Hammill Institute on Disabilities

Communications Disorders Quarterly
XX(X) 1–12

DOI: 10.1177/1525740112455053
http://cdq.sagepub.com
SAGE

# Assessing Spoken Language Competence in Children With Selective Mutism: Using Parents as Test Presenters

**Evelyn R. Klein, PhD[1], Sharon Lee Armstrong, PhD[1], and Elisa Shipon-Blum, DO[2]**

**Abstract**

Children with selective mutism (SM) display a failure to speak in select situations despite speaking when comfortable. The purpose of this study was to obtain valid assessments of receptive and expressive language in 33 children (ages 5 to 12) with SM. Because some children with SM will speak to parents but not a professional, another purpose was to explore the efficacy of employing parents to deliver test stimuli. Parents received training on presenting standardized test material and were monitored during testing by a professional who scored and interpreted the results. Professional-administered tests underestimated children's capabilities. However, even with parents, children's scores decreased as the tasks changed from receptive to expressive vocabulary and from narrative comprehension to telling a story on their own. Thus, although SM is typically classified as an anxiety disorder, an underlying expressive narrative language deficit was identified in 42% of children with SM using this new procedure.

**Keywords**

selective mutism, language assessment, anxiety disorders, children, parents

As specified by the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM-IV-TR*; American Psychiatric Association [APA], 2000) criteria, children with selective mutism (SM) exhibit a

> consistent failure to speak in specific social situations (in which there is an expectation for speaking, e.g., at school), despite speaking in other situations. [Further] the failure to speak is not due to a lack of knowledge of, or comfort with, the spoken language required in the social situation. [And] the disturbance is not better accounted for by a communication disorder. (pp. 125–127)

SM is considered a rare childhood disorder with prevalence rates between 0.47% and 0.76% (Viana, Beidel, & Rabian, 2009). Given that SM is a disorder with limited large-scale empirical evaluations and research investigating etiology, researchers have not settled on possible causes of SM (Cohan, Price, & Stein, 2006; Sharp, Sherman, & Gross, 2007). Further complicating the picture is the fact that patterns of not speaking in children with SM can vary: Some children may never talk outside the home, some may talk to a selected few, some may just whisper, and some may talk to someone whom they have never met.

In the following review, we offer a way to conceptualize the speech and language of children with SM in terms of *communication competence* versus *communication performance*. This conceptualization provides a framework for presenting the relevant literature. First, we address the performance factor in SM by presenting the research investigating anxiety as a contributing factor in SM. We also note that anxiety may have the effect of masking a child's true language competence, which may or may not be age appropriate. In the remainder of our literature review, we note that research findings suggest many children with SM exhibit communication difficulties. We also document the challenges of conducting speech-language assessments with this population and some of the worthwhile attempts that have been made to meet these challenges, which have not been entirely satisfactory.

[1]La Salle University, Philadelphia, PA, USA
[2]Selective Mutism Anxiety Research and Treatment Center, Jenkintown, PA, USA

**Corresponding Author:**
Evelyn R. Klein, Speech-Language-Hearing Science Program, La Salle University, 1900 W. Olney Ave., St. Benilde Tower, Philadelphia, PA 19141, USA
Email: klein@lasalle.edu

The purpose of this research is to present a new method of assessing children's language abilities that employs the accepted practice of using standardized, norm-referenced testing and using a child's own parent as a test presenter in an effort to minimize the provocation of anxiety. Furthermore, we include standardized measures of extended expressive discourse that require children to demonstrate more than language comprehension (which can be assessed nonverbally).

## The Competence–Performance Distinction and SM

According to Chomsky (1965), communication *competence* refers to people's knowledge of their language and their ability to use it under ideal circumstances, whereas communication *performance* refers to people's actual use of their language ability, which may be affected by contexts and internal states accompanying a speech event (e.g., states of fatigue, emotional arousal, or anxiety). This competence–performance distinction applies to both *speech* and *language*.

The APA restriction that diagnosis of SM should exclude children whose silence can be better accounted for by a communicative disorder can be understood vis-à-vis the competence–performance distinction as requiring that a SM diagnosis is warranted for only those children who exhibit *performance* problems in communication while their communicative *competence* remains intact. If a child who is a candidate for a SM diagnosis is observed to speak clearly and fluently, a *speech* competence disorder may be dismissed as a factor contributing to his or her silence, but one must also rule out a *language* disorder as a possible contributing factor before concluding that it is only performance deficits that are the precipitators of the SM. Unfortunately, ruling out a *language* disorder can be challenging. Although some deficits in language competence (e.g., aphasia, specific language impairment, dyslexia) may be apparent, other deficits (e.g., pragmatic and narrative language impairments) are not so obvious to either parents or teachers especially when the child does not speak. However, even when there are available samples of a child's speech, the identification of a language deficit requires the analysis of a speech-language-trained professional.

## Anxiety, Social Phobia, and Language Performance Problems in SM

Researchers and therapists need to move beyond purely diagnostic criteria to try to understand *why* there might be performance problems in communication for children with SM. Although the current criteria do not officially include anxiety (which is proposed to be included as a criterion in the *DSM-V*), it is argued to be the most likely underlying causal factor by a number of researchers. Anstendig (1999) suggested that children with SM are often diagnosed with an anxiety disorder, specifically social anxiety, and are most likely to speak normally in settings where they feel relaxed, comfortable, and secure, which is generally at home. Steinhausen and Juzi (1996) reported that in a clinical sample of 100 children with SM, 85% were identified as *shy* and 66% as *anxious*.

According to the *DSM-IV-TR* (APA, 2000), social phobia is listed as a social anxiety disorder with the essential feature being fear of social or performance-based situations. Sharp et al. (2007); Chavira, Shipon-Blum, Hitchcock, Cohan, and Stein (2007);Cunningham, McHolm, and Boyle (2006); Black (1996); and Black and Uhde (1995) confirmed the relationship of SM with social anxiety and social phobia. They found that individuals with SM exhibited the characteristic social phobia tendencies of avoiding situations they fear, especially if they were afraid of being evaluated or receiving undue attention. Interestingly, according to Black and Uhde (1995), 97% of children in their study diagnosed with SM were also diagnosed with social phobia. They also found that 70% of first-degree relatives of the children with SM had a history of social phobia and 37% had a history of SM. Supporting the claim of the role of anxiety in SM, Cohan et al. (2008) performed a latent profile analysis on survey data from 130 parents of children from 5 to 12 years of age and reported that children with SM fell into one of three (anxious) groups: (a) anxious–mildly oppositional (when pressured to speak), (b) exclusively anxious, and (c) anxious–communication delayed. In contrast, other researchers (e.g., Melfsen, Walitza, & Warnke, 2006) do not support SM as a manifestation of high social anxiety, basing their findings on significantly lower anxiety inventory scores in children with SM compared with socially phobic children.

Although social anxiety may be considered a pathway to the manifestation of SM, it does not explain why children withhold speech in particular (as opposed to exhibiting school refusal, obsessive-compulsive disorder, trichotillomania, panic disorder, etc.). Anstendig (1999) argued that children with language formulation problems, who also have a biological predisposition for behavioral inhibition and anxiety, are more likely to develop SM. As Anstendig stated, "The anxiety of interacting with others in social situations can be compounded by difficulty with speech and language. Therefore, withholding speech is a way to ease anxiety brought on by both social situations and difficulty in using language." (Anstendig, 1999, p. 431) Although this may be the case, not all children follow the same pattern. For example, Richards (2011) found that "The child with SM might surprise you and talk with you in the evaluation. That does not eliminate SM as a possibility." (p. 41). Recently, Kurtz (2011) reported that during an interview where children with SM had a brief rapport-building segment with an unknown examiner followed by questions, 27% of 56 children verbally responded to the first question the examiner

asked, 36% answered when the second question was posed, and 43% of the children verbally responded to the third question.

## Emerging Evidence for Language Competence Deficiencies in SM

The *DSM-IV-TR* specifically states that although children with SM "generally have normal language skills, there may occasionally be an associated Communication Disorder" (APA, 2000, p. 126). It should be noted that most children with SM do not show impairment in language comprehension. Most understand words and sentences at age-appropriate levels even though they do not speak in selected settings. However, according to Sharp et al. (2007), language-related difficulties appear to be a significant risk factor for the development of SM. Additional research supports the claim that speech and language deficits can be precipitating factors in the presentation of SM (Anstendig, 1999; Cantwell & Baker, 1985; Nowakowski et al., 2009; Steinhausen & Juzi, 1996; Tancer, 1992). Manassis et al. (2003) concluded that 43% of 14 children diagnosed with SM scored in the clinical range on at least one measure of speech-language assessment. McInnes, Fung, Manassis, Fiksenbaum, and Tannock (2004) evaluated 7 children with SM and 7 children with social phobia. Their study included narratives and pragmatic language skills such as turn taking, topic maintenance, and eye contact. The group of children with SM displayed normal nonverbal, cognitive, and receptive language abilities. However, they produced significantly shorter narratives with fewer details than the group with social phobia. Producing narratives requires cognitive and linguistic abilities that are more demanding than engaging in conversation, especially in a social setting with peers and teachers (Abbeduto, Benson, Short, & Dolish, 1995). It is not surprising then that McInnes et al. (2004) suggested that children with SM may have subtle expressive language deficits that affect academic performance because "higher level language skills . . . are critical for academic success" (p. 313).

In a study by Steinhausen and Juzi (1996), 38% of 100 children with SM were reported by their parents to have a speech or language disorder. In another study, Andersson and Thomsen (1998) determined that one third of 30 children with SM had delayed speech development. Other research findings indicate that 20% to 50% of children with SM may have some type of a communication problem (Cohan et al., 2008; Kolvin & Fundusis, 1981; Wilkins, 1985; Wright, 1968) or neurodevelopmental disorder (Kristensen, 2000) that may include deficits in auditory–verbal memory span (Kristensen & Oerbeck, 2006). Cleator and Hand (2001) evaluated 5 children at home using four standardized tests measuring receptive language and speech articulation, but no standardized and norm-referenced diagnostic tool was used to evaluate expressive language. Their results indicated that 4 children had communication disorders. The findings of these accumulated studies suggest that communication deficits may be more prevalent in children with SM than assumed. To further explore this possibility, standardized and norm-referenced testing is a valuable means for comparing children with SM to peers.

## Norm-Referenced Standardized Assessments of Expressive Language Competence

Few studies of children with SM have included standardized, norm-referenced measures of expressive language competence. Standardized measures provide a sound basis for comparing test results across settings with a variety of participants. Norm-referenced tests provide an estimate of a person's ability relative to his or her age cohort, which is essentially a built-in comparison group. Some researchers have begun to recognize the importance of incorporating such measures in studies, but the fact that children with SM often do not talk with professionals or in unfamiliar environments makes this difficult. Cleator and Hand (2001), described above, found it helpful to test children in their homes so they could feel more relaxed and also solicited the help of parents as needed. Other researchers have also recruited the assistance of parents in various ways, such as having parents carry out and monitor treatment protocols for generalization (Khanna & Kendall, 2009; Schill, Kratochwill, & Gardner, 1996). It may also be advantageous to use parents to support administration of standardized, norm-referenced tests that have clear instructions so that parents can learn to present them. The *Test of Narrative Language* (TNL; Gillam & Pearson, 2004) is a well-known and widely used instrument that satisfies these requirements. In fact, the National Association for the Education of Young Children (2002) supports assessment partnerships when undertaken with sound knowledge and sensitivity.

## Study Overview

The primary purpose of this study was to obtain valid assessment results of receptive and expressive vocabulary and connected discourse, primarily narrative language, in children with SM. Presently, there is a "need for a consistent definition and systematic assessment of SM that can be applied across settings" (Viana et al., 2009, p. 65). Standardized, norm-referenced speech and language assessment results can help determine whether a child has a speech and/or language deficit. Because some children with SM will speak to parents but not a professional, another purpose of this study was to explore the efficacy of employing parents to help deliver test stimuli to their children. Furthermore, given that children with SM typically speak with ease at home with their immediate family, the

authors hypothesized that assessments based on test administrations by the professional, an unfamiliar evaluator, would underestimate the children's true language capabilities in comparison with assessments based on test administrations supported by parents.

## Method

### Participants

A total of 33 participants were recruited from a practice in the Philadelphia area that specializes in the treatment of SM. Parents of children between the ages of 5 and 12 years were invited to take part in the study in which their children would receive a speech-language evaluation and report. In all, 33 children within this age range who met the *DSM-IV-TR* criteria for SM, and who passed vision and hearing screenings within the previous year took part in this study. Diagnosis was determined by a licensed psychologist, based on parents' completion of the *Behavior Assessment System for Children–Second Edition* (BASC-2), *Structured Developmental History* (SDH; Reynolds & Kamphaus, 2006), and a second structured therapist-created questionnaire based on the *DSM-IV-TR* criteria, specific to SM. Interviews with parents by the psychologist verified that the *DSM-IV-TR* criteria for SM were met with 100% agreement for the diagnosis. Exclusion criteria for the study included receiving full-time special education services for intellectual disability or autism spectrum disorder, bilingualism, or scoring below average on vocabulary measures. As it turned out, no children were excluded.

All children were Caucasian and were full-time students receiving regular education with 3 receiving additional part-time learning support services. The sample of 19 girls and 14 boys had a mean age of 7.25 years ($SD$ = 1.58 years) with 49% at ages 5 and 6, 33% at ages 7 and 8, and 18% at ages 9 to 12. The participants' family sizes ranged from 1 to 5 children with 52% of participants being the first born in their family. The primary language at home was English for all children. The majority of parents had a bachelor's degree (70% of mothers and 64% of fathers). Nearly 70% of the parents reported having a history of anxiety in their immediate family. In addition, 33% of the children were taking small doses of a selective serotonin reuptake inhibitor (SSRI) to reduce anxiety.

### Direct Assessment Measures of Communicative Competence

*The Peabody Picture Vocabulary Test–4 (PPVT-4)*. The PPVT-4 (Dunn & Dunn, 2007) assesses receptive vocabulary in single word form. It is a norm-referenced, standardized test to assess auditory comprehension for receptive vocabulary, and covers 20 categories of content and parts of speech. A finger-pointing response is required to indicate which of four pictures corresponds to the vocabulary word heard. Internal consistency reliability is high ($r$ = .93). The PPVT-4 is considered a valid vocabulary measure demonstrating good concurrent validity with the *Expressive Vocabulary Test–2* (EVT-2; Williams, 2007) and the *Clinical Evaluation of Language Fundamentals–4* (CELF-4; Semel, Wiig, & Secord, 2003) scales.

*EVT-2*. The EVT-2 (Williams, 2007) assesses expressive vocabulary for single words. It is a norm-referenced, standardized test designed to assess expressive vocabulary and word retrieval. This test covers 20 categories of content and parts of speech. Children name the picture seen after hearing the examiner's prompt. Internal consistency reliability is very high ($r$ = .93–.94) as is test–retest reliability ($r$ = .95). Concurrent validity with the CELF-4 is reported to be good. The EVT-2 is considered a valid vocabulary measure.

*TNL–Narrative Comprehension (TNL-C) and TNL–Oral Narration (TNL-O)*. To capture performance with language formulation, the TNL (Gillam & Pearson, 2004) considers the interactive nature of oral language within the context of functional discourse. The TNL is a norm-referenced, standardized test designed to measure children's ability to listen to stories, answer questions about stories, and tell stories. The TNL has two subscales: TNL-C and TNL-O. Each subscale has three types of stories: (a) a script with "no picture" cues, (b) five "sequenced pictures," and (c) a "single scene." The TNL-C, although considered a language comprehension measure, requires the child to verbalize his or her responses about orally presented stories. The TNL-O requires the child to listen to and retell a story, and, in addition, to generate his or her own stories about pictures. See Table 3 for a succinct review of the TNL tasks and the order of test administration.

*Scoring*. For the TNL-C, the manual requires posing 9 to 11 questions to the children for each of three component stories of the subscale. The manual also provides a list of acceptable responses. Children receive scores based on the number of acceptable responses they provide. For the TNL-O, scoring differs depending on the particular story within the subscale. For the first story, *McDonald's Retell* (where the children are asked to retell a story that they have just heard), children receive 1 point for each word they mention that the manual identifies as essential. In the second story, *Late for School* (where there are five sequenced pictures), the manual identifies essential content items for each picture. Children receive a point for each content item mentioned. They receive additional points for indicating temporal relationships, causal relationships, use of correct grammar, and story coherence and completeness. The third task, *Aliens Story*, requires the children to produce their own story. They receive points for including the story-grammar elements of setting, characters, problem and resolution, correct grammar, describing objects, and story coherence and completeness.

The raw scores for each component story are added to form a total raw score for each subscale, the TNL-C and TNL-O scores. These raw subscale scores are converted to age-determined standard scores, based on a mean of 10 and a standard deviation of 3. According to the manual, scores between 8 and 12 are considered average. Scores ranging from 13 to 20 are considered above average to very superior. Scores of 6 and 7 are considered below average, 4 and 5 poor, and 1 to 3 very poor.

Telling a story is more cognitively demanding than answering questions about a story, and narrative tasks are more demanding than vocabulary tasks (previously assessed). Using the TNL, both receptive and expressive language competence can be assessed in a more structured and thematic manner. Internal consistency is good ($r$ = .76–.88) as is test–retest reliability ($r$ = .85) and interrater reliability for scoring story transcripts (91%–98%). Validity is considered good with a positive prediction exceeding .85 for identifying language disorders, indicating good sensitivity (correct identification of those who have a language disorder) and specificity (correct identification of those who do not have a language disorder).

There was 100% interrater agreement on the TNL scores. All the testing sessions were videotaped, transcribed, and scored by two trained graduate students. A final check of the transcriptions and scoring was performed by a speech-language pathologist (SLP).

### Measures Provided by Parents

The BASC-2 SDH provided a format for gathering information regarding developmental, familial, health, social, and educational backgrounds.

### Procedure

After parents signed the Institutional Review Board informed consent form, the parents and their children began their participation in the study that comprised two 1.5-hr segments. A licensed SLP with more than 25 years experience, who was also a licensed psychologist (referred to as the *professional* throughout this article), performed the training and testing.

*Parent training*. For 32 of the 33 children, mothers presented the testing material. The professional accompanied parents to a separate room where they received instructions. Each test training took approximately 10 min and included a training video, information from the manual for following standardized procedures, and a protocol page with "dos and don'ts." Parents were instructed to read the exact wording of test items, how to position themselves and the materials, how much time to permit per item, where to begin and end test items, the number of permissible stimulus repetitions, how and when to deliver praise, and appropriate responses to children's questions. Parents were given starting points (basal) for each measure, and ceiling guidelines were written on each test. However, the professional who watched the testing session live on a video monitor (while it was being recorded) instructed the parents to present additional items if basal or ceiling criteria were not met. After the child completed each test, the parent met with the professional in another room to receive training for the next measure, while the child remained in the testing room.

The testing room was approximately 18 square feet with a large one-way mirror for viewing and videotaping, and shelves holding toys, games, and supplies. All administrations were video–audio recorded (using hidden equipment) with parent permission. Parents presented the testing stimuli by reading the directions to the child and noted their child's answers, but the professional was ultimately responsible for tracking responses and scoring.

*Testing order*. All tests were given in the following order: PPVT-4, EVT-2, TNL-C, and TNL-O. Vocabulary measures preceded narrative language measures, and within each of those categories, receptive testing preceded expressive testing. A total of 18 parents tested their children first (in the morning) followed by a lunch break. After lunch, the professional conducted the same tests on those children. The other 15 children were tested first by the professional in the morning and then by their parent in the afternoon. Assignment to testing by parent first or professional first was randomized based on a coin toss at the time the family arrived for the study.

## Results

### Quality of Parent Test Presentation

A random sample of one third of the 33 parents (11) was observed and rated on test presentation practices by a SLP and graduate student (who conducted these ratings as part of routine responsibilities in the clinic) using the General Test Administration Practices Checklist (GTAPC; Sattler, 2001). This checklist comprises 40 items (each ranging from 1 [*poor*] to 5 [*excellent*]) related to the test presenter's ability to communicate and follow standardized procedures. Table 1 presents the mean GTAPC scores assigned by the raters for parents' presentation of each language measure, indicating that parents were judged to have effectively presented stimuli from the standardized tests with 96% accuracy. The intraclass correlation coefficients, ranging from $r$ = .72 to $r$ = .88, indicate good agreement between the two raters about parental performance.

### Factors Related to Children's Verbalization During Testing

A chi-square test of independence for order of testing (parent or professional first) and whether the child spoke to

**Table 1.** Professional Raters' Scores of Parents on the General Test Administration Practices Checklist

| Test administered | SLP rating of parents (n = 11 parents) | | SLP graduate student rating of parents (n = 11 parents) | | Intraclass correlation coefficient |
|---|---|---|---|---|---|
| | *M*[a] | *SD* | *M*[a] | *SD* | *r* |
| PPVT-4 | 191.7 | 6.85 | 191.3 | 7.10 | .88 |
| EVT-2 | 192.2 | 6.78 | 192.6 | 5.52 | .77 |
| TNL | 191.5 | 7.88 | 192.6 | 5.28 | .72 |

Note: SLP = speech-language pathologist; PPVT = *Peabody Picture Vocabulary Test*; EVT = *Expressive Vocabulary Test*; TNL = *Test of Narrative Language*. Independent *t*-test results indicated no significant differences between raters.
[a]Based on a score of 200.

the professional was conducted. No significant relationship was found, $\chi^2(1) = 2.53$, $p > .05$.

All 33 children spoke to their parents during testing. On the PPVT-4, which only required pointing to a picture (no verbal response), all but 1 of the participants also responded nonverbally to the professional evaluator. When children were asked to name pictures from the EVT-2, the number of children who responded verbally to the professional dropped to 20 (61%). For question comprehension about stories (TNL-C), the same 20 children (61%) spoke to the professional. On the self-generated oral narration portion (TNL-O), 18 of these 20 children (55% of the total 33) spoke to the professional.

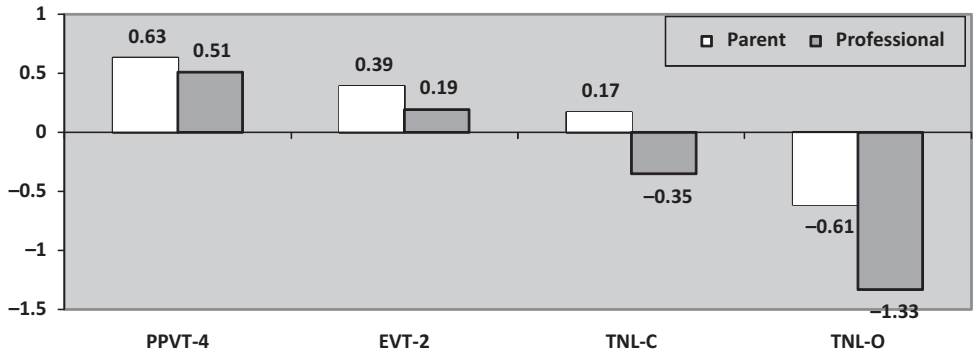
The fact that children in the study were all recruited from the same practice permitted analysis of the relationship between the number of previous treatment visits and speaking to the unfamiliar professional. Prior visits for the 33 children ranged from 0 to 28. A chi-square test of independence revealed no significant relationship between the number of prior therapy visits and whether the child spoke to the professional, $\chi^2(13, N = 33) = 16.46$, $p = .225$, Cramer's $V = .19$.

Additional analyses investigated the relationship between medication use for anxiety and whether the children spoke to the professional. A total of 11 of the 33 children were taking a SSRI at the time of testing. A chi-square test of independence indicated nonsignificance, $\chi^2(1, N = 33) = 0.550$, $p = .458$, Cramer's $V = .13$, for children's medication use and speaking to the professional.

## *Comparison of Scores in Children's Performance Between Parent and Professional Testing*

All test scores were transformed into a common *z* score format, and analyses were then performed using *z* scores. Prior to transformation, the PPVT-4 and EVT-2 were computed with a mean of 100 and a standard deviation of 15, and the TNL was computed with a mean of 10 and standard deviation of 3. A 2 × 4 repeated measures ANOVA was performed with the 18 children who participated in all four measures for both parent and professional. The within-participant factors were test presenter (parent, professional) and type of measure (PPVT-4, EVT-2, TNL-C, and TNL-O). There was a main effect for test presenter of $F(1, 17) = 8.80$, $p < .009$, $\eta^2_p = .34$, power = .80 with parents producing a significantly higher test score for children than the professional. There was also a main effect for type of measure with children performing significantly lower on the TNL-O measure than on the other three measures. Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(5) = 21.47$, $p = .001$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity, $\varepsilon = 0.54$, with $F(1.63, 27.65) = 13.72$, $p < .001$, $\eta^2_p = .45$, power = .97. The interaction of test presenter by type of measure did not reach significance. *Mauchly's test* indicated that the assumption of sphericity had been violated, $\chi^2(5) = 29.59$, $p < .001$; therefore, degrees of freedom were corrected using Greenhouse–Geisser estimates of sphericity, $\varepsilon = 0.49$, with $F(1.47, 24.93) = 2.946$, $p = .085$, $\eta^2_p = .15$, power = .44.

See Figure 1 for means and statistical test results for the language measures (all based on *z* scores) by parent and professional. Table 2 presents standard score means, standard deviations, and pairwise comparisons for each measure presented by parent and professional. Significant differences were found between test presenters on the EVT-2, TNL-C, and TNL-O. There was no significant difference on the PPVT-4. For the three statistically significant findings in which children performed better with their parents, effect sizes measured by Cohen's *d* indicated that two measures, the TNL-C and the TNL-O, were moderate. These effect sizes indicated that children scored more than half of a standard deviation higher when testing stimuli were delivered by parents than by the professional. Although the children scored significantly higher on the EVT-2 with parents, the effect size was small.




| | PPVT-4 | EVT-2 | TNL-C | TNL-O |
|---|---|---|---|---|
| | $n = 32$ | $n = 20$ | $n = 20$ | $n = 18$ |
| | $t(31) = 1.42$ | $t(19) = 2.24$ | $t(19) = 2.77$ | $t(17) = 2.43$ |
| | $p = .166$ | $p = .038$* | $p = .012$* | $p = .026$* |
| 95% CI | [-0.05, 0.29] | [0.01,0.39] | [0.13, 0.91] | [0.10, 1.35] |
| | Cohen's $d = 0.15$ | $d = 0.19$ | $d = 0.69$[b] | $d = 0.65$[b] |

**Figure 1.** Means (in *z* scores) and paired *t*-test results with 95% CIs for children who spoke to both parent and professional on language measures.
Note: CI = confidence interval.
[b]*d* = moderate effect size.
*$p < .05$.

**Table 2.** Paired *t*-Test Results With Means and Standard Deviations for Vocabulary and Narrative Language Tests With Parent and Professional Administering Test Stimuli

| Test administered | *n* | Scores for children by parent | | Scores for children by professional | | Significance | |
|---|---|---|---|---|---|---|---|
| | | *M (SD)* | *Range* | *M (SD)* | *Range* | *p* | Cohen's *d* |
| PPVT-4 | 32 | 109.47 (12.26) | 83–139 | 107.69 (12.24) | 83–136 | .166 | 0.15 |
| EVT-2 | 20 | 105.85 (16.51) | 60–135 | 102.85 (14.61) | 65–126 | .038* | 0.19 |
| TNL-C | 20 | 10.50 (2.28) | 8–16 | 8.95 (2.19) | 3–12 | .012* | 0.69[b] |
| TNL-O | 18 | 8.17 (3.07) | 3–14 | 6.00 (3.60) | 1–12 | .026* | 0.65[b] |

Note: PPVT = *Peabody Picture Vocabulary Test*; EVT = *Expressive Vocabulary Test*; TNL-C = *Test of Narrative Language–Narrative Comprehension*; TNL-O = *Test of Narrative Language–Oral Narration*. PPVT-4 and EVT-2 are based on a mean of 100 and standard deviation of 15. TNL-C (narrative comprehension) and TNL-O (oral narration) are based on a mean of 10 and standard deviation of 3.
[b]*d* = moderate effect size.
*$p < .05$.

## *Comparison of Performances Across Language Tests With Parent-Supported Administration*

See Figure 2 for means and statistical test results from all 33 children for the language measures (all based on *z* scores) when presented by parents. There was no significant difference in performance between receptive vocabulary (comprehension) and expressive vocabulary (production) when parents presented the test stimuli. However, when children were told a story and asked to answer questions aloud (to assess narrative comprehension), their narrative comprehension score was significantly lower than their expressive vocabulary scores. When asked to retell a story and create a story from a picture, the children's oral narration scores were the poorest of all performances. In fact, the drop in mean performance from the TNL-C to the TNL-O was almost one standard deviation as indicated by the effect size of $d = 0.91$ with parents presenting both test stimuli. The number of children (out of 33) who scored more than one standard

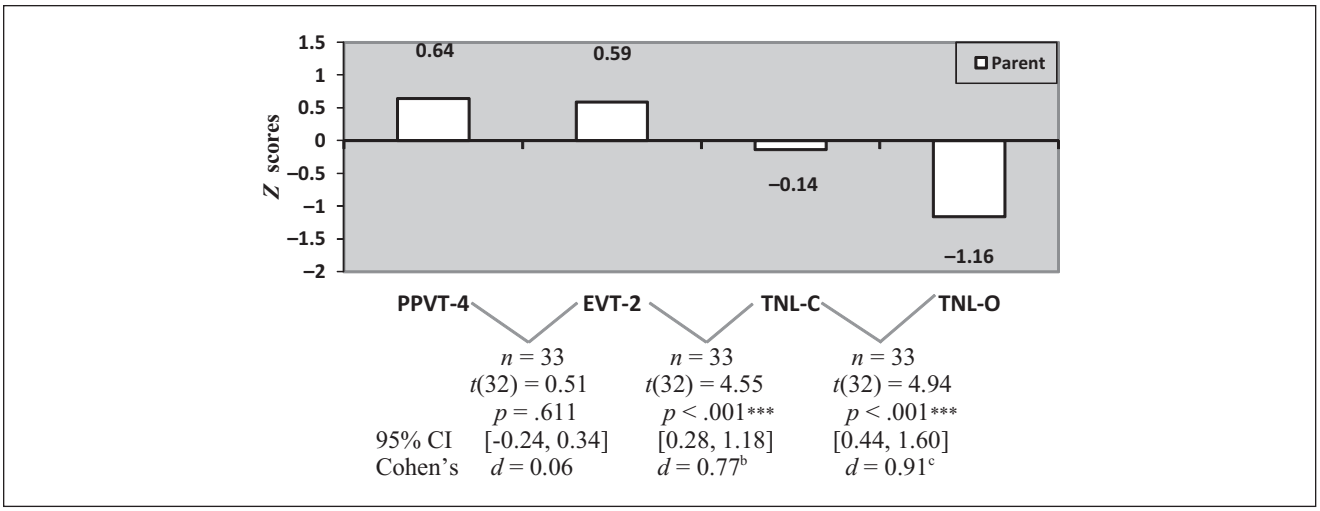



**Figure 2.** Means (in z scores) and paired post hoc t-test results with 95% CIs for children when parents presented different language measures.
Note: CI = confidence interval; PPVT-4 = *Peabody Picture Vocabulary Test–4;* EVT-2 = *Expressive Vocabulary Test–2;* TNL-C = *Test of Narrative Language–Narrative Comprehension;* TNL-O = *Test of Narrative Language–Oral Narration*.
[b]d = moderate effect size.
[c]d = large effect size.
***p < .001.

**Table 3.** Children's (*N* = 33) Mean Percentage Accuracy on Six Areas of the Test of Narrative Language

| Story | Narrative comprehension: *Shipwreck Story* | Narrative comprehension: *McDonald's Story* | Narrative comprehension: *Dragon Story* | Oral narration: *Late for School Story* | Oral narration: *Aliens Story* | Oral narration: *McDonald's Retell Story* |
|---|---|---|---|---|---|---|
| Task | Answer questions about story told with five sequenced pictures shown | Answer questions about story told without pictures | Answer questions about story told with one event scene shown | Tell own story when shown five sequenced pictures | Tell own story from one event scene shown | Retell story without pictures |
| Testing order | 3 | 1 | 5 | 4 | 6 | 2 |
| Mean % Correct | 74.8 | 57.3 | 57.0 | 35.1 | 33.2 | 22.7 |
| Number scoring below 60% correct | 6 | 20 | 13 | 29 | 27 | 30 |

deviation below the mean increased dramatically across the four tests: 1 child for the PPVT-4, 1 child for the EVT-2, 5 children for the TNL-C, and 18 children for the TNL-O.

Table 3 provides a breakdown of the two subscales of the TNL (the TNL-C for narrative comprehension [answering questions about a story] and TNL-O for oral narration [generating sentences to tell a story]). Although the TNL manual does not provide statistical information about the component stories of the subscales, we thought it might be useful to examine how our sample of children with SM performed on the separate tasks. Overall, retelling a story without pictures was most difficult for the 33 children with SM who obtained a mean of 22.7% accuracy with their parents providing the test stimuli. The easiest of the measures was answering questions about the shipwreck story given a sequence of five pictures to view (74.8% accuracy). Removing visual cues (pictures) had a negative effect on children's performance. Table 3 displays the percentages of children scoring less than or equal to 60% accuracy for each of the component stories of the two TNL subscales (TNL-C and TNL-O). Within each subscale, the largest number of children who scored less than or equal to 60% occurred in the McDonald's stories that had no pictures. It is also evident that the 60% accuracy performances were generally lower for the TNL-O. In all, 26 of the 33 children scored lower on the TNL-O than on the TNL-C, 4 children scored higher, and 3 scored the same.

## Discussion

The purpose of this study was to obtain accurate assessment results of receptive and expressive language competence for vocabulary and connected discourse in children with SM. In our efforts to develop a language measurement protocol, we recruited parents to present testing information that the professional scored and interpreted. Parent training for presenting the testing materials was highly successful as evidenced by accuracy on the GTAPC (Sattler, 2001). As expected, the children with SM talked to their parents but did not uniformly talk to the professional. Children with SM performed significantly better when parents delivered testing information that required verbal output. Having professionals administer test stimuli to children with SM, which is the standard practice, may underestimate their true language competence. In fact, if the professional had been the only source of gathering test data in this study, complete results from only 18 of the 33 cases would have been obtained.

An analysis of the performance of all 33 children when their parents presented the test stimuli revealed differences depending on the particular tests compared. Although children scored similarly on the receptive (pointing) and expressive (naming) vocabulary tests, a decrease was observed as the narrative tasks changed from narrative comprehension (answering questions about a story read to them) to oral narration (telling a story on their own). Children in this study scored within one standard deviation of the population mean on three of the four language tasks (PPVT-4, EVT-2, and TNL-C). However, even with parents, *z* scores on the test that assessed expressive language formulation (TNL-O) dropped more than one-half standard deviation below the mean of the participants' age cohort. This is an important finding because *z* scores convey the test-takers' performance relative to others in their age cohort. One would not expect a child's relative position in their age cohort to change unless he or she was exhibiting a specific deficit. Expressive language formulation specifically for connected discourse, as measured by the TNL-O, reflected children's difficulty retelling a story (*McDonald's Retell*), creating a story corresponding to a series of five pictures (*Late for School*), and creating a story corresponding to a pictured scene (*Aliens Story*). The decontextualized language production tasks of the TNL-O were more complex than the narrative comprehension tasks (TNL-C), and this is where many of the children with SM exhibited language competence deficits. Using the standard scores based on the two TNL subscales, children performed almost one standard deviation lower in expressing themselves using connected discourse than they did in answering questions about a story told to them, both verbal tasks. Based on their performance on the TNL-O, 42% of the participants exhibited expressive narrative language deficits, receiving a standard score of 5 or less, which placed them at or below the fifth percentile rank.

The following example illustrates how the performance of one 7-year-old child decreased as the language task changed from requiring a one-word response to requiring a formulated sentence on the McDonald's story. This story comprises 13 sentences that include the following characters: a mother; her two children, Lisa and Raymond; and a store clerk. When asked, "What kind of milkshake did Raymond want?" the child easily and correctly answered with a one-word utterance, "Vanilla." The same child had much difficulty formulating a response when asked to retell the story and stated, "Um, they (pause) went to McDonalds and they ordered and got food, and that all."

Given our findings of a disparity between receptive and expressive narrative competence, it cannot be assumed that all children with SM have full expressive language ability. Although previous research has revealed subtle language impairments in children with SM, the measures used generally included receptive language, nonverbal responses, parental reports, and functional assessments that did not require speaking (Dow, Sonies, Scheib, Moss, & Leonard, 1995; Dummit et al., 1997; Kern, Starosta, Cook, Bambara, & Gresham, 2007; Manassis et al., 2003) and thus have potentially underestimated the extent of language impairments in this population, specifically in connected discourse. It is interesting to note that in the current study, both the TNL measures (oral narration as well as narrative comprehension) required spoken responses. The issue for these children did not appear to be "talking" per se but narrative language formulation because their scores dropped below the normal range when they moved from answering questions verbally to developing a story.

Of further interest is the possibility that an expressive language formulation problem may be a contributing factor to the emergence of a child's SM. It is certainly possible that children with SM, who have difficulty with language formulation, may be anxious due to their perceived inadequacy in generating novel thoughts and converting them into more decontextualized spoken language. Children with these difficulties tend to do better with spontaneous verbal output in known contexts (as occurs at home) as compared with elicited language (where there is a prompt or question) or with decontextualized language (often at school; Peets, 2009).

### *Study Limitations*

Although this study employed a modest sample size of 33 children, the sample provided additional data about children with SM, a low-incidence population, which often is difficult to assess with individualized, direct assessment (Cleator & Hand, 2001; Kern et al., 2007; McInnes et al., 2004). With a modest sample, generalizability is limited

and more research is needed with a larger and more varied sample of children.

Although this study benefited from use of an office setting with a one-way mirror and videotaping equipment, unobtrusive to the study participants, which seemed to work well for the children in this study, some children with SM may not willingly participate in testing in an office setting. It is also possible that the children in our study were not as relaxed as they might be in a more familiar setting, and thus we cannot be sure that we have obtained every child's true language competence. It would be advantageous to have an independent measure of a child's level of arousal, such as a skin conductance monitor, incorporated into the testing procedure.

Although this study used standardized, norm-referenced language measures, there is more to communication that should be evaluated in this population, including speech analyses and assessments of more comprehensive grammatical and pragmatic competencies. It may also be useful to have additional measures of cognitive and behavioral functioning.

## Future Directions

*Implementing parent-supported testing*. This research demonstrated the efficacy and value of including parents in the process of assessing communicative competence in children with SM. It will be important for professionals adopting a parent-supported approach to safeguard the consistency and adequacy of the parent training process. In particular, parents must be literate and properly trained to present test stimuli to their children. This will require that a licensed professional train and monitor the entire process, and provide parents with practice opportunities and ongoing support, using video–audio monitoring throughout the testing situations. In cases where a child cannot perform in the office setting, special arrangements will be required to videotape the testing situation at home. Another possibility is to use videoconferencing (e.g., Skype) whereby the professional can view the parent-assisted testing session occurring in the child's home or school. Professionals will also need to review, score, and interpret all testing results.

*Implications for assessment and therapy*. We encourage use of our assessment model to help professionals determine whether a child with SM has an expressive language deficit affecting connected discourse. Our following model is suggested for how to proceed with a child who meets a minimum requirement for SM diagnosis: consistent failure to speak in specific social situations (in which there is an expectation for speaking, for example, at school) despite speaking in other situations. The model stipulates an orderly progression of assessments using standardized, norm-referenced measures and provides recommendations for whether to include language treatment (i.e., semantics, morphology, syntax, pragmatics).

Step 1: Assess vocabulary comprehension (e.g., PPVT, *Receptive One Word Picture Vocabulary Test*)
Step 2: Assess vocabulary production (e.g., EVT, *Expressive One Word Picture Vocabulary Test*)
Step 3: Assess narrative comprehension (e.g., TNL-C, *Test of Auditory Comprehension of Language*)
Step 4: Assess oral narration (TNL-O, *The Expressive Language Test*)

In the case where a child meets the requirements for SM and performs within the normal range on all of the above assessments, it would be advisable for the child to receive therapy modified specifically for SM. For example, using an approach that targets social communication anxiety reduction such as cognitive behavioral therapy including strategies such as reducing negative reinforcement of mutism, stimulus fading, contingency management, and desensitization may help reduce anxiety in various communication settings.

In the case where a child meets the requirements for SM and does not perform within the normal range on either vocabulary comprehension or production, it would be advisable for the child to receive a more comprehensive assessment for possible education support and speech-language support in addition to therapy that has been modified specifically for SM.

In the case where a child meets the requirements for SM and does not perform within the normal range on the TNL, it is advisable for the child to receive additional assessments and therapy that has been modified specifically for SM as well as receptive and expressive language support from a SLP.

When a child is competent with regard to the comprehension and production of words and sentences, it does not mean that he or she is proficient with language formulation of connected discourse. Although speech-language-related difficulties appear to be a significant risk factor for the development of SM (Sharp et al., 2007), it is important to determine whether a child with SM exhibits deficits in expressive language (connected discourse) so that professionals can provide appropriate language therapy simultaneously with treatment for SM.

Practically speaking, it seems likely that the more a child avoids speaking, the less likely he or she is to overcome anxiety in speaking situations, which may exacerbate the SM condition and which may prevent the practice needed to develop discourse skills. Therefore, we are suggesting that an appropriately credentialed professional conduct speech and language testing on children suspected of having SM using the parent-assisted method described in this article so that appropriate treatment may be initiated as early as

possible. Obtaining a comprehensive speech and language evaluation using standardized and norm-referenced measures is also beneficial for securing individualized educational plans from most public schools, which for many children is their primary opportunity for intervention.

## Acknowledgments

The authors wish to acknowledge Janice Gordon, Robin Kushner, and Ami Van Dine for their assistance on this research.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## References

Abbeduto, L., Benson, G., Short, K., & Dolish, J. (1995). Effects of sampling context on the expressive language of children and adolescents with mental retardation. *Mental Retardation, 33*(5), 279–288.

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

Andersson, C. B., & Thomsen, P. H. (1998). Electively mute children: An analysis of 37 Danish cases. *Nord PsykiatrTidsskr, 52*, 231–238.

Anstendig, K. D. (1999). Is SM an anxiety disorder? Rethinking its *DSM-IV* classification. *Journal of Anxiety Disorders, 13*, 417–434.

Black, B. (1996). Social anxiety and selective mutism. *American Psychiatric Press Review of Psychiatry, 15*, 469–495.

Black, B., & Uhde, T. (1995). Psychiatric characteristics of children with selective mutism: A pilot study. *Journal of American Academy of Child & Adolescent Psychiatry, 34*, 847–856.

Cantwell, D., & Baker, L. (1985). Speech and language: Development and disorders. In M. Rutter & L. Hersov (Eds.), *Child and adolescent psychiatry: Modern approaches* (2nd ed., pp. 531–533). London, England: Blackwell Scientific Publications.

Chavira, D., Shipon-Blum, E., Hitchcock, C., Cohan, S., & Stein, M. B. (2007). Selective mutism and social anxiety disorder: All in the family? *Journal of the American Academy of Child & Adolescent Psychiatry, 46*, 1464–1472.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Cleator, H., & Hand, L. (2001). SM: How a successful speech and language assessment really is possible. *International Journal of Language & Communication Disorders, 36*, 126–131.

Cohan, S. L., Chavira, D. A., Shipon-Blum, E., Hitchcock, C., Roesch, S. C., & Stein, M. B. (2008). Refining the classification of children with SM: A latent profile analysis. *Journal of Clinical Child & Adolescent Psychology, 37*, 770–784.

Cohan, S. L., Price, J. M., & Stein, M. B. (2006). Suffering in silence: Why a developmental psychopathology perspective on SM is needed. *Journal of Developmental & Behavioral Pediatrics, 27*, 341–355.

Cunningham, C. E., McHolm, A. E., & Boyle, M. H. (2006). Social phobia, anxiety, oppositional behavior, social skills, and self-concept in children with specific selective mutism, generalized selective mutism, and community controls. *European Child & Adolescent Psychiatry, 15*, 245–255.

Dow, S. P., Sonies, B. C., Scheib, D., Moss, S. E., & Leonard, H. L. (1995). Practical guidelines for the assessment and treatment of selective mutism. *Journal of American Academy of Child & Adolescent Psychiatry, 34*, 836–846.

Dummit, E. S., Klein, R. G., Tancer, N. K., Asche, B., Martin, J., & Fairbanks, J. A. (1997). Systematic assessment of 50 children with SM. *Journal of American Academy of Child & Adolescent Psychiatry, 35*, 653–660.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Minneapolis, MN: Pearson Assessments.

Gillam, R. B., & Pearson, N. A. (2004). *Test of narrative language: Examiner's manual*. Austin, TX: Pro-Ed.

Kern, L., Starosta, K. M., Cook, C. R., Bambara, L. M., & Gresham, F. R. (2007). Functional assessment-based intervention for SM. *Behavioral Disorders, 32*(2), 94–108.

Khanna, M. S., & Kendall, P. C. (2009). Exploring the role of parent training in the treatment of childhood anxiety. *Journal of Consulting and Clinical Psychology, 77*(5), 981–986.

Kolvin, I., & Fundusis, T. (1981). Elective mute children: Psychological, development and background factors. *Journal of Child Psychology and Psychiatry, 22*, 219–232.

Kristensen, H. (2000). SM and comorbidity with developmental disorder/delay, anxiety disorder, and elimination disorder. *Journal of American Academy of Child & Adolescent Psychiatry, 39*, 249–256.

Kristensen, H., & Oerbeck, B. (2006). Is SM associated with deficits in memory span and visual memory?: An exploratory case-control study. *Depression and Anxiety, 23*, 71–76.

Kurtz, S., & Comer, J. (2011, October). *Treating professional's guide to the functional behavioral assessment of selective mutism*. Paper presented at the Selective Mutism Group – Child Mind Institute Selective Mutism Conference, New York, NY.

Manassis, K., Fung, D., Tannock, R., Sloman, L., Fiksenbaum, L., & McInnes, A. (2003). Characterizing SM: Is it more than social anxiety? *Depression and Anxiety, 18*, 153–161.

McInnes, A., Fung, D., Manassis, K., Fiksenbaum, L., & Tannock, R. (2004). Narrative skills in children with SM: An exploratory study. *American Journal of Speech-Language Pathology, 13*, 304–315.

Melfsen, S., Walitza, S., & Warnke, A. (2006). The extent of social anxiety in combination with mental disorders. *European Child & Adolescent Psychiatry, 15*(2), 111–117.

National Association for the Education of Young Children. (2002). Assessment: Partnering with parents. *Scholastic Early Childhood Today, 17*, 15–16.

Nowakowski, M. E., Cunningham, C. C., McHolm, A. E., Evans, M. A.,Edison, S., St. Pierre, J., . . . Schmidt, L. A. (2009). Language and academic abilities in children with selective mutism. *Infant and Child Development, 18*, 271–290.

Peets, K. F. (2009). The effects of context on the classroom discourse skills of children with language impairment. *Language, Speech, and Hearing Services in Schools, 40*, 5–16.

Reynolds, C., & Kamphaus, R. (2006). *The Behavior Assessment System for Children* (2nd ed.). Minneapolis, MN: Pearson Assessments.

Richards, G. (2011). *Assessment for selective mutism. The source for selective mutism*. East Moline, IL: Lingui Systems.

Sattler, J. (2001). *General Test Administration Practices Checklist. Assessment of children: Cognitive applications* (4th ed.). Austin, TX: Pro-Ed.

Schill, M. T., Kratochwill, T. R., & Gardner, W. I. (1996). An assessment protocol for selective mutism: Analogue assessment using parents as facilitators. *Journal of School Psychology, 34*, 1–21.

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals* (4th ed.). San Antonio, TX: PsychCorp, A Division of Harcourt Assessment.

Sharp, W. G., Sherman, C., & Gross, A. M. (2007). Selective mutism and anxiety: A review of the current conceptualization of the disorder. *Journal of Anxiety Disorders, 21*, 568–579.

Steinhausen, H., & Juzi, C. (1996). Elective mutism: An analysis of 100 cases. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*, 606–614.

Tancer, N. (1992). Elective mutism: A review of the literature. *Advances in Clinical Child Psychopathology, 14*, 265–288.

Viana, A. G., Beidel, D. C., & Rabian, B. (2009). SM: A review and integration of the last 15 years. *Clinical Psychology Review, 29*, 57–67.

Wilkins, R. (1985). A comparison of elective mutism and emotional disorder in children. *British Journal of Psychiatry, 146*, 198–203.

Williams, K. T. (2007). *Expressive Vocabulary Test* (2nd ed.). Minneapolis, MN: Pearson Assessments.

Wright, H. (1968). A clinical study of children who refuse to talk in school. *Journal of the American Academy of Child Psychiatry, 7*, 603–617.