State of Practice for Language and Literacy Research: A Review of Methods

Carla Wood
Autumn McIlraith
Lisa Fitton
Florida State University, Tallahassee

The interests of scientists from a variety of disciplines overlap at the focal point of literacy development and intervention. Literacy has been targeted nationwide as a high priority area for external funding to promote improved outcomes in reading, understanding, and subsequent academic progress consistent with the No Child Left Behind Act (NCLB, 2003). This area of research is relevant to multiple fields such as psychology, education, special education, and speech-language pathology.

Given that a variety of disciplines conduct literacy research, a wide distribution of methods used in language and literacy research is to be expected. This diversity of methods is beneficial in that it provides a variety of angles from which to examine current theory and practice, which in turn strengthens our confidence in findings that are consistently replicated across different methods. However, in order to continue to contribute to the body of research, scholars must have a basic understanding of the wide range of methods available in order to...
recognize their influence on research findings and their implications for interpretation. In response to this need, we reviewed the current state of practice in language and literacy research in an effort to identify methodologies that are important for next-generation scholars who wish to continue to consume and produce research effectively.

Research evidence is considered one of the pillars of evidence-based practice (EBP) and is critical to clinical practice (American Speech-Language-Hearing Association [ASHA], 2005a; McCurtin & Roddam, 2012). EBP in speech-language pathology values both internal and external evidence, including consulting the published literature for the best available scientific evidence to support the use or disuse of specific speech and language intervention approaches. Among valued skills, speech-language pathologists (SLPs) are expected to demonstrate knowledge of various research practices and integration of research principles into EBP (ASHA, 2005b).

Although practitioners and researchers alike greatly value EBP, there is no general consensus on what constitutes best available scientific evidence and which methods and analyses are necessary or appropriate for future consumers and producers of research in our field. By definition, rigor is used in research to refer to strict precision; however, the criterion for acceptable precision may vary across fields. Varied standards and hierarchies are applied within and across disciplines (McCurtin & Roddam, 2012), but the focus of the present article is to examine methodology in language and literacy research from a perspective of pursuing diversity.

Different methods of research have different strengths and weaknesses that influence their utility to scholars. However, there is a need to address the suggestion that research in some fields, such as speech-language pathology, is not broadening to include newer advanced methods and statistics (e.g., Ioannidis, 2005). Given the national movement toward large data sets, randomization, and large-scale replication (Ioannidis, 2005), it is critical to examine the use of these methods within and outside of speech-language pathology research.

Ioannidis (2005), a widely acclaimed researcher, has called for a revolution in research practices, with a focus on replication by independent research teams with large data sets (Ebrahim et al., 2014). After extensive review of the research literature in medicine, Ioannidis made a convincing case for the need for better powered evidence from large studies. He concluded that small n studies without random assignment were underpowered and were at risk for leading practitioners to inaccurate conclusions. Ioannidis exposed a plethora of research findings with large effect sizes that could not be replicated, resulting in a call for more rigorous research methods and a focus on replication.

In addition to the use of larger data sets, the educational research community has increasingly incorporated statistics within the families of hierarchical linear modeling (HLM) and structural equations modeling (SEM). These techniques have expanded researchers’ abilities to answer more complex research questions through simultaneously examining interactions between predictor variables and their impacts on outcome variables. Recent research findings have suggested that the unique characteristics of individual participants interact considerably to influence observed results. For example, there have been repeated calls in bilingual literacy research to report more background characteristics such as age of first exposure to languages and type of classroom instruction because these characteristics have been found to be critically influential (e.g., August & Shanahan, 2006; Branum-Martin, Tao, & Garnaat, 2015). Although more traditional statistical models, including analyses of covariance (ANCOVAs) and multiple regression, allow for statistical inclusion of these background characteristics, they are limited in their ability to account for the interactions between variables. Rather, relationships between included variables are often evaluated individually in the order dictated by the researcher. This is appropriate when there is a strong theoretical basis for the chosen order, but poses a challenge when an established theoretical foundation for the research is lacking. In the latter case, it may be appropriate to consider statistical models such as HLM and SEM, which offer the benefit of examining relationships between all variables simultaneously (see Kline, 2015).

An additional advantage of multilevel techniques such as HLM is that they allow the statistical model to more accurately represent the structure of the data. For example, if children are presented with a set of words and are asked to read them, the individual responses to each item by each child are not independent. All responses from a single child are dependent in the sense that they originate from the same child, and certain child-related features potentially influence the responses of that child, such as the child’s previous exposure to those words or the child’s phonological awareness skills. All responses to a single word are similarly dependent, because word-related features such as regularity and frequency of occurrence may influence the accuracy of responses across all children. Using HLM, this dependence can be taken into account by “nesting” the individual responses within children and within items. Without accounting for this nesting in HLM or SEM to include both levels
of features in the same model, statistical analysis may inaccurately estimate effects attributable to each independent variable, resulting in inaccurate conclusions (Baayen, Davidson, & Bates, 2008). Multilevel modeling provides an excellent tool for researchers to use to scale statistical model complexity to match the structure of more complex data sets, resulting in more accurate estimates of effects (Compton, Miller, Gilbert, & Steacy, 2013).

Modeling Revolution

The rise in the use of multilevel modeling in many academic fields has led some researchers to refer to a “modeling revolution” (Rodgers, 2010). The term modeling is used to refer to “a set of assumptions together with implications drawn from them by mathematical reasoning” (Neimark & Estes, 1967, p. v). Evidence for an increase in the use of multilevel modeling was demonstrated by Reinhart, Haring, Levin, Patall, and Robinson (2013), who examined methods used in 275 empirical articles from five primary research journals that were published between 2000 and 2010. They found that the use of modeling increased from 15% of empirical studies in 2000 (nine out of 61) to 54% in 2010 (50 out of 93).

Challenges of Advanced Modeling and Methodology

Despite the development of innovative statistical techniques and Ioannidis’s (2005) call for larger sample sizes, exclusive reliance on traditional statistical practices persists among many scholars. Considering the abundant support in place for traditional systems and the lack of support and resources available for novel procedures, the adoption of new practices is difficult. In the field of speech-language pathology in particular, there are several notable challenges restricting the incorporation of newer and more complex methodologies into our research practice. A few barriers that will be discussed throughout the article include the prevalence of small n studies due to the low incidence of disabilities of interest, challenges to random assignment due to the inherent participant characteristic of presenting with a disability or not (e.g., researchers cannot randomly assign cochlear implants), and (c) ethical conflicts of random assignment to a treatment versus comparison control group when early intervention is warranted.

The prevalence of small n studies has been perpetuated by several factors. In speech-language pathology, the tendency toward small data sets may be partially attributed to our interest in low-incidence populations. Recruiting children and youth, particularly from low-incidence populations, presents challenges. Further, within the populations we research, it is difficult to employ true experimental designs with random assignment given that participant characteristics cannot be assigned and treatment often cannot be ethically withheld. Low-incidence populations of interest can also have high degrees of heterogeneity. This is exemplified in the area of aural rehabilitation for individuals who are deaf or hard of hearing. Among other variables, participants in such research commonly present with differences in age of onset, age of identification, underlying cause, severity, use of sensory device, comorbidity, communication method, and educational settings or approaches. Such an array of complex characteristics, coupled with the low incidence of deafness, presents methodological challenges for research in aural rehabilitation techniques, and many other specializations face similar complexities.

Perhaps equally as powerful as these methodological challenges is the current lack of infrastructure that would be required to support and sustain change or a revolution in methodology. Sharpe (2013) identified several infrastructure components that are necessary to support the inclusion of more complex methods in research practices, including (a) better access to continuing education in advanced methods and statistics, and (b) better use of mavens in the field. With regard to continuing education, SLPs and researchers in our field readily access continuing education in content across big areas (e.g., language, speech, hearing, fluency, voice, social communication, communication modalities, and cognition). However, continuing education regarding methodological or statistical innovations for consumers or producers of research is less available. One might argue that this level of continuing education is not feasible given that our field has multiple areas of content specialty to master. Sharpe also raised the notion of statistical mavens, or individuals who serve as liaisons between statistical innovators and content-area experts. The presence of mavens could serve to provide practitioners and researchers with the information and new skills needed to be effective consumers and producers of advanced methods research. Although a promising concept, using mavens to enhance research design and analyses has received limited attention in the field of speech-language pathology.

Without access to large data sets or infrastructure for continuing education on innovative research methodologies, research in speech-language pathology may be at risk for being “left behind” as new methods and statistical practices are developed and adopted by other fields in order to support more robust, generalizable research. Our resistance to including innovative techniques in our statistical toolboxes could come at great cost as external funding becomes increasingly
competitive and as related fields continue to publish research that was conducted using these complex methods. Researchers with access to large data sets and those who can leverage the most appropriate designs and employ random assignment have the upper hand in competing for external funding. When we consider the disorders that are studied within speech-language pathology, the heterogeneity of affected individuals, and the number of variables we know to be important, it is critical that we expand our methodological skills to include approaches that recognize and account for these levels of complexity.

Given the importance of rigor and of methodological diversity in research, we decided to examine the literature in language and literacy-related journals in order to foster discussion of the range and distribution of research methodologies that are currently being used. We decided to examine and describe general characteristics (e.g., number of participants, number of research institutions involved, population studied, methods, statistical analyses, and funding) of 2013–2014 publications in 10 journals related to child language and literacy based on a review of a subset of randomly selected articles. Specifically, we asked the following research questions:

- What were the average sample sizes and general characteristics of studies published in 2013–2014 in 10 journals that are recognized as publishing language and literacy research (based on a randomly selected sample of 120 articles)?
- What was the proportion of different types of analyses (e.g., qualitative; single-case; nonparametric; ordinary least squares [OLS] or linear least squares; and advanced statistics, such as SEM, HLM, item response theory) in 2013–2014 publications in 10 journals related to child language and literacy (based on a randomly selected sample of 120 articles)?
- To what extent was random assignment and replication used in publications in the sample of 2013–2014 articles in 10 journals?

Finally, a secondary interest of the current research (although not a specific research question) was to discuss the possible implications of these findings for current practices for scholars in higher education programs in communication disorders and specifically for those with interests in language and literacy.

**METHOD**

**Journals of Interest**

We selected 10 journals to review, including four journals of speech-language pathology: *Language, Speech, and Hearing Services in Schools (LSHSS)*; *American Journal of Speech-Language Pathology (AJSLP)*; *Journal of Speech, Language, and Hearing Research (JSLHR)*; and *Journal of Communication Disorders*. Six additional journals relevant to language and literacy were selected for inclusion due to their relevance to the aim of exploring journals that publish language and literacy research: *Journal of Education Psychology (JEP)*; *Scientific Study of Reading (SSR)*; *Reading Research Quarterly (RRQ)*; *Reading and Writing, Journal of Learning Disabilities*, and *Journal of Research for Educational Effectiveness (JREE)*.

*LSHSS*. *LSHSS* is a quarterly journal that is produced by ASHA. The journal focuses on research that is relevant to SLPs and audiologists in schools. The journal is designed to address the needs of researchers, clinicians, and students who are interested in school-based issues. In 2012, *LSHSS* reported an impact factor of 1.256 and a 5-year impact factor of 1.520. In 2013, the journal published four issues containing a total of 26 research articles and in 2014, 17 research articles, reporting research on typically developing children and children with communication disorders ranging in age from toddlers to high-school students and adults who serve as SLPs.

*AJSLP*. *AJSLP* is a quarterly journal that is also produced by ASHA. It is designed to disseminate research findings applicable to clinical practice in speech-language pathology. In 2012, *AJSLP* reported an impact factor of 2.448 and a 5-year impact factor of 2.897. Overlapping in school-based topics with *LSHSS*, *AJSLP* publishes on all aspects of clinical practice, not restricted to school-based topics. In 2013, the journal published four issues containing 27 research articles, and in 2014, published four issues containing 50 research articles.

*JSLHR*. *JSLHR* is a bimonthly journal that is also produced by ASHA. It is designed to disseminate basic and applied research that focuses on normal and disordered communication processes. The mission of *JSLHR* focuses on advancing evidence-based practices as well as providing new information and theoretical approaches that are relevant to speech, language, and hearing processes, assessment, and management. In 2012, *JSLHR* reported an impact factor of 1.971 and a 5-year impact factor of 2.745. In 2013, the journal published six issues containing 149 articles, and in 2014, published six issues containing 171 articles.

*Journal of Communication Disorders*. The *Journal of Communication Disorders*, published six times a year, disseminates articles with a focus on disorders of speech, language, and hearing. Although the journal does not exclusively publish language and literacy research, the assessment, diagnosis, and treatment of
reading disorders is a focus of the journal in addition to the reading-related implications of other communication disorders. The *Journal of Communication Disorders* reported an impact factor of 1.278 in 2015 and a 5-year impact factor of 1.864. In 2013, this journal published one volume (46) with six issues containing a total of 38 research articles; in 2014, the journal published six volumes (47–52) containing a total of 40 articles.

*Journal of Learning Disabilities*. The *Journal of Learning Disabilities*, a journal of the Hammill Institute on Disabilities by Sage, publishes six issues per year, with articles on the science of learning disabilities. The journal reports an impact factor of 1.901 and a ranking of 4 out of 39 in special education. In 2013, this journal produced one volume (46) with six issues containing 44 research articles. In 2014, the journal produced one volume (47) with six issues containing 43 research articles.

*JREE*. *JREE* is a quarterly publication of the Society for Research on Educational Effectiveness. Among the journal’s aims are to disseminate findings of intervention and evaluation studies or methodological studies that focus on the process and implementation of educational research, specifically related to problems in school classrooms. It was reported to have an impact factor of 3.154. In 2013, the journal published 15 research articles. In 2014, the journal published four issues with 16 research articles, including a special issue (number 3) on learning disabilities research studies reporting findings from projects funded by the National Institute of Child Health and Human Development. The introduction to a special issue and two commentaries were excluded from review.

*JEP*. *JEP* is a quarterly journal designed to disseminate research pertaining to education across the lifespan, from early childhood to geriatrics. Not exclusively focused on language and literacy research, the journal includes general research in the area of educational psychology. *JEP* identifies several key focus areas, including scholarship on learning, cognition, instruction, motivation, social issues, emotion, development, special populations, and individual differences. In 2012, *JEP* reported an impact factor of 3.08 and a 5-year impact factor of 4.93. In 2013, the journal published 81 research articles; in 2014, it published 79 research articles.

*SSR*. *SSR* is a bimonthly journal that is produced by the Society for the Scientific Study of Reading. It focuses on empirical studies related to language and literacy, although it also accepts literature reviews, papers on theory and constructs, and policy papers. The journal indicates that it places value on both theoretical and practical significance. *SSR* reports a 5-year impact factor of 3.124. In 2013, the journal published 26 research articles; in 2014, it published 21 research articles.

*RRQ*. *RRQ* is a quarterly journal that is produced by the International Reading Association. It is designed to facilitate connections between researchers in an effort to build a knowledge base in reading and literacy. The journal identifies empirical studies; multidisciplinary research; various modes of investigation; and diverse perspectives on teaching, practices, and learning as valued areas. *RRQ* reports an impact factor of 2.382. In 2013, the journal published 21 research articles; in 2014, it published 22 research articles.

*Reading and Writing: An Interdisciplinary Journal*. *Reading and Writing* is a quarterly journal published by Springer with a primary aim of disseminating scientific articles related to the process, acquisition, and loss of reading and writing skills. The journal description highlights that the focus of the journal spans several disciplines including neuropsychology, cognitive psychology, speech and hearing science, and education. Based on data from 2001 to 2005, the impact factor is reported to be 3.85. In 2013, this journal published nine issues containing a total of 69 articles. In 2014, the journal published nine issues containing 80 articles, 79 of which were research articles and one article that was an introduction to a special issue.

**Procedure**

In approaching the task of identifying and describing research designs and analyses used in each study, we first had to establish definitions of relevant terms across disciplines. For this review, we included only research studies and excluded general literature reviews, editorials, commentaries, tutorials, and position papers. We categorized studies by design using broad categories (refer to the Appendix for definitions and categorizations). Additionally, we identified basic characteristics such as number of participants, age range, number of research institutions involved, method(s) of analyses, presence of random assignment, and inclusion of replication.

We pulled a random sample of 12 articles from each journal, six from each of 2 years—2013 and 2014. Although six is a notably small proportion of articles for some of the journals, the quantity of six was selected in part because a few of the journals had a relatively small corpus of articles published in a given year (e.g., 17 articles in total for *LSHSS* and 15 for *JREE* in 2013). Albeit an arbitrary number, 12 was equivalent to bimonthly distribution. To complete this task, we entered each research article in a journal.
on a line in an Excel database, organized by year and journal title. Initially, we considered excluding articles that did not explicitly discuss literacy implications; however, upon further consideration, it was apparent that a case could be made that literacy components and implications of disorders are quite vast. For example, an article on phonetic processing during the acquisition of new words by children with cochlear implants may not explicitly discuss literacy, yet the research has implicit relevance to phonological awareness. As a result, we used an inclusive approach such that every research article in the relevant journals had an equal chance of being selected for review.

We entered the range of numbers associated with the rows of the Excel database into a random number generator (i.e., select six numbers between 1 and 132). The random numbers generated served to identify which articles would be reviewed for the journal. Six articles were selected in this manner for each year of the journal. In the event that the selected article did not qualify for the study (e.g., a literature review or editorial without data), the article was excluded and the random number generator was used to derive a new line number corresponding to another article. This occurred in three instances for JREE; twice for RRQ; and once for AJSLP, JLD, Journal of Communication Disorders, and SSR.

Each author independently coded four articles from each of the first six journals in an Excel database using mutually agreed-on criteria. A month into the process, we met to discuss the parameters for categorization and to adjust and refine the definitions for distinction of categories. The rules of coding were further defined to clarify the process based on our initial reviews. We agreed that in the event that there were multiple studies or experiments within the same scholarly article, the study with the most advanced statistical design would be included, and if the studies had the same design, which was the case in all instances, then the study with the largest number of participants would be included in the review. Of the articles reviewed, nine articles included multiple research studies within the scholarly article.

**Categorizing statistical analyses.** We coded each article with one of four categorical codes based on the type of research analysis included in it. The codes included four types: qualitative research, single-case methods or a case study, traditional estimation, or advanced statistics. **Qualitative studies** included research that employed qualitative analyses such as open-ended interviewing to describe and develop themes. The **single-case** code was assigned to studies that employed single-case design methods (e.g., A-B-A withdrawal, multiple baseline, or alternating treatments; Horner et al., 2005) or reported a case study. **Traditional estimation** included single-level analyses such as OLS approaches (e.g., analysis of variance [ANOVA], regression, \( t \) tests), which had single variance terms including correlations, descriptive statistics, and nonparametric analyses. **Advanced statistics** included methods that employed multiple levels of analysis (e.g., SEM or HLM) or considered multiple dimensions of sampling (e.g., subjects and time or subjects and items) such as growth curves and mixed models (Garson, 2013).

During the review process, we identified that several of the randomly selected studies used a meta-analysis. As a result, meta-analysis was added as an additional category. In addition to the assignment of categorical types, we also identified specific analyses used by name (e.g., hierarchical regression) to further describe the statistical methodology used in a study.

Randomization was coded with a binary code for the presence or absence of random sampling. This category was included because of the recent focus on randomization (Ioannidis, 2005); however, we recognized that the use of randomization may be less feasible or practical for studies examining individuals with communication disorders. Notably, randomization is not always possible in clinically relevant research in speech-language pathology. For example, researchers cannot randomize who has Down syndrome and who is typically developing, or which participants receive a cochlear implant at an early age versus those receiving treatment B (e.g., hearing aids).

**Agreement.** Of the articles reviewed, 14 were randomly selected to assess interrater agreement. Agreement was calculated by dividing the number of agreements by the total (disagreements plus agreements) \( \times 100 \). There was 100% agreement for the type of design (e.g., qualitative, single-case design/case study, traditional, or advanced) across coders for the six journals. For other data elements (e.g., number of participants, random assignment), there were three cases of confusion where one aspect of interest (e.g., replication) was not explicitly stated in the article. All instances of disagreement were discussed until we reached consensus.

**RESULTS**

Following completion of the review process and agreement consensus gathering, we conducted descriptive data analyses to identify trends in designs and analyses in order to address the research aims, which included describing (a) the average sample size and general characteristics, (b) the proportion of different types of statistical analyses used, and (c) the extent to which random assignment and replication was used.
General Characteristics
To address the first research aim, we aggregated data across the 12 articles selected from each journal and across the complete set of 120 articles. On average, the studies in the articles showed wide variability in the sample size, with a mean of 973 (SD = 4282). The extreme ends of the range were typically large n survey studies (n = 5) and single-case studies (n = 5). When the five survey studies were excluded from the average number of participants, the adjusted mean sample size was 699 (SD = 3313).

On average, the sample studies were conducted by two research institutes, with a range of one to five institutes. Large standard deviations are partially explained by the fact that the average sample size and number of research institutes varied across journals. Table 1 reflects detailed descriptive data by journal. The size of the participant pool tended to be smaller for journals that focused on populations with communication disorders (e.g., LSHSS, AJSLP, JSLHR, JCD) where the average number of participants ranged from 27 to 108. Journals that tended to include populations with typical development (e.g., JEP, SSR, RRQ, JREE) tended to have a larger average number of participants, with means ranging from 140 to 6,241. Three journals, Journal of Learning Disabilities, Reading and Writing, and JREE, appeared to be exceptions to that trend. These three journals included populations with learning disabilities and reading disorders but also had large mean sample sizes (215, 139, and 2,765, respectively).

Table 1. Summary of data set size based on 120 randomly selected articles reviewed.

<table>
<thead>
<tr>
<th>Journal</th>
<th>Universities</th>
<th>Participants</th>
<th>More than 50 participants</th>
<th>Range of sample size</th>
<th>Participants: surveys excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSHSS</td>
<td>1.70 0.89</td>
<td>108 147</td>
<td>50%</td>
<td>1–461</td>
<td>108 147</td>
</tr>
<tr>
<td>AJSLP</td>
<td>1.60 0.67</td>
<td>27 17</td>
<td>0%</td>
<td>3–48</td>
<td>27 17</td>
</tr>
<tr>
<td>JSLHR</td>
<td>1.75 0.75</td>
<td>54 40</td>
<td>33%</td>
<td>8–165</td>
<td>54 40</td>
</tr>
<tr>
<td>JCD</td>
<td>2.20 1.29</td>
<td>61 76</td>
<td>33%</td>
<td>4–250</td>
<td>61 76</td>
</tr>
<tr>
<td>JLD</td>
<td>2.20 1.21</td>
<td>215 292</td>
<td>92%</td>
<td>29–1,031</td>
<td>215 292</td>
</tr>
<tr>
<td>JREE</td>
<td>2.60 1.38</td>
<td>2765 4584</td>
<td>100%</td>
<td>114–13,803</td>
<td>2,765 4,584</td>
</tr>
<tr>
<td>JEP</td>
<td>2.50 1.31</td>
<td>6,241 11,746</td>
<td>92%</td>
<td>47–31,038</td>
<td>4,151 9,684</td>
</tr>
<tr>
<td>SSR</td>
<td>1.80 0.72</td>
<td>140 122</td>
<td>83%</td>
<td>40–466</td>
<td>140 122</td>
</tr>
<tr>
<td>RRQ</td>
<td>1.70 0.78</td>
<td>276 494</td>
<td>67%</td>
<td>1–649</td>
<td>143 187</td>
</tr>
<tr>
<td>RW</td>
<td>2.08 1.16</td>
<td>139 101</td>
<td>83%</td>
<td>28–386</td>
<td>139 101</td>
</tr>
</tbody>
</table>


*Refers to the percentage of studies reporting more than 50 participants; *Represents the mean number of participants without the studies that involved only surveyed participants.

The population of interest and the characteristics of the data set also varied by population type (Table 2). The articles in the journals either focused on typical populations exclusively (60%), focused on participants with atypical development or disorders (20%), or included both groups of participants with typical and atypical characteristics (20%). The studies largely reported on children in grades K–12 (57%), but the age group of focus was somewhat distributed: infant-toddler (9%), preschool (6%), college attendees (9%), and adults (17%). A large portion of the studies reported that the research was grant funded (65%), although some authors may not have reported grant funding as it did not appear to be a customary practice for all of the journals.

Types of Analyses
The 120 articles reviewed used six types of design, including qualitative (n = 4), single-case design or case study (n = 7), traditional quantitative (n = 69), advanced statistics (n = 37), meta-analysis (n = 2), and one simulation that was not specified in our original coding scheme. The prevalence of advanced statistics varied considerably across journals and particularly between journals that focused on communication disorders and other journals (refer to Table 3). Based on the descriptive data in Table 3, six of the 10 journals employed advanced statistics on 25% or more of the randomly selected articles (RRQ, RW, JLD, JREE, SSR, and JEP). Journals that focused more exclusively on participants with communica-
tion disorders (LSHSS, AJSLP, JSLHR, JCD) demonstrated a lower proportion of advanced statistics but employed a variety of designs (i.e., qualitative, single case, traditional, and advanced). The majority of the sampled articles employed traditional methods (58%–92% of the time), most commonly reporting descriptive data, ANOVAs, or regression analyses.

Randomization and Replication

Random assignment was not a predominate characteristic of the research studies reviewed, with only 22% using random assignment. Most of the studies did not report that the research was an attempt to replicate a previous study or did not include replication in their methods (92%). Of the studies in the review pool that addressed replication (n = 10), nine built replication into their design in order to replicate their own findings in a sequence of studies and one reported that a primary aim was to replicate another existing study.

DISCUSSION

Key Findings

Based on trends in the randomly selected sample of 120 journal articles, multidimensional methods of analysis were commonly used, particularly in articles published in JREE, JLD, JEP, RW, RRQ, and SSR. The common use of multidimensional methods seen here is consistent with trends reported in the literature, as noted by one methodologist: “Multilevel and
hierarchical modeling through various types of linear mixed models has rapidly become a required asset in the statistical toolkit of researchers worldwide” (Garson, 2013, p. 23). Trends in the current review suggest that the data sets used in the studies varied largely in terms of size, with 62% including 50 or more participants. Articles from journals in speech-language pathology showed lower average sample sizes.

Across all journals, random assignment was the exception rather than the norm. Studies used a wide variety of types of design and statistical method. Advanced statistical methods that considered multiple levels of analysis and dimensions of sampling were more prevalent in journals that were not specifically within the field of speech-language pathology (e.g., JREE, JEP, RRQ, JLD, RW, and SSR).

It is not surprising that studies from selected journals in speech-language pathology showed smaller sample sizes on average. The focus on low-incidence populations in speech-language pathology may partially explain the tendency toward small data sets. Consistent with this explanation, the out-of-field journals showed a higher percentage of articles pertaining to typical populations. Recruiting children and youth, particularly from low-incidence populations, may require innovative collaborations across institutions in order to access larger data sets. Further, the nature of our interest in populations with unique characteristics may impact the types of methods used, in that employing true experimental designs with random assignment may be challenging given that participant characteristics cannot be assigned and treatment often cannot be ethically withheld.

Mechanisms of Change

One of the intentions of this article was to generate discussion of ways to enhance leadership training in order to prepare next-generation scholars to ensure rigor in research in our field. The results illuminate trends, similarities, and differences in the methods used in scholarly journals related to language and literacy. The discrepancies between journals highlight the need to ensure that scholars in speech-language pathology are poised to be consumers and producers of research who employ diverse methods in order to be competitive for external funding.

Although the identification of levers of change is beyond the scope of this article, one of the aims was to generate such discussion. There are many possible mechanisms of change; among them, professional development, combining forces to establish large data sets, and Big Data at a multistate level. Although professional development is widely available in communication disorders programs, it is perhaps less commonly focused on furthering research competencies for consumers or producers of evidence-based practices. Further, there are not many examples of large open-access or multistate data sets available in the extant literature; however, the concept is aligned with Heilmann, Miller, and Nockerts’ (2010) description of the establishment of large language-sample databases. In this example, the establishment of multistate data using common language-sample protocols allowed for replication and ultimately a discriminant function analysis to validate the use of language sample measures in classifying children’s language status.

Other potential change mechanisms noted in the literature include (a) fostering the establishment of mavens with content knowledge in our field, (b) providing social support such as setting up a special interest group on advanced statistics and methodology, and/or (c) creating blogs on statistical design (Sharpe, 2013). Leaders in the field of speech-language pathology have noted the critical need for building partnerships between highly qualified researchers and school-based SLPs to conduct well-designed high-quality intervention studies focused on improving children’s language and literacy needs (Nippold, 2015).

Becoming producers or consumers of rigorous research using more complex statistics and methodology is difficult, if not impossible, without access to comprehensible professional development. It would not be surprising to find that well-intentioned lifelong learners have purchased an advanced methods book hoping for professional guidance, only to find that they cannot easily digest it without a guide or interpreter. In response, increased offerings of short courses or research translation sessions may be considered that focus on innovative statistics and methodology with relevant application to our field and practices.

There may be no immediate or simple solutions that emerge from the literature. In brainstorming possible options to overcome challenges of working with low-incidence populations, it would seem important for us to leverage incentives to build a useful body of evidence. National organizations could offer incentives for individual clinicians to register item responses on commonly administered assessments in order to facilitate the formation of national data sets that might have utility for difficult-to-answer research questions. It may also be beneficial for national organizations to have designated mavens or information brokers who are skilled at bundling evidence into meaningful, manageable information packages to facilitate use in clinical practices, as findings of rigorous research may otherwise go ignored without translation (Mullen, 2005).
Although the role of SLPs in language and literacy research may be highly regarded, our field risks being left behind if there are lags in the time it takes us to incorporate novel statistics and methodology into our research repertoire. Some authors have suggested that preparation for scientific rigor begins at the undergraduate curriculum so as to optimally prepare students for scientific careers in our field (Koehnke, McNeil, Chapman, Folsom, & Nunez, 2014). Research experiences and course work in methodology and statistics should begin early in students’ career paths but also be present as a constant focus throughout training and beyond.

Each decade also brings waves of innovations in research practices. SEM, HLM, and advances in statistical software bring new opportunities for rigorous research designs. In response, continuing education opportunities are needed to help CSD faculty compete in the research-funding climate. Within a relatively short time of obtaining a degree or even a terminal degree comes the sweeping realization that nothing is terminal about one’s understanding of research design, methodology, and statistics. Indeed, whether one’s desire is to produce high-quality research or to be a wise consumer of it, knowledge of best practices in research can quickly become archaic without continuing education opportunities to stay current.

The end goal in advocating for reliable evidence is to better inform policy and practice. With their sails set on random assignment, massive data sets, and replication, researchers in educational and CSD research may benefit from strategic preparation to poise themselves to take full advantage of the new options to promote rigorous research.

**Study Limitations**

It cannot be assumed that the randomly selected sample of journal articles is representative of the collective set of articles in each journal or representative of the typical proportion of articles using each type of design. We may have derived different trends if we had characterized every article in each issue of each journal. Also, it should be noted that the current review is not an exhaustive review of all journals pertaining to language and literacy. Notably, there are numerous other journals that could be included to expand the review; however, we felt these journals to be of interest as flagship journals.

**Implications**

Despite limitations, the trends in the findings of the review support the need to prepare future faculty and scholars in speech-language pathology to be proficient consumers and producers of a variety of research designs and statistical methodology. Doctoral programs in speech-language pathology may want to carefully consider the research-related competencies to include a range of qualitative, single-case method, and multilevel models. Based on the trends, particularly in JEP, JREE, RRQ and SSR, there is regular use of SEM and hierarchical methods. The current review suggests that studies published in ASHA journals may be shifting a bit in the use of advanced methods as well, as noted by the use of multilevel models in some of the articles reviewed. It may be difficult to consider intensive systems change in preparing doctoral students for use of a range of methods when doctoral students in CSD programs are already in short supply. In response, mechanisms of change may need to be considered.

**References**


Ioannidis, J. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), 696–701. doi:10.1371/journal.pmed.0020124


Contact author: Carla Wood, Florida State University, Communication Disorders, 201 W Bloxham, Tallahassee, FL 32309. Email: carla.wood@cci.fsu.edu
Appendix. Description of Constructs for Coding

**Advanced** was used to describe quantitative studies that utilized more complex statistical analyses including path analysis, item response theory, and approaches that used multiple dimensions of sampling, such as structural equation modeling and multilevel modeling, and meta-analyses.

**Atypical sample** was defined as a sample selected to target specifically members of an atypical population. Samples that were not specified to target specifically atypical populations were labeled as **typical samples**.

**Design** was used to describe the methodological procedures employed to address the purpose of the article and was categorically labeled. Categories included single group: single case/case study; single group: manipulation; single group: observation; multiple group: observation; multiple group: manipulation without random assignment; and multiple group: manipulation with random assignment.

**Manipulation** was defined as the experimenter changing or controlling some variable(s) in the research design.

**Most advanced analysis** was defined as the most advanced statistical analysis employed in the study. Possible classifications were: 1) Qualitative; 2) Single Case/Case Study; 3) Traditional; 4) Advanced.

**Multiple group** was used to describe any design with two or more groups.

**Multiple studies** was the label assigned to research articles including more than one explicitly identified research study.

**Number of groups** was defined as the number of clusters or groups in which participants belonged or were placed.

**Observation** was defined as the experimenter obtaining data without changing or controlling some variable in the research design.

**Population** was categorically labeled and was determined by the age of the participants included in the sample. Population categories included: infant/toddler, Pre-K, K-12, college students, adults, geriatric (ages 60+), or mixed.

**Qualitative** was used to describe studies that employed only qualitative methodology, such as open-ended interviewing to describe and develop themes, rather than quantitative methodology.

**Quantitative** was used to describe studies that employed any form of quantitative methodology, using probability statistics to make inferences or draw conclusions. Studies that employed only surveys for data collection were excluded from this category.

**Random assignment** was defined as statistically-random assignment of groups or manipulations.

**Replication** was defined as the explicitly identified repeating of a previous research investigation with the intent to cross-examine conclusions obtained in that previous study.

**Research** was defined as any design including systematic evaluation of data, excluding meta-analyses. **Single-Case or Small n** was used to describe any single-group design with less than 5 participants or where participants served as their own controls.

**Research entities** were counted as the separate institutions listed within each research article as having participated in the investigation through supporting or employing one or more of the authors.

**Sample size** was defined as the total number of individuals participating in the investigation.

**Single case/case study** was used to describe quantitative studies that examined one or a few subjects using either single case research design or case study design; **Single group** was defined as a research design with one group for the entire duration of the research project.

**Survey** was used to describe studies that employed only questionnaires or script-based interviews for data collection.

**Traditional** was used to describe quantitative studies that utilized only OLS-based or equivalent nonparametric statistical analyses, such as regression, multiple regression, ANOVA, t test, Mann-Whitney, Pearson product–moment correlation, Spearman’s rank order correlation.

**Type of study** was defined categorically as the type of research conducted. Categories included: qualitative, quantitative, survey, and mixed.

**Typical sample** was defined as a sample following a conventional and predictable pattern based on the majority of the general population.