Hammill Institute on Disabilities

Journal of Learning Disabilities
2017, Vol. 50(3) 252–260

DOI: 10.1177/0022219415617165
journaloflearningdisabilities.sagepub.com
SAGE

Empirical Implications of Matching Children With Specific Language Impairment to Children With Typical Development on Nonverbal IQ

F. Sayako Earle, MA[1], Erica L. Gallinat, MA, CCC-SLP[1], Bernard G. Grela, PhD, CCC-SLP[1], Alexa Lehto, MA[1], and Tammie J. Spaulding, PhD, CCC-SLP[1]

Abstract

This study determined the effect of matching children with specific language impairment (SLI) and their peers with typical development (TD) for nonverbal IQ on the IQ test scores of the resultant groups. Studies published between January 2000 and May 2012 reporting standard nonverbal IQ scores for SLI and age-matched TD controls were categorized into those that matched and did not match children with SLI and TD on nonverbal IQ. We then compared the nonverbal IQ scores across matching criterions within each diagnostic category. In studies that matched children on nonverbal IQ, children with SLI scored significantly higher on nonverbal IQ tests relative to children with SLI in studies that did not match on this criterion. Therefore, it appears that the nonverbal IQ performance of children with SLI is not comparable across studies that do and do not match samples on nonverbal IQ. This suggests that the practice of nonverbal IQ matching may have unintended consequences for the generalization of research findings to the broader SLI population.

Keywords

disorders, language, intelligence, IQ tests, research method

Specific language impairment (SLI), otherwise known as language learning impairment or developmental aphasia, is a subtype of specific learning disabilities (Individuals with Disabilities Education Improvement Act of 2004, 34 CFR 300.307) that affects approximately 7% of children (Tomblin et al., 1997). SLI refers to the condition in which children fail to acquire language in an age-appropriate manner, despite the absence of some other primary deficit to which the language deficit can be attributed (such as hearing loss, socioemotional disorder, frank neurological damage or dysfunction, or intellectual disability). Because of the debate surrounding the etiology and diagnosis of SLI, SLI is considered a research label rather than a diagnostic category (see American Speech and Hearing Association, 2012). However, this population remains of great interest to researchers and educators, as roughly 50% of children observed with SLI in early childhood later develop reading disability (Catts, Fey, Tomblin, & Zhang, 2002) and 45% of children with attention-deficit/hyperactivity disorder are described as having comorbid SLI (Tirosh & Cohen, 1998). Consequently, individuals with a history of SLI often struggle throughout compulsory education (Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998), and often attain lower levels of education compared to their peers with similar socioeconomic backgrounds (Conti-Ramsden, Durkin, Simkin, & Knox, 2009).

In investigations of children with SLI, children with typical development (TD) commonly serve as the comparison, with their level of performance on the outcome variable(s) of interest serving as the benchmark of expectation. The interpretations that can be drawn from experimental outcomes depend on the matching procedure and the consequent comparison groups; in other words, careful consideration needs to be given to the matching protocol used in any investigation. For example, studies that match children with SLI and TD on chronological age are designed to determine if children with SLI perform differently than age-based expectations. In contrast, studies that also match

[1]Department of Speech, Language, and Hearing Sciences, University of Connecticut, Storrs, USA

Corresponding Author:
F. Sayako Earle, Department of Speech, Language, and Hearing Sciences, University of Connecticut, 850 Bolton Rd., Unit 1085, Storrs, CT 06269, USA.
Email: Frances.earle@uconn.edu

the two samples on nonverbal IQ do so to ensure that differences on outcome variables cannot be attributable to cognitive differences between these two groups of children. However, low average nonverbal IQ appears to be part of the profile of children with SLI (e.g., Gallinat & Spaulding, 2014); therefore, matching across diagnostic groups for nonverbal IQ could result in samples of children with SLI with unrepresentatively high and/or TD children with unrepresentatively low nonverbal IQ test scores. The benefit of nonverbal IQ matching may therefore come at a significant cost to the generalization of the research findings to the respective SLI and TD populations. As a first step in determining if there is evidence for this potential concern, we compiled empirical data to evaluate the effect of nonverbal IQ matching in previous studies on the nonverbal cognitive test performance of the resulting SLI and TD groups.

Concerns With Matching Children With SLI and TD for Nonverbal IQ

One concern regarding the practice of matching on nonverbal IQ is the heavy burden that is placed on the measure used to determine nonverbal IQ. Standard scores derived from norm-referenced test performance often serve as the measure by which children are matched for nonverbal cognitive ability. While some norm-referenced tests of nonverbal cognition do not include subtests (e.g., *Test of Nonverbal Intelligence–Second Edition*; Brown, Sherbenou, & Johnsen, 1990b), the majority of nonverbal cognitive tests are composed of a set of subtests, each designed to measure a discrete parameter of nonverbal cognition (e.g., *Kaufman Assessment Battery for Children–Second Edition*; Kaufman & Kaufman, 2004a; *Wechsler Preschool and Primary Scale of Intelligence–Third Edition*; Wechsler, 2002). Therefore, a composite standard score on these tests reflect a summation of different cognitive skills; however, prior research has demonstrated that the nonverbal IQ profiles of children with SLI are not flat (Swisher & Plante, 1993; Swisher, Plante, & Lowell, 1994). Swisher and Plante (1993) investigated preschool children's performance on two tests of nonverbal intelligence, and found that relationships between performance on subtests or groups of subtests were different both within the population with SLI and between the SLI and TD groups. Likewise, when investigating older school-age children, Swisher et al. (1994) found that the SLI and TD groups differed in their respective correlations between performance on particular item types and overall performance, with performance profiles of the SLI group were more variable with respect to the TD group. Therefore, while the practice of matching for nonverbal IQ may increase our confidence that both groups have the same cognitive ability, the common practice of matching on composite scores may mask different profiles of strengths and weaknesses in nonverbal skills between SLI and TD populations (Swisher & Plante, 1993). Therefore, unlike the transparency of chronological age as the matching criterion, composite IQ scores are opaque as to what behaviors/abilities are actually being equated when children across groups are matched on this criterion.

Furthermore, different nonverbal cognitive tests assess different aspects of nonverbal cognition. Therefore, when we consider the variability in performance across different subtests of nonverbal IQ in children with SLI, we should expect that children with SLI perform differently depending on which nonverbal cognitive test is administered (e.g., DeThorne & Watkins, 2006; Miller & Gilbert, 2008; Swisher & Plante, 1993). To illustrate, Miller and Gilbert (2008) administered the *Wechsler Intelligence Scale for Children–Third Edition* (WISC-III; Wechsler, 1991) and the *Universal Nonverbal Intelligence Test* (UNIT; Bracken & McCallum, 1998) to four groups of children, including those with SLI and children who were TD. While the children with TD obtained correlated scores between the two tests, the children with SLI did not. Given the variability of the SLI group, 16% of the children meeting the criteria for SLI on the WISC-III failed to meet the criteria for SLI on the UNIT. Furthermore, 74% of children whose nonverbal cognitive skills were too low to be characterized as having SLI on the WISC-III met the classification for SLI on the UNIT. Therefore, the nonverbal cognitive test administered not only resulted in variable scores for children with poor language skills, but it influenced the membership of diagnostic groups. This finding suggests that a proportion of children with SLI matched for nonverbal cognition to TD children based on the scores on one test may not meet the criteria for inclusion in the SLI group if given another nonverbal IQ test. Because different research investigators use different nonverbal cognitive tests (see Gallinat & Spaulding, 2014), this is also an issue when attempting to control for between-group differences in nonverbal IQ.

The third concern with the practice of nonverbal IQ matching specific to SLI stems from the positive relationship between language and nonverbal cognitive skills. Traditionally, the rate of cognitive development has been argued to constrain language acquisition as not to exceed available nonverbal capacities such as processing speed, memory, and attentional control (e.g., Bloom, 1970; Mehler, 1971; Piaget, 1959). Alternatively, the medium of language as a communicative device has prompted theory that language acquisition shapes the development of nonverbal cognition (e.g., Boroditsky, 2001; Choi, 2006; Semin, 2000; see Whorf, 1956, for the linguistic determinism hypothesis). Experimental evidence supports that there exists a temporal relationship between cognitive and linguistic development in typical children (e.g., Gopnik & Meltzoff, 1992; Mervis & Bertrand, 1994; Purcell et al., 2001). For example, Gopnik and Meltzoff (1992) found that sorting ability is positively associated with vocabulary size in

18-month-olds. Similarly, Mervis and Bertrand (1994) found that categorization skill is associated with fast-mapping ability in 16- to 20-month-olds. Such relationships have also been confirmed in children with SLI (e.g., Botting, 2005; Conti-Ramsden, Botting, Faragher, 2001; Trouton, Spinath, & Plomin, 2002; Wetherell, Botting, & Conti-Ramsden, 2007). For example, Wetherell et al. (2007) found that nonverbal IQ scores were predictive of narrative skills in adolescents with SLI. Therefore, we might expect that children with SLI (who by definition are impaired in language) obtain lower scores on norm-referenced tests of nonverbal cognition relative to those whose language skills are intact. Supporting this, Gallinat and Spaulding (2014) conducted a meta-analysis of studies investigating children with SLI and their age-matched peers with TD, and found that children with SLI scored on average 0.695 standard deviations below children with TD on tests of nonverbal cognition.

Consequently, by matching children with SLI to children with TD by nonverbal IQ test performance, the sample of the SLI population chosen for a particular study may represent higher scorers on average than would be truly representative of the population. Likewise, the sample selected to represent children with TD may be taken from those that perform lower on tests of nonverbal cognition with respect to the general population. If nonverbal IQ matching results in skewed samples, then the findings for both the children with SLI and TD children may not accurately represent the abilities of their respective populations. In addition to these concerns of external validity, there are potential internal validity implications as well. Matching for nonverbal IQ is done because the investigator assumes a confounding relationship between nonverbal IQ and the dependent variable(s) under investigation. If this practice influences the average nonverbal cognitive performance to be uncharacteristically high within the SLI sample and low within the TD sample, this decreases the likelihood that between-group differences would be found on the dependent measures that potentially exist as disparities in the general SLI and TD populations. Beyond this risk of making a Type II error, an attenuated effect size is a potential consequence of this practice.

In summary, in an effort to identify the strengths and weaknesses of children with SLI, investigators often employ matching procedures to control for differences between the SLI and TD subjects on factors aside from the experimental variables. However, documented differences in nonverbal cognition between the two populations raise several issues with nonverbal IQ matching protocols that can potentially affect experimental outcomes. First, the nonverbal cognitive profiles of children with SLI tend to be dissimilar to TD peers even when their resulting composite or average scores are similar (e.g., Swisher et al., 1994; Swisher & Plante, 1993). Second, the practice suggests an implicit confidence on the reliability on the instrument used to measure nonverbal cognition. This point is particularly concerning given prior work has shown that while TD children tend to perform similarly across nonverbal cognitive tests, SLI children do not (e.g., DeThorne & Watkins, 2006; Miller & Gilbert, 2008). Finally, the tendency for children with SLI to obtain lower nonverbal IQs than their peers with TD may skew an SLI sample to have higher, and TD sample to have lower, scores than would be representative of their respective populations. The purpose of the current study therefore was to determine if matching for nonverbal IQ within research investigations has an effect on the nonverbal IQ scores of the resultant SLI and TD groups. Specifically, we predicted two main effects of nonverbal IQ matching on the SLI and TD samples. First, we proposed that children with SLI reported in studies that employ a nonverbal IQ matching criterion have higher nonverbal IQ scores relative to those in studies that do not. This prediction is based on the assumption that a nonverbal IQ matching criterion will require children with SLI that obtain higher nonverbal IQ scores to be matched to children with TD. Conversely, children with TD reported in studies that employ a nonverbal IQ matching criterion are predicted to have lower nonverbal IQ scores relative to those in studies that do not.

Method

Selection of Studies

Four electronic databases (PsycINFO, Scopus, PubMed, Dissertation Abstracts International) were searched in June 2012 using the following search terms: *specific language impairment*, *primary language impairment*, *language impairment*, *language disorder*, *developmental aphasia*, *developmental dysphasia*, *language learning impairment.* The databases were set to yield studies published between 2000 and May 2012, in English, in peer-reviewed journals or as published dissertations, and that included human participants. We chose 2000 because, first, we wished to gather at least one decade's worth of articles so as to ensure adequate statistical power. Second, to our knowledge, the practice of nonverbal IQ matching was not common prior to 2000. In addition, filters were applied within each database to rule out further studies that were not relevant to the current research questions, including those that focused on non-English-speaking or bilingual participants, subjects with degenerative diseases, and participants with speech disorders such as stuttering and apraxia of speech. This resulted in 2,621 studies. Using EndNOTE software to remove duplicate entries reduced the list to 2,042 studies.

The remaining 2,042 titles were organized by publication date. Each study was examined to determine if it met the following additional inclusionary criteria: (a) included both SLI and TD child participants, (b) specified that the SLI and TD groups were age matched per chronological

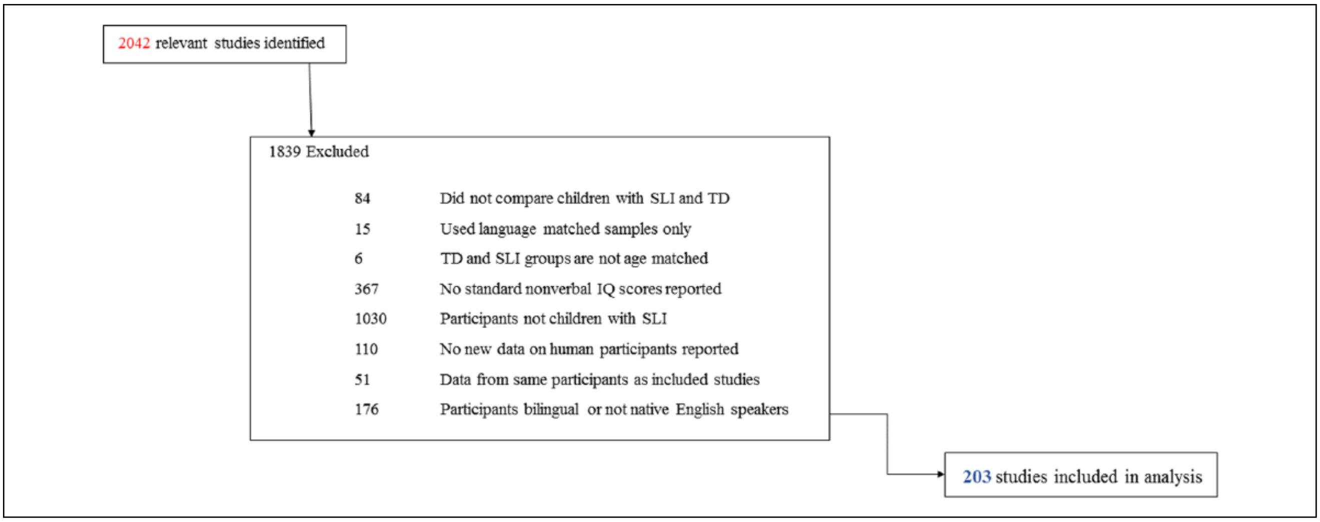

Figure 1. Flowchart of excluded studies.

age, and (c) reported either standard scores or scores that could be converted to standard scores (e.g., percentiles) from a norm-referenced test of nonverbal IQ for SLI and/or TD participants. We note that, to avoid the potential confounds of ability (linguistic or nonverbal) influenced by multilingualism, difference in language background, or diagnostic criteria, samples included in the analysis were of monolingual English speaking children only. See Figure 1 for reasons for exclusion of individual studies. A total of 203 studies met the inclusionary criteria and were selected for the present study, including a total of 4,324 children with SLI and 5,306 TD participants. Of these, 33 studies which included 66 samples of children (33 samples of 678 children with TD, and 33 samples of 653 with SLI) stated that subjects were also matched on nonverbal IQ; samples from these studies served as our "age and nonverbal IQ matched" cohort. The remaining 170 studies, including 323 samples of children (153 samples of 4,628 children with TD, and 170 samples of 3,671 with SLI), made no such statement and the samples from these studies served as our "age-matched" cohort. We note that the discrepancy in number of TD and SLI samples is because not all studies within the age-matched cohort reported nonverbal IQ information on their TD participants. In addition, 42 of these studies included two TD controls matched on age and language ability. For these, only TD data for the age-matched participants were included. See Table 1 for a summary of participant ages included in the analyses. For a summary of nonverbal IQ tests used in the included studies, see Table 2.

Table 1. Number of Participants Included in the Analyses by Age Range.

	NVIQ and CA matched		CA only	
Age	TD	SLI	TD	SLI
36–47 months	0	0	60	32
48–59 months	0	0	675	685
60–71 months	98	97	350	319
72–83 months	15	15	790	537
84–95 months	19	19	264	256
96–107 months	92	76	811	577
108–119 months	195	186	395	365
120–131 months	131	128	206	183
132–143 months	0	0	96	90
144–155 months	8	8	44	37
156–167 months	52	51	378	121
168–179 months	0	0	153	88
180–191 months	34	35	140	155
>192 months	26	30	134	128
No report/no range reported	8	8	132	98
	678	653	4628	3671

Note. CA = chronological age; NVIQ = nonverbal IQ; SLI = specific language impairment; TD = typical development. Numbers of studies are listed according to cohort (NVIQ and CA matched and CA only).

Procedure

The data collected from each study included the number of participants per group, the mean nonverbal IQ test score of each group, the quantitative metric used to report nonverbal IQ scores, and if the participants were age only matched or matched for age as well as nonverbal IQ. If all the participants in both groups were given more than one nonverbal cognitive test, the combined mean of both tests were calculated. If studies were longitudinal in nature, a combined weighted mean was used based on the number of participants at each time point.

Table 2. IQ Tests Used by Studies Included in the Analyses.

	Number of studies	
IQ TEST used	NVIQ + CA	CA only
BAS	1	1
CMMS	1	52
Hiskey–Nebraska DEQ	0	1
KABC	0	18
KABC II	0	11
KBIT	4	6
KBIT II	0	2
Leiter-R	3	5
LIPS	4	6
McCarthy	0	1
RCPM	8	10
Stanford–Binet	0	1
TONI	2	22
TONI II	0	4
TONI III	3	7
UNIT	0	2
WASI	3	9
WISC	0	2
WISC-III	4	7
WISC-R	0	1
WPPSI	0	4
WPPSI III	0	1
Total	33	173[a]

Note. CA = chronological age; NVIQ = nonverbal IQ. Numbers of studies are listed according to cohort (NVIQ and CA matched and CA only). Test names in alphabetical order: BAS = *British Ability Scales* (Elliott, Smith, & McCulloch, 1985); CMMS = *Columbia Mental Maturity Scale* (Burgemeister, Blum, & Lorge, 1972); Hiskey–Nebraska DEQ = *Nebraska Test of Learning Aptitude for Young Deaf Children* (Hiskey, 1941); KABC = *Kaufman Assessment Battery for Children* (Kaufman & Kaufman, 1983; KABC II: Kaufman & Kaufman, 2004a); KBIT = *Kaufman Brief Intelligence Test* (Kaufman & Kaufman, 1990; KBIT II: Kaufman & Kaufman, 2004b); LIPS = *Leiter International Performance Scale* (Leiter & Porteus, 1936; Leiter-R: Roid & Miller, 2011); McCarthy = *McCarthy Scales of Children's Abilities* (McCarthy, 1972); RCPM = *Raven's Coloured Progressive Matrices* (Raven, Raven, & Court, 1962); Stanford–Binet = *Stanford–Binet Intelligence Scale* (Thorndike, Hagen, & Sattler, 1986); TONI = *Test of Nonverbal Intelligence* (Brown, Sherbenou, & Johnsen, 1988; TONI-II: Brown, Sherbenou, & Johnsen, 1990a; TONI-III: Brown, Sherbenou, & Johnsen, 1997); UNIT = *Universal Nonverbal Intelligence Test* (Bracken & McCallum, 1998); WASI = *Wechsler Abbreviated Scale of Intelligence* (Wechsler, 1999); WISC = *Wechsler Intelligence Scale for Children* (Wechsler, 1949; WISC-R: Wechsler, 1974; WISC-III: Wechsler, 1991); WPPSI = *Wechsler Preschool and Primary Scale of Intelligence* (Wechsler, 1967; WPPSI-III: Wechsler, 2002).
[a]Three of the studies in the age-matched (CA) cohort used two different tests.

We performed a cumulative analysis of norm-referenced test scores of nonverbal cognition from the studies that investigated children with SLI relative to their age-matched peers with TD. We categorized the studies into those that matched for age-and nonverbal IQ and those that matched for age but not nonverbal IQ. Test performances that were reported as scaled scores or percentiles were converted to a mean of 100 and a standard deviation of 15. We then compared the nonverbal IQ scores within the SLI and TD groups, between studies that matched for both age and nonverbal IQ across groups, and those that only matched across groups on age, using methods outlined by Bland and Kerry (1998) for conducting a weighted comparison of means.

Reliability

A trained undergraduate student conducted a separate search for studies that met the inclusionary criteria in fall 2012 and identified one additional study not originally identified for inclusion. This study is included in the selection process specified above. The undergraduate student also examined 36 studies and recorded data on the number of participants included in the samples, the mean nonverbal IQ test scores of the participants with SLI and TD, the metric used for reporting nonverbal IQ scores, and if the groups were matched for age only or nonverbal IQ and age. In summary, 213 out of 216 data points were consistent with the data set originally compiled by the first author, resulting in a point-to-point reliability of .99. Discrepancies were reviewed, and the outcome determined, by the first and last authors.

Analysis and Results

Separate independent *t* tests were conducted for the SLI and TD Groups. The independent variable was matching criterion (nonverbal IQ + age, age only), and the dependent variable was the mean standard score on the nonverbal IQ norm-referenced test. Weighted means and weighted standard deviations (Bland & Kerry, 1998) were calculated for each diagnostic group in each match criterion and were used for comparison purposes.

We conducted two independent samples *t* tests to compare nonverbal IQ scores across matching criterion within each Group. Confidence intervals were calculated using Bonferroni correction for multiple comparisons to control for family-wise error rate at .05. This resulted in a significant effect of matching criterion for the SLI group. Nonverbal IQ scores of the participants with SLI were significantly higher in studies which matched for both age and nonverbal IQ relative to studies which matched for age but not nonverbal IQ, $t(201) = 3.23$, $p = .001$, 95% CI = 1.578, 7.228. The weighted means and weighted standard deviations of the SLI group were 101.43 (6.23) and 97.40 (6.62) for the nonverbal IQ + age-match studies and the age-match-only studies, respectively. In the TD group, participants in studies which matched for both nonverbal IQ + age obtained comparable scores on norm-referenced tests of nonverbal cognition relative to TD samples in studies that matched for age only, $t(184) = 1.54$, $p = .128$; 95% CI = −1.31, 7.00. The weighted means and standard deviations of the TD group were 103.4 (5.25) in nonverbal IQ + age-match studies relative to 106.24 (10.28) in the age-match-only studies. See Figure 2 for a summary of this pattern.

Discussion

In our introduction section, we addressed our rationale for suspecting that the nonverbal cognitive test scores of children who are matched for nonverbal IQ and age will likely differ from those children who are matched for age but not nonverbal IQ. Our investigation was motivated by the common practice in SLI research to match participants between groups on factors other than age. Such factors may include gender, socioeconomic status, maternal education level, language ability, and/or nonverbal IQ. This procedure is based on the assumption that failing to adequately control for potential confounds will introduce noise in the data, which in turn will reduce the significance of the measured effects and/or invalidate the experimental results. Our findings from the current investigation, however, suggest that the matching criterion employed may in itself introduce the experimental confound that affects the outcome.

Specifically, our first aim was to determine if children with SLI who were matched for age and nonverbal IQ to children with TD performed differently on tests of nonverbal IQ with respect to children with SLI who were matched to TD subjects for age only. We predicted that a nonverbal IQ matching protocol would result in a sample of children with SLI with a relatively high nonverbal IQ. Our results supported this prediction: In studies that used a nonverbal IQ and age match procedure, children with SLI scored 4 points higher on average (range: 2 to 7) on nonverbal cognitive tests relative to those in studies that did not employ this matching procedure.

Our second aim was to determine if children with TD who were matched for age and nonverbal IQ to children with SLI perform differently on tests of nonverbal IQ as compared to TD children who were matched to SLI subjects for age only. We predicted that a nonverbal IQ matching procedure would result in a group of participants with TD that obtain a relatively low nonverbal IQ. While the numerical trend of our data supported these predictions, the effects were not significant. According to the 95% confidence interval, children with TD scored 3 points lower on average (range = −1 to 7) on nonverbal cognitive tests in studies that matched them to children with SLI, with respect to those who were matched on age only.

Given the current findings, the broader implications point to the external validity of studies matching participants on nonverbal IQ. Specifically, it appears that studies that employ nonverbal IQ matching employ an SLI sample that subtly overestimates the nonverbal cognitive ability of the SLI population. The results based on group comparisons with such samples may not accurately reflect the nature of the relationships that exist between the two populations. In addition to these external validity concerns, this practice compromises our ability to draw conceptual insights about children with SLI more broadly. Samples that overestimate the nonverbal skills of the SLI population may lead experimenters to fail to detect a difference on experimental measures where one exists, or to underestimate the magnitude of this difference.

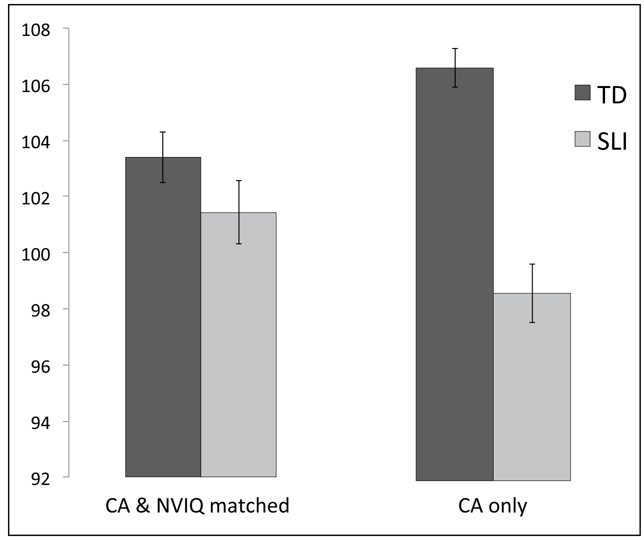

Figure 2. Nonverbal IQ scores by group, by matching criterion.
Note. Effect of matching criterion nonverbal IQ standard scores (y-axis): chronological age (CA) and nonverbal IQ (NVIQ) matched as compared to chronological age only. Error bars denote standard errors of the mean.

For example, Miller, Kail, Leonard, and Tomblin (2001) examined differences in the processing speed of children with SLI and TD on a variety of tasks. The SLI and TD samples were selected based on similarity in performance IQ scores on the *Wechsler Preschool and Primary Scale of Intelligence–Revised* (Wechsler, 1989). Several of the subtests (e.g., Block Design, Object Assembly) were timed subtests that took into account both accuracy and processing speed in the score determinations. As the authors note, in matching subjects based on test scores on an instrument that include processing speed amongst its measures, between group differences in processing speed, the dependent variable under investigation, would be technically more difficult to find. Although the authors still found between group differences in processing speed despite controlling for nonverbal IQ test scores, they indicate that the magnitude of the difference in processing speed between the SLI and TD samples was likely smaller than what would be characteristic of differences between these two populations. The authors further point to previous data from Kail and Salthouse (1994), in which the effect size for the group difference observed in processing speed was found to be larger between the SLI and TD groups than in their study.

We acknowledge that the practice of controlling for between group disparities in nonverbal IQ through matching procedures is an attempt to equate these two groups on nonlinguistic ability with the potential to influence the outcome measure. However, instead of assuming that there is an association between nonverbal IQ and the dependent variable(s) a priori, a more direct alternative is to first verify that that relationship exists. If an association is determined, one has the option to statistically control for the variance in performance on the dependent variables that is attributable to nonverbal IQ. This can be accomplished in several ways. One option is to make statistical adjustments using nonverbal IQ as a covariate. For example, in a prior study in our lab (Spaulding, 2010), differences between children with SLI and TD were compared in their ability to resist distractor interference and their response inhibition. The results supported a significant group effect observed on the dependent measures, a significant group difference on scores obtained on the nonverbal IQ test, and a positive relationship between nonverbal IQ scores and the dependent variables under investigation. In our case, we conducted an additional analysis of covariance (ANCOVA) with nonverbal IQ serving as a covariate for each of the dependent measures. Such adjustments can be made in regression analyses (including those assessing nonparametric data) by adding a term for nonverbal IQ in the regression equation. A further alternative is to explicitly define the normative relationship between nonverbal IQ and the outcome measure(s). This would require the recruitment of a large number of participants with TD, whose performance on the outcome measures would then be regressed against their nonverbal IQ scores. The performance of children with SLI could then be assessed relative to this distribution, to investigate and quantitatively measure if there is a discrepancy between performance expectations given their nonverbal cognitive skill and their actual performance on the outcome measure. In other words, the standard error in the regression estimate of the TD sample could be used to standardize the performance of the SLI participants indexed by cognitive ability (note: this procedure assumes a linear distribution in performance by the TD sample; if this is not the case, nonlinear regression approaches would need to be used).

Regardless of the statistical options that are available as the alternative to matching procedures, it is further important to note that neither one of these approaches may be appropriate to use when investigating children with SLI. In considering that nonverbal IQ is a characteristic of the disorder (Gallinat & Spaulding, 2014), investigating between-group differences after controlling for nonverbal IQ may be myopic in practice (Dennis et al., 2009). The interaction during development of language and cognition is particularly difficult to disentangle: For example, a longitudinal analysis by Botting (2005) suggests that children with SLI fall up to 20 normative IQ points over the course of development. Consequently, Spaulding, Plante, and Vance (2008) found that group disparities in the sustained selective attention skills of children with SLI and their TD peers disappeared when either nonverbal cognition or linguistic skills was independently entered into the analysis as a covariate. Therefore, controlling for nonverbal cognition to isolate the specific contributions of language on the outcome variables of interest may not necessarily achieve the intended isolation of the linguistic from the nonlinguistic effects.

Finally, while the implications of this study are primarily intended to guide research practice, there may also be implications for educators working with children who struggle to attain age-appropriate levels of language. First, to the careful consumer of literature concerning this population, it is important to point to the subtle differences in research practice (such as the matching criterion employed) that may affect the applicability of study findings to each particular student in the classroom. From a broader perspective, this work joins a growing literature that describes SLI as a population in which linguistic deficits are not domain-specific, but likely reflects general deficits in cognitive processes (e.g., Gathercole & Baddeley, 1990; Ullman & Pierpont, 2005) that may be reflected in their nonverbal IQ (Gallinat & Spaulding, 2014). As such, school psychologists and educators may find that children with a learning disability primarily in language may benefit from programs that target cognitive processing in addition to language-specific goals.

In conclusion, the consequence of matching children with SLI to those with TD appears to be that the resultant samples with SLI are not representative of the population with SLI. We have argued furthermore that controlling for differences in nonverbal cognition across the two groups may obscure potential insights into the disorder. Taken together, we propose that, in most circumstances, matching children with SLI to children with TD on nonverbal IQ is ultimately an unproductive research practice.

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) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by a New Investigator Grant from the American Speech and Hearing Foundation (ASHF) to Tammie J. Spaulding, and also by the ASHF New Century Doctoral Scholarship to F. Sayako Earle.

References

American Speech and Hearing Association. (2012, June). *ASHA's recommended revisions to the DSM-V. Position statement submitted by Shelly Chabon*. Retrieved from http://www.asha.org/uploadedFiles/DSM-5-Final-Comments.pdf

Bland, J. M., & Kerry, S. M. (1998). Weighted comparison of means. *British Medical Journal*, *316*, 129.

Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, *43*, 1–22.

Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*, 317–326.

Bracken, B. A., & McCallum, S. (1998). *Universal Nonverbal Intelligence Test*. Chicago, IL: Riverside.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1988). *Test of Nonverbal Intelligence*. Austin, TX: PRO-ED.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1990a). *Test of Nonverbal Intelligence: A language-free measure of cognitive ability* (TONI-2). Austin, TX: PRO-ED.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1990b). *Test of Nonverbal Intelligence* (2nd ed.). Austin, TX: PRO-ED.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *TONI-3, Test of Nonverbal Intelligence: A language-free measure of cognitive ability*. Austin, TX: PRO-ED.

Burgemeister, B. B., Blum, L. H., & Lorge, I. (1972). *Columbia Mental Maturity Scale*. New York, NY: Harcourt Brace Jovanovich.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research*, *45*, 1142–1157.

Choi, S. (2006). Influence of language-specific input on spatial cognition: Categories of containment. *First Language*, *26*, 207–232.

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*, 741–748.

Conti-Ramsden, G., Durkin, K., Simkin, Z., & Knox, E. (2009). Specific language impairment and school outcomes. I: Identifying and explaining variability at the end of compulsory education. *International Journal of Language and Communication Disorders*, *44*, 15–35.

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, *15*, 331–343.

DeThorne, L. S., & Watkins, R. V. (2006). Language abilities and nonverbal IQ in children with language impairment: Inconsistency across measures. *Clinical Linguistics and Phonetics*, *20*, 641–658.

Elliott, C., Smith, P., & McCulloch, K. (1985). *British Ability Scales*. Chiswick, UK: NFER-Nelson.

Gallinat, E., & Spaulding, T. J. (2014). Differences in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, *57*, 1363–1382.

Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*, 336–360.

Gopnik, A., & Meltzoff, A. (1992). Categorization and naming: Basic level sorting in eighteen month-olds and its relation to language. *Child Development*, *63*, 1091–1103.

Hammill, D. D., Pearson, N. A., & Weiderholt, J. L. (2009). *Comprehensive Test of Nonverbal Intelligence* (2nd ed.). Austin, TX: PRO-ED.

Hiskey, M. S. (1941). *Nebraska Test of Learning Aptitude for Young Deaf Children*. Lincoln: University of Nebraska, Department of Educational Psychology and Measurements.

Individuals With Disabilities Education Improvement Act of 2004, 20 U.S.C. § 1400 et seq. (2004).

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199–225.

Kaufman, A. S., & Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. New York, NY: John Wiley.

Kaufman, A. S., & Kaufman, N. L. (1990). *Manual for the Kaufman Brief Intelligence Test*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004a). *Kaufman Assessment Battery for Children–Second Edition* (KABC-II). Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004b). *Kaufman Brief Intelligence Test–Second Edition*. New York, NY: John Wiley.

Leiter, R. G., & Porteus, S. D. (1936). *The Leiter International Performance Scale* (No. 13). Wood Dale, IL: Stoelting.

McCarthy, D. (1972). *McCarthy Scales of Children's Abilities*. New York, NY: Psychological Corporation.

Mehler, J. (1971). Studies in language and thought development. In R. Huxley & E. Ingram (Eds.), *Language acquisition: Models and methods* (pp. 201–225). New York, NY: Academic Press.

Mervis, C., & Bertrand, J. (1994). Acquisition of novel name-nameless category (N3C) principle. *Child Development*, *65*, 1646–1662.

Miller, C. A., & Gilbert, E. (2008). Comparison of performance on two nonverbal intelligence tests by adolescents with and without language impairment. *Journal of Communication Disorders*, *41*, 358–371.

Miller, C. A., Kail, R., Leonard, L. B., & Tomblin, J. B. (2001). Speed of processing in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *44*, 416–433.

Piaget, J. (1959). *The language and thought of the child* (Vol. 5). New York, NY: Psychology Press.

Purcell, S., Eley, T. C., Dale, P. S., Oliver, B., Petrill, S. A., Price, T. S., . . . Plomin, R. (2001). Comorbidity between verbal and nonverbal cognitive delays in 2-yr-olds: A bivariate twin analysis. *Developmental Science*, *4*, 195–208.

Raven, J. C., Raven, J. C., & Court, J. H. (1962). *Coloured Progressive Matrices*. London, UK: HK Lewis.

Roid, G. H., & Miller, L. J. (2011). *Leiter International Performance Scale–Revised (Leiter-R)*. Madrid, Spain: Psymtec.

Semin, G. R. (2000). Agenda 2000—Communication: Language as an implementational device for cognition. *European Journal of Social Psychology*, *30*, 595–612.

Spaulding, T. J. (2010). Investigating mechanisms of suppression in preschool children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *53*, 725–738.

Spaulding, T. J., Plante, E., & Vance, R. (2008). Sustained selective attention skills of preschool children with specific language impairment: Evidence for separate attentional capacities. *Journal of Speech, Language, and Hearing Research*, *51*, 16–34.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research*, *41*, 407–418.

Swisher, L., & Plante, E. (1993). Nonverbal IQ tests reflect different relations among skills for specifically language-impaired and normal children: Brief report. *Journal of Communication Disorders*, *26*, 65–71.

Swisher, L., Plante, E., & Lowell, S. (1994). Nonlinguistic deficits of children with language disorders complicate the interpretation of their nonverbal IQ scores. *Language, Speech, and Hearing Services in Schools*, *25*, 235–240.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford–Binet Intelligence Scale*. Chicago, IL: Riverside.

Tirosh, E., & Cohen, A. (1998). Language deficit with attention-deficit disorder: A prevalent comorbidity. *Journal of Child Neurology*, *13*, 493–497.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech Language Hearing Research*, *40*, 1245–1260.

Trouton, A., Spinath, F. M., & Plomin, R. (2002). Twins Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behavior problems in childhood. *Twin Research*, *5*, 444–448.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*, 399–433.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York, NY: Psychological Corporation.

Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence–WPPSI*. New York, NY: Psychological Corporation.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children, Revised*. New York, NY: Psychological Corporation.

Wechsler, D. (1989). *WPPSI-R, manual: Wechsler Preschool and Primary Scale of Intelligence, Revised*. New York, NY: Psychological Corporation.

Wechsler, D. (1991). *WISC-III: Wechsler Intelligence Scale for Children: Manual*. New York, NY: Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. New York, NY: Psychological Corporation.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence–III*. New York, NY: Psychological Corporation.

Wetherell, D., Botting, N., & Conti-Ramsden, G. (2007). Narrative skills in adolescents with a history of SLI in relation to nonverbal IQ scores. *Child Language Teaching and Therapy*, *23*, 95–113.

Whorf, B. (1956). Language, mind, and reality. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (pp. 246–270). Cambridge, MA: MIT Press.