting		MEHRIT	5.28 (7.03)	11.52 (11.69)	4.572	.160
		CT	5.62 (8.20)	5.12 (5.76)	$p = .015$	

Table 5 continued

Major category	Sub-category	Group	Mean (sd)		F	η_p^2
			Time 1	Time 2		
	Protest other's action or communicative turn	MEHRIT	2.36 (3.96)	5.24 (5.46)	1.615	.032
		CT	1.46 (3.43)	1.77 (2.03)	$p = .210$	
	Reject the suggested idea/object	MEHRIT	2.92 (4.72)	6.28 (6.89)	9.33	.160
		CT	4.15 (5.25)	3.35 (4.51)	$p = .004^{\dagger}$	
Social conventions		MEHRIT	3.00 (4.80)	3.28 (3.82)	.580	.024
		CT	4.31 (5.30)	6.58 (17.33)	$p = .57$	
	Scripted social convention	MEHRIT	.44 (.92)	1.04 (1.49)	.910	.018
		CT	.31 (.68)	.62 (1.33)	$p = .345$	
	Scripted social routine	MEHRIT	2.56 (4.40)	2.24 (3.69)	.236	.005
		CT	4.00 (5.26)	5.96 (17.24)	$p = .63$	
Spontaneous social expression		MEHRIT	.16 (.47)	.60 (1.08)	3.97	.075
		CT	.31 (1.19)	.12 (.33)	$p = .05$	

[a] Indicates a significant pre-treatment difference ($p < .05$)

[†] Indicates a value that is significant when adjusted using Holm's (1979) sequential Bonferroni adjustment for multiple comparisons (separate adjustments were conducted for major categories and sub-categories). To correct for skew, scores were log10 transformed after adding .5 to all scores to eliminate zeros

That parents' interaction skills might influence those of the child, however, does not lessen the central point of this paper; namely, that it is important to consider both formal and functional analyses of language when working with children diagnosed with autism. It also does not nullify the value of a social-interactional treatment approach such as MEHRIT since increasing a child's functional communication during social interaction, whether through parental skills or some other method, provides critical opportunities for the child to learn, much needed input, and practice at interaction.

Acknowledgments This research was made possible by the generous support of the Harris Steel Foundation and the Harris family, which made it possible to create the Milton and Ethel Harris Research Initiative. We are also grateful for the support we have received from the Unicorn Foundation, Cure Autism Now, the Public Health Agency of Canada, and the Templeton Foundation. We also thank the families involved in the study as well as team members Narmilee Dhayanandan, Nadia Noble, and Alicia Alison for their hard work and dedication. We thank Eunice Lee, Ashley Abergel, Riti Dass, Heba Mouniemneh, Jasmine Dzeko, Farah Vali, and Lauren Mitchell for their perseverance in coding and transcribing the mountain of raw data, and we thank John Hoffman for his outstanding assistance with research and preparation of this manuscript. Finally, we thank the editor and three anonymous JADD reviewers whose comments substantially improved the quality of this manuscript.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.).

Austin, J. L. (1962). *How to do things with words*. Cambridge, MA: Harvard University Press.

Baker, G. P., & Hacker, P. M. S. (1980). *Wittgenstein: Understanding and meaning*. Oxford: Blackwell.

Bates, E. (1974). *Language and context: Studies in the acquisition of pragmatics*. New York, NY: Academic Press.

Bayley, N. (2005). *Bayley scales of infant and toddler development* (3rd ed ed.). New York: Pearson.

Bloom, L., & Lahey, M. (1978). *Language development and language disorders*. Somerset, NJ: John Wiley and Sons Inc.

Botting, N., Conti-Ramsden, G., & Crutchley, A. (1997). Concordance between teacher/therapist opinion and formal language assessment scores in children with language impairment. *International Journal of Language and Communication Disorders, 32*(3), 317–327.

Bruinsma, Y., Koegel, R. L., & Koegel, L. K. (2004). Joint attention and children with autism: A review of the literature. *Mental Retardation and Developmental Disabilities Research Reviews, 10*(3), 169–175.

Bruner, J. S. (1975). The ontogenesis of speech acts. *Journal of Child Language, 2*(1), 1–19.

Capps, L., Kehres, J., & Sigman, M. (1998). Conversational abilities among children with autism and children with developmental delays. *Autism: The International Journal of Research and Practice, 2*(4), 325–344.

Carter, A., Messinger, D., Stone, W., Celimli, S., Nahmias, A., & Yoder, P. (2011). A randomized controlled trial of Hanen's 'More than Words' in toddlers with early autism symptoms. *Journal of Child Psychology and Psychiatry, 52*(7), 741–752.

Casenhiser, D. M., Shanker, S. G., & Stieben, J. (2013). Learning through interaction in children with autism: Preliminary data from a social-communication-based intervention. *Autism, 17*(2), 220–241.

Condouris, K., Meyer, E., & Tager-Flusberg, H. (2003). The relationship between standardized measures of language and measures of spontaneous speech in children with autism. *American Journal of Speech-language Pathology, 12*(3), 349–358.

Dennis, M., Lazenby, A., & Lockyer, L. (2001). Inferential language in high-function children with autism. *Journal of Autism and Developmental Disorders, 31*(1), 47–54.

Dore, J. (1974). A pragmatic description of early language development. *Journal of Psycholinguistic Research, 3*(4), 343–350.

Dore, J. (1975). Holophrases, speech acts and language universals. *Journal of Child Language, 2*(1), 21–40.

Dunn, M., Flax, J., Sliwinski, M., & Aram, D. (1996). The use of spontaneous language measures as criteria for identifying children with specific language impairment. *Journal of Speech and Hearing Research, 39*, 643–654.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65–70.

Jaswal, V., & Markman, E. (2001). Learning proper and common names in inferential versus ostensive contexts. *Child Development, 27*(3), 768–786.

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.

Lahey, M. (1988). *Language disorders and language development*. New York, NY: Macmillan.

Lake, J. K., Humphreys, K. R., & Cardy, S. (2011). Listener vs. speaker-oriented aspects of speech: Studying the disfluencies of individuals with autism spectrum disorders. *Psychonomic Bulletin and Review, 18*(1), 135–140.

Lord, C., Risi, S., Lambrecht, L. K., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (1999). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205–223.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685.

Loveland, K. A., Landry, S. H., Hughes, S. O., Hall, S. K., & McEvoy, R. E. (1988). Speech acts and the pragmatic deficits of autism. *Journal of Speech and Hearing Research, 31*(4), 593–604.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Mundy, P., Delgado, C., Block, J., Venezia, M., Hogan, A., & Seibert, J. (2003). *A manual for the abridged early social communication scales (ESCS)*. Coral Gables, FL: University of Miami.

Ninio, A., & Wheeler, P. (1986). A manual for classifying verbal communicative acts in mother infant interaction. *Transcript Analysis, 3*(1), 1–82.

Rogers, S. J., Hayden, D., Hepburn, S., Charlifue-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*(8), 1007–1024.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. London: Cambridge University Press.

Tager-Flusberg, H. (1994). Dissociations in form and function in the acquisition of language by autistic children. In H. Tager-Flusberg (Ed.), *Constraints on language acquisition: Studies of atypical children* (pp. 175–194). Hillsdale, NJ: Erlbaum.

Tager-Flusberg, H. (2000). The challenge of studying language development in children with autism. In L. Menn & N. Bernstein Ratner (Eds.), *Methods for studying language production* (pp. 313–331). Mahwah, NJ: Erlbaum.

Tager-Flusberg, H., & Anderson, M. (1991). The development of contingent discourse ability in autistic children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 32*(7), 1123–1134.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*(236), 433–460.

Venker, C., McDuffie, A., Weismer, S. E., & Abbeduto, L. (2011). Increasing verbal responsiveness in parents of children with autism: a pilot study. *Autism, 16*(6), 568–585.

Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence* (3rd ed.). San Antonio, TX: Pearson.

Wilkinson, K. M. (1998). Profiles of language and communication skills in autism. *Mental Retardation and Developmental Disabilities Research Reviews, 4*(2), 73–79.

Wittgenstein, L. (1953). *Philosophical investigations*. G. E. M. Anscombe & R. Rhees (Eds.), (G. E. M. Anscombe, Trans.). Oxford: Blackwell.

Yoder, P., & Stone, W. (2006). A randomized comparison of the effect of two prelinguistic communication interventions on the acquisition of spoken communication in preschoolers with ASD. *Journal of Speech, Language, and Hearing Research, 49*, 698–711.

Zimmerman, I., Steiner, V., & Pond, R. (2006). *Preschool language scale-4*. New York: The Psychological Corporation and Harcourt Brace Jovanovich.