

# Research and Practice in the Schools

---

The Official Journal of the Texas Association of School Psychologists

---

Volume 4, Issue 1 ♦ October 2016



**Texas Association of School Psychologists**

*A partner on the path to student success*

**Research and Practice in the Schools:  
The Official Journal of the Texas Association of School Psychologists**

**Volume 4, Issue 1      October 2016**

**Editors:**

Jeremy R. Sullivan, *University of Texas at San Antonio*  
Arthur E. Hernandez, *University of the Incarnate Word*

**Graduate Student Section Editor:**

Ashley Doss, *Stephen F. Austin State University*

**Editorial Review Board:**

A. Alexander Beaujean, *Baylor University*  
Crystal Carbone, *University of Houston – Clear Lake*  
Felicia Castro-Villarreal, *University of Texas at San Antonio*  
Christy Chapman, *Texas Tech University*  
Sarah Conoyer, *Texas A&M University – Commerce*  
Krystal T. Cook, *Texas A&M University*  
Lisa Daniel, *Lewisville ISD*  
Kathy DeOrnellas, *Texas Woman’s University*  
Norma Guerra, *University of Texas at San Antonio*  
Elise N. Hendricker, *University of Houston – Victoria*  
Laurie Klose, *Trinity University*  
Jennifer Langley, *Sheldon ISD*  
Coady Lapierre, *Texas A&M University – Central Texas*  
Ron Livingston, *University of Texas at Tyler*  
William G. Masten, *Texas A&M University – Commerce*  
Daniel McCleary, *Stephen F. Austin State University*  
Anita McCormick, *Texas A&M University*  
Ryan J. McGill, *College of William & Mary*  
Kerri P. Nowell, *University of Houston*  
Sarah Ochs, *University of Houston*  
Nancy Peña Razo, *University of Texas Rio Grande Valley*  
Brook Roberts, *Texas Tech University*  
Billie Jo Rodriguez, *Springfield Public Schools (Springfield, OR)*  
Thomas Schanding, *University of Houston – Clear Lake*  
Andrew Schmitt, *University of Texas at Tyler*  
Jennifer Schroeder, *Texas A&M University – Commerce*  
Tara Stevens, *Texas Tech University*  
Victor Villarreal, *University of Texas at San Antonio*  
Shannon Viola, *Cypress-Fairbanks ISD*

**Graduate Assistant:**

Colleen Pirog, *University of Texas at San Antonio*

---

**Publication Information:**

*Research and Practice in the Schools* is a peer-reviewed, online journal published by the Texas Association of School Psychologists. ISSN: 2329-5783

**Disclaimer:**

The views and opinions expressed by contributors to the journal do not necessarily reflect those of the publisher and/or editors. The publication of advertisements or similar promotional materials does not constitute endorsement by the publisher and/or editors. The publisher and editors are not responsible for any consequences arising from the use of information contained in the journal.

# **Research and Practice in the Schools: The Official Journal of the Texas Association of School Psychologists**

*Research and Practice in the Schools* is a publication of the Texas Association of School Psychologists (TASP). It is an online, peer-reviewed journal that provides TASP members with access to current research that impacts the practice of school psychology. The primary purpose of *Research and Practice in the Schools* is to meet the needs of TASP members for information on research-based practices in the field of school psychology. To meet this need, the journal welcomes timely and original empirical research, theoretical or conceptual articles, test reviews, book reviews, and software reviews. Qualitative and case-study research designs will be considered as appropriate, in addition to more traditional quantitative designs. All submissions should clearly articulate implications for the practice of psychology in the schools.

## **Instructions for Authors**

### **General Submission Guidelines**

All manuscripts should be submitted in electronic form to either of the co-editors ([jeremy.sullivan@utsa.edu](mailto:jeremy.sullivan@utsa.edu) or [aeherna8@uiwtx.edu](mailto:aeherna8@uiwtx.edu)) as an email attachment. Manuscripts should be submitted in MS Word format and labeled with the manuscript's title.

It is assumed that any manuscript submitted for review is not being considered concurrently by another journal. Each submission must be accompanied by a statement that it has not been simultaneously submitted for publication elsewhere, and has not been previously published.

Authors are responsible for obtaining permission to reproduce copyrighted material from other sources. IRB approval should have been obtained and should be noted in all studies involving human subjects. Manuscripts and accompanying materials become the property of the publisher. Upon acceptance for publication, authors will be asked to sign a publication agreement granting TASP permission to publish the manuscript. The editors reserve the right to edit the manuscript as necessary for publication if accepted.

Submissions should be typed, double-spaced with margins of one inch. All articles should meet the requirements of the *APA Publication Manual, 6<sup>th</sup> ed.*, in terms of style, references, and citations. Pages should be numbered consecutively throughout the document. Illustrations should be provided as clean digital files in .pdf format with a resolution of 300 dpi or higher. Tables and figures may be embedded in the text. A short descriptive title should appear above each table with a clear legend and any footnotes below.

### **The Review Process**

After receiving the original manuscript, it will be reviewed by the Editors and anonymously by two or more reviewers from the Editorial Board or individuals appointed on an *ad hoc* basis. Reviewers will judge manuscripts according to a specified set of criteria, based on the type of submission. Upon completion of the initial review process, feedback will be offered to the original (primary) author with either (a) a preliminary target date for publication; (b) a request for minor editing or changes and resubmission; (c) significant changes with an invitation for resubmission once these changes are made; or, (d) a decision that the submission does not meet the requirements of *Research and Practice in the Schools*.

## **Call for Special Issue Proposals**

We invite proposals for special issues of the journal, with the goal of publishing one special issue each year in addition to the general issue. Special issues will include collections of papers related to some cohesive theme in the field of School Psychology, and will be edited by Guest Editors who will take the lead in soliciting contributions and coordinating the peer review process. In addition to special issues that focus on research and scholarship in School Psychology, we welcome special issues that cover important practical and applied issues in the field.

Special issue proposals should include a brief description of the theme to be covered by the issue, approximate number of articles to be included, qualifications and expertise of those who will serve as Guest Editors of the issue, and a plan for soliciting manuscripts and conducting the reviews. Proposals for special issues, and questions about the process, should be sent to [jeremy.sullivan@utsa.edu](mailto:jeremy.sullivan@utsa.edu).

## **Graduate Student Section**

Ashley Doss, doctoral student at Stephen F. Austin State University, serves as Graduate Student Section Editor for the journal. The Graduate Student Section is devoted to publishing the work of graduate students, including research studies, comprehensive literature reviews on relevant topics, and reviews of books or psychological/educational tests published within the past two years. As with all submissions to the journal, graduate student manuscripts should highlight implications for practice in the schools. If you are a graduate student and you have questions about how you can best contribute to the journal (as an author, reviewer, or both), please email Ashley at: [dossan2@jacks.sfasu.edu](mailto:dossan2@jacks.sfasu.edu).

Please note: all manuscripts submitted to the Graduate Student Section must include either a faculty co-author or a faculty sponsor who provides the student with mentorship on the process of preparing and submitting their work for peer review. When submitting their manuscripts for review, student authors should include a cover letter verifying that their work has been vetted by a faculty co-author or sponsor.

## **Self-Archiving Policy**

Authors retain the right to self-archive the final, accepted manuscript of their submission on their own websites or deposit this version of the manuscript in any repository, provided it is only made publically available one calendar year (12 months) after publication or later. The archived version should be the final typeset article as it appears in the online issue of the Journal and the archive should include the appropriate citation and link to the Journal issue in which it appeared.

Article

---

## History of Credentialing in School Psychology in Texas: Keynote Address from the 2016 TASP Annual Convention

Daniel C. Miller  
*Texas Woman's University*

Gail M. Cheramie  
*University of Houston-Clear Lake*

Donna L. Black  
*Member of Texas State Board of Examiners of Psychologists*

Amidst a movement of education reform to the Texas education system, the Texas Education Code was completely overhauled in 1995 with the passage of Senate Bill 1 (SB1). Specific to the profession of school psychology, SB1 mandated professionals working in schools to be licensed by their respective appropriate boards/agencies. For school psychology, this meant licensure by the Texas State Board of Examiners of Psychologists. However, a license specific to school psychology was not in existence at that time, thus leading to the creation of the Licensed Specialist in School Psychology (LSSP) license. This article presents the history of the credentials required for the practice of school psychology in Texas, changes in and refinements of the rules governing the LSSP and practice of school psychology over the past two decades, and possible trends for the future.

**Keywords:** Licensure, credentialing, psychology

With the passage of Senate Bill 1 in the 1995 session of the Texas Legislature, the Texas State Board of Examiners of Psychologists (TSBEP) became the regulatory agency for the practice of school psychology in the State of Texas (Miller, 2008a, 2008b). In 1996, the TSBEP created the requirements for the credential and the initial rules of practice for the Licensed Specialist in School Psychology (LSSP) (Miller, 2009a, 2009b). Since 2016 marks the 20th anniversary of the creation of the LSSP, it seems fitting to look back on the history of school psychology credentialing in Texas. This article will briefly review the credentialing of school psychology in Texas prior to 1995, the development of the TSBEP rules for the LSSP, and propose possible future directions for the regulation of school psychology in the state.

**Author Note:** Daniel C. Miller, Ph.D., ABPP is the Executive Director of the Woodcock Institute for the Advancement of Neurocognitive Research and Applied Practice and Professor Emeritus in the Department of Psychology and Philosophy at Texas Woman's University.

Gail M. Cheramie, Ph.D., is an Associate Professor in the specialist-level NASP-Approved School Psychology program in the Department of Clinical, Health and Applied Sciences at the University of Houston-Clear Lake.

Donna Black, M.A., is a member of the Texas State Board of Examiners of Psychologists and has served as a board member since April, 2006. She also is an Assessment Consultant at Pearson Clinical Assessment.

Correspondence for this article should be addressed to: Dr. Daniel C. Miller, Woodcock Institute for the Advancement of Neurocognitive Research and Applied Practice, Texas Woman's University, Box 425249, Denton, TX 76204 Contact: dmiller@twu.edu

### School Psychology Credentialing Prior to 1995

Prior to 1995, the Texas Education Agency (TEA) was the credentialing body for school psychology. There were multiple paths for obtaining the required credential for delivering psychological services in the schools. Four levels of certification specific to school psychology were in existence through TEA: Professional School Psychologist (Doctoral level), Professional Associate School Psychologist (Master's Level), Intermediate School Psychologist (Doctoral Level), and Intermediate Associate School Psychologist (Master's Level). The Intermediate certifications allowed individuals with incomplete training (e.g., specific coursework, specific experience) to practice in schools while working to complete all requirements. At that time, the TSBEP had two levels of licensure that were germane to the practice of school psychology: the Licensed Psychologist (LP) and the Licensed Psychological Associate (LPA).

Technically, if an individual sought certification by TEA, that individual was first licensed by TSBEP as an LP or LPA and then applied to TEA to obtain the appropriate certification. In other words, the individual must first be licensed by the psychology board to practice psychology and then certified by the TEA to work in the schools. The TEA certification required the individual to take four additional courses specific to education: organization, administration, and financing of public schools; curriculum and instruction; education of handicapped children; and multicultural education (19 TAC Section 141.26). Between 1976 and 1991, if a practitioner who was delivering school psychological services was asked what credential was needed to work in that setting, the response would be quite varied. Basically, practitioners were dually credentialed, holding a license and certificate, or in the case of the NCSP, dually certified.

The minimum entry-level credential to work in the schools delivering school psychological services was the Licensed Psychological Associate (LPA) from the TSBEP. The LPA required 42 hours of graduate training in psychology; however, no specialized coursework in school psychology was required. Therefore, in the 1970s through the early

1990s, there were many practitioners in Texas who were working in the schools delivering school psychological services and they had no specific training in, or professional identity with, school psychology. In fact, an individual could be credentialed to work in schools in Texas as a school psychologist or associate school psychologist and never have completed a practicum or internship in schools.

In 1989, the National Association of School Psychologists (NASP) introduced the Nationally Certified School Psychologist (NCSP) credential. By 1991, TEA allowed practitioners who held the NCSP credential to bypass the TSBEP licensure process and become directly certified by TEA (Texas Register, Volume 16, Number 88, November, 1991), thus allowing full credentialing of school psychology to reside with the TEA. The recognition of the NCSP credential in Texas was a major milestone. Left unresolved in 1991-1993 was the issue of where the NCSP holder fell within the continuum of service delivery. Could a TEA certified school psychologist with a NCSP work independently in the schools without supervision by a licensed psychologist, given that prior to this path to credentialing an individual must have either an LPA or LP *and* a TEA credential? This was the proverbial "\$64,000 question" that created a great deal of tension among the wide variety of practitioners who were delivering school psychological services within the schools of Texas. The debate about the recognition of the NCSP credential as evidence for entry level competency for independent practice in the schools created a firestorm of political debate that soon thrust Texas into the vortex of the long-standing national debate regarding the doctoral/non-doctoral qualifications for independent practice.

In 1994, the U.S. Department of Education, Office of Special Education made a ruling that states could not have dual levels of credentialing for school psychologists if the only difference between the levels was the terminal degree. This OSEP regulation put the TEA dual credentialing of Professional School Psychologist (Doctoral) or Professional Associate School Psychologist (non-Doctoral) in jeopardy. However, this problem was never addressed by TEA because of the major legislative change that occurred in 1995.

In Texas, the legislature only meets every other year to enact legislation and pass a biennium budget. The political climate in 1995, both nationally and in Texas, reflected governmental downsizing and greater local control. In 1995, the Texas Education Code was due for revision, as required for all regulatory statutes every ten years. As a result of this mandate, the first bill of the 1995 legislative session was Senate Bill 1 (SB1), which proposed an extensive revision to the Texas Education Code, Title 2.

SB1 included a radical revision of how educational personnel were to be certified for work in the schools. Professionals such as school psychologists were to be *licensed* by professional practice boards rather than *certified* by TEA. The specific language in the Texas Education Code (TEC), Chapter 21, Subchapter A, Section 21.003(b) was as follows:

*"...a person may not be employed by a school district as an audiologist, occupational therapist, physical therapist, physician, nurse, school psychologist, associate school psychologist, licensed professional counselor, marriage and family therapist, social worker, or speech language pathologist unless the person is licensed by the state agency that licenses that profession and may perform specific services within those professions for a school district only if the person holds the appropriate credential from the appropriate state agency."* (TEC, Chapter 21, Section 21.003 (a))

The unresolved and politically charged question in Texas suddenly became: what constitutes "the appropriate credential" to practice school psychology in the state? The passage of SB1 shifted the credentialing of school psychology from TEA to the TSBEP, but the two principal licenses granted by the TSBEP were the LP at the doctoral level and the LPA at the master's level. There was no license available specific to the specialist-level of training and specific to the practice of school psychology. The proposed law potentially left school psychology unregulated in the state.

During the state senate hearings on the proposed legislation, the professional organizations that had a vested interest in the practice of psychology in the schools and representatives of political action

groups with vested interests were present and provided testimony. The polarization within the field along the doctoral and non-doctoral constituencies was very apparent to the senate committee members. The chief legislative aide for the senator who sponsored the bill talked with Dr. Miller, the TASP President, and a few other representatives after the hearing and instructed the groups to come to a consensus about language that would "fix the problem" and bring recommendations to the senator's office.

A coalition of association (TPA, TAPA, and TASP) presidents, or their representatives, and spokespersons for the various political action groups were invited to come together at a meeting to draft legislative language that would create a new license issued by TSBEP for the practice of psychology in the schools. After intense discussion, consensus language for the creation of the Licensed Specialist in School Psychology (LSSP) was presented to the senator's office for insertion into SB1.

The result was as follows:

*Senate Rule 12.03(4) is suspended to permit the committee in Section 51 of the conference committee report to add Section 26, Psychologists' Certification and Licensing Act (Article 4512c, Vernon's Texas Civil Statutes), to read as follows:*

*Sec. 26. LICENSED SPECIALIST IN SCHOOL PSYCHOLOGY. (a) The board by rule shall adopt the license classification of "licensed specialist in school psychology." A license issued under this section constitutes the appropriate credential for a person to provide psychological services in a school in this state as required by Section 21.003(b), Education Code.*

*(b) The board shall set the standards for qualification of a license issued under this section. The standards must include:*

- (1) minimum recognized graduate degree requirements;*
- (2) completion of graduate course work at a regionally accredited institution of higher education in the following areas:*
  - (A) psychological foundations;*
  - (B) educational foundations;*
  - (C) interventions;*



- (D) assessments; and*
- (E) professional issues and ethics;*
- (3) completion of a minimum of 1,200 hours of supervised experience;*
- (4) receipt of a passing score on a nationally recognized qualifying examination determined to be appropriate by the board and on any other examination determined to be necessary by the board; and*
- (5) satisfaction of the requirements imposed under Section 11(d) of this Act.*

*(c) The rules of practice for a licensed specialist in school psychology must comply with nationally recognized standards for the practice of school psychology.*

*Explanation: This change is necessary to require the setting of standards for persons who provide psychological services in schools.*

The title Licensed Specialist in School Psychology was a major compromise between the TASP representatives who were advocating for NASP training standards and the TPA representatives who were advocating for APA training standards. TASP, with the backing of NASP, advocated that the title “School Psychologist” should be used for this new level of licensure since it was the title being used across the country for both non-doctoral (Master’s and Specialist Levels of Training) and doctoral individuals. However, this was not acceptable to TPA. In a conversation with Dr. Emily Sutter (May 10, 2016) who was Chair of the TSBEP Rules Committee at that time, she noted that TSBEP could not assign the title of “psychologist” as this was a protected title for doctoral-level individuals. The dilemma was how to create the new title for school psychology given that “school psychologist” was used in almost all states in the country, but “psychologist” was a protected title in Texas. Thus, the LSSP was the compromise that acknowledged the specialist level of training.

The compromise language that created the LSSP was submitted to the senate in March 1995. Afterward, during the first week of May, TASP representatives received calls from the senator’s

office who sponsored SB1 indicating that the support for the creation of the LSSP was falling apart. The TPA was withdrawing their support in favor of requiring all providers of independent school-based psychological services to have a doctorate and be licensed as a psychologist. Despite having the LSSP submitted as a compromise, TPA decided to withdraw their support. At that point, the fate of the LSSP in SB1 was not clear. However, the LSSP language was reinserted into the bill at the last minute and was approved by both the senate and house. So what happened to tip the scales? It was reported to Miller (2008b) by the key legislator’s aide that at the last minute, the Texas Council of Administrators of Special Education (TCASE) came to realize the critical importance of the LSSP license and had done some last minute lobbying on behalf of the LSSP.

A great deal can be learned from the process of the creation of the LSSP. Tremendous controversy and dissention led to important advocacy work and attention to details was critically important. One lesson learned was that a single word matters when crafting legislation. In the inclusion of the LSSP language in SB1, there was controversy over what qualifying term should appear before “appropriate credential.” Was the LSSP to be “an” appropriate credential or “the” appropriate credential? Because OSEP had issued the letter in 1994 about dual levels of credentialing, TASP was able to ensure that the term “the” was included which meant that all practitioners, including those with doctoral degrees, had to obtain the LSSP in order to provide psychological services in the schools. Another lesson learned was that the leaders/representatives of organizations are the most effective advocates for the profession when they possess a vast knowledge base of the profession they represent. For school psychology this means knowledge not only of psychology but of education, related laws and decisions that affect the practice of psychology in schools. A third lesson learned was the importance of forming and maintaining coalitions with related professional organizations. Professional coalitions are powerful partners in advocacy for any profession. The coalition support from TCASE was instrumental in the passage of the LSSP language in SB1. Finally, the most important lesson learned through the process was that dissention and

disagreement can be managed through problem solving and compromise. Without the hard work of those involved in developing the compromise language, the LSSP would never have been established.

With the passage of SB1, the LSSP was created and the TSBEP was charged with developing rules and adopting minimum academic standards for the license. The TSBEP rules would not only govern who could obtain the license, but would set the stage for the future practice of school psychology in Texas.

### **The LSSP: 1996-2005**

The actual birthdate of the LSSP Rule is April 10, 1996 (TSBEP, 465.38) and it went into effect on September 1, 1996. As noted above, Dr. Emily Sutter was Chair of the Rules Committee for the TSBEP at the time of the creation of the LSSP rules. Although the title dilemma noted above was resolved by the creation of the LSSP, the other dilemmas regarding supervision and training standards remained. SB1 instructed the TSBEP to adopt rules that would comply with nationally recognized standards for the practice of school psychology, so the TSBEP adopted the NASP training and practice standards, but added a Texas twist.

In an interview on May 10, 2016, Dr. Emily Sutter reported that the TSBEP was very concerned about having to create rules for the LSSP because the TSBEP did not have authority to regulate any government agency, such as the public school system. Until this point in time, the purpose of TSBEP was to regulate the practice of psychology in the private sector, not public agencies. Dr. Sutter, as chair of the rules committee, was responsible for drafting the LSSP rules. She noted that the committee consulted rules related to school psychology in other states and national standards created by NASP. In consultation with TPA, TASP, and TAPA the rules were developed; but not without extreme controversy again. Doctoral psychologists were adamantly opposed to a non-doctoral individual working without supervision and believed that all non-doctoral individuals were meant to work under supervision. It should be noted that the one-year supervision for a fully licensed LSSP was another compromise between those who

wanted supervision forever and those who advocated for the NASP model that did not require formal supervision once an individual was fully credentialed. Dr. Sutter pointed out that in order for this rule to work, a conceptualization had to be formulated of how school psychology was different than practice in the private sector since school based practice is regulated, in part, by federal laws. To accomplish this conceptualization, Dr. Sutter included the distinction of school psychology as a unique profession within the rule itself. TSBEP Rule 465.38, Psychological Services in the Schools, reads as follows:

*(a) This rule acknowledges the unique difference in the delivery of school psychological services in the public schools from psychological services in the private sector. The Board recognizes the purview of the State Board of Education and the Texas Education Agency in safeguarding the rights of public school children in Texas. The mandated multidisciplinary team decision making, hierarchy of supervision, regulatory provisions, and past traditions of school psychological service delivery both nationally and in Texas, among other factors, allow for rules of practice in the public schools which reflect these occupational distinctions from the private practice of psychology.*

The TSBEP held statewide hearings and gathered public testimony prior to the adoption of the rules. The polarization within the field that manifested itself in the SBOE Special Education Rule change hearings in 1994 and in the passage of the LSSP law in 1995 had not dissipated in 1996 as these LSSP rules were being drafted. The doctoral-only constituencies asserted that the TSBEP should draft LSSP rules of practice that followed the APA training standards, while TASP and other non-doctoral groups advocated for the NASP training standards to be adopted. Level heads prevailed and the TSBEP adopted the new LSSP rules of practice in April of 1996.

The key elements of the LSSP rules were modeled after NASP Graduate Preparation Standards, State Credentialing Standards, Principles for Professional Ethics, Standards for Practice and the requirements for the Nationally Certified School

Psychologist credential. In addition, applicants for licensure had to pass the state’s jurisprudence exam. The rules allowed for a streamlined application process for NCSP holders and included a definition of school psychological practice designed to prohibit non-school psychologists from delivering the full range of psychological services in the schools.

According to TSBEP Board Rules, the Licensed Specialist in School Psychology was defined as *...a person who is trained to address psychological and behavioral problems manifested in and associated with educational systems by utilizing psychological concepts and methods in programs or actions which attempt to improve the learning, adjustment and behavior of students, including the assessment of emotional or behavioral disturbance, for educational purposes, using psychological techniques and procedures.*

The LSSP law and TSBEP rules went into effect on September 1, 1996. A liberal grandparenting provision was written into the TSBEP rules that allowed current practitioners to obtain the LSSP, but all providers of the broad range of school psychological services had to have the LSSP no later than September 1, 1997.

In March 1996, the TSBEP mailed to all licensees a *Memorandum to Interested Parties* (dated March 18, 1996). The memorandum contained detailed answers to a set of frequently asked questions about the new licensing requirement. Applicants who were eligible for the license under the grandparenting provisions included those who were retained by a public school district (on or after 9/1/92 but before 9/1/96) to provide psychological services and were also

licensed by TSBEP, and/or held the NCSP, or were credentialed by TEA. Individuals who met this requirement had until 9/1/1997 to apply for the LSSP under the grandparent provision. It was also clear in this memorandum that the LSSP was the credential required to offer the broad range of school psychological services, whether employed full-time or contractually, and that the LSSP was a license for use only in public schools and did not apply in the private sector. During the grandparenting period (1996-1997), approximately 1000 LSSPs were licensed.

One of the most important features of the TSBEP Board Rules for the LSSP was the mandate that a practitioner who worked in the schools delivering comprehensive school psychological services must have specific training in school psychology. This provision alone has dramatically improved the quality of school psychological services in Texas.

Within the first decade of the LSSP, specific rules were written and adopted to address the changes in requirements for the delivery of school psychological services in Texas public schools. There were two general rules that addressed the LSSP, one involving the license requirements (originally was Rule 463.32, but later was changed to Rule 463.9) and one involving the practice of school psychology (Rule 465.38). These rules were amended several times, and the amendments for the first decade are reflected in Table 1. Basically, the rule changes during the first decade refined and expanded the scope of practice for LSSPs, established that the LSSP was the correct title, addressed supervision, and ensured that internships were completed as part of a formal university program.

**Table 1** *LSSP Rule Changes from 1996-2005*

<b>Board Rule §465.32 Licensed Specialist in School Psychology, Requirements for Licensure</b>		
<b>Date</b>	<b>Status</b>	<b>Significant Amendments or Additions</b>
April 1996	Adopted	Pursuant to Section 21.003(b) of the Education Code, the Board was authorized to set rules for the LSSP in order to replace the school psychologist and associate school psychologist certificates previously issued by the Texas Education Agency for providers of school psychological

**Table 1 continued**

		services. This rule established the training qualifications (i.e., coursework, practicum experience, and internship experience,) along with examination requirements, additional requirements, temporary licensure requirements, and grandparenting provisions for individuals seeking the LSSP.
October 1996	Amended	Reformatted sections of the rule – No significant changes
November 1997	Amended	(a) Training qualifications. <ul style="list-style-type: none"> <li>Clarified that applicants must have completed a graduate degree in psychology from a regionally accredited academic institution and obtained 60 graduate level semester credit hours that were obtained from a regionally accredited academic institution.</li> </ul> (b) Completion of internship or experience. <ul style="list-style-type: none"> <li>Changed language to state that individuals must be designated as “interns” and removed the option of being designated as a Trainee. Also clarified that internships must include “direct intern application” in the areas of experience that were itemized in the rule.</li> </ul>
October 1999	Repealed	<ul style="list-style-type: none"> <li>Rule was repealed and replaced by Board rule §463.9</li> </ul>
<b>Board Rule §463.9 LICENSED SPECIALIST IN SCHOOL PSYCHOLOGY</b>		
<b>Date</b>	<b>Status</b>	<b>Significant Additions or Amendments</b>
October 1999	Adopted	New rule established requirements for licensure that included <ul style="list-style-type: none"> <li>Application requirements</li> <li>Training qualifications requirements</li> <li>Completion of internship/experience requirements</li> <li>Examination requirements (i.e., National School Psychology Exam, TSBEP Jurisprudence Exam)</li> <li>Trainee requirements</li> </ul>
May 2000	Amended	(c) <b>Completion of internship.</b> <ul style="list-style-type: none"> <li>Added requirement that internships completed after July 1, 2001 must be completed as part of a formal course of supervised study from a regionally accredited institution of higher education in which the applicant was enrolled, or be obtained in accordance with requirements for obtaining the LP license.</li> <li>Additional language was added for applicants whose internships began before the specified date. The additional language clarified that either a formal internship or experience would meet this requirement.</li> </ul>
July 2001	Amended	(b) <b>Training qualifications.</b> <ul style="list-style-type: none"> <li>Clarified that, in addition to training qualifications being met, internship qualifications also were met if</li> </ul>

**Table 1 continued**

		candidates for licensure held a current and valid NCSP credential or if they graduated from a training program approved by NASP or accredited in School Psychology by APA.
June 2003	Amended	<p><b>(c) Completion of internship.</b></p> <ul style="list-style-type: none"> <li>Removed language in the rule that applied to applicants whose internships began before, on or after July 1, 2001 since it was no longer necessary to provide transitional requirements for any group of licensees (i.e., all current applicants would be required to meet the current standards for completion of the internship.)</li> </ul>
August 2004	Amended	<p><b>(b) Training Qualifications.</b></p> <ul style="list-style-type: none"> <li>Added the requirement that applicants for licensure who did not hold a currently valid NCSP or did not graduate from a NASP-approved training program or a program accredited in School Psychology by APA, must have completed a graduate degree in psychology from a regionally accredited academic institution and 60 graduate level semester credit hours, also from a regionally accredited academic institution.</li> </ul>
<b>Board Rule §465.38 PSYCHOLOGICAL SERVICES IN THE SCHOOLS</b>		
<b>Date</b>	<b>Status</b>	<b>Significant Amendments or Additions</b>
April 1996	Adopted	<ul style="list-style-type: none"> <li>Defined the type and scope of practice permitted by individuals providing psychological services in public schools as a LSSP, as well as the minimum level of competency that all practitioners must possess in order to obtain the LSSP license.</li> </ul>
October 1996	Amended	<p><b>(3) Supervision.</b></p> <ul style="list-style-type: none"> <li>Revised the requirements for supervision of interns, trainees, first year licensees, and LSSPs who provide services outside his/her area of training or experience.</li> <li>Specified an exemption from supervision for those LSSPs who received the license under the grandparenting provision.</li> </ul>
November 1997	Amended	<p><b>(2) Titles.</b></p> <ul style="list-style-type: none"> <li>Established the correct title to be used by persons holding this license (i.e., <i>Licensed Specialist in School Psychology</i> or <i>LSSP</i>.)</li> </ul> <p><b>(3) Providers of School Psychological Services.</b></p> <ul style="list-style-type: none"> <li>Clarified that individuals holding a temporary license issued by the Board could also provide school psychological services</li> </ul> <p><b>(4)(a)(ii) Supervision.</b></p> <ul style="list-style-type: none"> <li>Required individuals who had met the training requirements, had applied for licensure, and were still required to be supervised for one year to be designated</li> </ul>

Table 1 continued

		<p>as trainees.</p> <p><b>(5) Supervisor Qualifications.</b></p> <ul style="list-style-type: none"> <li>Stipulated that supervision could only be provided by a LSSP with three years of experience providing school psychological services (this included anyone who was licensed under the grandparenting provision) and an individual could count his/her intern or trainee year as one of the qualifying years.</li> </ul>
June 1998	Amended	<p><b>(4)(A)(ii) Qualified Applicants.</b></p> <ul style="list-style-type: none"> <li>Clarified that individuals were qualified to practice as a trainee only after they were notified by the Board that their applications were complete as defined by §463.5(6).</li> </ul>
October 1999	Amended	<p><b>(3) Providers of Psychological Services.</b></p> <ul style="list-style-type: none"> <li>Removed temporary license holders as qualified providers.</li> </ul> <p><b>(4)(A)(ii) Supervision requirements.</b></p> <ul style="list-style-type: none"> <li>Clarified that individuals were considered to have met the training requirements if they passed the National School Psychology Examination at the Texas cutoff score or above and had been notified in writing by the Board.</li> <li>Stipulated that these individuals could practice under supervision for no more than one calendar year.</li> </ul> <p><b>(4)(A)(iii) Supervision requirements following licensure.</b></p> <ul style="list-style-type: none"> <li>Specified that individuals must continue to be supervised for one academic year following licensure, unless they also are licensed as a Licensed Psychologist.</li> </ul> <p><b>(B) Supervision of individuals who were licensed under the grandparenting provisions.</b></p> <ul style="list-style-type: none"> <li>This section was no longer needed, therefore was eliminated.</li> </ul>
March 2000	Amended	<p><b>(2) Titles.</b></p> <ul style="list-style-type: none"> <li>Clarified that only individuals who met the requirements of §465.6 (relating to Listings, Public Statements and Advertisements, Solicitation, and Specialty Titles) could refer to themselves as School Psychologists.</li> </ul>
July 2000	Amended	<p><b>(6) Conflict Between Laws and Board Rules</b></p> <ul style="list-style-type: none"> <li>Added language stating that state or federal statutes take precedence whenever there is a conflict between the statutes and the Board's rules.</li> </ul>
March 2001	Amended	<p><b>(6) Conflict Between Laws and Board Rules</b></p> <ul style="list-style-type: none"> <li>Removed the examples of laws that were provided</li> </ul> <p><b>(7) Compliance with Applicable Education Laws.</b></p> <ul style="list-style-type: none"> <li>Added language specifying that LSSPs are required to</li> </ul>

**Table 1 continued**

		<p>comply with applicable state and federal laws affecting the practice of school psychology.</p> <p><b>(7) (A)-(E) Applicable Laws</b></p> <ul style="list-style-type: none"> <li>Listed those laws that may be applicable to the practice of school psychology.</li> </ul>
July 2001	Amended	<p><b>(B) Definition.</b></p> <ul style="list-style-type: none"> <li>Removed the part of the definition specifying that activities provided by LSSPs included only the assessment of emotional or behavioral disturbance for educational purposes.</li> <li>Added language to clarify that activities provided by LSSPS were not only limited to addressing special education eligibility, but also included other activities, such as conducting manifestation determinations and assisting with the development and implementation of individual educational programs.</li> </ul> <p><b>(3) Providers of School Psychological Services.</b></p> <ul style="list-style-type: none"> <li>Added language to clarify that school districts can contract for specific types of psychological services that are not readily available from the LSSP, so long as the contracting is on a short-term or part-time basis and does not involve the broad range of school psychological services listed in part (1)(B) of this rule.</li> </ul>
July 2002	Amended	<p><b>(3) Providers of School Psychological Services.</b></p> <ul style="list-style-type: none"> <li>Added language to clarify that an LSSP who contracts with a school district to provide school psychological services may not permit another individual who does not hold the LSSP to perform any of the contracted services.</li> </ul>

**The LSSP: 2006-Present**

The TSBEP rules have continued to evolve over the last decade. Table 2 presents major amendments in the rules within the past 10 years. Some noteworthy changes include the use of the NCSP credential, requirements related to supervision, and further alignment of the LSSP with school-based practice (e.g., informed consent, forensic exclusion, and termination of employment).

When the initial rules were developed and refined, holders of the NCSP credential automatically met training and internship requirements, but there was no provision for use of the NCSP. This of course led to controversy, more testimony before the TSBEP, and significant

advocacy work by TASP with the full support of NASP. In 2012, specific language was added to allow use of the NCSP along with the LSSP. This controversy was reminiscent of the doctoral versus non-doctoral issues and use of the term psychologist. Clearly, this remains an important issue in Texas.

In recent rule changes, supervision requirements were overhauled and two particular changes are noteworthy. Prior to April of 2016, a supervisor had to have three years of experience, with the internship and trainee years allowed to count in the three-year period. Thus, a supervisor’s three years of experience (internship, trainee, and first-year LSSP) could be met during the time period when that individual was under supervision.

This changed, however; and now the three years of required experience to provide supervision can occur only after full licensing as an LSSP. The other major change removed the one-year post-licensure requirement for supervised practice, which as noted before was a compromise. As the LSSP has become more closely identified with education and the national model for school-based practice, current rules reflect that due process hearings are not the

practice of forensic psychology; that informed consent, if obtained under federal laws, is necessary and sufficient for services, thus, alleviating the need for additional consent to perform school psychological duties; and that the LSSP is restricted to practice in the public schools. All TSBEP Rules can be accessed through the website [www.tsbep.texas.gov](http://www.tsbep.texas.gov).

**Table 2** *LSSP Rule Changes from 2006-Present*

<b>Board Rule §463.9 LICENSED SPECIALIST IN SCHOOL PSYCHOLOGY</b>		
<b>Date</b>	<b>Status</b>	<b>Significant Additions or Amendments</b>
June 2008	Amended	<p><b>(g) Provision of psychological services in the public schools by unlicensed individuals.</b></p> <ul style="list-style-type: none"> <li>This new section was added to clarify that individuals who had completed an internship and were no longer enrolled in a formal program of study could not legally provide psychological services in the public schools of Texas until they had (1) passed the National School Psychology Exam, (2) applied for licensure with the Board, <b>and</b> (3) been issued a LSSP trainee status letter from the Board.</li> </ul>
February 2009	Amended	<p><b>(g) Provision of psychological services in the public schools.</b></p> <ul style="list-style-type: none"> <li>Added language to clarify that unlicensed individuals could provide psychological services in schools if they were under supervision and if they (1) were enrolled in an internship, practicum or other site based training in a school psychology program in a regionally accredited institution of higher education, <b>or</b> (2) had completed an internship in a school psychology program in a regionally accredited institution of higher education, <b>and</b> had an application for licensure pending before the Board, <b>and</b> had not been notified by the Board that his/her training requirements had not been met, <b>or</b> (3) had been issued a trainee status letter by the Board.</li> </ul> <p><b>(h) Completed internship.</b></p> <ul style="list-style-type: none"> <li>Clarified that once an internship was completed and the National School Psychology Exam had been taken and passed, a licensee must then apply to the Board for the LSSP license. After the Board</li> </ul>



Table 2 continued

		reviewed and approved the application, a trainee status letter would be issued to the applicant allowing him/her to practice as a trainee.
April 2016	Repealed & Re-Adopted	<p><b>The Board repealed the previous version of the rule and replaced it with a newly adopted rule which can be accessed at the TSBEP website noted previously. A summary of key provisions in the new rule is as follows:</b></p> <p><b>Licensure requirements:</b></p> <ul style="list-style-type: none"> <li>• Application and required documentation – includes university transcripts, proof that the applicant obtained the NCSP from the National School Psychologists’ Certification Board, proof of internship, score on the National School Psychology Exam, and reference letters.</li> <li>• Training qualifications – includes evidence that an applicant either holds the NCSP or graduated from a training program approved by the NASP or accredited in School Psychology by the APA. If not, then the applicant must submit evidence that the required graduate level coursework (as specified in the rule) was completed.</li> <li>• Completion of internship – includes specific criteria related to the number of hours and where hours can be obtained, requirements for enrollment in a formal course of study at a regionally accredited institution of higher education, supervision requirements, maximum length of the internship, and areas of required experience.</li> <li>• Additional requirements – includes requirements that are specified in the Psychologists’ Licensing Act.</li> <li>• Examination requirements – includes criteria for taking and passing the National School Psychology Exam and the Board’s Jurisprudence Exam.</li> <li>• Trainee status – specifies how an applicant becomes designated as a trainee and how long the applicant is allowed to practice under supervision in the public schools.</li> <li>• Provision of school psychological services by unlicensed individuals – provides criteria under which unlicensed individuals are allowed to practice in public schools.</li> </ul>
<b>Board Rule §465.38 PSYCHOLOGICAL SERVICES IN THE SCHOOLS</b>		
<b>Date</b>	<b>Status</b>	<b>Significant Amendments or Additions</b>
February 2008	Amended	<b>Amended the title of the rule (formerly <i>Psychological Services in the Schools</i>, now</b>

Table 2 continued

		<p><i>Psychological Services for Public Schools.)</i></p> <p><b>(5) Supervisor Qualifications.</b></p> <ul style="list-style-type: none"> <li>Removed outdated language stating that a qualified supervisor included an individual who obtained licensure by grandparenting.</li> <li>Added language that a licensee could be qualified to supervise in Texas public schools if he/she had obtained the necessary experience in this or another state and could document that experience from the authority that regulated the provision of school psychological services in that state. He/She also would be required to provide documented proof from the public school where he/she provided such services.</li> </ul>
November 2010	Amended	<p><b>(4)(A)(iii) Supervision period.</b></p> <ul style="list-style-type: none"> <li>Added language that allowed for waiver of the supervision requirement for individuals who legally provided full-time, unsupervised school psychological services in another state for a minimum of three academic years before applying for licensure in Texas. Qualified individuals requesting this waiver must hold the Nationally Certified School Psychologist (NCSP) credential or have graduated from a training program approved by the National Association of School Psychologists (NASP) or accredited in school psychology by the American Psychological Association (APA.)</li> </ul>
June 2012	Amended	<p><b>(2) Titles.</b></p> <ul style="list-style-type: none"> <li>Added language stating that a LSSP who had achieved the NCSP could use this credential along with the license title of LSSP.</li> </ul>
September 2012	Amended	<p><b>(3) Providers of School Psychological Services.</b></p> <ul style="list-style-type: none"> <li>Added language stating that persons who were seeking to fulfill licensing requirements for the LPA, PLP, and LP could also provide school psychological services while seeking these licenses.</li> </ul>
March 2013	Amended	<p><b>(1)(A) Definition.</b></p> <ul style="list-style-type: none"> <li>Clarified that a person using the LSSP could not provide psychological services in any context or capacity outside of his/her contract or employment with public schools.</li> </ul>
April 2016	Repealed & Re-Adopted	<p><b>(a)(1) Scope of Practice.</b></p> <ul style="list-style-type: none"> <li>Expanded the list of activities performed by LSSPs to include behavioral assessments and the designing and implementing of behavioral</li> </ul>

**Table 2 continued**

		<p>interventions and supports.</p> <p><b>(g) Informed consent</b></p> <ul style="list-style-type: none"> <li>Added language to clarify that LSSPs must follow the rules and requirements of IDEIA and the U.S. Department of Education for obtaining informed parental consent in delivering school psychological services in public schools. In doing so, LSSPs will have met the Board’s requirements for informed consent and no additional consent is necessary.</li> </ul>
<b>Board Rule §465.21 TERMINATION OF SERVICES</b>		
<b>Date</b>	<b>Status</b>	<b>Significant Additions or Amendments</b>
September 2013	Amended	<p><b>(g)(1)-(2) Termination of employment with public schools.</b></p> <ul style="list-style-type: none"> <li>Requires LSSPs to provide written resignation to a public school no later than the 45<sup>th</sup> day before the first day of instruction of the following school year.</li> <li>Specifies that LSSPs who are under an employment contract with a public school can terminate the contract at any time if there is just cause, or if the school agrees by providing written consent to the LSSP.</li> </ul>
<b>Board Rule §465.1 DEFINITIONS</b>		
<b>Date</b>	<b>Status</b>	<b>Significant Additions or Amendments</b>
April 2016	Amended	<p><b>(3) Forensic psychological services</b></p> <ul style="list-style-type: none"> <li>Language was added to clarify that forensic psychological services does not include evaluations, proceedings, or hearings under the Individuals with Disabilities Education Improvement Act (IDEIA).</li> </ul>
<b>Board Rule §463.11 LICENSED PSYCHOLOGIST</b>		
<b>Date</b>	<b>Status</b>	<b>Significant Additions or Amendments</b>
January 2015	Amended	<p><b>(c)(1)(N) Pertaining to the use of the LSSP during a doctoral internship</b></p> <ul style="list-style-type: none"> <li>Language was added to allow individuals holding the LSSP to use the LSSP title during a doctoral internship, so long as the supervised experience was taking place within the public schools and those receiving psychological services were clearly informed that the LSSP was under the supervision of an LP, who also is an LSSP.</li> </ul>
<b>Board Rule §465.2 SUPERVISION</b>		

Table 2 continued

Date	Status	Significant Additions or Amendments
April 2016	Repealed & Re-Adopted	<p><b>This rule consolidates the requirements of supervision for all licensees into one rule with specific requirements for each category of licensees contained in separate sections.</b></p> <p><b>(a) Supervision in General</b> (applies to all supervisory relationships.)</p> <ul style="list-style-type: none"> <li>Contains language from the original rule, but includes additional requirements for supervision through remote or electronic means; specifies that this method should only be used if there are difficulties in providing full-time supervision in person and doing so would place an undue burden on the delivery of psychological services; states that no more than fifty percent of the supervision can take place through this method.</li> </ul> <p><b>(d) Supervision of LSSP interns and trainees</b> (<i>applies to all supervisory relationships involving LSSPs, including interns and trainees.</i>)</p> <ul style="list-style-type: none"> <li>Establishes the amount of supervision required, the qualifications of LSSP supervisors and the requirements for signing of educational documents (i.e., student evaluation reports or similar reports provided to consumer, other professionals or other audiences, and progress reports for which the supervisee is providing psychological or counseling services.)</li> </ul>
		<ul style="list-style-type: none"> <li>Clarifies that it is not a violation of this rule if the supervisor does not sign documents related to committee deliberations of an educational meeting which the supervisee attended and participated, unless the supervisor also attended and participated in the meeting.</li> <li>Specifies what must be documented in all supervision sessions and what the supervisor is responsible for (i.e., contracts or service agreements between the public school district, the university training program, and the supervisee; professional liability insurance coverage of the supervisee; training logs required by the school psychology training program; and the supervisee’s trainee or licensure status.)</li> <li>Requires that individuals completing any portion of an internship under Board rules are provided a written agreement of expectations, duties, and responsibilities of each party, including the total hours to be performed by the intern, benefits and</li> </ul>

**Table 2 continued**

		<p>supports provided to the supervisee, and the process for evaluating and supervising the intern.</p> <ul style="list-style-type: none"> <li>• Requires supervisors to ensure that supervisees are provided with a process for addressing serious concerns regarding the supervisee’s performance.</li> </ul> <p><b>(e) Various parts of the rule.</b> Clarifies that specific provisions of the rule will take precedence over general provisions whenever there is a conflict between the two.</p>
August 2016	Amended	<p><b>(d)(3) Supervisors must sign educational documents.</b></p> <ul style="list-style-type: none"> <li>• Removes the requirement for supervisors to sign progress reports for which the supervisee is providing psychological or counseling services.</li> </ul>

**Future Directions**

Since its inception in 1996, the number of LSSPs has grown steadily. At the close of the 2015 fiscal year, there were 3,350 LSSPs. This number constitutes 35% of all TSBEP licensees (Licensed Psychologist = 4826, Provisionally Licensed Psychologist = 231, Licensed Psychological Associate = 1105).

There is no doubt that the LSSP has come a long way since its initial creation and that the TSBEP has come to recognize and more fully appreciate what was written in 1996 – that the practice of school psychology is unique. The many amendments and new rules over the past 20 years have shown a positive response to specific school-based issues and a forward thinking approach to not only maintain, but to refine and develop this license and associated practice.

In 2013, HB 646 amended Section 501.260 of the Occupations Code to add a provision to the Psychologist’s Licensing Act. This provision requires a member of the TSBEP Board to hold the LSSP license. While this is a major accomplishment, it is incomplete since the holder of the LSSP may also hold another credential. Future efforts should include advocacy for a TSBEP member for whom the LSSP is the only license required.

The Texas Education Code that created the LSSP required the LSSP to be modeled after national standards. TASP continues to support the TSBEP decision to use the NASP standards as the model for the LSSP. However, the title issue remains. Current TSBEP rules and regulations are not entirely consistent with current standards. The NASP standards use the title “school psychologist” for those professionals who meet the NASP standards. Thus, the title of LSSP could become SP (School Psychologist). Because of the history of controversy and compromise between interested groups in Texas, advocacy efforts might be more effective in focusing on school psychology practice, not title.

The LSSPs in this state reflect highly qualified professionals who have collectively and individually made substantial improvements to the education and mental health of students in Texas public schools. We are proud of our accomplishments and know that the only way to continue promoting the license and practice is to have more involvement in both state and national associations. LSSPs should be extremely proud of their profession and get involved. There is a great deal more to be done in the next 20 years.

### References

19 TAC Section 141.26

House Bill 646, 2013, 83<sup>rd</sup> Texas State Legislature.

Miller, D. C. (2008a). The Texas licensing and certification experience: Separating fact from fiction, an eyewitness account: Part 1. *The Texas School Psychologist*, 18(3), 12-13.

Miller, D. C. (2008b). The Texas licensing and certification experience: Separating fact from fiction, an eyewitness account: Part 2. *The Texas School Psychologist*, 18(4), 4-6.

Miller, D. C. (2009a). The Texas licensing and certification experience: Separating fact from fiction, an eyewitness account: Part 3. *The Texas School Psychologist*, 19(1), 4-6.

Miller, D. C. (2009b). The Texas licensing and certification experience: Separating fact from fiction, an eyewitness account: Part 4. *The Texas School Psychologist*, 19(1), 4-6.

Senate Bill 1, 1995, 74<sup>th</sup> Texas Legislature.

Texas Register, Volume 16, Number 88, November, 1991

Texas State Board of Examiners of Psychologists – all rules accessed at [www.tsbep.texas.gov](http://www.tsbep.texas.gov)

Article

---

# Review of the Pattern of Strengths and Weaknesses Approach in Specific Learning Disability Identification

Laura Phipps  
*Baylor University*

A. Alexander Beaujean  
*Baylor University*

The 2004 revision of the Individuals with Disabilities Education Act allowed states to have flexibility in determining the criteria they used to determine if a student had a specific learning disability (SLD). Currently, the three major approaches to identifying a SLD are ability-achievement discrepancy, response to intervention, and finding a pattern of strengths and weaknesses (PSW). School psychologists are relatively familiar with the first two approaches, but the PSW approach is much newer. Consequently, there is currently variability in school psychologists' understanding of the PSW approach. The purpose of this article is to explain the PSW approach to identifying SLDs and to provide a review of the literature on the pros and cons of its usage.

**Keywords:** Specific Learning Disability, Pattern of Strengths and Weaknesses

The diagnosis of a specific learning disability (SLD) is a complex endeavor (Fletcher, Francis, Morris, & Lyon, 2005). Although it is the most common form of disability seen in schools (National Center for Education Statistics, 2015), SLD is a heterogeneous condition whose manifestation varies across individuals (National Association of School Psychologists, 2011). Thus, the criteria for identifying a SLD are broad and the process schools use to identify a SLD varies greatly both across and within states (Zirkel & Thomas, 2010). Moreover, training in identifying a SLD varies greatly between school psychology programs, with many relegating learning about this area to practicum instead of having explicit instruction in identification (Reynolds, Fisher, & Morlock, 2014). Most scholars agree that the central component of a SLD is that of underachievement, indicating that individuals are not attaining expected levels of academic achievement (Taylor, 2014). Measurement of this underachievement, however, is an area where there are many disagreements

(Fletcher, Denton, & Francis, 2005). The 2004 revision of the Individuals with Disabilities Education Act (IDEA, 2011), and subsequent regulations, noted this disagreement by allowing school-based professionals freedom to approach the SLD identification process using different methods.

Currently, there are three major approaches that states allow to identify a SLD: ability-achievement discrepancy, response to intervention, and examination of the pattern of cognitive strengths and weaknesses within an individual (Lichtenstein, 2014; Maki, Floyd, & Roberson, 2015). Of the three approaches, the ability-achievement discrepancy method has the longest history in SLD identification (Hallahan & Mercer, 2002). It requires the comparison of scores from a

---

**Author Note:** Correspondence concerning this article should be addressed to A. Alexander Beaujean, Department of Educational Psychology, One Bear Place #97301, Waco, TX 76798-7301. Phone: 254.710.1548. Fax: 254.710.3265. Email: Alex\_Beaujean@baylor.edu

standardized intellectual ability measure and standardized measure of academic achievement. If there is at least one academic achievement score that is well below the same student's intellectual ability score, then the student may have a SLD. While this approach to diagnosing a SLD has traditionally been very common, research on it has shown that it is not very accurate in distinguishing between those who have a SLD and those who have academic difficulties but do not have a SLD (Stuebing, Fletcher, LeDoux, Lyon, Shaywitz, & Shaywitz, 2002). Moreover, this approach tends to overlook individuals who have academic difficulties but do not manifest a large enough discrepancy between their intellectual ability and academic achievement scores (Dombrowski, Kamphaus, & Reynolds, 2004; Francis, Fletcher, Stuebing, Lyon, Shaywitz, & Shaywitz, 2005).

In the late 1990s/early 2000s, an alternative to using ability-achievement discrepancies was proposed that focused on measuring response to quality instruction as a key component to identifying a SLD (Gresham, 2002; President's Commission on Excellence in Special Education, 2002). This response to intervention (RTI) framework goes beyond just SLD identification, as it is a general approach to address student difficulties within a multi-tiered system that involves early identification of problems and using outcome data to make decisions about the effectiveness of a structured intervention (Kratochwill, Clements, & Kalymon, 2007). When specifically applied to identifying a SLD, the RTI approach often involves using a dual discrepancy model (Hauerwas, Brown, & Scott, 2013; for alternative models, see Fuchs, 2003). The two elements of the dual discrepancy criterion are that (a) the student exhibits a level of academic performance that is below expectations for the instructional setting, and (b) the student's rate of learning is behind that of peers, despite appropriate general education instruction and interventions (Fuchs, Fuchs, & Speece, 2002). A key aspect in the RTI approach is repeated measurement of academic performance using curriculum-based measures.

While RTI is a very popular method for SLD identification, critics have argued that neither the general RTI process nor its implementation via curriculum-based measures was designed to make

specific diagnoses. For example, Reynolds and Shaywitz (2009a,b) argued against using RTI to make diagnostic decisions because there is no widely accepted criterion for identifying how inadequate a response to intervention needs to be for a SLD diagnosis. Moreover, they wrote that the operationalization of a SLD in RTI fundamentally alters the concept of SLD from being endogenous to an individual to being contextual (i.e., an inadequate response to appropriate instructional methods). Consequently, critics argue that while the RTI approach may be a way to prevent or remediate early academic difficulties for some students, an accurate SLD diagnosis requires comprehensive psychological evaluation (Kavale, Holdnack, & Mostert, 2006).

Given the noted problems with the ability-achievement discrepancy and RTI models, some have argued for a "third approach" to identify SLDs. This proposed method examines if the pattern of strengths and weaknesses (PSW) in an individual's cognitive processing are consistent with any academic deficits, which PSW advocates argue is the only way that "makes the most empirical and clinical sense" to identify a SLD (Hale et al., 2010, p. 225). There is not a single PSW approach; instead, there are a variety of ways to implement this method. We discuss these in more detail in the *Patterns of Strengths and Weaknesses Approach* section.

As with the RTI model, the PSW approach is currently growing in popularity, and has even made its way into some recent standardized test manuals (e.g., Wechsler, 2014). Although some states currently allow the use of PSW for SLD diagnosis, states are not consistent in their regulations or guidance on how to implement it (e.g., as a stand alone method, in conjunction with RTI) or on what constitutes a PSW (Maki et al., 2015).

For example, Texas allows for the use of PSW to identify a SLD (Texas Adaptations for Special Education, 2015). The Texas Education Agency (TEA, n.d.) interpreted the PSW approach as the evaluation of "specific areas of cognitive function, academic achievement, or both and comparing those results against each other or in contrast to other measures of student performance" (p. 2) and recommend that the PSW method be used in conjunction with curriculum-based or criterion-



referenced measures. TEA's definition is purposefully broad as it includes both the traditional ability-achievement discrepancy method as well as the more current method of examining cognitive profiles.

Like Texas, Oregon also allows for using the PSW method to identify a SLD (Oregon Administrative Rules for Special Education, 2011). Oregon's interpretation of PSW, however, is more detailed and specific. In the Oregon School Psychologist Association's technical assistance paper to the Oregon Department of Education, they described the PSW approach in much detail and noted the PSW method has no need for a composite IQ score except in identifying intellectual disability (Hanson, Sharman, & Esparza-Brown, 2008).

Given the differing definitions and manifestations of the PSW approach to SLD identification, there is a need for school psychologists to understand what exactly constitutes a PSW. Moreover, there is also a need for school psychologists to have some acquaintance with the literature that supports and criticizes the approach. The current article aims to do both. Hopefully, this will aid practitioners in developing an informed opinion on the PSW approach and aid them as they advocate for the use of best practices in their schools.

### **Pattern of Strengths and Weaknesses Approach**

#### **Definition**

The overarching goal of the PSW method is to identify underlying cognitive processing deficits that directly relate to the SLD. The PSW method assumes an individual with a SLD has certain cognitive processing deficits that are not only causally and predictably linked to their learning difficulties, but that these deficits are relevant to educational planning (Hale, Kaufman, Naglieri, & Kavale, 2006). Consequently, the fundamental components used to diagnose a SLD from a PSW perspective involve finding data that show academic deficits (typically measured by both curriculum-based and standardized tests) are related to a discrepancy between strengths and weaknesses in the student's cognitive processing as measured by standardized tests of cognitive ability.

#### **Determining a Pattern of Strengths and Weaknesses**

There is not a single PSW approach. Instead, there are three common PSW operationalizations: (a) consistency-discrepancy model (Naglieri, 2011), (b) the concordance-disconcordance or cognitive hypothesis testing model (Hale & Fiorello, 2004), and (c) cross-battery or ability-achievement consistency model (Flanagan, Alfonso, & Mascolo, 2011). While the three major PSW methods have their own idiosyncrasies, they also have many similarities. Specifically, all three PSW operationalizations are characterized by the following features: (a) collect data from multiple sources across multiple time points using a variety of assessment tools and strategies, (b) analyze the data to look for patterns, (c) rely on predictive and treatment validity literature, and (d) use logical and empirical evidence to guide decision making (Flanagan, Fiorello, & Ortiz, 2010; Schultz, Simpson, & Lynch, 2012).

All three PSW approaches require the collection and integration of at least three sources of data in the area of the suspected SLD: (a) informal information, (b) non-standardized test scores, and (c) standardized test scores. Informal information includes things such as archival records, classroom observations, classroom work samples, and parent and teacher information. Non-standardized tests typically come from criterion-referenced assessments designed to assess academic functioning, such as benchmark test scores on grade-level curriculum, curriculum-based measures, current classroom records, and criterion-referenced tests. Standardized test scores typically come from norm-referenced, individually administered measures of academic achievement and cognitive ability. Usually, only non-global index scores from these tests are used (Fiorello, Hale, Holdnack, Kavanagh, Terrell, & Long, 2007).

Once the data are collected, they need to be analyzed to determine if there are patterns in the data that indicate a SLD is present. This requires multiple steps (Flanagan, Alfonso, & Mascolo, 2011; Schultz et al., 2012). The first step is to examine the trustworthiness of the data, which involves a judgment of how much confidence should be placed in the data given the source of the information and the process involved in obtaining it.

The next step is *triangulation*, which involves using data from multiple sources (i.e., informal information, non-standardized test scores, standardized test scores) to determine that the student has academic difficulty. Most of the information should converge in showing the student has difficulties learning in a specific area. If multiple sources of information show a learning difficulty, the next step is to examine if there are reasons other than a SLD that might be causing the difficulty (i.e., exclusionary factors) such as English language proficiency, educational opportunity, cultural differences, or other disabilities (e.g., visual, hearing, motor, intellectual, emotional).

If a student has an academic difficulty and it is not primarily due to exclusionary factors, then the last step is to determine if there is a pattern in the data consistent with a SLD. This involves

examining each piece of academic achievement and cognitive ability information and determining if it is a strength, weakness, or neither. This is a difficult task as there are no universal guidelines for determining a strength or weakness. For quantitative data, PSW advocates suggest making the decisions based on frequency of occurrence in the comparison sample (Flanagan, Alfonso, & Mascolo, 2011; Hale, Wycoff, & Fiorello, 2011), although some argue that relative strengths and weaknesses should also be considered (Naglieri, 2011). Some examples are given in the top section of Table 1. For qualitative data, it is primarily up to the diagnostic team to determine the criteria for whether the information constitutes a strength or weakness. Some example criteria are given in the bottom section of Table 1.

**Table 1.** Example Criteria for Determining Strengths and Weaknesses.

Data Type	Strength	Weakness
Quantitative Data		
Standardized test score (norm- or criterion-referenced)	≥ 84 %ile	≤ 16th %ile
State achievement tests	Meets or exceeds standard	Does not meet standard
Curriculum-Based Measurement Benchmarks	At <i>benchmark</i> level or higher	At <i>at-risk</i> level or lower
Progress monitoring	Consistently meeting or exceeding the aimline	Consistently falling below aimline
Grades	A or B	D or F
Qualitative Data		
Teacher Reports	Teacher judges student to be at or above the level expected for grade	Teacher judges student to be below the level expected for grade
Academic Observations	Demonstrates understanding of academic content at a level typical to, or greater than, other students in the classroom	Demonstrates understanding of academic content well below other students in the classroom
Records Review	Documented history of strengths in academic content	Documented history of weaknesses in academic content

After classifying the collected data, strength and weakness classifications need to be synthesized to determine if there are patterns. This involves examining if:

- A. academic weaknesses in the referral areas are found in other sources of data;
- B. there is a connection in the literature between the areas of academic weaknesses and the pattern of strengths and weaknesses in the cognitive assessment scores;
- C. there are data that would provide a different explanation of the academic weakness other than SLD; and
- D. other information is needed before making a diagnosis.

Only if the answers to areas A and B are *yes* and the answers to areas C and D are *no* is SLD a viable diagnosis, although there is room for some clinical discretion in these decisions.

### **Integrating the PSW approach with RTI**

Most PSW advocates do not advise that using a PSW approach should completely replace RTI in identifying a SLD. In fact, what Hale et al. (2006) meant when they called the PSW method a "third approach," was that it incorporated the best parts of the ability-achievement discrepancy and RTI approaches. Consequently, PSW advocates often argue that examining strengths and weaknesses should be part of the RTI framework and some have even suggested that the PSW process can be thought of as Tier IV in RTI (Flanagan et al, 2010). That is, integrating cognitive assessment with RTI should provide a more comprehensive understanding of why some individuals have academic difficulties than either approach alone. Moreover, by integrating both approaches school psychologists can potentially avoid conducting time-intensive psychological evaluations with students who can adequately respond to Tier I or Tier II interventions (Hale et al., 2006).

## **Support and Criticisms of the PSW Approach**

### **Support for the PSW Approach**

Proponents of the PSW approach argue that it has a key component lacking in other SLD identification methods. Specifically, the PSW approach appropriately measures cognitive processing weaknesses (Hale & Fiorello, 2004). To support this, Hale et al. (2010) surveyed learning disability experts and reported that there was a consensus among them that measuring different aspects of cognitive ability is required to make a SLD diagnosis. Such measures are required because having weaknesses in some area of cognitive ability in the presence of otherwise typical cognitive functioning is a core feature of a SLD and is what diagnostically differentiates it from global cognitive deficits. Further, Hale et al. argued that if a pattern of cognitive strengths and weaknesses exists within an individual, it could be used to develop individualized interventions instead of only relying on learning techniques that work with the general population.

Those who advocate for a PSW approach typically cite two types of research for support. The first type of research is based on John Carroll's (1993) seminal meta-analysis showing that the general structure of human cognitive abilities is comprised of a general factor as well as multiple non-global factors (although this model is often misinterpreted, Beaujean, 2015). The second type of research is comprised of studies showing links between non-global cognitive ability factors and academic achievement (for a review, see Flanagan et al., 2011). Typically, these second type of studies involve administering one or more standardized measures of cognitive ability and one or more standardized measures of academic achievement to a large group of individuals. Then, using either multiple regression or structural equation models, the authors examine what non-global components of the cognitive ability measures are related to the different areas of academic achievement.

There have been a few empirical articles published that ostensibly support the PSW approach, but the methods employed are questionable. For example, Fiorello, Flanagan, and Hale (2014) edited a special issue of the journal *Learning Disabilities:*

A *Multidisciplinary Journal* that was devoted to studies supporting the PSW approach to SLD identification. All the articles' authors claimed to show that a PSW approach was able to differentiate separate types of SLD, but their data analysis and reporting are not aligned with best practice so their conclusions may be premature.

As an example, Feifer, Nader, Flanagan, Fitzer, and Hicks (2014) and Kubas, Schmid, Drefs, Poole, Holland, and Fiorello (2014) both used cognitive ability instrument subtests to predict scores on academic achievement subtests. In both studies, they wanted to see if cognitive ability differentially predicted academic achievement for students who had different PSW profiles. Typically, this question is best answered through comparing a hierarchical sequence of regression models with and without interaction/moderation effects (Aiken & West, 1991). Feifer et al. and Kubas et al., however, chose to fit separate regression models in each of the different PSW groups and then compare the standardized coefficients--approaches that methodologists have long argued should not be used (Judd, McClelland, & Culhane, 1995). Moreover, they did not address any of the assumptions that accompany regression models, the sample sizes for their PSW groups were often small (across both studies the group sizes ranged from 21-123, with the average size being 58), and they did not report any standard errors or confidence intervals for their statistics, which is contrary to many different reporting guidelines (American Psychological Association, 2009; Kelley & Maxwell, 2010; Wilkinson, et al., 1999). Consequently, it is impossible to determine if the studies' results actually support the PSW approach being a robust method for identifying SLD.

### **Criticisms of the PSW Approach**

Despite the growing popularity of the PSW approach to identifying SLD--especially when combined with RTI--there have been some criticisms of the approach. In what follows, we provide an overview of some of the most common criticisms.

#### **Profiles of cognitive ability are unstable.**

Perhaps the most common criticism of the PSW approach is that it requires the creation of cognitive profiles (Fletcher, et al., 2005). Developing a profile

from patterns of scores on standardized tests has a long history in psychology (Watkins, 2000), and previous research has shown it to have two major problems: stability and utility.

The stability problem is that an individual's profile of strengths and weaknesses tends to differ both across time and across instruments (Miciak, Taylor, Denton, & Fletcher, 2015). Watkins and Smith (2013) compared Wechsler Intelligence Scale for Children-IV scores from students who took the test two times, approximately 3 years apart. A sizable portion of the sample (29%-44%) had differences in the index scores of 10 points or more, with some scores differing by as much as 41 points. One reason for instability is likely due to different individuals administering the tests (McDermott, Watkins, & Rhoad, 2014). Another reason for instability is likely that the scores used in the PSW determination come from non-global composites, which have lower reliability estimates than global scores (Sandip, Gautam, & Shelby, 2010). The reliability issue is exacerbated within some PSW approaches (e.g., consistency-discrepancy model) because they require an examination of score differences. If scores from two tests are positively correlated, which is generally the case for measures of cognitive ability and academic achievement, then their difference will be less reliable than their constituent scores (Rogosa, & Willett, 1983). This introduces even more instability in the profiles.

The utility problem is that it is difficult to make diagnostic decisions, especially SLD, from profiles of test scores (Canivez & Watkins, in press).<sup>1</sup> Part of this is due to the nature of SLD. Lichtenstein (2014) argued that learning disabilities are dimensional (i.e., fall on a continuum) rather than being categorically distinct entities. Thus, it is difficult to distinguish between a weakness, a difficulty, and a disorder/disability. Another part of the utility problem is that non-global scores on tests of cognitive ability are not pure measures of their underlying constructs (Canivez, in press; Floyd, Reynolds, Farmer, & Kranzler, 2013). Instead, the scores are influenced by general intelligence, random error, and the

---

<sup>1</sup> Test utility is the information a test's scores provide to make accurate diagnostic or placement decisions (Wasserman & Bracken, 2003).

unique non-global component that is not shared with any other score (Carroll, 1995). Moreover, most of the variance in the non-global scores is accounted for by general intelligence, not the unique non-global component (e.g., Gignac & Watkins, 2013). Thus, Kranzler and Floyd (2013) go so far as to say that little can be gained by examining subtest or composite score differences on intelligence tests when making diagnostic decisions.

Recently, there has been some empirical investigation documenting the limited utility of the PSW approach. Stuebing, Fletcher, Branum-Martin, and Francis (2012) conducted a simulation study examining the three major PSW operationalizations. They found that all three methods were good at identifying observations that did not have a SLD (i.e., high specificity and negative predictive values), but were not good at identifying observations that had a SLD. Specifically, there were large differences in sensitivity (i.e., classifying a SLD observation as having SLD) across PSW methods and SLD base-rates, and the positive predictive values (i.e., probability that students diagnosed with a SLD actually have the disability) were all low to moderate. The combination of high specificity and low positive predictive values resulted in many false positives. Thus, many observations classified as having a SLD did not actually have the disability. Similarly, Miciak et al. (2014) examined the utility of the concordance-disconcordance and the cross-battery PSW methods on a sample of students who were inadequate responders to RTI Tier 2 interventions. Not only did the two PSW methods result in very different numbers of students diagnosed with a SLD (17.3% vs. 47.5%), but also there were no robust differences between groups that met and did not meet criteria according to the two PSW methods. Consequently, the authors questioned the idea that these two PSW operationalizations are interchangeable as well as if either operationalization had sufficient utility to be used in making SLD diagnostic decisions.

**Cognitive profiles do not provide information about appropriate interventions.** PSW proponents advocate that knowing an individual's cognitive strengths and weaknesses can help with planning interventions, but others disagree. For example, Fletcher (2012) argued that the literature showing relationships among different

aspects of cognitive ability and academic achievement is inconsistent, and that there is little evidence that instruction addressing strengths and weaknesses in cognitive skills is related to intervention outcomes. Moreover, two recent meta-analyses on specific cognitive skills training have indicated that it does not appear to result in improved academic achievement for most students (Kearns & Fuchs, 2013; Melby-Lurvag & Hume, 2013).

**There is not a consensus from either the federal government or SLD experts for the use of the PSW approach.** Although Hale et al. (2010) wrote that the PSW approach is consistent with the language in IDEA, Zirkel (2013) argued that this interpretation of IDEA is wrong. While technically it is allowed as part of the "other alternative research-based procedures," the pattern of strengths and weakness language in IDEA simply refers to SLD identification methods other than RTI. This is buttressed by the US Department of Education's comments on IDEA's 2004 revision.

The Department [of Education] does not believe that an assessment of psychological or cognitive processing should be required in determining whether a child has an SLD. There is no current evidence that such assessments are necessary or sufficient for identifying SLD. Further, in many cases, these assessments have not been used to make appropriate intervention decisions. However, Sec. 300.309(a)(2)(ii) permits, but does not require, consideration of a pattern of strengths or weaknesses, or both, relative to intellectual development, if the evaluation group considers that information relevant to an identification of SLD. In many cases, though, assessments of cognitive processes simply add to the testing burden and do not contribute to interventions. (Assistance to States for the Education of Children With Disabilities and Preschool Grants for Children With Disabilities, 2006, p. 46651)

In a similar fashion, although Hale et al. (2010) wrote that there was a consensus among SLD experts for the use of PSW, the Consortium for



Evidence-Based Early Intervention Practices (2010) wrote that no such consensus exists. Instead,

schools . . . may be distracted, at best, by efforts to ensure specific learning disabilities (SLD) eligibility based on a pattern of cognitive strengths and weaknesses (PSW) without sufficient evidence of efficacy. At worst, we believe that schools could be mandated to invest considerable time and financial resources into implementation of unproven PSW models for SLD identification and intervention planning" (pp. 2-3).

Thus, dissention was echoed by the National Center for Learning Disabilities RTI Action Network (2011), who stated that there is no consensus on the extent that cognitive assessment should be included in a comprehensive SLD evaluation.

### Clinical Implications

Assessing for a SLD is a very complex endeavor. Not only is the disability a complex entity, but there is no consensus on the best method for diagnosis. While the ability-achievement discrepancy used to be the "go to" method, it is no longer the default SLD identification method used by most school personnel. In fact, Maki et al. (2015) reported that only 34 states currently allow for use of this approach, and 10 states prohibit its use altogether. In its place, most states now either require RTI be used—either alone or in conjunction with the discrepancy or the PSW approaches. In either scenario, however, there is considerable variability in state guidelines for their use.

Approximately 25% of states currently allow for PSW to be used (Maki et al., 2015). Unfortunately, the guidance for its implementation is minimal (e.g., Texas Education Agency, n.d.). When this is coupled with the fact that there are three popular PSW methods--which, despite appearing to be similar, do not appear to produce interchangeable results--it likely means that there is significant variability in the SLD identification results between districts that use PSW to identify students with a SLD. This is unfortunate as it goes

against the principles of evidence-based assessment (Hunsley & Mash, 2011).

It appears that more investigation needs to be done directly on the PSW approach before school districts widely adopt it. This idea is supported by the variability in state directions for implementing PSW, the lack of stability and utility in cognitive profiles, the U.S. Department of Education not regarding PSW as being essential to identifying a SLD, and there not being a consensus among SLD experts about its use in the SLD diagnostic process. If a school district has already adopted the PSW approach, then school psychologists should be cognizant of both the strengths and limitations of the PSW approach before using it, and make sure they have collected ample supporting evidence to be able to assure parents or other professionals who might be wary of the diagnostic conclusions.

Due to their strong training in assessment, school psychologists are well positioned to be involved in conversations at the state, district, and school level regarding evidenced-based practices—especially those involving disability identification. We hope that this article can aid in those conversations and the advocacy of the use of best practices.

### References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- American Psychological Association. (2009). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Assistance to States for the Education of Children With Disabilities and Preschool Grants for Children With Disabilities, 71 Fed. Reg. 46540-46845. (August 14, 2006).
- Beaujean, A. A. (2015). John Carroll's views on intelligence: Bi-factor vs. higher- order models. *Journal of Intelligence*, 3, 121-136. doi: 10.3390/jintelligence3040121
- Canivez, G. L. (in press). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In C. DiStefano & K. Schweizer (Eds.), *Principles and methods of test construction: Standards and recent advancements*: Hogrefe.
- Canivez, G. L., & Watkins, M. W. (in press). Review of the Wechsler Intelligence Scale for Children-Fifth Edition: Critique, commentary, and independent analyses. In A. S.

- Kaufman, S. E., Raiford, & D. L. Coalsen (Eds.), *Intelligent testing with the WISC-V*. Hoboken, NJ: Wiley.
- Consortium for Evidence-Based Early Intervention Practices. (2010). *Response to the Learning Disabilities Association of America (LDA) white paper on specific learning disabilities (SLD) identification*. Retrieved from [www.rtinetwork.org/images/content/articles/learn\\_about\\_RTI/LDAResponsefinal.pdf](http://www.rtinetwork.org/images/content/articles/learn_about_RTI/LDAResponsefinal.pdf)
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Carroll, J. B. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research*, *30*, 429-452. doi: 10.1207/s15327906mbr3003\_6
- Dombrowski, S. C., Kamphaus, R. W., & Reynolds, C. R. (2004). After the demise of the discrepancy: Proposed learning disabilities diagnostic criteria. *Professional Psychology: Research and Practice*, *35*, 364-372. doi: 10.1037/0735-7028.35.4.364
- Feifer, S., Nader, R. G., Flanagan, D., Fitzer, K., & Hicks, K. (2014). Identifying specific reading disability subtypes for effective educational remediation. *Learning Disabilities: A Multidisciplinary Journal*, *20*(1). Retrieved from <http://js.sagamorepub.com/ldmj/article/view/5151>
- Fiorello, C. A., Flanagan, D. P., & Hale, J. B. (Eds.). (2014). The utility of the pattern of strengths and weaknesses approach [Special issue]. *Learning Disabilities: A Multidisciplinary Journal*, *20*(1)
- Fiorello, C. A., Hale, J. B., Holdnack, J. A., Kavanagh, J. A., Terrell, J., & Long, L. (2007). Interpreting intelligence test results for children with disabilities: Is global intelligence relevant? *Applied Neuropsychology*, *14*, 2-12. doi: 10.1080/09084280701280338
- Flanagan, D. P., Alfonso, V. C., & Mascolo, J. T. (2011). A CHC-based operational definition of SLD: Integrating multiple data sources and multiple data-gathering methods. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification*, (pp. 233-298). Hoboken, NJ: Wiley.
- Flanagan, D. P., Fiorello, C. A., & Ortiz, S. O. (2010). Enhancing practice through application of Cattell-Horn-Carroll theory and research: A "third method" approach to specific learning disability identification. *Psychology in the Schools*, *47*, 739-760. doi: 10.1002/pits.20501
- Fletcher, J. M. (2012). Classification and identification of learning disabilities. In B. Wong & D. L. Butler (Eds.), *Learning about learning disabilities* (pp. 1-25). Waltham, MA: Academic Press.
- Fletcher, J. M., Denton, C., & Francis, D. J. (2005). Validity of alternative approaches for the identification of learning disabilities: Operationalizing unexpected underachievement. *Journal of Learning Disabilities*, *38*, 545-552. doi: 10.1177/00222194050380061101
- Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *34*, 506-522. doi: 10.1207/s15374424jccp3403\_7
- Floyd, R. G., Reynolds, M. R., Farmer, R. L., & Kranzler, J. H. (2013). Are the general factors from different child and adolescent intelligence tests the same? Results from a five-sample, six-test analysis. *School Psychology Review*, *42*, 383-401.
- Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., & Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities*, *38*, 98-108. doi: 10.1177/00222194050380020101
- Fuchs, L. S. (2003). Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research & Practice*, *18*, 172-186. doi: 10.1111/1540-5826.00073
- Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly*, *25*, 33-45. doi: 10.2307/1511189
- Gignac, G. E., & Watkins, M. W. (2013). Bifactor modeling and the estimation of model-based reliability in the WAIS-IV. *Multivariate Behavioral Research*, *48*, 639-662. doi: 10.1080/00273171.2013.804398
- Gresham, F. M. (2002). Responsiveness to intervention: An alternative approach to the identification of learning disabilities. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 467-519). Mahwah, NJ: Lawrence Erlbaum.
- Hale, J. B., & Fiorello, C. A. (2004). *School neuropsychology: A practitioner's handbook*. New York: Guilford Press.
- Hale, J., Alfonso, V., Berninger, V., Bracken, B., Christo, C., Clark, E., Yalof, J. (2010). Critical issues in response-to-intervention, comprehensive evaluation, and specific learning disabilities identification and intervention: An expert white paper consensus. *Learning Disability Quarterly*, *33*, 223-236. doi: 10.2307/25701450
- Hale, J. B., Kaufman, A., Naglieri, J. A., & Kavale, K. A. (2006). Implementation of IDEA: Integrating response to intervention and cognitive assessment methods. *Psychology in the Schools*, *43*, 753-770. doi: 10.1002/pits.20186
- Hale, J. B., Wycoff, K. L., & Fiorello, C. A. (2011). RTI and cognitive hypothesis testing for identification and intervention of specific learning disabilities: The best of both worlds. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 173-201). Hoboken, NJ: Wiley.
- Hallahan, D. P., & Mercer, C. D. (2002). Learning disabilities: Historical perspectives. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 1-68). Mahwah, NJ: Lawrence Erlbaum.
- Hanson, J., Sharman, L., A., & Esparza-Brown, J. (2008). *Pattern of strengths and weaknesses in specific learning disabilities: What's it all about? Technical assistance paper to the Oregon Department of Education*. Portland, OR: Oregon School Psychologists Association.

- Hauerwas, L. B., Brown, R., & Scott, A. N. (2013). Specific learning disability and response to intervention: State-level guidance. *Exceptional Children, 80*, 101-120. doi: 10.1177/001440291308000105
- Hunsley, J., & Mash, E. J. (2011). Evidence-based assessment. In D. H. Barlow (Ed.), *The Oxford handbook of clinical psychology* (pp. 76-97). New York: Oxford University Press.
- Individuals With Disabilities Education Act, 20 U.S.C. § 1400 (2011).
- Judd, C. M., McClelland, G. H., & Culhane, S. E. (1995). Data analysis: Continuing issues in the everyday analysis of psychological data. *Annual Review of Psychology, 46*, 433-465. doi:10.1146/annurev.ps.46.020195.002245
- Kavale, K. A., Holdnack, J. A., & Mostert, M. P. (2006). Responsiveness to intervention and the identification of specific learning disability: A critique and alternative proposal. *Learning Disability Quarterly, 29*, 113-127. doi: 10.2307/30035539
- Kearns, D. M., & Fuchs, D. (2013). Does cognitively focused instruction improve the academic performance of low-achieving students? *Exceptional Children, 79*, 263-290.
- Kelley, K., & Maxwell, S. E. (2010). Multiple regression. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social and behavioral sciences* (pp. 281-297). New York: Routledge.
- Kranzler, J. H., & Floyd, R. G. (2013). *Assessing intelligence in children and adolescents: A practical guide*. New York: Guilford.
- Kratochwill, T. R., Clements, M. A., & Kalymon, K. M. (2007). Response to intervention: Conceptual and methodological issues in implementation. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention* (pp. 25-52). New York: Springer.
- Kubas, H., Schmid, A., Drefs, M., Poole, J., Holland, S., & Fiorello, C. A. (2014). Cognitive and academic profiles associated with math disability subtypes. *Learning Disabilities: A Multidisciplinary Journal, 20*(1), 31-44. Retrieved from <http://js.sagamorepub.com/ldmj/article/view/5152>
- Lichtenstein, R. (2014). Best practices in identification of learning disabilities. In P. L. Harrison & A. Thomas (Eds.), *Best practices in school psychology: Foundations* (pp. 331-354). Bethesda, MD: National Association of School Psychologists.
- Maki, K. E., Floyd, R. G., & Roberson, T. (2015). State learning disability eligibility criteria: A comprehensive review. *School Psychology Quarterly, 30*, 457-469. doi: 10.1037/spq0000109
- McDermott, P. A., Watkins, M. W., & Rhoad, A. M. (2014). Whose IQ is it? Assessor bias variance in high-stakes psychological assessment. *Psychological Assessment, 26* (1), 207-214. DOI: 10.1037/a0034832
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*, 270-291. doi: 10.1037/a0028228
- Miciak, J., Fletcher, J. M., Stuebing, K. K., Vaughn, S., & Tolar, T. D. (2014). Patterns of cognitive strengths and weaknesses: Identification rates, agreement, and validity for learning disabilities identification. *School Psychology Quarterly, 29*, 27-37. doi: 10.1037/spq0000037
- Miciak, J., Taylor, W. P., Denton, C. A., & Fletcher, J. M. (2015). The effect of achievement test selection on identification of learning disabilities within a patterns of strengths and weaknesses framework. *School Psychology Quarterly, 30*, 321-334. doi: 10.1037/spq0000091
- Naglieri, J. A. (2011). The discrepancy/consistency approach to SLD identification using the PASS theory. In D. P. Flanagan & V. C. Alfonso (Eds.), *Essentials of specific learning disability identification* (pp. 145-172). Hoboken, NJ: Wiley.
- National Association of School Psychologists. (2011). *Identification of students with specific learning disabilities*. Bethesda, MD: Author.
- National Center for Education Statistics. (2015). Fast facts: How many students with disabilities receive services? Retrieved from <https://nces.ed.gov/fastfacts/display.asp?id=64>
- National Center for Learning Disabilities RTI Action Network. (2011). *Position statement on determination of specific learning disabilities*. Retrieved from <http://www.rtinetwork.org/about-us/position-statement-on-determination-of-specific-learning-disabilities>
- Oregon Administrative Rules for Special Education, Oregon Administrative Rules §581-015-2170 (2011)
- President's Commission on Excellence in Special Education. (2002). *A new era: Revitalizing special education for children and their families*. Washington, DC: U.S. Department of Education.
- Reynolds, C. R., & Shaywitz, S. E. (2009a). Response to intervention: Prevention and remediation, perhaps. Diagnosis, no. *Child Development Perspectives, 3*, 44-47. doi: 10.1111/j.1750-8606.2008.00075.x
- Reynolds, C. R., & Shaywitz, S. E. (2009b). Response to Intervention: Ready or not? Or, from wait-to-fail to watch-them-fail. *School Psychology Quarterly, 24*, 130-145. doi: 10.1037/a0016158
- Reynolds, J. L., Fisher, S. D., & Morlock, L. A. (2014). School psychologist training in specific learning disability identification. *Trainer's Forum: Journal of the Trainers of School Psychologists, 33*(1), 50-65.
- Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement, 20*, 335-343. doi: 10.1111/j.1745-3984.1983.tb00211.x
- Sandip, S., Gautam, P., & Shelby, H. (2010). Reporting diagnostic scores in educational testing: Temptations, pitfalls, and some solutions. *Multivariate Behavioral Research, 45*, 553-573.
- Schultz, E. K., Simpson, C. G., & Lynch, S. (2012). Specific learning disability identification: What constitutes a pattern of strengths and weaknesses. *Learning Disabilities: A Multidisciplinary Journal, 18*, 87-97.



- Stuebing, K. K., Fletcher, J. M., Branum-Martin, L., & Francis, D. J. (2012). Evaluation of the technical adequacy of three methods for identifying specific learning disabilities based on cognitive discrepancies. *School Psychology Review, 41*, 3-22.
- Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal, 39*, 469-518. doi: 10.3102/00028312039002469
- Taylor, A. E. B. (2014). *Diagnostic assessment of learning disabilities in childhood: Bridging the gap between research and practice*. New York: Springer.
- Texas Adaptations for Special Education, Texas Administrative Code, §89.1040 (2015)
- Texas Education Agency. (n.d.). *Response to intervention and learning disability eligibility*. Austin, TX: Author. Retrieved from <http://tea.texas.gov/WorkArea/linkit.aspx?LinkIdIdentifier=id&ItemID=25769817951&libID=25769818054>
- Wasserman, J. D., & Bracken, B. A. (2003). Psychometric characteristics of assessment procedures. In J. R. Graham & J. A. Naglieri (Eds.), *Handbook of psychology (Vol 10): Assessment psychology* (pp. 43-66). Hoboken, NJ: John Wiley and Sons.
- Watkins, M. W. (2000). Cognitive profile analysis: A shared professional myth. *School Psychology Quarterly, 15*, 465-479. doi: 10.1037/h0088802
- Watkins, M. W., & Smith, L.G. (2013). Long-term stability of the Wechsler Intelligence Scale for Children- fourth edition. *Psychological Assessment, 25*(2), 477-483. doi: 10.1037/a0031653
- Wechsler, D. (2014). *Wechsler intelligence scale for children* (5th ed.). San Antonio, TX: The Psychological Corporation.
- Wilkinson, L., & American Psychological Association Science Directorate Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604. doi: 10.1037/0003-066X.54.8.594
- Zirkel, P. A. (2013). The Hale position for a “third method” for specific learning disabilities identification: A legal analysis. *Learning Disability Quarterly, 36*, 93-96. doi: 10.1177/0731948713477850
- Zirkel, P. A., & Thomas, L. B. (2010). State laws for RTI: An updated snapshot. *Teaching Exceptional Children, 42*, 56-63.

Article

---

# The Validity and Diagnostic Accuracy of a Computer Adaptive Test of Reading

Sarah Ochs, Milena A. Keller-Margulis, and Sam McQuillin  
*University of Houston*

John Jones  
*Deer Park Independent School District*

Universal screening is a widely used practice in schools. In the past, reading curriculum-based measurement (R-CBM) was the primary technology used for universal screening of reading ability. However, computer adaptive tests (CATs) have recently become more commonly used in education. There is emerging research suggesting CATs may offer improved accuracy over traditional screening options. The purpose of this study was to examine the validity and diagnostic accuracy of a CAT of reading as it relates to a statewide, standardized assessment. In a sample of about 1,200 third through fifth grade students, we found moderate to high correlations between the statewide test and a CAT of reading ( $p < .01$ ). In addition, results suggest that the CAT may offer adequate accuracy for identifying student performance on the standardized assessment. A discussion of how cut scores impact diagnostic accuracy is presented as well as practical implications of this work and future directions for research.

**Keywords:** Universal screening, multi-tiered systems of support, computer adaptive tests, diagnostic accuracy

Universal screening is a vital part of a prevention-oriented approach to service delivery such as Multi-tiered System of Supports (MTSS). Reading curriculum-based measurement (R-CBM), a measure of oral reading fluency that is related to statewide standardized assessments (Buck & Torgesen, 2003; Hintze & Silbergliitt, 2005; McGlinchey & Hixon, 2004; Shapiro, Keller, Lutz, Santoro, & Hintze, 2006; Vander Meer, Lentz, & Stollar, 2005), is typically used in the context of MTSS for screening and progress monitoring. Although R-CBM has strengths as an assessment tool, there are limitations associated with its use. The results of existing studies indicate that R-CBM demonstrates moderate accuracy in classifying students as at-risk (Johnson, Jenkins, & Petscher, 2010; McGlinchey & Hixson, 2004; Stage & Jacobsen, 2001). Insufficient accuracy may result in unacceptable levels of false negatives and false positives, meaning that children at-risk may be

missed in screening. Computer adaptive tests (CATs) may produce higher levels of diagnostic accuracy when compared with CBM (Shapiro & Gebhardt, 2012), but this has only been studied in the areas of math and early literacy. The purpose of this study was to examine the utility of a CAT in reading when used as a universal screening measure.

## Universal Screening with CBM

On-going student assessment through screening and progress monitoring are core components of

---

**Author Note:** Correspondence concerning this paper should be addressed to Milena A. Keller-Margulis, Department of Psychological, Health, and Learning Sciences, University of Houston, 487 Stephen Power Farish Hall, Houston, TX 77204.  
E-mail: [mkmargulis@central.uh.edu](mailto:mkmargulis@central.uh.edu)

MTSS (Glover & DiPerna, 2007; Ikeda, Nessen, & Witt, 2007). Screening results are typically used to make a decision about student risk status in a particular skill area (Ikeda, Nessen, & Witt, 2007; Johnson, Jenkins, & Petscher, 2010). CBM is one of the most commonly used measures for screening in the context of MTSS and numerous studies have explored the use of CBM as a universal screener in schools (Fuchs & Deno, 1991; Fuchs, Fuchs, & Hamlett, 2007; Wayman, Wallace, Wiley, Ticha, & Espin, 2007). One approach to the study of diagnostic accuracy and validity of CBM is examination of the relationship with statewide achievement measures. Researchers have consistently found a positive, significant relationship between CBM and statewide achievement tests (Buck & Torgesen, 2003; Hintze & Silbergliitt, 2005; McGlinchey & Hixon, 2004; Shapiro et al., 2006; Vander Meer, Lentz, & Stollar, 2005). Correlational analyses between R-CBM and statewide achievement tests in Pennsylvania, Florida, and Ohio indicated strong, positive relationships (Buck & Torgesen, 2003; Shapiro et al., 2006; Vander Meer et al., 2005) with moderate to strong relationships in Michigan and Oregon (Crawford, Tindal, & Stieber, 2001; McGlinchey & Hixon, 2004). Additionally, recent meta-analytic studies of the validity of R-CBM with statewide measures have indicated consistently positive, moderate relationships between these measures (e.g. Yeo, 2010).

The ability of R-CBM to accurately identify students who are likely to fail a criterion, like statewide tests of achievement, is typically studied using diagnostic accuracy. This approach includes the four statistics of sensitivity, specificity, positive predictive power, and negative predictive power (Johnson, Jenkins, & Petscher, 2010; Swets, Dawes, & Monahan, 2000). Sensitivity is the proportion of students who failed a criterion and were identified as at-risk on the screener and specificity is the proportion of students who passed the criterion and were expected to do so based on screener performance. Positive and negative predictive power are the proportion of students identified as at-risk or not on the screener, who then failed or passed the criterion, respectively (see Swets, Dawes, & Monahan, 2000 for review). Diagnostic accuracy studies of R-CBM for identifying students' performance on statewide tests find sensitivity levels

that range from .55 to .77 and specificity levels of .74 to .92 (Buck & Torgesen, 2003; Hintze & Silbergliitt, 2005; McGlinchey & Hixon, 2004; Stage & Jacobsen, 2001). Shapiro et al. (2006) identified R-CBM cut points across fall, winter, and spring screening time points that maximized sensitivity for detecting potential failure on the spring statewide test. Sensitivity and specificity results were above .72 across two districts for the winter CBM time point, with some results exceeding .80 (Shapiro et al., 2006). Results also indicated, however, that although accuracy was adequate, there were students who were inaccurately classified based on the screener.

Recommended sensitivity levels for CBM in MTSS vary considerably (Kilgus, Methé, Maggin, & Tomasula, 2014) with some suggesting sensitivity levels of .90 are necessary (Compton, Fuchs, Fuchs, & Bryant, 2006; Jenkins et al., 2007). Results of many R-CBM diagnostic accuracy studies fail to meet this criterion, resulting in the risk of false negatives in the screening process (e.g. Buck & Torgesen, 2003; McGlinchey & Hixon, 2004). False negatives are students who were in need of intervention but were not identified using a single screening point (Johnson, Jenkins, & Petscher, 2010). Although some researchers suggest prioritizing high sensitivity levels in screening, most suggest .80 to be acceptable (Kilgus et al., 2014). Regardless of the diagnostic accuracy levels selected, the purpose of screening is to accurately and efficiently identify students at-risk for poor performance. Although computer adaptive tests (CATs) have been in existence for some time, they have more recently been considered for use in universal screening and may be an option for achieving desired levels of diagnostic accuracy.

### **Computer Adaptive Tests**

The theoretical foundation of CATs is item response theory (IRT). IRT focuses on item-level examination, as opposed to test-level examination used in classical test theory. Instead of providing all students with identical questions delivered in the same order, items are presented according to the unique individual response. Based on the examinee's estimated ability level, an optimal start point item is selected from a pool of available test items (Thissen & Mislevy, 2000). After each question is answered,

the ability estimate is recalculated and a new item is selected from the pool based on the examinee's updated ability level. Each item has a difficulty level that is a measure of the percentage of students answering that item correctly, with higher p-values indicating easier items or a greater percentage of students answering the item correctly (Renaissance Learning, 2013). Items that better discriminate between low and high performing students are more related to the total test score. Technical adequacy data for CATs indicate a reliable and valid format with an administration of roughly 50% fewer items compared to conventional vocabulary and math tests (e.g. Mardberg & Carlstedt, 1998; Moreno & Segall, 1997). Unlike traditional paper-and-pencil tests, CATs offer greater security and more rapidly available test scores (Parshall, Spray, Kalohn, & Davey, 2002). On the other hand, the use of CATs assume a certain level of comfort with and knowledge of computer technology, require a much higher initial purchase cost, and the availability of necessary equipment. The design of CATs may make them efficient options for use as a universal screening measure for some districts; however there are only a few studies that examine them in this context.

### **CATs as Universal Screeners**

Few studies have examined the use of a CAT as a universal screening tool. McBride, Ysseldyke, Milone, and Stickney (2010) examined the relationship between STAR Early Literacy (SEL), a CAT, and three other early literacy measures at kindergarten through second grade. McBride et al. (2010) studied the convergent validity of each tool for measuring five critical components of reading: phonemic awareness, phonics, vocabulary, comprehension, and fluency. Correlation results varied based on grade level. In kindergarten, SEL was weakly correlated to all of the reading components except for phonics where the relationship reached a moderate level ( $r = .42$ ). SEL was more strongly related to each reading domain in first and second grades, with the strongest correlation for reading fluency at first grade ( $r = .69$ ). Researchers concluded that SEL was the most cost-effective of all four measures studied, suggesting the potential feasibility of CATs for early reading

assessment (McBride et al., 2010).

A second study examined the degree to which fall, winter, and spring kindergarten SEL scores predict end of kindergarten and end of first grade reading skills (Clemens et al., 2015). Researchers also analyzed the additional variance explained by paper-based measures and the accuracy of SEL relative to paper-based assessments. Ninety-eight kindergarten students participated, with 71 remaining at the end of first grade one year later. SEL scores came from routine fall, winter, and spring kindergarten administration. Students completed paper-based measures of letter-naming fluency (LNF), letter-sound fluency (LSF), word identification (WID), word attack (WAT), and word reading fluency (WRF) in May of kindergarten and WID, WAT, and R-CBM in May of first grade. A model of end-of-year word-reading skills was measured by WID, WAT, and WRF at kindergarten and two latent factors at first grade: reading accuracy factor (WID and WAT) and reading fluency (R-CBM). Fall, winter, and spring SEL scores were significant predictors of kindergarten, end-of-year word-reading skills, with the fall SEL accounting for 35% of the variance and winter and spring accounting for 38%. When SEL, LNF, and LSF were combined and used as a predictor, the proportion of explained variance increased to 58%. Kindergarten, spring SEL was a significant predictor of both first grade factors, accounting for 37% of variance in reading accuracy, and 33% of the variance in reading fluency. This is a smaller proportion of variance when compared to WRF, which when used as a predictor, accounted for 43% and 54% of the variance in end-of-year word-reading factors, respectively.

Clemens et al. (2015) also used ROC curves to examine the accuracy of SEL in predicting later reading performance (above or below the 40th percentile) on the WID, WAT, and WRF in kindergarten and (above or below the 40th and 15th percentile) on the WID, WAT, and R-CBM in first grade (Clemens et al., 2015). For the kindergarten analysis, AUC ranged from .81 to .94. Sensitivity levels were generally low when using the SEL cut score associated with the 40th percentile alone (.54 for fall, .62 for winter, and .46 for spring) compared to the paper-based measures (.92 for LNF and LSF alone). Combining all three measures (LNF, LSF,

and SEL) as the predictor resulted in the same level of sensitivity as the paper-based measures alone (.92). When examining the accuracy of SEL scores from spring of kindergarten year for predicting word-reading skills at the end of first grade, sensitivity levels were poor. Sensitivity levels of the paper-based measures were higher for both criteria ranging from .90 to 1.0. These findings support the practice of combining measures to increase the accuracy of identification.

Shapiro and Gebhardt (2012) examined the validity and diagnostic accuracy of a CAT math measure (STAR Math [STAR-M]; Renaissance Learning, 2011) and math CBMs (i.e., math computation [M-CBM], and math concepts and applications [M-CAP]) with statewide test performance. Data from students in first through fourth grades, collected in the context of universal screening, were used in the study (Shapiro & Gebhardt, 2012). Results included significant correlations across grades and administration time points for most STAR-M and M-CBM relationships, and for all STAR-M and M-CAP relationships. First grade STAR-M correlations with M-CBM ranged from .50 to .56, second grade from .18 (not significant) to .45, third grade from .18 (not significant) to .48, and fourth grade from .12 (not significant) to .58. Second grade STAR-M correlations with M-CAP ranged from .35 to .61, third grade from .36 to .64, and fourth grade from .40 to .54. Correlations between STAR-M, M-CBM, M-CAP and the Pennsylvania statewide achievement test were reported for fall and winter administrations for third and fourth grade only. Across third and fourth grades, M-CBM and M-CAP demonstrated weak to moderate relationships with PSSA (.29 and .36, respectively) and the average correlation between STAR-M and PSSA was moderate (.60).

Diagnostic accuracy was examined using the 16<sup>th</sup> percentile as the screening cut score due to its correspondence with proficiency on the statewide achievement test (Shapiro & Gebhardt, 2012). Consistently high levels of specificity across measures and times for all grades (.86 - .97) were achieved while sensitivity levels varied. M-CBM sensitivity levels ranged from .28 to .54, M-CAP from .28 to .48. Sensitivity of STAR-M for fall and winter ranged from .68 to .72, suggesting STAR-M may more accurately identify students who later fail

the statewide test. Levels of sensitivity were highest for all measures for the winter time point. These results suggest marginally higher levels of sensitivity, while maintaining levels of specificity consistent with other measures (Shapiro & Gebhardt, 2012); however, the sensitivity levels did not reach the thresholds identified for screening adequacy. Although there is preliminary evidence of the utility of CATs as screeners, additional work is needed, particularly in the area of reading.

### Current Study

The purpose of this study was to examine the validity and diagnostic accuracy of scores on the STAR-Reading test, a CAT measure of reading, when used as a universal screener. In addition, the change in diagnostic accuracy associated with different cut scores was examined. The Texas statewide standardized assessment, the State of Texas Assessments of Academic Readiness (STAAR), was used as the criterion and is referred to as the *statewide achievement test* to avoid confusion with the STAR-R. The following research questions were addressed.

1. What is the relationship between STAR-R at fall, winter, and spring time points with the spring statewide achievement test for third through fifth grade students?
2. What is the diagnostic accuracy of fall and winter STAR-R in third through fifth grades, including the levels of sensitivity, specificity, positive predictive power, and negative predictive power when using cut scores identified via ROC curves when sensitivity levels are .90?
3. What is the diagnostic accuracy of fall and winter STAR-R scores in third through fifth grades when both sensitivity and specificity are optimized?

**Table 1:** *Participant Demographics*

	Grade 3		Grade 4		Grade 5	
	N	%	N	%	N	%
Gender						
Male	574	47.20	644	49.61	608	48.41
Female	644	52.80	654	50.39	648	51.59
Ethnicity						
American Indian/Alaska	4	.33	3	.23	3	.24
Asian	14	1.15	27	2.08	25	1.99
Black/African American	19	1.56	21	1.62	16	1.27
Hispanic	446	36.62	491	37.83	497	39.57
White	696	57.14	723	55.70	692	55.10
Two or More Races	39	3.20	33	2.54	23	1.83
Participated in Special Education	61	5.25	73	5.62	65	5.18
Total N	1218		1298		1256	

## Method

Participants in this project were third through fifth graders in a suburban school district outside a large metropolitan area in the southern United States. The data set included 1,218 students in third grade, 1,298 in fourth grade, and 1,256 in fifth grade during the 2011-2012 and 2012-2013 school years. Additional demographic information including gender, ethnicity, and special education participation is included in Table 1.

## Measures

**STAR-Reading (Renaissance Learning, 2013).** STAR-Reading (STAR-R) is one type of CAT designed to provide an assessment of reading skills efficiently, that can be compared to national norms and used to track growth (Renaissance Learning, 2013). STAR-R is designed to assess 36 reading skills across five domains (e.g. Word Knowledge and Skills, Analyzing Literary Text) and administration can easily occur in a large group setting. The STAR-R is a fixed-length CAT with 25 questions and a typical completion time of 15 minutes (Renaissance Learning, 2013). Each STAR-R administration is unique, which eliminates a practice effect.

During the item calibration, 2,133 items were narrowed to 1,409 with items being eliminated if the item discrimination was  $< .30$ , other answer options had high item discrimination, a small sample of students attempted the item ( $< 300$ ), the item did not fit the Rasch IRT model, or the item was too difficult or too easy. Reported technical adequacy of the STAR-R is adequate with an overall reliability coefficient across all grades of .95 and ranged from .89 for third and fourth grade to .90 for fifth grade (Renaissance Learning, 2013). Concurrent, predictive, and construct validity coefficients were all reported to be greater than .72. More recent studies of validity and diagnostic accuracy of the STAR-R with various statewide achievement tests conducted by the test publisher suggest technical adequacy (Renaissance Learning, 2014). Predictive validity coefficients reported with other state assessments as the criterion ranged between .33 and .63 for third through fifth grades. An analysis of classification accuracy from STAR-R data collected three to eleven months prior to state test administration yielded an average sensitivity level of .78 and specificity of .81 across states (Renaissance Learning, 2013).

**Statewide Achievement Test.** The State of Texas Assessment of Academic Readiness (STAAR) is the statewide assessment used in Texas to assess skills across three general categories: Understanding across Genres, Understanding and Analysis of Literary Texts, and Understanding and Analysis of Informational Texts. For the 2012 administration, two cut points divided scores into three categories. Level I: Unsatisfactory Academic Performance describes students performing below a scaled score of 1331 for third grade, 1410 for fourth grade, and 1458 for fifth grade. Level II: Satisfactory Academic Performance delineates students scoring at or above Level I scaled scores, but below 1523, 1633, and 1667, respectively. Lastly, students performing above Level II scaled score are Level III: Advanced Academic Performance. The percentage of students across third through fifth grades statewide identified as performing at Level II or above ranged from 76-77% for spring of 2012 and 72-79% for spring 2013. The above represents adequate passing for the 2012 administration, but the state was gradually shifting the performance standard with phase-in scoring. For the purpose of this project, the final or recommended performance standard was used to increase generalizability and practical utility of results. Although this performance standard was not in place when these data were collected, it will be the final performance standard used for the lifespan of this statewide test. Using this criterion, a scaled score of 1468 and above is satisfactory performance for third grade, 1550 and above for fourth grade, and 1582 and above for fifth grade. Reliability data for the 2011-2012 STAAR includes a stratified alpha for the overall Reading test of .89 (Texas Education Agency, 2013).

The statewide reading achievement test requires students to read several selections and then choose the best answer to multiple-choice questions. Higher grade level tests have passages that contain more information, such as a written passage as well as a diagram or graphic. Number of passages and questions increase with grade level. Students have four hours to finish the STAAR test and take the test in late March in fifth grade and late April in third and fourth grade.

## **Procedures**

Existing screening data from a local school district in the southern United States were used for this study. STAR-R data were collected during fall 2011 through spring 2013 from students in third through fifth grades across seven elementary campuses. If a student data point was not available during the screening time point, the next proximal data point within the window of fall (September through November), winter (January through March), or spring (April through June) was used. These time points align with traditional R-CBM benchmark time points. The first data point within each time period was selected for use as the STAR-R benchmark. Most students were assessed during traditional benchmark time periods, but adopting this larger window allowed for the inclusion of students who might have participated just outside of that time period. The data were collapsed across two years and any students who were retained in a grade were removed.

Students took one practice session to become familiar with STAR-R administration procedures. When taking the test, students sat down at a computer with their whole class or as a small group while teachers or other staff circulated and were available to provide administration assistance but were not directly involved with test administration. This study used statewide achievement testing (i.e., STAAR) data obtained from the district's April 2012 and 2013 administration for third, fourth, and fifth grade students. All test administrators were professionals holding valid education credentials or are under supervision of such professionals. Test administrators follow standardized procedures such as how to record student breaks, actively monitor rooms, answer student questions, and set up the testing room.

## **Data Analysis**

Data were screened for missing values and outliers. Outliers were determined using Mahalanobis Distance. The 5.8% of total missing data were addressed using the Multiple Imputation function in SPSS which avoids assuming that data are missing at random and minimizes the likelihood of artificially raising the mean. Descriptive statistics including mean, standard deviation, and range were reviewed along with skewness and kurtosis.

Correlations were used to examine validity in research question one. Diagnostic accuracy analyses used receiver operating characteristic curves (ROC) to model all possible cut scores on the STAR-R at both fall and winter time points. The spring screening time point was not examined because it occurred after the administration of the statewide achievement test. Two approaches to selecting cut scores were used. First, the cut-score associated with a sensitivity of .90 was used, a decision that was based on recommendations within the MTSS and universal screening literature (Compton et al., 2006; Jenkins, 2003; Jenkins et al., 2007; Johnson, Jenkins, & Petscher, 2010). In practice, accurately identifying the students that are likely to fail the criterion of interest and therefore need intervention (sensitivity) is a higher priority than identifying those that do not need intervention and are anticipated to perform successfully. In addition, data analyses included examination of the diagnostic accuracy statistics that result when a cut score is selected that maximizes sensitivity with minimal decrease in specificity, consistent with previous research in this area (e.g., Keller-Margulis, Shapiro, & Hintze, 2008) where cut

scores that resulted in both sensitivity and specificity at .70 or higher were selected. The diagnostic accuracy statistics of interest included sensitivity, specificity, positive (PPP) and negative predictive power (NPP), and classification accuracy results of positive and negative predictive power as well as the area under the curve (AUC) generated through the ROC analyses. The false positive and false negative rates were also examined.

**Results**

Data were screened for outliers and there were 15 STAR-R cases at third grade, 23 cases at fourth grade, and 21 cases at fifth grade that were potentially outliers. After review of the data, it was determined that these cases were actual student scores and they were therefore retained in the analyses. Appropriate degrees of skewness were displayed by nearly all variables at less than  $\pm 1.96$  (Madansky, 1988). Descriptive statistics and sample sizes across grades and measures are presented in Table 2.

**Table 2:** *Descriptive Statistics for STAR-R and Statewide Assessment (STAAR)*

	N	Mean	SD	Range	Skewness	Kurtosis
Third Grade						
STAR-R Fall	1218	338.07	127.77	71 - 865	.47	.54
STAR-R Winter	1218	385.91	131.14	64 - 959	.39	1.04
STAR-R Spring	1218	428.17	129.46	71 - 966	.34	1.09
STAAR	1218	69.56	16.94	1080 - 1911	-.51	-.33
Fourth Grade						
STAR-R Fall	1298	436.83	134.89	51 - 999	.56	1.51
STAR-R Winter	1298	479.00	145.93	61 - 1245	.79	2.53
STAR-R Spring	1298	610.56	155.09	72 - 998	.31	.88
STAAR	1298	68.51	17.24	1127 - 2002	-.50	-.40
Fifth Grade						
STAR-R Fall	1256	540.90	171.78	54 - 1302	1.00	1.85
STAR-R Winter	1256	581.06	181.98	82 - 1331	.92	1.38
STAR-R Spring	1256	622.96	193.52	67 - 1328	.82	.93
STAAR	1256	71.91	14.94	1284 - 2021	-.58	-.11

*Note.* STAAR – State of Texas Assessment of Academic Readiness



**Correlations**

The relationship between the spring standardized assessment (STAAR) and fall, winter, and spring administrations of STAR-R are presented in Table 3. All STAR-R correlations are statistically

significant ( $p < .01$ ). STAR-R demonstrated consistently positive correlations across all grades and time points ranging from .60 to .70.

**Table 3:** *Correlations Between STAR-R and Statewide Assessment (STAAR)*

	STAR-R			
	N	Fall	Winter	Spring
Third Grade	1218	.66	.70	.68
Fourth Grade	1298	.60	.64	.60
Fifth Grade	1256	.63	.62	.65

*Note.* STAR-R – STAR Reading Measure; STAAR – State of Texas Assessments of Academic Readiness. All correlations significant at  $p < .01$ .

**Table 4:** *Diagnostic Accuracy with Sensitivity Levels of .90*

	N	Cut Score	Sensitivity	Specificity	PPP	NPP	AUC	False Negative	False Positive
Fall STAR-R									
Grade 3	1218	404	.90	.50	.69	.80	.83**	.20	.31
Grade 4	1298	500	.90	.51	.72	.79	.82**	.21	.28
Grade 5	1256	601	.90	.54	.72	.80	.83**	.20	.28
Winter STAR-R									
Grade 3	1218	443	.90	.60	.74	.83	.86**	.17	.26
Grade 4	1298	535	.90	.60	.76	.81	.85**	.19	.24
Grade 5	1256	646	.90	.50	.72	.78	.84**	.22	.28

*Note.* STAR-R - STAR-Reading; STAAR - State of Texas Assessment of Academic Readiness; PPP – positive predictive power; NPP – negative predictive power; AUC – area under the curve. \*\*  $p < .01$

**Diagnostic Accuracy**

Diagnostic accuracy statistics were first examined using ROC curves to identify a cut score that offered sensitivity of .90. Results from this analysis are presented in Table 4. Corresponding cut score as well as sensitivity, specificity, positive (PPP) and negative predictive power (NPP), and area under the curve (AUC) are reported. An area under the curve of 1 represents a perfect test. The area under the curve was statistically significant ( $p < .01$ ) for all grade levels with the highest statistic yielded by winter STAR-R at fifth grade and all grades and time points exceeding .80. Specificity levels were consistently above .50 with marginally higher values reported for third and fourth grade at the winter time point. STAR-R yielded relatively high levels of both PPP and NPP with all values exceeding .72 with the exception of PPP at the fall time point in third grade (.69). Examining the false positive and false negative rates indicates that when the sensitivity level is set to .90, as was the case in this analysis,

there was a slightly higher level of false positives than false negatives.

The diagnostic accuracy statistics generated when a cut score was selected that maximized both sensitivity and specificity (i.e.,  $\geq .70$ ) resulted in lower cut scores across all grades for both fall and winter time points (see Table 5). In addition, sensitivity and specificity levels almost all exceeded .75. Slightly higher positive predictive power was observed with most values above .80. The rate of false negatives was higher in this analysis than when the sensitivity levels were set to .90, which is expected given the lower cut scores. With a lower cut score, there will be more students who perform above the cut score on the screener but who are later unsuccessful on the criterion measure. When using cut scores that maximize both sensitivity and specificity, the result is a lower false positive rate.

**Table 5:** *Diagnostic Accuracy Results with Balanced Sensitivity and Specificity*

	N	Cut Score	Sensitivity	Specificity	PPP	NPP	AUC	False Negative	False Positive
Fall STAR-R									
Grade 3	1218	339	.74	.74	.78	.70	.83**	.30	.22
Grade 4	1298	446	.76	.77	.82	.69	.82**	.31	.18
Grade 5	1256	528	.75	.75	.80	.69	.83**	.30	.20
Winter STAR-R									
Grade 3	1218	395	.79	.76	.80	.75	.86**	.25	.20
Grade 4	1298	481	.76	.77	.82	.69	.85**	.31	.18
Grade 5	1256	563	.76	.76	.81	.70	.84**	.30	.19

*Note.* STAR-R - STAR-Reading; STAAR - State of Texas Assessment of Academic Readiness; PPP – positive predictive power; NPP – negative predictive power; AUC – area under the curve. \*\*  $p < .01$

**Discussion**

Current research indicates that R-CBM lacks adequate levels of classification accuracy (Johnson, Jenkins, & Petscher, 2010; McGlinchey & Hixson, 2004; Stage & Jacobsen, 2001). High levels of accuracy are necessary for schools to properly allocate finite resources to supplemental instruction or intervention. The purpose of this study was to examine the relationship between one CAT, the STAR-R, used as a universal screening measure and

a statewide, standardized test and the diagnostic accuracy it provides as a screening measure. We found that the average STAR-R score increased across the year and across grades. There was a positive, moderate to strong relationship between the STAR-R and the statewide measure of achievement, similar to that seen for other screening measures, such as R-CBM (e.g., Yeo, 2010) and were similar if not slightly higher than those reported

for math CATs (Shapiro & Gebhardt, 2012). For the most part, there was little variation in this relationship across grades and years with most results in the .60 range.

The diagnostic accuracy of the STAR-R was examined in two ways, first to examine the accuracy when sensitivity levels are set to .90 and, second, to examine diagnostic accuracy rates when the key indicators of sensitivity and specificity are balanced. This approach was taken to illustrate the changes in cut score as well as false positive and false negative rates. STAR-R demonstrated a relatively high accuracy overall for discriminating between passing and failing students in this sample as indicated by the high and significant AUC values for each grade. When the analyses were conducted to maximize sensitivity levels, the result was a higher false positive rate than false negative rate. This indicated that at the cut scores used to achieve high sensitivity, there were students identified as at-risk for poor performance on the criterion measure who were ultimately successful. In contrast, there were fewer students identified as not at-risk who later failed (i.e., false negatives). The implications of this are that there were more students who were identified as potentially at-risk than there were students at-risk who were missed by the screening. Although this may be taxing on available resources for the provision of intervention, the result of casting a wider net in screening (i.e., using higher sensitivity levels that result in higher cut scores) is that more students are provided with intervention.

The changes in cut scores and diagnostic accuracy results when sensitivity and specificity are both maximized were in the expected directions. Maximizing both diagnostic accuracy metrics resulted in slightly lower cut scores, as well as lower false positive rates and higher false negative rates. From a practical perspective, the result would be more students identified as not being at-risk (i.e., false negatives) who therefore did not receive intervention, but who ultimately failed the criterion measure.

### **Implications for Research and Practice**

The purpose of this study was to expand on recent studies of CATs by examining the validity and diagnostic accuracy of a CAT in reading with a statewide achievement test. Results provide a

replication of findings presented by Renaissance Learning (2014) but also present changes in diagnostic accuracy when different cut scores are examined using ROC curves. The results presented here also demonstrate diagnostic accuracy changes across benchmark time points. Results demonstrate moderate to high levels of diagnostic accuracy, suggesting that it may be a viable alternative for universal screening in reading and further confirming that the use of ROC curves is a flexible and effective way to achieve desired levels of sensitivity for screening. This study adds to the small body of literature regarding the validity of CATs for universal screening. Given the group administration format and the efficiency of CATs as a function of their design, they may be a strong option for universal screening. While R-CBM is still a strong and valuable measure, CATs are more diagnostic in the breadth of information that they gather and range of skills tested and may be useful in the direction that they can provide for instruction and intervention.

The results of this study provided a valuable illustration of the impact of selecting various cut scores on the diagnostic accuracy results. This study adds the practical application of providing cut scores associated with different levels of diagnostic accuracy and different cut score selection procedures. For instance, the lowest cut score needed across grades in this sample in order to reach a sensitivity level of .90 versus a balanced sensitivity and specificity is provided. In this study as well as others, most students in the sample were accurately classified as either passing or failing the criterion measure, however, decisions about cut scores and the resulting false positive and false negative rates are most challenging. These decisions are often driven by external constraints, such as funding available to provide supplemental intervention or the desire to improve passing rates on the statewide achievement test. The changes in false positive and false negative rates associated with changes in the cut score selected illustrated in this study demonstrate this challenge. If resources for intervention are plentiful, then a school or district may elect to set sensitivity levels to .90 thus maximizing the number of students identified as potentially at-risk even though some of those students may be successful on the criterion regardless. A related result is that there will be fewer

students who are missed by the screening, that is, they perform above the cut point but are then unsuccessful on the criterion. For example, the results of the winter of third grade, where the cut score of 443 on the STAR-R was identified, yielded a false positive rate of .26 but a false negative rate of only .17.

These findings may assist districts in making decisions about screening measures to use at the elementary level as well as a means for setting criteria on screeners that will facilitate identification of students at-risk for failure on high stakes outcomes. Specifically, the STAR-R may offer improved screening accuracy, particularly at the upper elementary grades where measures of fluency alone tend to be less useful (e.g., Keller-Margulis et al., 2008). Additional considerations not explored by this study include cost of screening measures, resources required to train staff, administration time, ease of scoring, and usability of data. These are important issues that are relevant to the selection of measures for use in the context of MTSS that require future study, particularly as it relates to CAT measures for screening.

### **Limitations and Future Directions**

There are some limitations associated with this study. One limitation is the lack of generalizability of results. In particular, diagnostic accuracy levels reported reflect one specific criterion in a specific sample of students and these statistics are impacted by the base rate of failure in the sample (VanDerHeyden, 2011). One way to address this limitation is to conduct replications of this study using the same standardized state test with a different student population. A second limitation of this study is the use of existing data. The researcher had no control over ensuring the fidelity of administration or scoring due to the previous collection of data, however systems were in place within the district to minimize the likelihood of compromising the integrity of administration. Although these conditions are common for applied research, they may affect the results. For instance, students were primarily administered the STAR-R in whole-class groups but sometimes used small-group administration in the case of absences. Although the teacher has no interaction with students during

administration, the teacher-to-student ratio may affect student performance. Third, use of fall and winter STAR-R scores impacts interpretation of diagnostic accuracy statistics. After winter data collection, instruction continues from winter to spring, before administration of the state assessment. Therefore, this analysis, like many other similar studies, does not consider learning that occurred during this time. Additionally, the statewide assessment measure is administered at slightly different times in the spring with the fifth grade test occurring in late March and the third and fourth grade reading test occurring a month later in late April. Ultimately, this study as well as others regarding diagnostic accuracy and CATs can only serve as a demonstration of the utility of these measures in context given the influence that base rates of performance have on these statistics as well as the limitations to generalizability as a function of the CAT item bank used and criterion measure studied.

The results of this study and the limitations noted point to numerous directions for additional work. It is necessary to further study the technical adequacy and practical utility of CATs in general across states with unique standardized assessments. These studies should continue to examine the relationship and diagnostic accuracy of other CATs with standardized criterion measures. It is common for screening measures to also be used for progress monitoring. Based on the results of the current study, it seems as though performance on the CAT increases within the year for students and across grades, however, research is needed to examine the use of CATs for progress monitoring student response to instruction and intervention. It may also be useful to investigate the effect of using STAR-R and R-CBM scores together to predict performance on state assessments. Alternatives to single point assessments for screening must also be studied. CATs may function well as a first phase to screening given the feasibility of their use and the information generated regarding direction for instruction.

There is a need for improved screening measures and the results of this study suggest that CATs offer one potential alternative, but additional studies are required to address the stated limitations in this study and replicate the findings. The need for

screening measures that can be used to reliably and accurately identify students who are at-risk for failure on criterion measures of interest remains a critical issue for further study.

## References

- Buck, J., & Torgesen, J. (2003). *The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test* (Technical Report 1). Tallahassee: Florida Center for Reading Research.
- Clemens, N. H., Hagan-Burke, S., Luo, W., Cerda, C., Blakely, A., Frosch, J., Jones, M. (2015). The predictive validity of a computer-adaptive assessment of kindergarten and first-grade reading skills. *School Psychology Review, 44*, 76-97.
- Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98*, 394-409. doi:10.1037/00220663982394
- Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment, 7*, 303-323.
- Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57*, 488-500.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (2007). Using curriculum-based measurement to inform reading instruction. *Reading and Writing, 20*, 553-567.
- Glover, T. A., & DiPerna, J. C. (2007). Service delivery for response to intervention: Core components and directions for future research. *School Psychology Review, 36*, 526-540.
- Hintze, J. M., & Silbergliitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review, 34*, 372-386.
- Ikeda, M., Nessen, E., & Witt, J. (2007). Best practices in universal screening. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology* (pp. 103-114). Bethesda, MD: National Association of School Psychologists Publishers.
- Jenkins, J. R. (2003, December). *Candidate measures for screening at-risk students*. Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention symposium, Kansas City, MO. Retrieved March 3, 2013, from <http://www.nrclid.org/symposium2003/jenkinsindex.html>
- Jenkins, J. R., Hudson, R. F., & Johnson, E. S. (2007). Screening for at-risk readers in a response to intervention framework. *School Psychology Review, 36*, 582-600.
- Johnson, E. S., Jenkins, J. R., & Petscher, Y. (2010). Improving the accuracy of a direct route screening process. *Assessment for Effective Intervention, 35*, 131-140. doi:10.1177/1534508409348375
- Keller-Margulis, M.A., Shapiro, E.S., Hintze, J.M. (2008). Long term diagnostic accuracy of curriculum-based measures in reading and mathematics. *School Psychology Review, 37*(3), 374-390.
- Kilgus, S. P., Methe, S. A., Maggin, D. M., & Tomasula, J. L. (2014). Curriculum-based measurement of oral reading (R-CBM): A diagnostic test accuracy meta-analysis of evidence supporting use in universal screening. *Journal of School Psychology, 52*, 377-405.
- Madansky, A. (1988). *Prescriptions for working statisticians*. New York: Springer-Verlag.
- Mardberg, B., & Carlstedt, B. (1998). Swedish enlistment battery: Construct validity and latent variable estimation of cognitive abilities by the CAT-SEB. *International Journal of Selection and Assessment, 6*, 107-114. doi: 10.1111/1468-2389.00079
- McBride, J. R., Ysseldyke, J., Milone, M., & Stickney, E. (2010). Technical adequacy and cost benefit of four measures of early literacy. *Canadian Journal of School Psychology, 25*, 189-204.
- McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*, 193-203.
- Moreno, K. E., & Segall, O. D. (1997). Reliability and construct validity of CAT-ASVAB. In W.A. Sands, B. K. Waters, & J. R. McBride (Eds.), *Computerized adaptive testing: From inquiry to operation* (p. 169-179). Washington, DC: American Psychological Association.
- Parshall, C. G., Spray, J. A., Kalohn, J. C., & Davey, T. (2002). *Practical considerations in computer-based testing*. New York, NY: Springer.
- Pearson. (2012). *AIMSweb reading curriculum based measurement administration and scoring guide*. Available from <http://www.aimsweb.com/>
- Renaissance Learning. (2011). *STAR Math: Understanding reliability and validity*. Wisconsin Rapids, WI: Renaissance Learning. Retrieved from <http://research.renleam.com/research/pdfs/135pdf>
- Renaissance Learning. (2013). *STAR Reading: Technical manual*. Wisconsin Rapids, WI: Renaissance Learning. Available by request to [research@renaissance.com](mailto:research@renaissance.com)
- Renaissance Learning. (2014). *Relating STAR Reading and STAR Math to the state of Texas assessments of academic readiness (STAAR) performance*. Wisconsin Rapids, WI: Renaissance Learning. Retrieved from <http://doc.renlearn.com/KMNet/R004573515 K64DE.pdf>
- Shapiro, E. S., & Gebhardt, S. N. (2012). Comparing computer-adaptive and curriculum based measurement methods of assessment. *School Psychology Review, 41*, 295-305.
- Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests:

- Reading and math performance in Pennsylvania. *Journal of Psychoeducational Assessment*, 24, 19-35.  
doi: 10.1177/0734282905285237
- Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review*, 30,407-419.
- Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, 1, 1-26. doi: 10.1111/1529-1006.001
- Texas Education Agency. (2013). *Technical Digest 2012-2013*. Available from tea.texas.gov
- Thissen, D., & Mislevy, R. J. (2000). Testing algorithms. In Wainer, H. (Ed.), *Computerized adaptive testing: A primer* (2<sup>nd</sup> ed., pp. 101-134). Mahwah, NJ: Lawrence Erlbaum Associates.
- Vander Meer, C. D., Lentz, E. F., & Stollar, S. (2005). *The relationship between oral reading fluency and Ohio proficiency testing in reading* (Ohio Technical Report). Eugene, OR: University of Oregon.
- VanDerHeyden, A. M. (2011). Technical adequacy of response to intervention decisions. *Exceptional Children*, 77, 335-350. doi: 10.1177/001440291107700305
- Wayman, M. M., Wallace, T., Wiley, H. I., Ticha, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *Journal of Special Education*, 41, 85-120.
- Yeo, S. (2010). Predicting performance on state achievement tests using curriculum-based measurement in reading: A multilevel meta-analysis. *Remedial and Special Education*, 31, 412-422.  
doi:10.1177/0741932508327463

Article

---

## Girls with Autism: Exploring Gender Disproportionality in Special Education

Rachel H. Fein, Christie Brewton, Katherine Bergez, and Sarah S. Mire  
*University of Houston*

Males are four-to-five times more likely than females to be identified with autism spectrum disorder (ASD) in the general population in the U.S. Similarly, gender disproportionality exists within some special education categories (e.g., Emotional Disturbance [ED], Specific Learning Disability [SLD]), suggesting males with these educational classifications receive more school-based services than their female peers. However, potential gender disproportionality within the autism (AU) special education category has not been investigated. The current study aimed to investigate gender proportions in the AU category in Texas by academic year and across school district regions. Data were analyzed from 171,434 school-age children from Texas with AU special education classification between the 2008–2009 and 2012–2013 academic years. Results indicated gender proportions in the AU category in Texas matched the ratio found in the general population (i.e., 4:1 to 5:1). However, differences in these proportions were found in some areas of the state. Implications on research and practice, as well as limitations and future directions, are discussed.

**Keywords:** gender disproportionality; autism; special education; Texas

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by persistent deficits in social communication and social interaction across multiple domains, as well as by patterns of restricted, repetitive behavior, interests or activities (American Psychiatric Association [APA], 2013). Families of children with ASD rely heavily on school-based services obtained through the special education system within public schools (Mire, Raff, Brewton, & Goin-Kochel, 2015), either in conjunction with or instead of services received through the private sector. Having a clinical (i.e., DSM-5) diagnosis of ASD, however, does not automatically entitle a student to special education supports and services. For a student to receive services through U.S. public schools, they must qualify under the Individuals with Disabilities Education Improvement Act (IDEIA; IDEIA, 2004) as a student who has both a disabling condition *and* demonstrates a related educational need (i.e., disability adversely affects educational performance). For children with ASD, many qualify for services within the IDEIA (2004) Autism (AU) disability category, though some

students with ASD may be served under other disability classifications (e.g., Specific Learning Disability [SLD]) or be determined by the school-based evaluation team not to meet special education classification criteria at all.

The prevalence of ASD diagnoses in the U.S. general population has increased from 1:110 in 2006 to 1:68 in 2010 (CDC, 2012). Concurrent with increases in ASD diagnostic prevalence, the number of students served under the AU special education classification in the U.S. public school systems has risen from the 2004–2005 academic year (i.e., 191,000 students) to the 2011–2012 year (i.e., 455,000 students), which represents an increase of 0.4 to 0.9 percent of the total number of students who receive special education services in the U.S. (U.S. Department of Education, 2015).

The prevalence rates of ASD diagnosis vary between males and females (Kirkovski, Enticott, & Fitzgerald, 2013); in fact, males are four to five

times more likely than females to have a diagnosis of ASD (APA, 2013). Some have suggested that there are differences in the clinical presentation of ASD symptoms in females, which contributes to the gender differences in diagnostic prevalence of ASD (Attwood, 2007). Key differences include those in cognitive ability, restricted interests/repetitive behaviors, social interaction (i.e., play behavior and developing friendships), and communicative deficits (Kirkovski et al., 2013). In regards to cognitive ability, Dworzynski et al. (2012) found that females with cognitive impairments are more likely than those without deficits in cognitive ability to be diagnosed with ASD. Furthermore, Fombonne (2003) found that, with regard to cognitive test scores (i.e., intelligence quotient [IQ]), the male-to-female ratio is 5.75:1 in cognitively high-functioning children (i.e., full-scale IQ  $\geq 70$ ) versus 1.9:1 in low-functioning ASD (i.e., full-scale IQ  $< 70$ ). Some researchers have proposed that the differences in ASD symptom profiles between males and females should be reflected in diagnostic instruments; that is, gender-specific thresholds for an ASD diagnosis should be created (Constantino & Charman, 2012).

In terms of restricted interests and repetitive behaviors, Solomon, Miller, Taylor, Hinshaw, and Carter (2012) found that males with ASD experience greater impairment. This was further supported by Szatmari et al. (2012) who reported large discrepancies between males and females for items on tests measuring unusual preoccupations, circumscribed interests, and repetitive use of objects or interest in parts of objects (e.g., males were found, on average, to have greater impairment). When considering differences in social interaction, various studies have suggested that females with ASD are more likely to engage in pretend play (Knickmeyer, Wheelwright, & Baron-Cohen, 2008) and may imitate others' actions more often than males (Kirkovski et al., 2013). Regarding gender differences in communication abilities, research is mixed. For example, some studies reported greater *impairment* of communication abilities among females with ASD (Hartley & Sikora, 2009) (e.g., more echolalic speech among females [Andersson, Gillberg, & Miniscalco, 2012]), while others indicated greater *strengths* in social communication abilities among affected females compared to males

(Lai et al., 2012; McLennan, Lord, & Schopler, 1993). Although methodology and level of impairment and/or strengths of participants with ASD among these studies vary, taken together, research regarding gender differences in presentation suggests that the differences in symptom profiles between females and males with ASD might lead to delayed or missed diagnoses among females in the general population (Van Wijngaarden-Cremers et al., 2014).

### **Gender Disproportionality in the Special Education Population**

In general, disproportionality or disproportionate representation within a given population refers to the representation of a particular group of individuals at a rate different than what is found in the general population. When discussing disproportionality within special education, researchers typically are referring to comparisons between groups of students who are receiving special education services by racial and ethnic backgrounds, socioeconomic status, national origin, and English proficiency (U.S. Department of Education, 2007). For instance, if students from a particular ethnic/racial background (e.g., students who are Black) are identified with a disabling condition at a different rate (i.e., higher or lower) than those from another ethnic/racial background (e.g., students who are White), students who are Black would be considered "disproportionately represented" in that special education category (Williams, 2007). Furthermore, within special education categories, a particular group can be referred to as overrepresented (i.e., identified at a greater rate than other groups) or underrepresented (i.e., identified at a lesser rate than other groups). As states are only required to collect and examine data regarding disproportionality as it relates to race/ethnicity (U.S. Department of Education, 2007), there is limited research investigating potential *gender* disproportionality of students within special education settings, and even less research has focused on gender disproportionality of students classified under AU.

It is well known that many childhood disorders (i.e., ASD, Intellectual Disability [ID], Attention-Deficit/Hyperactivity Disorder [ADHD], Specific Learning Disability [SLD]) are more prevalent



among males compared to females (APA, 2013). However, even when taking into consideration that these childhood disorders are more likely to occur in males, research on gender disproportionality within special education settings has revealed that female students are more likely to be underrepresented across a variety of special education categories (e.g., OHI/ADHD, LD, ID; Achilles et al., 2007; Coutinho & Oswald, 2005). Achilles, McLaughlin, and Croninger (2007) examined a national sample of students receiving special education and found male-to-female ratios for the following IDEIA special education eligibility categories: (a) 2:1 for Specific Learning Disability (SLD); (b) 4:1 for Emotional Disturbance (ED); and (c) 3:1 for ADHD, for which students often are captured within the Other Health Impairment (OHI) category. Similarly, Coutinho and Oswald (2005) found an underrepresentation of females in every state for three disability categories: ID, ED, and SLD, with ED being the most disproportionate and ID being the least disproportionate.

In regard to regional differences within the U.S., the Southern states had higher gender disproportionality (i.e., underrepresentation of females) for SLD than any other region. Further, Oswald et al. (2003) investigated the possibility of gender disproportionality among different racial/ethnic groups. Surprisingly, Oswald, Best, Coutinho, and Nagle (2003) found similar gender disproportionality across racial/ethnic groups, regardless of disability condition (i.e., ID, ED, LD), which suggests that the influences of gender disproportionality impact racial/ethnic groups similarly. As a result of these findings, the authors suggested state policies and procedures be closely monitored for various instances of faulty implementation practices including potential bias (i.e., in referral and identification processes) that may lead to gender disproportionality in special education.

### **Gender Disproportionality in the AU Population**

Although some research has examined gender disproportionality within the context of special education in general, to date researchers have only examined educational placement of students identified under the AU IDEIA category, without

clearly investigating gender disproportionality. White, Scahill, Klin, Koenig, and Volkmar (2007) conducted a study with 92 male and 11 female students with a clinical diagnosis of ASD and identified as AU in order to examine how child factors (e.g., cognitive ability, gender, age) influenced educational placement. The authors found no gender differences with regards to educational placement of male and female students who had ASD diagnoses. Of note, all the students in this sample had “high-functioning ASD.” Given the difference in identification rates between males and females with a full scale IQ greater/less than 70 (Fombonne, 2003), it is plausible that the IQ of the participants in the study could have impacted placement. Moreover, the authors did not specify from which state they gathered data, resulting in a lack of generalizability of findings.

Some international researchers have also examined the educational placements of students with ASD. Yianni-Coudurier et al. (2008) examined characteristics that influenced placement within regular education classes across 66 male and 11 female students with ASD in France. In contrast to White and colleagues’ 2007 U.S. study, Yianni-Coudurier et al. (2008) discovered that female students had higher rates of placement in regular education classes compared to males. It was hypothesized that the female students may have engaged in less disruptive behaviors compared to the male students. Although these two studies contribute to understanding gender disproportionality in special education, both had small sample sizes (e.g.,  $n = 113$  [White et al., 2007],  $n = 77$  [Yianni-Coudurier et al., 2008]), with a particularly small proportion of female students.

May, Cornish, and Rinehart (2014) explored differences in symptom presentation between male ( $N = 32$ ) and female ( $N = 32$ ) students with ASD in Melbourne, Australia and how this could impact educational placement. The authors found significant gender differences in the number of children receiving a one-on-one classroom aid. Specifically, even for students with similar cognitive profiles, only 4 of 25 female students with ASD (16%) received a classroom aid, compared to 13 out of 25 male students with ASD (52%). The authors suggested that teacher concerns which prompt special education referrals and/or

programming may reflect gender differences (i.e., more externalizing behaviors such as hyperactivity in male students; more internalizing symptoms like anxiety in female students). Oswald et al. (2003) highlighted that males may be more likely to demonstrate problem behaviors, which may prompt teachers and other professionals to make referrals to special education across disability type. Similarly, Younger et al. (2005) suggested that male students might engage in behaviors that display apathy towards school, as well as risk-taking behaviors that they might foster an identity acceptable to other male students. Because these behaviors often conflict with the expectations of schools, these male students may be more likely to be referred for a special education evaluation and granted services compared to female students.

### **Current Study Purpose and Aims**

As mentioned previously, research has highlighted gender disproportionality within special education categories across various states and countries and outlined some potential contributing factors to different rates of services between males and females. However, similar reports for the AU educational classification, specifically, are scarce. Further research is needed in this area to help determine whether female students are underrepresented in U.S. public school systems, a situation which would result in less access to needed special education services and supports for girls. Since families of children with autism rely heavily on school-based services (Mire et al., 2015), this is of particular relevance within the ASD population. Moreover, research is also lacking regarding potential fluctuations in gender disproportionality across academic years, an important area of investigation considering that the prevalence rates of ASD diagnosis are steadily increasing across time.

Unfortunately, there is no federal or state requirement to examine special education AU data in terms of gender disproportionality. Further, not every state uses the federal AU criteria (see Florida); instead, states are tasked with determining how they classify students through special education, including students with clinical diagnoses of ASD. Texas, however, uses the IDEIA's (2004) AU criteria. This permits

comparisons between how many children are diagnosed with ASD (i.e., outside of the school system), and how many students are identified as AU (i.e., through the school system) then receive special education services.

The current study had two primary aims. Using the CDC's (2012) male-to-female ASD ratio as a parameter (i.e., 4:1 to 5:1), the first aim of our study was to investigate changes in gender proportions of students educationally classified under the AU category within public schools in Texas across a five-year period. Given that Texas is a culturally and geographically diverse state, the second aim of the study was to explore whether there were differences between rates of male and female students classified under AU across the 20 Texas Education Agency (TEA) regions. Because data indicate that U.S. students identified under AU continues to increase (U.S. Department of Education, 2015), it was hypothesized that there would be greater gender disproportionality across Texas for the 2012–2013 academic year compared to the 2008–2009 academic year. Additionally, due to the size of the state and diversity of its residents, it was hypothesized that there would be differences in gender disproportionality across TEA Regions.

## **Method**

### **Participants**

Data were analyzed for Texas public school students between the ages of 3 and 21 who had an IDEIA (2004) AU educational classification (i.e., receiving special education services in Texas under this category). Data were analyzed for the 2008–2009 through the 2012–2013 academic years, inclusive ( $N = 171,434$ ). Data included 5,926 school districts, which represented the number of districts with data available for the academic years used in analyses (i.e., 2008–2013). Due to the rezoning of districts from 2008–2013, 22 districts that existed in previous years did not exist in later years and were not included in the final sample. Thus, the final sample included a total sample of 144,174 students from 1,196 districts.

### **Procedures**

Prior to data analysis, appropriate IRB approval was obtained to use TEA public record data from

the 2008–2013 academic school years. All databases are publicly available and requests for data can be made through [http://tea.texas.gov/Reports\\_and\\_Data/](http://tea.texas.gov/Reports_and_Data/). The following data were requested: number of male and female students, by district, who were being served under each of the 13 IDEIA special education

eligibility classifications for each academic year between 2008–2009 and 2012–2013. Upon receipt of the data, we organized the district-level data to align with the 20 TEA Regions (see Figure 1), which allowed exploration of differences across geographical areas (and major cities) of Texas.

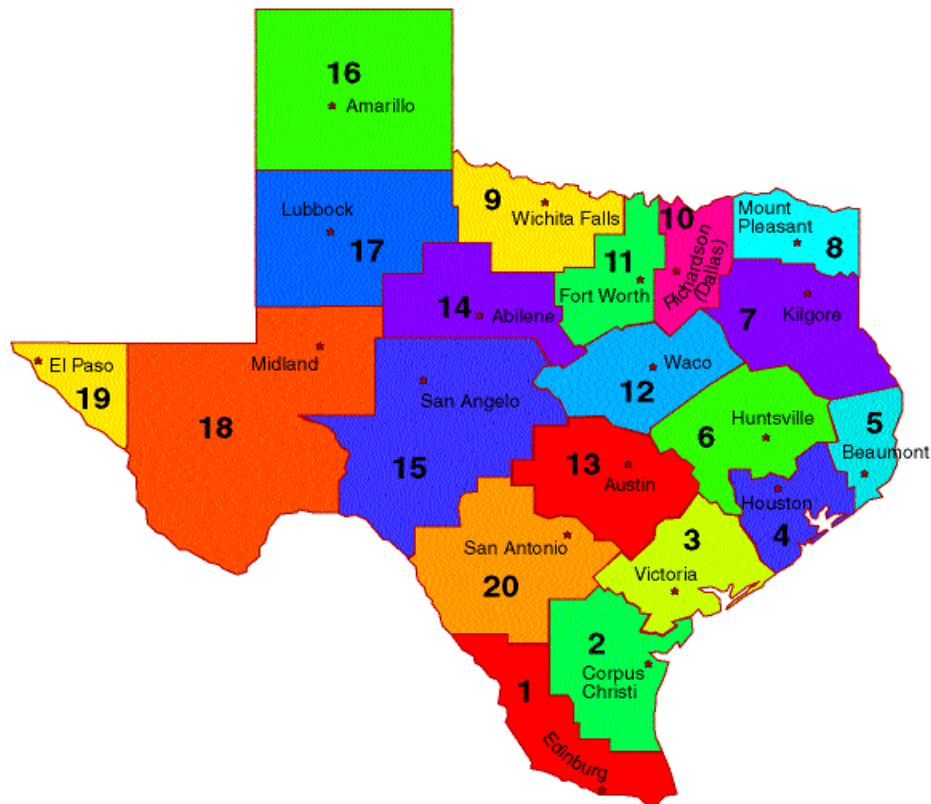


Figure 1. Map of Texas Education Agency (TEA) Regions

We used a combination of descriptive and inferential statistics. Before addressing the primary aims of the study (i.e., gender proportion differences across years; gender proportion differences between TEA Regions), we used descriptive statistics to calculate: (a) male-to-female ratios of students with AU in Texas from 2008–2013 (see Table 1), and (b) male-to-female ratios of students with AU in Texas across TEA Regions from 2008–2013 (see Table 2). To address the two primary aims of the study, chi-square goodness of fit tests were conducted to determine

whether a ratio of male-to-female students with AU differed across academic school years and/or between TEA Regions. For interpretation of chi-square results, we used a Bonferroni correction to reduce the likelihood of Type I errors arising from multiple comparisons, such that  $\alpha = .01$  for examining the ratio of male-to-female students with AU across five academic years and  $\alpha = .0025$  for examining the ratio of male-to-female students with AU in each of the 20 TEA Regions across five academic years.

**Table 1** Male-to-female Ratios of Students with AU in Texas from 2008–2012

Academic Year	Male-to-female AU Ratio in Texas
2008–2009	5.0:1
2009–2010	5.1:1
2010–2011	5.2:1
2011–2012	5.2:1
2012–2013	5.2:1

**Table 2** Male-to-female Ratios of Students with AU in Texas across TEA Regions from 2008–2012

TEA Region	2008–2009	2009–2010	2010–2011	2011–2012	2012–2013
1	4.58:1	4.80:1	4.65:1	4.72:1	4.62:1
2	4.86:1	4.61:1	4.37:1	4.88:1	5.08:1
3	3.78:1	3.23:1*	3.01:1*	2.94:1*	3.60:1
4	5.48:1	5.44:1	5.52:1	5.53:1	5.42:1
5	4.38:1	4.56:1	4.99:1	4.97:1	5.05:1
6	4.77:1	4.64:1	4.78:1	5.07:1	4.62:1
7	4.15:1	4.60:1	4.25:1	4.30:1	4.27:1
8	4.61:1	4.66:1	4.10:1	4.73:1	4.81:1
9	4.77:1	5.30:1	5.17:1	4.49:1	3.91:1
10	5.45:1	5.51:1	5.57:1	5.55:1	5.52:1
11	5.45:1	5.42:1	5.40:1	5.53:1	5.40:1
12	4.61:1	4.57:1	4.98:1	4.87:1	4.87:1
13	4.75:1	4.90:1	5.01:1	5.03:1	5.12:1
14	3.41:1	4.85:1	5.42:1	4.52:1	4.75:1
15	6.68:1	5.70:1	6.49:1	8.98:1	8.23:1
16	4.32:1	3.59:1	4.23:1	4.40:1	4.77:1
17	4.49:1	4.46:1	4.52:1	4.82:1	4.86:1
18	5.16:1	5.14:1	5.46:1	6.23:1	5.35:1
19	5.44:1	5.91:1	6.08:1	6.54:1	6.19:1
20	4.65:1	4.83:1	5.00:1	4.92:1	4.89:1

Note: \* = statistical significance

## Results

### Aim 1: Differences in Identification across Time

When examining the ratio of male-to-female students with AU across academic years, our findings revealed an overall male-to-female ratio of 5.0:1 of students identified in Texas under AU in the 2008–2009 academic year and a 5.2:1 ratio in the 2012–2013 academic year (see Table 1). A chi square goodness of fit test was performed to determine whether the ratio of male-to-female students with AU differed across years. The male-to-female ratio *did not* differ across the five academic school years: 5.0:1 ratio in 2008-2009, ( $\chi^2$

(1,  $N = 27035$ ) = 5.66,  $p = 0.02$ ); 5.1:1 ratio in 2009-2010, ( $\chi^2$  (1,  $N = 30645$ ) = 3.37,  $p = 0.07$ ); 5.2:1 ratio in 2010-2011, ( $\chi^2$  (1,  $N = 34120$ ) = 0.96,  $p = 0.33$ ); 5.2:1 ratio in 2011-2012, ( $\chi^2$  (1,  $N = 37762$ ) = 0.08,  $p = 0.77$ ); and 5.2:1 ratio in 2012-2013, ( $\chi^2$  (1,  $N = 41648$ ) = 1.23,  $p = 0.27$ ). This indicates that there were no significant differences between the male-to-female ratio of students identified under AU from the 2008–2009 to the 2012–2013 academic years.

Consistent with higher rates of ASD clinical diagnoses in males, more male students were served under the AU category in Texas during all academic years examined. Moreover, when considering the

state as a whole, the male-to-female ratios remained relatively stable across all five years, and the results are comparable to the 4:1 to 5:1 male-to-female ratio of individuals with ASD in the general population (CDC, 2012). In regard to the percentage of female students classified with AU in Texas for the 2008–2009 academic year, 16.53% of students with an AU classification were female, with similar percentages seen in the 2009–2010 (16.38%), 2010–2011 (16.19%), 2011–2012 (16.05%), and 2012–2013 (16.20%) academic school years.

### **Aim 2: Differences across TEA Regions and Time**

Variation in male-to-female ratios of AU classification was further examined to determine whether the observed ratio of males-to-females with AU in each TEA Region differed significantly from Texas's overall ratio of males-to-females with AU (i.e., consistent with the CDC [2012] ratio, as described in Aim 1). For each Region, chi square goodness of fit tests were performed for each academic school year by Region to examine this difference, which refers to whether the observed frequencies are statistically different from the expected frequencies.

For the 2008-2009 and the 2012-2013 academic school years, there were no statistically significant differences between any Regions' male-to-female ratio of students with AU and the ratio for the state of Texas. However, there *were* significant differences found within the academic years of 2009-2010, 2010-2011, and 2011-2012 (see Table 2). Within each of these years, chi-square results indicated that the male-to-female ratio of students with AU in TEA Region 3 was significantly *lower* than the expected (i.e., State's) ratio: 3.23:1 ratio in 2009-2010, ( $\chi^2(1, N = 244) = 10.50, p < .0025$ ); 3.01:1 ratio in 2010-2011, ( $\chi^2(1, N = 257) = 15.29, p < .0025$ ); and 2.94:1 ratio in 2011-2012, ( $\chi^2(1, N = 266) = 17.40, p < .0025$ ). This indicates that within TEA Region 3 for these three academic years, males with AU may have been disproportionately *underrepresented* and/or females with AU were *overrepresented*. For other Regions of the State, no statistically significant differences were found

between the Regions in terms of male-to-female ratios in AU.

### **Discussion**

The purpose of this study was to examine gender representation among students receiving special education supports and services under the AU educational classification, over time and across TEA Region. Data from a large and diverse state (Texas) were analyzed across 20 TEA Regions for the 2008–2009 through the 2012–2013 academic school years. Prior studies have only examined gender disproportionality for other IDEIA (2004) special education eligibility categories such as ID, ED, and LD (Couthino & Oswald, 2005) without including AU. Although the current study focused exclusively on students classified under AU in a single state (Texas), this study was the first to examine the issue of gender representation among students classified with AU in any U.S. state. Findings suggest that, overall, AU classification in Texas has risen proportionally among male and female students and that the male-to-female ratio of AU educational classification is, overall, comparable to the gender ratio found in the general population (i.e., 4:1 to 5:1; CDC, 2012). In other words, the results of this study suggest that at least in one state (e.g., Texas) the male-to-female ratio of students classified under AU is similar to the diagnostic prevalence of males and females with ASD. Interestingly, this conflicts with our initial hypothesis suggesting there would be greater disproportionality across Texas for the 2012-2013 academic school year. Further, this implies that female and male students classified under AU in Texas are comparably likely to be eligible for special education supports and services.

Across the five years examined, 95% of TEA Regions in Texas maintained male-to-female ratios of AU that did not differ significantly from the expected male-to-female ratio (i.e., Texas's AU gender ratio, which was consistent with the CDC's [2012] rates). The only statistically significant differences were found in one TEA Region. In Region 3 (near Victoria; see Figure 1), the difference in male-to-female ratio of students with AU was statistically significantly different from that of the State's. Although this finding supports our

hypothesis that there would be some gender disproportionality across TEA Regions, it was surprising that males may have been *underrepresented* and/or females were *overrepresented* in this TEA Region during three out of the five academic school years that were examined. Further, it suggests that more females are receiving special education services in this Region compared to males.

Importantly, visual inspections of the male-to-female ratios across Regions and school years yield figures that appear to deviate from the overall Texas ratio (which is consistent with the CDC's rates). For instance, as seen in Table 2, TEA Region 15 (San Angelo) and TEA Region 19 (El Paso) maintained male-to-female ratios across five academic school years that appear to be different than the 4:1 to 5:1 ratio (CDC, 2012), but inferential analyses did not indicate that these deviations reached statistical significance. Nonetheless, these patterns may indicate something of clinical (i.e., practical) significance in terms of a need to monitor male-to-female ratios and attend carefully to eligibility processes to identify factors that may influence disproportionality. For these reasons, our findings suggest that the considerations offered below may be applicable not only to Regions where *statistically* significant differences were found, but also Regions where potentially *clinically* significant differences were noted.

The difference in some Regions' AU gender ratios from the States' ratios may be due, in part, to a variety of factors that also influence disproportionality in special education in general, including race/ethnicity, linguistic barriers, as well as socioeconomic status of students in different areas of Texas. Texas is a large state, comprised of 20 different TEA Regions, each with distinct populations; these factors may potentially influence special education eligibility evaluation proceedings such as gathering parent and teacher report, which are critical to AU identification. More specifically, within special education settings, raters (teachers and parents) may vary in their reports of ASD symptoms of male and female students as well as students from various cultural groups (Blacher et al. 2014), which may contribute to the differences seen in some TEA Regions.

In terms of race/ethnicity, research has found

fewer students identified under AU in Texas school districts that served high percentages of Hispanic students (Palmer et al., 2010). Further, considering that deficits in social communication is a hallmark characteristic of ASD, students recently learning English may be at-risk for being misidentified with AU due to clinicians misinterpreting deficits of expressive and receptive language skills. Research has demonstrated that students are more likely to be identified with AU when their primary language was not English (Estrem & Zhang, 2010). Of note, ASD is one of the few developmental disabilities where a positive correlation exists between socioeconomic status (SES) and diagnostic prevalence (Durkin et al., 2010). Families with low SES are less likely to have health insurance, less access to regular medical care, and are more likely to have difficulty accessing special care (Palmer et al., 2010). Moreover, decades of research highlight how students from low SES groups are overrepresented in the special education population (Morrier & Gallagher, 2012). In conjunction with differences in how females with ASD manifest symptoms, these factors may not explain the cause of gender disproportionality; however, they may exacerbate instances of potential gender disproportionality and the AU classification within certain TEA Regions.

Further, some authors have suggested that students' age and state-specific policies may impact gender disproportionality in special education, as well. For example, Phipps (1982) found that among children ages 5 to 11, special education referral rates of male students were significantly higher than for female students, whereas no gender differences were found among referral rates for children 4-years-old and younger and 12-years-old and older. With regards to state policies, though IDEIA (2004) is a federal law, states create implementation regulations (i.e., state-level interpretation of IDEIA's AU criteria may vary). Further, the referral, identification, and placement policies and procedures often vary within the same state (i.e., by TEA Region) or even by school district (Merrell & Walker, 2004). Regardless of the potential reasons behind gender disproportionality, overrepresentation of males and/or underrepresentation of females in special education

can translate into more resources being allocated to male students.

### **Limitations**

Several limitations should be considered with regard to the findings of this study. First, the data analyzed did not differentiate whether students who met special education eligibility for AU carried only this classification or carried one or more additional classifications (i.e., ID, ED, OHI). Second, the Family Educational Rights and Privacy Act (FERPA) requires school districts to protect students' confidentiality by not reporting data when a special education eligibility category contains less than 5 students; in these instances, districts enter the median number of students available (i.e., 2.5 students), resulting in a potential under-estimation or over-estimation of male-to-female ratios across Texas and TEA Regions. Lastly, data obtained from TEA included students 3 to 21 years of age but were not broken down by grade level; it may be that age-related differences in identification of females and males with AU exist.

### **Future Directions**

This study represents an initial investigation of special education gender disproportionality within the classification of AU in a single state. As this was the first study of its kind to address gender disproportionality within the AU population as well as the AU population within Texas, we hoped to use this study as a springboard to generate research questions and hypotheses for future studies related to gender disproportionality and students identified as AU. Though overall results indicated alignment of AU identification (in Texas) with clinical diagnostic prevalence of ASD, it is plausible that different results would be found across other U.S. regions and particularly between U.S. states. Additionally, child-level data may allow researchers to better understand the potential impact of individual characteristics on special education eligibility (and ultimately placement, service, and supports). More in-depth investigation into how

race/ethnicity, language barriers and socioeconomic status impact students classified within the AU category is needed in order to determine if there is a differential effect on how male and female students are classified within this special education category. Another area to consider is exploring disproportionality across ages and/or grades, particularly investigating whether differences in gender of students identified under AU varies across grade levels and age ranges (i.e., preschool, elementary school, middle school, high school). Specifically, this information would aid in determining if gender disproportionality of certain TEA regions is limited to certain developmental time points. As ASD is a life-long disorder, district personnel may be in need of more training in screening children at specific ages and developmental milestones. These gaps in the gender disproportionality research have yet to be investigated.

With regards to clinical practice, school districts can support gender-equitable education by, first, being aware of the possibilities of imbalance, and second, by evaluating current practices to ensure their referral and evaluation practices do not contribute to gender disproportionality in students deemed eligible for school-based services. As part of the identification process, school district personnel could review the research regarding the different ways in which ASD symptoms manifest differently in females (Kirkovski, 2013; see Table 3). For instance, play behavior typically presents as atypical in males with ASD. However, females may present with play behavior that appears more typical, including restricted interests in the domain of social behavior and echolalia of social actions. Additionally, persons contributing to school-based evaluations should be aware of potential cultural and linguistic biases evaluators and raters (i.e., teachers, parents) may bring into the assessment, as well as ensure that selected instruments have been validated in a representative sample.

**Table 3** *Examples of Differences of ASD Symptomatology of Males and Females*

Domain	Males	Females
Communication	Inconsistent findings; suggested greater impairment in non-verbal communication (Park et al., 2012)	Inconsistent findings; suggested greater overall communication deficits (Hartley & Sikora 2009; Carter et al., 2007); suggested increased use of echolalia (Kopp & Gillberg, 1992; Andersson et al., 2012); suggested greater socio-communicative skills (McLennan et al., 1993; Lai et al., 2012)
Play Behavior	Play behavior typically presents as atypical	More appropriate compared to males (Lord et al., 1982; McLennan et al., 1993; Tsai & Beisler, 1983); heightened imaginative play ability (Knickmeyer et al., 2008); can demonstrate RI of social behaviors including echolalia of actions which appears as appropriate play behavior—this masks true social deficits (Kopp & Gillberg, 1992; Attwood, 2007)
Developing Friendships	Present with deficits in social function more readily identifiable by diagnostic criteria and measures	Greater impairment in social functioning and interaction (Billstedt et al., 2007; Lord et al., 1982) greater difficulty developing and maintaining friendships, particularly for older females (McLennan et al., 1993; Holtmann et al., 2007); present with a phenotype different than the diagnostic criteria
Restricted and Repetitive Behaviors and Interests	More severely impacted in this domain, RRBI symptoms more clinically identifiable than those exhibited by females; demonstrate less socially appropriate RI	Females with an average DQ less likely to exhibit RRBI (Sipes et al., 2011); symptoms less clinically identifiable than those exhibited by males (Kopp & Gillberg, 1992); may develop RI in the realm of social interaction (Attwood, 2007; Kopp & Gillberg, 1992)

Note: DQ=Developmental Quotient; RRBI=repetitive, restricted behavior and interests; RI=Restricted Interests

### Implications

Several important implications of these findings should be considered. As noted earlier, IDEIA AU educational classification and clinical diagnosis of ASD are results of distinct processes: AU classification focuses on the presence of a disabling condition and a related educational need,

with the goal of service determination within the public school system; whereas clinical diagnosis of ASD is driven by determining whether symptoms manifested are consistent with the criteria outlined in the DSM-5 (APA, 2013). Findings within this study indicate, however, that despite different criteria and goals, gender ratios of students with AU



in Texas, overall, are consistent with gender ratios of ASD in the general population. These results are encouraging and suggest that female and male students classified under AU in Texas, overall, are likely to receive appropriate special education eligibility and related services.

Importantly, however, our study did identify gender disproportionality among students with AU classifications in some areas of the state. This is particularly important when considering the demographic characteristics of students within the various Regions where disproportionality was indicated. That socioeconomic, language, and/or racial/ethnic diversity could be contributing factors to gender disproportionality in special education has social justice implications (i.e., potential bias in evaluation/eligibility processes). Regardless of the reasons, however, gender disproportionality in special education may result in an imbalance in allocation of school-based resources to male and female students who are identified under AU.

## References

- Achilles, G. M., McLaughlin, M. J., & Croninger, R. G. (2007). Sociocultural correlates of disciplinary exclusion among students with emotional, behavioral, and learning disabilities in the SEELS national dataset. *Journal of Emotional and Behavioral Disorders, 15*(1), 33–45.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.
- Andersson, G. W., Gillberg, C., & Miniscalco, C. (2012). Pre-school children with suspected autism spectrum disorders: Do girls and boys have the same profiles? *Research in Developmental Disabilities, 34*, 413–422.
- Attwood, T. (2007). *The complete guide to asperger's syndrome*. London: Jessica Kingsley Publishers.
- Billstedt, E., Gillberg, I. C., & Gillberg, C. (2007). Autism in adults: Symptom patterns and early childhood predictors. Use of the DISCO in a community sample followed from childhood. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 48*(11), 1102–1110.
- Blacher, J., Cohen, S. R., & Azad, G. (2014). In the eye of the beholder: Reports of autism symptoms by Anglo and Latino mothers. *Research In Autism Spectrum Disorders, 8*(12), 1648–1656. doi:10.1016/j.rasd.2014.08.017
- Carter, A. S., Black, D. O., Tewani, S., Connolly, C. E., Kadlec, M. B., & Tager-Flusberg, H. (2007). Sex differences in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*(1), 86–97.
- CDC. Prevalence of autism spectrum disorders– Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. MMWR 2012: 61 (No. SS–3).
- Constantino, J. N., & Charman, T. (2012). Gender bias, female resilience, and the sex ratio in autism. *Journal of The American Academy of Child & Adolescent Psychiatry, 51*(8), 756–758. doi:10.1016/j.jaac.2012.05.017
- Coutinho, M. J., & Oswald, D. P. (2005). State variation in gender disproportionality in special education findings and recommendations. *Remedial and Special Education, 26*(1), 7–15.
- Durkin, M. S., Maenner, M. J., Meaney, F. J., Levy, S. E., DiGuseppi, C., Nicholas, J. S., & ... Schieve, L. A. (2010). Socioeconomic inequality in the prevalence of autism spectrum disorder: Evidence from a U.S. cross-sectional study. *Plos ONE, 5*(7), 1–8. doi:10.1371/journal.pone.0011551
- Dworzynski, K., Ronald, A., Bolton, P., & Happe', F. (2012). How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders? *Journal of the American Academy of Child and Adolescent Psychiatry, 51*, 788–797.
- Estrem, T. t., & Shiju, Z. (2010). Prevalence and disproportionality of autism spectrum disorders among English language learners in Minnesota. *Multiple Voices For Ethnically Diverse Exceptional Learners, 12*(2), 5–20.
- Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders, 33*(4), 365–382.
- Hartley, S. L., & Sikora, D. M. (2009). Sex differences in autism spectrum disorder: An examination of developmental functioning, autistic symptoms, and coexisting behavior problems in toddlers. *Journal of Autism and Developmental Disorders, 39*(12), 1715–1722. doi:10.1007/s10803-009-0810-8
- Holtmann, M., Bo'lte, S., & Poustka, F. (2007). Autism spectrum disorders: Sex differences in autistic behavior domains and coexisting psychopathology. *Developmental Medicine and Child Neurology, 49*(5), 361–366.
- Individuals With Disabilities Education Act, 20 U.S.C. § 1400 (2004).
- Keen, D. V., Reid, F. D., & Amone, D. (2010). Autism, ethnicity, and maternal immigration. *British Journal of Psychiatry, 196*, 274–281.
- Kirkovski, M., Enticott, P. G., & Fitzgerald, P. B. (2013). A review of the role of female gender in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 43*(11), 2584–2603.
- Knickmeyer, R. C., Wheelwright, S., & Baron-Cohen, S. B. (2008). Sex-typical play: Masculinization/defeminization in girls with an autism spectrum condition. *Journal of Autism and Developmental Disorders, 38*, 1028–1035.
- Kopp, S., & Gillberg, C. (1992). Girls with social deficits and learning problems: Autism, atypical asperger syndrome or a variant of these conditions. *European Child and Adolescent Psychiatry, 1*(2), 89–99.
- Lai, M. C., Lombardo, M. V., Ruigrok, A. N. V., Chakrabarti, B., Wheelwright, S. J., Auyeung, B., et al. (2012).

- Cognition in males and females with autism: Similarities and differences. *PLoS ONE*, 7(10), e47198.
- Lord, C., Schopler, E., & Revicki, D. (1982). Sex differences in autism. *Journal of Autism and Developmental Disorders*, 12(4), 317–330.
- May, T., Cornish, K., & Rinehart, N. (2014). Does gender matter? A one-year follow-up of autistic, attention, and anxiety symptoms in high-functioning children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44, 1077–1086.
- McLennan, J. D., Lord, C., & Schopler, E. (1993). Sex differences in higher functioning people with autism. *Journal of Autism and Developmental Disorders*, 23, 217–227.
- Merrell, K. W., & Walker, H. M. (2004). Deconstructing a definition: Social maladjustment versus emotional disturbance and moving the EBD field forward. *Psychology in the Schools*, 41, 899–910.
- Mire, S. S., Raff, N. S., Brewton, C. M., & Goin-Kochel, R. P. (2015). Age-related trends in treatment use for children with autism spectrum disorder. *Research In Autism Spectrum Disorders*, 15–16, 29–41. doi:10.1016/j.rasd.2015.03.001
- Morrier, M. J., & Gallagher, P. A. (2012). Racial disparities in preschool special education eligibility for five southern states. *The Journal Of Special Education*, 46(3), 152–169. doi:10.1177/0022466910380465
- Oswald, D. P., Best, D. P., Coutinho, M. J., & Nagle, H. A. L. (2003). Trends in the special education identification rates of boys and girls: A call for research and change. *Exceptionality*, 11(4), 223–237.
- Palmer, R. F., Walker, T., Mandell, D., Bayles, B., & Miller, C. (2010). Explaining low rates of autism among Hispanic school children in Texas. *American Journal of Public Health*, 100(2), 270–272.
- Phipps, P. M. (1982). The LD learner is often a boy—why? *Academic Therapy*, 17(4), 425–430.
- Sipes, M., Matson, J. L., Worley, J. A., & Kozlowski, A. M. (2011). Gender differences in symptoms of autism spectrum disorders in toddlers. *Research in Autism Spectrum Disorders*, 5(4), 1465–1470.
- Solomon, M., Miller, M., Taylor, S. L., Hinshaw, S. P., & Carter, C. S. (2012). Autism symptoms and internalizing psychopathology in girls and boys with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 42, 48–59.
- Szatmari, P., Liu, X. Q., Goldberg, J., Zwaigenbaum, L., Paterson, A.D., Woodbury-Smith, M., et al. (2012). Sex differences in repetitive stereotyped behaviors in autism: Implications for genetic liability. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 159B(1), 5–12.
- Texas Autism Supplement, Commissioners Rules TAC §89.1055 (2007).
- Tsai, L., & Beisler, J. (1983). The development of sex differences in infantile autism. *The British Journal of Psychiatry*, 142(4), 373.
- U.S. Department of Education. (2007). *IDEA Regulations: Disproportionality and Overidentification*, 2007. Retrieved from: <http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C7%2C>
- U.S. Department of Education, National Center for Education Statistics. (2015). *Digest of Education Statistics, 2013* (NCES 2015–011), Table 204.30. Retrieved from: <https://nces.ed.gov/fastfacts/display.asp?id=64>
- Van Wijngaarden-Cremers, P. M., van Eeten, E., Groen, W. B., Van Deurzen, P. A., Oosterling, I. J., & Van der Gaag, R. J. (2014). Gender and age differences in the core triad of impairments in autism spectrum disorders: A systematic review and meta-analysis. *Journal Of Autism And Developmental Disorders*, 44(3), 627–635. doi:10.1007/s10803-013-1913-9
- White, S. W., Scahill, L., Klin, A., Koenig, K., & Volkmar, F. R. (2007). Educational placements and service use patterns of individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37(8), 1403–1412.
- Williams, P. (2007, January). Disproportionality and overrepresentation (Module 5). *Building the Legacy: IDEA 2004 Training Curriculum*. Washington, DC: National Dissemination Center for Children with Disabilities. Available online at: [www.nichcy.org/training/contents.asp](http://www.nichcy.org/training/contents.asp)
- Yianni-Coudurier, C., Darrou, C., Lenoir, P., Verrecchia, B., Assouline, B., Ledesert, B., Michelon, C., Pry, R., Aussilloux, C. & Baghdadli, A. (2008). What clinical characteristics of children with autism influence their inclusion in regular classrooms? *Journal of Intellectual Disability Research*, 52(10), 855–863.
- Younger, M., Warrington, M. Gray, J., Rudduck, J., McLellan, R., Bearne, E., Kershner, R. & Bricheno, P. (2005). *Raising boys' achievement. Research report RR636*. Annesley, Nottingham: Department for Education and Skills.

Article

---

## Scholarly Productivity of School Psychology Faculty Members in Specialist-Level Programs: 2002-2011

Jeff Laurent and Elizabeth Runia  
*Western Illinois University*

The scholarly productivity of school psychology faculty members in specialist-level only programs was examined. Information was gathered from the School Psychology Program Information portion of the website for the National Association of School Psychologists. A total of 136 specialist-level only school psychology programs were identified. Authorship credit was computed for faculty members within each program based on journal articles published from 2002-2011. The 25 individuals with the highest authorship credit were identified. The number of journal publications across program faculty members was averaged, and the top 25 programs were identified. Finally, journals in which the top 25 faculty members published articles were identified. Observations about issues that arose while collecting data are presented. The study presents a snapshot of scholarship among specialist-level only programs and their faculty members during the 10-year period from 2002-2011.

**Keywords:** school psychology, scholarly productivity, specialist-level programs

An area of academic school psychology that has been surveyed from time to time has been the scholarly productivity of faculty. As Joy (2006) notes, "Scholarly productivity is an important determinant of academic success, utilized in crucial personnel decisions such as hiring, promotion, and awarding tenure, as well as in determining an academic's prestige among disciplinary peers" (p. 346). In addition to benefitting individuals, the scholarly productivity of faculty also affects programs and their reputation. For example, Carper and Williams (2004) suggested that scholarly productivity might influence the decision-making process of students and potential faculty members considering a position with a program. In effect, scholarly productivity might act as a recruitment tool. These authors also suggested that information regarding scholarly productivity could help school psychology programs seek institutional support from their universities, and act as a yardstick for program improvement.

Early researchers examined faculty scholarly productivity in terms of the school psychology literature. For example, Webster, Hall, and Bolen (1993) examined the institutional affiliations of

authors who published in five school psychology journals from 1985-1991. Three were well-established journals (i.e., *Journal of School Psychology*, *School Psychology Review*, *Psychology in the Schools*), while the other two were relatively new (i.e., *Professional School Psychology* [now *School Psychology Quarterly*], *Journal of Psychoeducational Assessment*). Although these investigators considered agencies and school districts, the top 50 rankings that Webster et al. reported represented universities.

---

**Author Note:** The authors thank Kelli Abell, Felicia Austin, Rachel Bush, Haley Humes, Brian Mendoza, Ashley Nelson, Melissa Nemece, Kaitlyn O'Riley, Nicole Palumbo, Paige Peterson, and Jessica Satterlee for their assistance in data collection.

A version of this paper was presented as a poster at the 48th Annual Convention of the National Association of School Psychologists, New Orleans, LA, February, 2016. Correspondence concerning this manuscript should be addressed to Jeff Laurent, Department of Psychology (WG 100), Western Illinois University, 1 University Circle, Macomb, IL, 61455. Contact: jl-laurent@wiu.edu.

The rankings were presented based on number of publications and on a program productivity formula.

Levinson, Barker, and Lillenstein (1994) took the work of Webster et al. (1993) one step further. These researchers examined the degree to which the 50 universities identified by Webster et al. supported and rewarded faculty scholarship. Specifically, Levinson et al. categorized the 50 universities ranked by Webster et al. using classification systems devised by the Carnegie Foundation and the American Association of University Professors. They then ranked universities within these classification systems based on number of publications and on a program productivity formula. Levinson et al. reported that doctoral institutions had higher productivity ratings than did institutions classified as comprehensive universities, at least when it came to publishing in school psychology journals.

Little (1997) expanded this line of research when he added a sixth school psychology journal to the mix, *School Psychology International*. He collected author data from 1987-1995, and reported number of publications, number of first-authored articles, and authorship credit ratings. Among the findings reported were rankings of the top 50 authors in the school psychology literature for 1987-1995, and rankings of the degree granting universities of the top authors. Subsequently, a series of studies extended the work of Little. For example, Davis, Zanger, Gerrard-Morris, Roberts, and Robinson (2005) reviewed the productivity and collaboration of authors in four school psychology journals from 1991-2003. The journals they reviewed were *Journal of School Psychology*, *School Psychology Review*, *Psychology in the Schools*, and *School Psychology Quarterly*. These investigators defined productivity by the number of articles authored. For the 20 most productive authors, Davis et al. then collected information concerning collaboration (i.e., publications with more than one author, including graduate student authors). In a second study, Roberts, Davis, Zanger, Gerrard-Morris, and Robinson (2006) used data collected from 1996-2005, and excluded the *Journal of Psychoeducational Assessment* from the original list of school psychology journals examined by Little, resulting in five journals (*Journal of School Psychology*, *School Psychology Review*,

*Psychology in the Schools*, *School Psychology International*, and *School Psychology Quarterly*). Roberts et al. examined the number of publications and authorship credit ratings, and generated a list of top 50 authors for both variables. As one might suspect, the lists generated by Little (1997) and Roberts et al. (2006) reflect the differences that occur naturally over time as people retire or develop in their careers.

Several researchers provide insight into the scholarship of faculty members affiliated with school psychology doctoral programs accredited by the American Psychological Association (APA). For example, Carper and Williams (2004) examined the record of article publications of faculty at APA-accredited doctoral programs in school psychology from 1995-1999; programs accredited as of September 2000 were included in the study. Information for core school psychology faculty members from 53 programs was collected using the PsycINFO database. These authors divided journals into two categories, primary (i.e., *Journal of School Psychology*; *School Psychology International*; *School Psychology Review*; *School Psychology Quarterly*; *Psychology in the Schools*) and secondary (i.e., all other educational/psychological journals indexed in the PsycINFO database). The rankings that Carper and Williams presented of programs based on authorship credit ratings and number of publications provided a snapshot of faculty productivity during the late 1990's.

Wagner, Lail, Viglietta, and Burns (2007) extended the work of Carper and Williams by examining the scholarly productivity of faculty members at 56 APA-accredited school psychology programs from 2000-2005. These researchers used the PsycINFO database to collect data on number of faculty publications. Wagner et al. presented the rankings of the top 20 faculty members by number of articles published, and the top 10 programs based on median publications by faculty. In addition, these investigators noted the mean number of publications by Carnegie Foundation research classification.

Kranzler, Grapin, and Daley (2011) also replicated the work of Carper and Williams (2004) in their examination of the productivity of the core school psychology faculty members of 59 APA-accredited programs from 2005-2009; programs

accredited as of May 2010 were included in the study. In order to compare findings with those reported by Carper and Williams, Kranzler et al. examined authorship credit ratings and number of publications for the 59 programs using the PsycINFO database. Like Wagner et al. (2007), this group then examined the scholarly productivity from 2005-2009 of individual faculty members of the 59 APA-accredited school psychology programs (Grapin, Kranzler, & Daley, 2013). Specifically, Grapin et al. provided rankings of the top 25 individuals based on authorship credit ratings and number of publications. Together, the work of Carper and Williams (2004), Wagner et al. (2007), and Kranzler, Grapin, and Daley (Grapin et al., 2013; Kranzler et al., 2011) provides another snapshot of school psychology, specifically, of faculty and program productivity from 1995-2009.

Summarizing the research conducted on the scholarly productivity of faculty members in school psychology, one sees that early studies included a wide range of participants (i.e., all school psychology faculty), but a narrow range of journals (i.e., 4-6 school psychology journals). More recent studies reversed this trend, focusing on a narrow range of participants (i.e., faculty of APA-accredited programs), but a broad range of journals (i.e., those included in the PsycINFO database). The current research examined a group of school psychology faculty members that has not been studied specifically – those who are employed at institutions whose only school psychology program is at the specialist level. Although recent interest in scholarly productivity of school psychology faculty members has focused on APA-accredited doctoral programs, specialist programs outnumber these programs and doctoral programs, in general. Of the 238 school psychology programs in the United States listed by Miller (2008), 138 were specialist-only programs, 22 were doctoral-only programs, and 78 had both doctoral and specialist programs. The distribution of programs listed on the National Association of School Psychologists (NASP) website is very similar: 140 specialist-only, 25 doctoral-only, 79 doctoral and specialist programs ([apps.nasponline.org/standards-and-certification/graduate-education/index.aspx](http://apps.nasponline.org/standards-and-certification/graduate-education/index.aspx)). The APA website lists 64 APA-accredited doctoral programs ([apa.org/ed/accreditation/programs/index.aspx](http://apa.org/ed/accreditation/programs/index.aspx)).

The fact that institutions offering specialist-only programs have not been considered in recent studies of scholarly productivity (e.g., Grapin et al., 2013; Kranzler et al., 2011), even though they are the most numerous type of school psychology program, suggests it is time to examine these programs.

Also, no previous study has focused specifically on scholarly productivity of faculty members in specialist-only programs. Although Webster et al. (1993) initially cast a broad net, the current trend is to examine APA-accredited programs exclusively (Grapin et al., 2013; Kranzler et al., 2011; Wagner et al., 2007). The focus on APA-accredited doctoral programs excludes information on the scholarly productivity of hundreds of school psychology faculty members. Clearly, there are differences between specialist and doctoral programs that impact scholarly productivity. The curriculum of specialist-only programs results in a three-year program with tight timelines for completion. As an outcome, these programs tend to emphasize practice over research. On the other hand, doctoral programs allow more flexibility with respect to completion of the curriculum. This allows not only for specialization in various areas of training (e.g., neuropsychology, autism, etc.), but more expectations and opportunities for research. Graduate students may choose their doctoral programs based on faculty members' research interests. This creates a mentoring relationship that benefits both faculty and graduate student, in terms of scholarship. Doctoral students interested in pursuing a career in academia are socialized in the importance of publications. Faculty members' teaching loads in doctoral programs typically reflect attempts by administration to accommodate pursuit of research programs, mentoring graduate student research, and supervision of dissertations with the expectation that reduced teaching loads will result in increased scholarly productivity. These same resources, expectations, and opportunities may not exist for faculty members in specialist-only programs, and this may be reflected in their scholarly productivity. We examined this issue by gathering information from specialist-only programs.

The current study was interested in many of the same questions asked by previous research in the

area. Specifically, who were the most productive faculty members, in terms of journal articles, among those in school psychology programs that only offered the specialist-level degree? Which programs were the most productive, in terms of journal articles? In what journals did the most productive faculty members publish? All of these questions were asked in the context of the years 2002-2011.

## Method

### Sample

Information was gathered on specialist-only programs in school psychology from the School Psychology Program Information portion of the NASP website during the Fall 2012 semester. For the purposes of the current study, a specialist-only program was a school psychology program located in an academic unit where the terminal school psychology degree involved two years of training on-campus and an internship during the third year. The exact degree title might vary from program to program, but the basic structure of the programs was the same. Additionally, there was no school psychology doctoral program available at the institution. Using these criteria, a total of 136 specialist-only school psychology programs were identified.

### Data Collection Procedures

Data collection occurred as part of a research project for undergraduate psychology majors who expressed an interest in school psychology during the 2012 fall semester and the 2013 spring semester. Twelve students participated during the 2012 fall semester. Eight of the 12 students continued on the project during the 2013 spring semester, and one new student joined the project.

During the 2012 fall semester, the first stage of the project, 2-3 students were given the task of identifying specialist-only programs within assigned states from the School Psychology Program Information portion of the NASP website. Students went to the “Programs Offered and Program Approvals” tab for each program listed on the NASP website. It was each student’s task to identify, independently, the institutions that did *not* have doctoral programs. These institutions became the potential pool of specialist-only school psychology

programs; 144 programs were identified (when applying the criteria presented previously; the final number of programs was 136). In addition, the names of the faculty members listed under the “Faculty” tab on the NASP website were recorded for each program. The links to the program websites were accessed in order to check the consistency between the faculty members listed on the NASP website and those listed on the program website.

Next, students searched the ERIC and PsycINFO databases to obtain citations for each faculty member’s journal publications from 2002-2011. Students were instructed to obtain citations for journal articles only. Excluded from consideration were books, book chapters, book reviews, test reviews, commentaries or replies, obituaries or memoriam, online journals, newsletters (i.e., *Communique* [NASP], *The School Psychologist* [APA Division 16]), reports, and ERIC ED documents. This approach was similar to that used in previous research cited earlier.

During the 2013 spring semester, the second stage of the project, pairs of students reviewed the information gathered the previous semester. Specifically, these pairs of students would select a state (e.g., California) and compare the programs that had been identified and the faculty members listed for programs by the individuals who collected the information during the fall semester. If discrepancies existed, the pair of students went back to the NASP website or on occasion to the program website in an attempt to resolve the differences. Resolution of the discrepancies might have required additional searches of the ERIC and PsycINFO databases. Once discrepancies were resolved and additional searches completed, students assigned the journal publications of each faculty member to one of three broad categories based on the sample employed in the publication. The first category, *P-12*, was used to indicate studies that employed children, teachers, parents, or administrators of infant through high school-aged youths. The second category, *College*, was used to indicate studies that employed a college sample. The third category, *Other*, was for articles that did not fit the other two categories. Assignment of an article to a category was done after reviewing the title and abstract.

The final stage occurred during the 2013-2014 academic year and involved the authors verifying

the information gathered in the second stage of the project. Specifically, the authors directly accessed the articles found for each faculty member and reviewed the initial classifications assigned by the teams during the second stage. At this point, the authors decided to further divide the *Other* category into *General Other* and *Other School Psychology* categories. The latter consisted of a broad array of articles that would interest school psychologists, but did not have a sample of participants. For example, articles dealing with legal issues related to the profession, describing the response-to-intervention model, presenting historical aspects of school psychology, or reviewing the literature in an area fell within the *Other School Psychology* category (e.g., Crespi, 2009; Flanagan & Esquivel, 2006). The *General Other* category represented articles that typically would not draw the attention of those working with children in school or clinical settings because they employed adult samples, were about psychometric methods, or were on topics somewhat unusual for school psychology (e.g., Beck, 2010; Rojahn, Wilkins, Matson, & Boisjoli, 2010; Schmitt & Sass, 2011). Normally, the authors reviewed and verified or reclassified articles in tandem during weekly meetings. If the verification process occurred individually, questions that arose were discussed and determination through consensus of classifications occurred as part of the weekly meetings.

As part of the final stage of the project, order of authorship on each journal article for each faculty member was noted. Authorship credit was then computed using the formula first presented by Howard, Cole, and Maxwell (1987):  $Credit = (1.5^n - 1) / (\sum 1.5^{i-1})$ . This formula is commonly used in productivity research, and was employed by the studies cited earlier that did not merely count number of publications. In effect, the formula weights the order of authorship; the authorship credit for an article always equals 1.00. The higher author receives proportionally more credit than subsequent authors. A list of the 25 individuals with the highest authorship credit was created; individuals had to have at least one publication at their current institution. The program websites for those individuals were searched in order to obtain information from their vitae regarding the university from which they obtained their doctoral degrees.

Also, a list of the 25 programs with the highest mean number of publications was created. This was done by summing the number of journal articles across faculty members affiliated with the program and dividing by the number of faculty members. If there were multiple authors from the same institution on an article, the article was only counted once for the university. The Carnegie Classification of Institutions of Higher Education (i.e., Carnegie classification) was determined for the institutions of both the top individuals and top programs (Carnegie Foundation for the Advancement of Teaching, 2010). A Post-Baccalaureate Comprehensive institution awards master's degrees in the humanities, social sciences, and science, technology, engineering and math (STEM) fields, as well as degrees in one or more professional fields. The Post-Baccalaureate with Arts and Sciences/Education Dominant category represents universities that award master's degrees in both arts and sciences and professional fields; the field with the largest number of graduate degrees is education. The Single Doctoral/Other Field classification includes institutions that award research doctoral degrees in a single field other than education; they may have more extensive offerings at the master's or professional level. Institutions classified as Comprehensive Doctoral No Medical/Veterinary according to the Carnegie system award research doctoral degrees in the humanities, social sciences, and STEM fields, and offer professional education in fields such as business, education, engineering, law, public policy, social work, or health professions other than medicine, dentistry, or veterinary medicine. Universities classified as Comprehensive Doctoral with Medical/Veterinary award doctorates in the humanities, social sciences, and STEM fields, graduate or professional degrees in one or more professional fields, and medical or veterinary doctoral degrees. Institutions classified as Doctoral, STEM Dominant award most of their doctorates in STEM fields. Universities classified as Doctoral, Professional Dominant award research doctoral degrees in a variety of areas with the largest number of doctorates in professions other than engineering, such as education, health professions, public policy, or social work. They may also offer professional education in law or medicine.

Finally, a list of the journals in which the articles of the top 25 individuals appeared was created.

### Results

Table 1 contains the names and other information for the 25 individuals who obtained the highest authorship credit ratings for the years 2002-2011. These top 25 individuals had authorship credit ratings ranging from 13.85 to 4.91. Gary Canivez of Eastern Illinois University (EIU) was ranked as the faculty member with the highest authorship credit rating during this 10-year time span. Canivez published 25 articles that were cataloged in the ERIC and/or PsycINFO databases during this time; 4 were solo-authored publications and 13 were multiple-author publications where he was listed as first author. Twenty-three of these publications involved samples that fell within the *P-12* category. A closer examination revealed that several of the studies were psychometric in nature and employed normative data sets from commercially-published assessment instruments. Canivez earned his doctoral degree in Educational Psychology with an emphasis in School Psychology and Counseling from Southern Illinois University-Carbondale, a department that no longer offers degrees in school psychology. The institution at which he was employed during the time period covered by the current study, EIU, is identified as a Post-Baccalaureate Comprehensive university, using the Carnegie classification system. All 25 of Canivez's articles represent his affiliation with EIU. This was not the case for all those listed in Table 1. For example, T. Steuart Watson of Miami University had a total of 19 articles, 9 published at Miami and 10 published while at Mississippi State University, an institution that offers both doctoral and specialist degrees in school psychology.

Further examination of Table 1 revealed that several of the faculty members affiliated with school psychology programs whose scholarly productivity led these individuals to be ranked in the top 25 did not receive degrees in school psychology. Doctoral degrees in other areas included experimental psychology, measurement and statistics, applied developmental psychology, and clinical/clinical child psychology. Also noted in

Table 1 is the fact that 2 individuals had no publications that fell within the *P-12* or *Other School Psychology* categories; 2 other individuals had only one publication in either of these two categories.

With respect to university affiliation, schools identified by the Carnegie classification system criteria as Post-Baccalaureate Comprehensive universities employed 8 individuals. One individual worked at a university that fell within the Post-Baccalaureate with Arts and Sciences/Education Dominant category. Schools that were classified as Single Doctoral/Other Field (2) or Single Doctoral/Education (1) employed three individuals total. Of the 25 individuals listed in Table 1, institutions identified as Comprehensive Doctoral No Medical/Veterinary according to the Carnegie classification system employed 7. The remaining 6 individuals were from universities classified as Doctoral, Professional Dominant.

Table 2 presents the 25 universities with specialist-only programs that obtained the highest mean number of articles published for the years 2002-2011. Only faculty members' publications while at their current institution were counted. Using T. Steuart Watson as an example again, his 9 articles published while at Miami University were tallied for that university; his 10 articles published while at Mississippi State University were not counted as part of the total for Miami University. On the other hand, in the case of Lea Theodore of the College of William & Mary, 7 of her 25 publications were tallied for that school, and 9 were tallied for CUNY-Queens College (the 5 articles published while at Hofstra and 4 published as a graduate student at University of Connecticut did not contribute to any school, because of their doctoral status). The mean number of articles published by the top 25 ranked programs ranged from 14.25 to 3.00. In every case where a faculty member no longer at an institution had contributed to the tally, like Theodore, there was a "newer" faculty member with no publications. Therefore, when computing program means, the former faculty member was "substituted" for the newer faculty member.

Brigham Young University was ranked as the program with the highest mean number of articles published during the 10 years included in the



**Table 1** Top 25 Faculty Members Based on Authorship Credit

Rank	Name/University	Doctoral Training	Carnegie Classification	Publication Categories				Number of Pubs / Solo / 1st	Number of Pubs at Curr/Prev	Author Credit
				SP/P-12	SP/Other	College	Other			
1	Gary Canivez/ Eastern Illinois University	Educational Psychology (School Psychology & Counseling); Southern Illinois University - Carbondale	Postbac – Comp	23	2			25/4/13	25/0	13.85
2	Richard Beck/ Abilene Christian University	Experimental Psychology; Southern Methodist University	Postbac – Comp			14	4	18/9/6	18/0	12.90
3	Oliver Edwards/ University of Central Florida	Educational Psychology (School Psychology); University of Florida Educational Leadership (Administration & Supervision); Florida International University	CompDoc/NMedVet	5	14			19/4/13	19/0	12.29
4	Johannes Rojahn/ George Mason University	Unspecified/University of Vienna	CompDoc/NMedVet	13	2	2	15	32/0/12	31/1	10.41
5	Jennifer Bonds-Raacke/ Fort Hays State University	Cognitive/Human Factors; Kansas State University	Postbac – Comp			11	3	14/5/5	3/11	9.02
6	Daniel Sass/ University of Texas – San Antonio	Educational Psychology (Measurement & Statistics); University of Wisconsin - Milwaukee	CompDoc/NMedVet	6	1	1	10	18/2/8	11/7	8.95
7	Jeremy Sullivan/ University of Texas – San Antonio	Educational Psychology (School Psychology); Texas A&M University	CompDoc/NMedVet	5	2	3	7	17/2/10	11/6	8.14
8	Tony Crespi/ University of Hartford	Student Development (School Psychology); University of Massachusetts - Amherst	S-Doc/Other		10	1	1	12/3/8	12/0	8.07
9	Timothy Smith/ Brigham Young University	Combined Program Clinical, Counseling, School; Utah State University	Doc/Prof	1		7	12	20/0/11	20/0	7.87

Table 1 continued

Rank	Name/University	Doctoral Training	Carnegie Classification	Publication Categories				Number of Pubs / Solo / 1st	Number of Pubs at Curr/Prev	Author Credit
				SP/P-12	SP/Other	College	Other			
10	Jon Lasser/ Texas State University	Educational Psychology (School Psychology); University of Texas – Austin	Doc/Prof	3	8		2	13/1/10	13/0	7.27
11	Paul McCabe/ CUNY – Brooklyn College	Clinical-School Psychology; Hofstra University	Postbac – Comp	5	8			13/3/4	11/2	7.13
12	Stefan Dombrowski/ Rider University	Educational Psychology (School Psychology); University of Georgia	Postbac – A&S/Ed	5	12			17/0/11	17/0	6.95
13	Kristin Powers/ California State University – Long Beach	Educational Psychology; University of Minnesota	Postbac – Comp	9	5			14/2/7	14/0	6.51
14	Rosemary Flanagan/ Touro College	School-Community Psychology; Hofstra University	Doc/Prof	1	9			10/3/5	4/6	6.46
15	Sylvia Ramirez/ University of Texas – Pan American	Educational Psychology (School Psychology); University of Wisconsin - Madison	Doc/Prof	2	6	2	4	14/1/6	14/0	6.12
16	T. Steuart Watson/ Miami University	Educational Psychology (School Psychology); University of Nebraska – Lincoln	CompDoc/NMedVet	11	5	2	1	19/1/2	9/10	6.09

Table 1 continued

Rank	Name/University	Doctoral Training	Carnegie Classification	Publication Categories				Number of Pubs / Solo / 1st	Number of Pubs at Curr/Prev	Author Credit
				SP/P-12	SP/Other	College	Other			
17	Frank Gardner/ Kean University	Clinical Psychology; Hofstra University	Postbac – Comp		1	3	6	10/2/4	3/7	6.00
18	Ellie Young/ Brigham Young University	Educational and Psychological Studies (School Psychology); University of South Florida	Doc/Prof	11	9		1	21/0/8	21/0	5.69
19	Lea Theodore/ College of William and Mary	School Psychology; University of Connecticut	CompDoc/NMedVet	12	13			25/0/6	7/18	5.61
20	Susanne Denham/George Mason University	Applied Developmental Psychology; University of Maryland – Baltimore County	CompDoc/NMedVet	6	6			12/2/5	12/0	5.59
21	Alexander Beaujean/ Baylor University	Educational, School & Counseling Psychology (School Psychology) and (Statistics & Measurement); University of Missouri	CompDoc/NMedVet	1	1	4	5	11/1/8	8/3	5.54
22	Melissa Heath/ Brigham Young University	Educational Psychology (School Psychology); Texas A&M University	Doc/Prof	3	16		1	20/0/8	20/0	5.37
23	Stephen Brock/ California State University - Sacramento	Education (Psychological Studies); University of California – Davis	S-Doc/Other	1	10			11/1/5	11/0	5.32

**Table 1 continued**

<u>Rank</u>	<u>Name/University</u>	<u>Doctoral Training</u>	<u>Carnegie Classification</u>	<u>Publication Categories</u>				<u>Number of Pubs / Solo / 1st</u>	<u>Number of Pubs at Curr/Prev</u>	<u>Author Credit</u>
				<u>SP/P-12</u>	<u>SP/Other</u>	<u>College</u>	<u>Other</u>			
24	Jeremy Jewell/ Southern Illinois University – Edwardsville	Educational Psychology (School Psychology); University of Texas – Austin	Postbac – Comp	7	2	2	1	12/0/9	12/0	4.95
25	Stephen Hupp/ Southern Illinois University – Edwardsville	Clinical Child Psychology; Louisiana State University	Postbac – Comp	9	3	2	1	15/0/5	13/2	4.91

*Note.* Carnegie Classifications: Postbac-Comp = Post-Baccalaureate Comprehensive; CompDoc/NMedVet = Comprehensive Doctoral No Medical/Veterinary; S-Doc/Other = Single Doctoral Other Field; DocProf = Doctoral, Professional Dominant; Postbac-A&S/Ed = Post-Baccalaureate with Arts & Sciences Education Dominant.

Publication Categories: SP/P-12 = school psychology-oriented articles with samples from Grades Preschool – 12; SP/Other = articles that are school psychology-related topics.

Number of Pubs/Solo/1st = Total number of publications/number of solo authored articles/number of publications on which the person was listed as first author.

Number of Pubs at Curr/Prev = Number of publications at the author’s current institution/number of publications at previous institutions including as a graduate student.

**Table 2 Top 25 Programs Based on Average Number of Journal Articles Published by Faculty Members**

---

<u>Rank</u>	<u>University</u>	<u>Carnegie Classification 2010</u>	<u>Number of Pubs</u>	<u>Number of Faculty</u>	<u>Average</u>	<u>Range</u>
1	Brigham Young University	Doc/Prof	57	4	14.25	7 – 19
2	University of Central Florida	CompDoc/NMedVet	38	3	12.67	9 - 18
3	George Mason University	CompDoc/NMedVet	46*	4	11.50	0 – 31
4	Eastern Illinois University	Postbac – Comp	27	3	9.00	0 – 25
5	Rider University	Postbac – A&S/Ed	17	2	8.50	0 – 17
6	University of Texas–San Antonio	CompDoc/NMedVet	30	5	6.00	0 – 11

Table 2 continued

<u>Rank</u>	<u>University</u>	<u>Carnegie Classification 2010</u>	<u>Number of Pubs</u>	<u>Number of Faculty</u>	<u>Average</u>	<u>Range</u>
7	Southern Illinois University–Edwardsville	Postbac – Comp	22	4	5.50	1 – 13
8	California State University–Long Beach	Postbac – Comp	16	3	5.33	0 – 12
9	Eastern Kentucky University	Postbac – Comp	15	3	5.00	2 – 13
10	California State University–Sacramento	S-Doc/Other	14	3	4.67	0 – 10
10	University of Nebraska–Omaha	Doc/Prof	14	3	4.67	2 – 9
12	Arkansas State University	Doc/Prof	9	2	4.50	3 – 6
13	Miami University	CompDoc/NMedVet	19*	5	3.80	0 – 9
14	California State University–San Bernardino	Postbac – Comp	11	3	3.67	2 – 6
14	University of Northern Iowa	Doc/Prof	11*	3	3.67	0 – 6
14	Fort Hays State University	Postbac – Comp	11*	3	3.67	3 – 4
14	University of Dayton	Doc/STEM	11	3	3.67	0 – 9
18	City University New York–Queens College	Postbac – Comp	18*	5	3.60	0 – 9
19	University of Hartford	S-Doc/Other	14	4	3.50	1 – 10
20	City University New York–Brooklyn College	Postbac – Comp	20	6	3.33	0 – 11
20	Abilene Christian University	Postbac – Comp	20	6	3.33	0 – 17
20	Oswego State University of New York	Postbac – Comp	10	3	3.33	0 – 10
23	University of Texas–Pan American	Doc/Prof	13	4	3.25	0 – 11
24	Florida International University	CompDoc/NMedVet	12	4	3.00	0 – 7
24	New Mexico State University	CompDoc/NMedVet	12*	4	3.00	1 – 5
24	University of Toledo	CompDoc/MedVet	6	2	3.00	0 – 6

*Note.* Ranking based on average number of publications. \* Indicates publications included from a faculty member no longer at the institution.

Carnegie Classifications: Postbac-Comp = Post-Baccalaureate Comprehensive; CompDoc/MedVet = Comprehensive Doctoral with Medical/Veterinary; CompDoc/NMedVet = Comprehensive Doctoral No Medical/Veterinary; S-Doc/Other = Single Doctoral Other Field; DocProf = Doctoral, Professional Dominant; Postbac-A&S/Ed = Post-Baccalaureate with Arts & Sciences Education Dominant; Doc/STEM = Doctoral, Science, Technology, Engineering and Math Dominant.

current study. The four faculty members at BYU published 57 articles during this time period or a mean number of articles per faculty member of 14.25. The number of articles published by individual BYU faculty members ranged from 7 to 19. Table 2 also contains the Carnegie classification for each program. Ten schools were identified as Post-Baccalaureate Comprehensive universities. One university fell within the Post-Baccalaureate with Arts and Sciences/Education Dominant category. Two programs were classified as Single Doctoral/Other Field. Six institutions were identified as Comprehensive Doctoral No Medical/Veterinary according to the Carnegie classification system. Five programs were located within universities classified as Doctoral, Professional Dominant. The remaining 2 programs were identified as Doctoral, STEM Dominant (1) or Comprehensive Doctoral with Medical/Veterinary (1).

It is worth noting that 9 of the 25 individuals listed in Table 1 come from 4 programs: Brigham Young University (3); University of Texas-San Antonio (2); George Mason University (2); and Southern Illinois University-Edwardsville (2). Also, these 4 programs were ranked among the top 7 listed in Table 2. Three of the 4 programs were at doctoral-granting institutions; only 1, Southern Illinois University-Edwardsville, was located within a Post-Baccalaureate Comprehensive university. Also, at 16 of the 26 institutions listed in Table 2, there was at least one faculty member with no publications, as defined in the current study.

Finally, Table 3 provides a partial list of journals in which the identified top 25 faculty members published articles. These individuals published 412 articles in 152 different journals during the time period examined. Of the top 10 journals, 6 are familiar to school psychologists: *Psychology in the Schools*, *Journal of Psychoeducational Assessment*, *School Psychology Quarterly*, *Journal of Applied School Psychology* (formerly *Special Services in the Schools*), *School Psychology International*, and *Contemporary School Psychology* (formerly *The California School Psychologist*). The remaining journals in the top 10 were perhaps less familiar within traditional school psychology circles: *Journal of Psychology and Theology*, *Journal of Developmental and Physical*

*Disabilities*, *Journal of Instructional Psychology*, *Journal of Evidence-Based Practices for Schools*, and *Journal of School Violence*. The journal with the highest frequency of articles from these faculty members was *Psychology in the Schools*. A closer examination of publications in this journal revealed that 11 individuals accounted for 39 articles; the highest number of articles by a single faculty member was 14. Note that *Psychology in the Schools* published 6-10 issues each year from 2002-2011, whereas many journals published 4 issues per year. As a result, there were more opportunities for articles to appear in this journal. Next in line was *Journal of Psychoeducational Assessment*, where 7 faculty members contributed to a total of 17 articles; one faculty member published 7 articles in this journal. There were instances where a single faculty member accounted for virtually all the publications within a journal. For example, 9 of the 10 articles published in *Journal of Psychology and Theology* were from a single faculty member.

## Discussion

The current study adds to the periodic snapshots of faculty productivity within the field of school psychology. Because recent studies focused on faculty members of APA-accredited doctoral programs (Grapin et al., 2013; Kranzler et al., 2011; Wagner et al., 2007), we were interested in faculty who were affiliated with specialist-only school psychology programs. Employing the commonly used metric of journal authorship credit, we identified the 25 individuals who were most productive in the years 2002-2011.

Previous studies used various lengths of time in their examination of scholarly productivity, so direct comparisons between data from the current study and past research is difficult. However, there was one study that also examined scholarly productivity over a 10 year period, as was done in the current study. Roberts et al. (2006) reported on the top contributors to the school psychology literature from 1996-2005. We considered a broader range of journals, and did not limit ourselves to 5 school psychology journals. Nevertheless, it was informative to examine the range on authorship credit ratings for the Roberts et al. study and our data. The range of authorship credit in school

**Table 3** *Journals in Which the Top 25 Faculty Members Published Articles*

<u>Journal</u>	<u>Number of</u> <u>Articles</u>	<u>Number of</u> <u>Faculty</u>	<u>Most by an</u> <u>Author</u>
Psychology in the Schools	39	11	14
Journal of Psychoeducational Assessment	17	7	7
School Psychology Quarterly	14	5	4
Journal of Applied School Psychology (Special Services in the Schools)	13	7	5
Journal of Psychology and Theology	10	2	9
School Psychology International	10	6	3
Journal of Developmental and Physical Disabilities	9	4	5
Journal of Instructional Psychology	9	4	3
Contemporary School Psychology (California School Psychologist)	8	2	5
Journal of Evidence-Based Practices for Schools	8	4	4
Journal of School Violence	8	4	4
Behavior Modification	6	3	4
School Psychology Review	6	5	2
College Student Journal	5	4	2
Intelligence	5	3	2
Journal of Early Childhood and Infant Psychology	5	1	5
Journal of Psychology and Christianity	5	1	5
Professional Psychology: Research and Practice	5	3	3
Professional School Counseling	5	5	1
Canadian Journal of School Psychology	4	2	2
Child & Family Behavior Therapy	4	2	3
Educational and Psychological Measurement	4	3	2
Journal of Child and Family Studies	4	2	3
Journal of Clinical Sport Psychology	4	1	4
Journal of Counseling Psychology	4	2	3
Personality and Individual Differences	4	1	4
Psychological Reports	4	3	2
Research in Developmental Disabilities	4	2	3

*Note.* Number of Faculty = number of faculty members who have a publication in the journal; Most by an Author = the highest number of articles by a single faculty member.

psychology journals for the top 25 faculty members in the Roberts et al. study was 11.72 to 4.96. That compares to the range of 13.85 to 4.91 in the current study that did not restrict publications to only school psychology journals. The range in number of articles published by the top 25 contributors in the Roberts et al. study was 39 to 12, compared to 32 to 10 in the current study. Although potentially interesting, these comparisons should be viewed cautiously, because the limited number of journals included in the Roberts et al. study likely deflated the authorship credits for the top 25 contributions in their data set. In other words, it is likely that at least some, if not all, of the individuals in the Roberts et al. study published in other journals than those included in their study.

Grapin et al. (2013) did not restrict publications to school psychology journals, similar to the current study, but looked only at faculty members in APA-accredited programs over a 5-year time span. The authorship credit for their top 25 faculty members ranged from 18.88 to 6.37, with number of articles ranging from 40 to 11. In the context of the current study, it may be tempting, but would be too simplistic to merely “multiply by 2” to compare our authorship credit ratings over a 10 year period to those of Grapin et al. over a 5 year time span. Alternatively, authorship credit ratings could be compared by calculating yearly averages for the Grapin et al. study and the current study. The ranges that resulted, 3.78 – 1.27 (Grapin et al., 2013) and 1.39 – 0.49 (current study), need to be viewed cautiously, due to the differences between APA-accredited doctoral programs versus specialist-only programs.

Perhaps more than anything else, comparing findings from the Grapin et al. and Roberts et al. studies demonstrates how restricting school psychology faculty members’ scholarly productivity to school psychology journals likely results in an underestimation of scholarship. Future research should recognize that school psychology faculty members publish in journals other than those in the discipline, and consider the broad spectrum of scholarly outlets that exist in clinical, child clinical, developmental and other areas of psychology and education.

Are faculty members in doctoral programs more productive than faculty members in specialist-

only programs in terms of journal publications? Unfortunately, differences in time spans and databases examined across published studies on the scholarly productivity of school psychology faculty members and the current study makes this question difficult to answer. However, observations made by Joy (2006) regarding scholarly productivity of academic psychologists, in general, may be applicable to school psychology. He notes that faculty members at research universities and doctoral institutions publish the most and tend to continue to publish throughout their careers. Joy further states that faculty members at master’s universities publish less. After receiving tenure, Joy noted that faculty members at master’s universities tend to publish even less or stop publishing altogether. It should be noted that Joy’s research was limited to universities in the northeastern part of the United States. Nevertheless, Joy’s work suggests that institutional expectations/climate may contribute to scholarly productivity. This reinforces the findings of Levinson et al. (1994) that school psychology programs located in doctoral institutions had higher productivity ratings than those located in comprehensive universities.

An examination of Table 1 revealed a fairly equal distribution of top faculty at doctoral (52%) versus non-doctoral (48%) institutions when the Carnegie classification of Single Doctoral was collapsed with Post-Baccalaureate Comprehensive. Data in Table 2 revealed a 50%-50% distribution of the top ranked programs across doctoral and non-doctoral institutions, again, combining Single Doctoral and Post-Baccalaureate Comprehensive institutions. As noted by Joy (2006) and Levinson et al. (1994), institutional factors may provide opportunities that increase scholarly productivity. Many assume that a faculty position at a doctoral institution provides the opportunity for or requires more scholarly productivity. Nonetheless, the top 2 individuals during the period examined (i.e., Canivez and Beck) were from Post-Baccalaureate Comprehensive universities, suggesting that there are individual variables that likely also influence a faculty member’s scholarly productivity (e.g., motivation, desire to contribute to the field, etc.). Future researchers might consider examining personality characteristics of those who are among the most productive scholars in the field. In fact,



Martínez, Floyd, and Erichsen (2011) examined the responses of highly productive school psychology scholars to a questionnaire regarding research strategies. Among the seven categories into which they placed responses was one labeled, “personal character traits that foster productivity.” The top two responses in that category were: “Persistence, discipline, and really hard work,” and “Interest, curiosity, flexibility, creativity, and passion.” Research like this might tease out environmental versus individual characteristics that contribute to scholarly productivity. In terms of environmental influences, it is possible that in a department that offers doctoral degrees in areas other than school psychology (e.g., BYU), there is the expectation that school psychology faculty members serve on dissertation committees. Serving on a dissertation committee may or may not lead to collaborative publication opportunities that are not available in institutions that do not have doctoral programs. Similarly, school psychology programs located in a Post-Baccalaureate Comprehensive university where a thesis is required also may lead to publication opportunities that are not available in programs in Post-Baccalaureate Comprehensive universities that do not require a thesis to meet research requirements. Future research might consider whether there are program and/or department characteristics that promote scholarly productivity (e.g., reduced teaching loads and/or service expectations for faculty, student research requirements).

Along with the findings from the current study, we made a number of observations during data collection that might benefit others interested in conducting similar research. For example, the advent of the School Psychology Program Information portion of the NASP website greatly facilitated the data collection process. Annually, school psychology programs are encouraged to provide updated information about faculty members. Reminders are provided on the school psychology trainers’ listserv. Despite best efforts, it is not clear whether all possible school psychology programs are included on this website. Also, whether a program’s information actually is the most current available may depend on when the website is consulted. It is useful to verify the information on the NASP website by going directly to the program

website. As part of the program information, the NASP website typically contains a link to the program page at their university. Although not all links on the NASP website were correct, it was fairly easy to find the websites for universities. As might be expected, some university and school psychology program websites were easier to navigate than others.

Another issue that arose quickly was determination of what constituted a school psychology faculty member. Does one have to be trained as a school psychologist to be a faculty member of a school psychology program? The NASP website listed the faculty members provided by the programs. It was evident that some programs listed only core school psychology faculty members, while other programs listed all faculty members who taught courses taken by school psychology students. Our intent was to examine the scholarly productivity of core school psychology faculty members. Consulting program handbooks, when available on program websites, often was helpful in differentiating core school psychology faculty members from faculty members affiliated with a program through teaching a specialized class like research methods or statistics. Ultimately, however, we allowed programs to define their faculty members, especially if there was no clear differentiation between core faculty and affiliated faculty. Perhaps the way in which those who coordinate the NASP website ask the question of programs could clarify this issue in the future. Asking programs to identify core school psychology faculty, rather than a more generic listing of school psychology faculty, might clarify this issue.

When searching the PsycINFO and ERIC databases, it became evident that some faculty members contribute to the school psychology literature through books and book chapters rather than journal articles. Traditionally, only journal articles have been considered when examining scholarly productivity. Generally, the peer-review process associated with journal articles is viewed as lending some degree of quality control to the publication process. Unfortunately, that viewpoint negates the contributions made by faculty authors of well-done books or book chapters in school psychology or related areas. Increasingly, peer-review is used for publications like the *Best*

*Practices in School Psychology* series. Likewise, peer-review also is used when considering contributions to school psychology newsletters such as the *Communique* (NASP) and *The School Psychologist* (APA Division 16). Some articles in these newsletters look more and more like journal articles. This was especially true of *The School Psychologist* before changes that occurred when APA became involved in production of this newsletter; this publication is returning to its earlier format, so it may again become a viable outlet. The *Communique* is included in the ERIC database. Another publication, *Trainers' Forum*, is evolving into its own journal, and regional journals, such as *Research and Practice in the Schools*, that provide online access are being developed. Journal articles have been the focus of research on scholarly productivity, but it may be time to broaden the definition to include additional forms of scholarship. Future researchers should consider expanding the traditional approach that relies solely on journal articles with respect to scholarly productivity.

We also discovered that databases are not always accurate. Whenever possible, we attempted to match the output from our database searches with individual vitae on programs' or faculty members' websites. On some program websites, faculty members listed representative publications or presentations. On others, there were links to complete vitae of faculty members. When complete vitae were available, it was obvious that some faculty members were more diligent than others in providing up-to-date information. We also discovered that some publications in well-known journals did not always appear in our searches of authors' names. If we discovered a journal publication on a faculty member's vita within the target time period, we reran the search on PsycINFO and ERIC using the article title, and the article would appear. The fact that an article might not appear when searching by author's name, but would appear when searching by article title listed on a vita, suggests that the PsycINFO and ERIC databases are not 100% reliable when it comes to identifying all possible publications. In the current study, we attempted to be as thorough and accurate as possible by comparing information from the databases and websites. It also should be remembered that databases do limit the journals

they cover. Therefore, by examining only the PsycINFO and ERIC databases, perhaps not all journals in which school psychology faculty members might publish were captured in our study. Both database accuracy and coverage issues could influence authorship credit and resulting rankings of faculty members and programs.

There were other observations that occurred through the data collection process. For example, some names changed with marital status. Sometimes these changes were in the form of hyphenated last names. Other times, the last name changed completely with change in marital status. In the case of hyphenated last names, the databases generally were good at providing publications that included maiden names as part of the last name. When last names changed as a result of change in marital status, the task of identifying articles became more challenging. Again, if individuals included information about publications at their programs' websites or on vitae that included their maiden names, searches were then done with both their maiden names and married names. Occasionally, we were able to discover maiden names through articles with collaborators, information contained in notes to articles, or other serendipitous methods. As with the database accuracy and coverage issues, our ability to link married to maiden names could influence authorship credit and resulting rankings of faculty members and programs.

Also, the time span chosen for a project such as the current study can influence the results. A longer time span allows more articles to be considered, potentially boosting authorship credit and resulting rankings. Previous studies of scholarly productivity among school psychology faculty members have ranged from 5 years (Carper & Williams, 2004; Grapin et al., 2013) to 13 years (Davis et al., 2005). To be truly accurate, all faculty members would need to be employed as academics during the entirety of the time span considered. This is more likely to occur over shorter time spans, but even in these situations, it is inevitable that levels of experience will fluctuate. In any given time span, it is likely that you will be examining individuals at various stages in their careers. In other words, some individuals may be in the prime of their careers, while others may be transitioning to retirement. Still

other faculty members will be entering academia and in the initial stages of establishing themselves and navigating the balance of teaching and scholarship. In the case of the current study, that might mean that some individuals were active all 10 years covered, others may have been publishing articles initially during the time span, but were less focused on this aspect of their careers as they transitioned to retirement, and new faculty members may have been contributing articles to the literature for as little as 1 year. Obviously, those who were active during all 10 years covered had the potential for higher authorship credit and resulting rankings.

A final comment deals with the approach taken in determining scholarly productivity for individuals versus programs. We examined an individual's scholarly productivity for the entire period from 2002-2011, regardless of university, which resulted in 11 of the top 25 individuals with publications at multiple institutions. Examining the scholarly productivity across institutions becomes problematic, because resources can vary. As noted earlier, Joy (2006) and Levinson et al. (1994) found that expectations typically differ at different level institutions. Because research is typically emphasized more at doctoral-level research universities, course loads tend to be less than at comprehensive master's degree granting institutions. Funding also varies across universities, with those granting doctoral degrees typically attracting more outside funding. This potential confound was controlled to some degree in the approach to ranking programs, where we counted publications at current institutions, also crediting previous non-doctoral institutions with publications by faculty members who had relocated. In the case of Theodore, this allowed both her current program at the College of William & Mary and the previous program at CUNY-Queens College where she was employed to receive credit for her scholarship. Of course, attempts at portraying program scholarship are influenced by the career development issues mentioned earlier.

The aforementioned observations might suggest limitations of the current study or at least issues to be considered in future investigations of the topic of faculty scholarly productivity. Every attempt was made to gather complete information on each program and faculty members within

programs. Nevertheless, as noted, issues related to the databases chosen and the accuracy of those databases, and accuracy of information on the NASP, university and program websites may have influenced findings. Some might take issue with the use of authorship credit ratings as a meaningful indicator of scholarly productivity. This is the metric commonly used when examining scholarly productivity with respect to school psychology. Authorship credit ratings proportion credit to reflect contribution to an article, but it does not reflect impact of the article. Future research might consider incorporating indirect information about impact through inclusion of "times cited in this database" from PsycINFO or the "cited by" statistic for articles in the Google Scholar database.

Despite the issues noted, the current study contributes to the 20-plus years of research on the scholarly productivity of school psychology programs and/or faculty members. Like the studies before it, the current study provides a snapshot of scholarly productivity, in this case for the years 2002-2011. As such, the current study contributes to the historical scrapbook of academic school psychology.

## References

- Beck, R. (2010). The moral and psychological landscape of chit chat: Toward a theological hermeneutic of everyday conversation. *Journal of Psychology and Christianity, 29*, 41-45.
- Carnegie Foundation for the Advancement of Teaching. (2010). Carnegie Classification of Institutions of Higher Education. Retrieved from <http://carnegieclassifications.iu.edu>
- Carper, R. M., & Williams, R. L. (2004). Article publications, journal outlets, and article themes for current faculty in APA-accredited school psychology programs: 1995-1999. *School Psychology Quarterly, 19*, 141-165. doi: 10.1521/scpq.19.2.141.33312
- Crespi, T. D. (2009). Group counseling in the schools: Legal, ethical, and treatment issues in school practice. *Psychology in the Schools, 46*, 273-280. doi: 10.1002/pits.20373
- Davis, K. S., Zanger, D., Gerrard-Morris, A., Roberts, G., & Robinson, D. H. (2005). Productivity and collaboration of authors in school psychology journals, 1991-2003. *The School Psychologist, 59*, 129-133.
- Flanagan, R., & Esquivel, G. B. (2006). Empirical and clinical methods in the assessment of personality and psychopathology: An integrative approach to training. *Psychology in the Schools, 43*, 513-526.

- doi: 10.1002/pits.20165
- Grapin, S. L., Kranzler, J. H., & Daley, M. L. (2013). Scholarly productivity and impact of school psychology faculty in APA-accredited programs. *Psychology in the Schools, 50*, 87-101. doi: 10.1002/pits.21658
- Howard, G. S., Cole, D. A., & Maxwell, S. E. (1987). Research productivity in psychology based on publication in the journals of the American Psychological Association. *American Psychologist, 42*, 975-986. doi: 10.1037/0003-066X.42.11.975
- Joy, S. (2006). What should I be doing, and where are they doing it? Scholarly productivity of academic psychologists. *Perspectives on Psychological Science, 1*, 346-364. doi: 10.1111/j.1745-6916.2006.00020.x
- Kranzler, J. H., Grapin, S. L., & Daley, M. L. (2011). Research productivity and scholarly impact of APA-accredited school psychology programs: 2005-2009. *Journal of School Psychology, 49*, 721-738. doi: 10.1016/j.jsp.2011.10.004
- Levinson, E. M., Barker, W., & Lillenstein, D. (1994). Publication productivity in selected school psychology journals: A reconsideration based on institutional mission. *Psychology in the Schools, 31*, 120-127. doi: 10.1002/1520-6807(199404)31:2<120::AID-PITS2310310206>3.0.CO;2-W
- Little, S. G. (1997). Graduate education of the top contributors to the school psychology literature: 1987-1995. *School Psychology International, 18*, 15-27. doi: 10.1177/0143034397181002
- Martínez, R. S., Floyd, R. G., & Erichsen, L. W. (2011). Strategies and attributions of highly productive scholars and contributors to the school psychology literature: Recommendations for increasing scholarly productivity. *Journal of School Psychology, 49*, 691-720. doi: 10.1016/j.jsp.2011.10.003
- Miller, D. C. (2008). Appendix VII: School psychology training programs. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. clv-cxcviii). Bethesda, MD: National Association of School Psychologists.
- Roberts, G. A., Davis, K. S., Zanger, D., Gerrard-Morris, A., & Robinson, D. H. (2006). Top contributors to the school psychology literature: 1996-2005. *Psychology in the Schools, 43*, 737-743. doi: 10.1002/pits.20183
- Rojahn, J., Wilkins, J., Matson, J. L., & Boisjoli, J. (2010). A comparison of adults with intellectual disabilities with and without ASD on parallel measures of challenging behavior: The Behavior Problems Inventory-01 (BPI-01) and Autism Spectrum Disorders-Behavior Problems for Intellectually Disabled Adults (ASD-BPA). *Journal of Applied Research in Intellectual Disabilities, 23*, 179-185. doi: 10.1111/j.1468-3148.2009.00519.x
- Schmitt, T. A., & Sass, D. A. (2011). Rotation criteria and hypothesis testing for exploratory factor analysis: Implications for factor pattern loadings and interfactor correlations. *Educational and Psychological Measurement, 71*, 95-113. doi: 10.1177/013164410387348
- Wagner, A. R., Lail, K. E., Viglietta, E., & Burns, M. K. (2007). Faculty publications in APA-accredited school psychology training programs between 2000 and 2005. *The School Psychologist, 61*, 16-21.
- Webster, R. E., Hall, C. W., & Bolen, L. M. (1993). Publication productivity in selected school psychology journals: 1985-1991. *Psychology in the Schools, 30*, 136-142. doi: 10.1002/1520-6807(199304)30:2<136::AID-PITS2310300206>3.0.CO;2-K

Article

---

# Evidence-based Treatments for Youth with Externalizing Behaviors Associated with ADHD

Jeffrey D. Shahidullah  
*Rutgers University*

Justin A. Barterian  
*Ohio State University*

John S. Carlson  
*Michigan State University*

The three evidence-based treatments for use with children demonstrating externalizing ADHD symptoms, such as hyperactivity-impulsivity, are (1) stimulant medication, (2) behavioral treatment implemented across home and school, and (3) a combination of the two. Despite the compelling research support for each of these treatments, there remains variability in the outcome literature regarding a number of parameters, such as long-term outcomes, type and intensity of treatment, setting, and individual differences in the size and topography of treatment response. This paper reviews findings from a number of seminal studies pertaining to the treatment of externalizing ADHD symptoms. The discussion includes those studies that the American Academy of Child and Adolescent Psychiatry, the American Academy of Pediatrics, and the American Psychological Association include as empirical support in the development of their practice parameters. Implications for school psychologists are discussed.

**Keywords:** ADHD, youth, medication, behavioral health treatment, psychiatry

## Evidence-Based Treatment of Externalizing ADHD Problems in Children

The three evidence-based treatments for use with children demonstrating externalizing ADHD symptoms (i.e., hyperactivity-impulsivity) are (1) stimulant medication, (2) behavioral treatment implemented across home and school, and (3) a combination of the two (Barkley, 2014). Despite demonstrated short-term effectiveness with each treatment, there is variability in the literature regarding long-term outcomes, type and intensity of treatment, setting, and individual differences in the size and topography of treatment response (Barkley, 2014). The following sections review findings from a number of seminal studies (e.g., MTA 1999a, 1999b) as well as literature focusing on treatments delivered in primary care. The discussion includes studies that professional organizations (AAP, 2011; APA, 2006; AACAP, 2007; hereafter referred to as

practice parameters) include as empirical support for the development of their treatment guidelines. It is noteworthy that while the construct of interest in the present study is externalizing behaviors associated with ADHD, there is a paucity of literature that differentiates treatment response between inattentive and hyperactive-impulsive symptoms (herein referred to as “externalizing symptoms”). Further, of the literature that reports on externalizing symptoms, few studies differentiate

---

**Author Note:** Correspondence concerning this paper should be addressed to Jeffrey D. Shahidullah, Rutgers-The State University of New Jersey, Graduate School of Applied and Professional Psychology, 152 Freylinghuysen Road, Room A357, Piscataway, NJ 08854; E-mail: jeff.shahidullah@rutgers.edu

between core (i.e., hyperactivity, impulsivity) and peripheral (aggression, oppositionality, conduct problems) symptoms. Therefore, the following literature review discusses ADHD as a global construct. However, specific studies are emphasized that have an acute focus on core or peripheral externalizing symptoms.

**Pharmacological treatment.** Stimulant medications (e.g., methylphenidate, amphetamine compounds), which act on the central nervous system, are common treatments for core externalizing symptoms of ADHD in children. Specifically, they are often prescribed to increase on-task behavior and decrease fidgeting at school, while improving parent-child interactions at home (Reiff & Stein, 2011). However, their efficacy has been better demonstrated in improving core rather than peripheral externalizing symptoms (MTA, 1999a). Several stimulant medications exist containing slightly different formulations that vary in pharmacokinetic effects. Though some nonstimulants (e.g., atomoxetine, guanfacine) are FDA-approved (Mohammadi & Akhondzadeh, 2007), this medication class remains understudied and is typically prescribed only when children are unresponsive to stimulants or unable to tolerate their side-effects. Therefore, only studies evaluating stimulants are presented hereafter.

Double-blind, placebo-controlled studies have consistently found stimulants to be superior to placebo in remediating externalizing ADHD symptoms. In a meta-analysis, Faraone and Buitelaar (2010) investigated results of over 20 double-blind, placebo-controlled studies of amphetamine (Adderall) and methylphenidate (Concerta, Ritalin) use and found most effect size measures to be statistically significant (0.99 for amphetamine and 0.72 for methylphenidate). However, they only examined core externalizing symptoms. In a different meta-analysis of 115 studies, Crenshaw and colleagues (1999) reported effect sizes of 0.46 for academic impairments and 0.72 for externalizing behavioral symptoms. Nevertheless, for peripheral externalizing symptoms, effect sizes were smaller (0.61), though still robust.

The largest ADHD treatment study is the Multimodal Treatment for ADHD (MTA, 1999a, 1999b) trial. Its multisite, 14-month randomized

clinical trials assigned 579 children, aged 7 to 9.9 years, with ADHD to a medication, behavioral treatment, combined treatment, or a community care comparison group. Children in the community care group were not treated in the study, but were referred to their community primary care physician (PCP) to receive care-as-usual. Over two-thirds ( $n = 97$  [67.4%]) of community care participants received stimulant medication. At a 14-month follow-up, the use of MTA-provided medication (i.e., methylphenidate) decreased core externalizing symptoms, both alone and in conjunction with behavioral treatment. Further, participants receiving MTA-prescribed medication displayed greater symptom improvement, including peripheral externalizing symptoms such as aggression and oppositionality (Jensen et al., 2001) than those receiving community-provided medication. Also, MTA-medication management yielded slightly greater improvement in hyperactive-impulsive symptoms than inattentive symptoms on parent and teacher ratings.

While most participants in the MTA-provided medication and community care groups received stimulant treatment, there were marked differences in titration for determining therapeutic dosage. The more rigorous protocol for monitoring treatment response for the MTA-provided medication included longer (30 vs. 18 min.) and more frequent (8.8 vs. 2.3 visits per year) follow-up visits (Jensen et al., 2001). Symptom improvement discrepancies between the two groups confirm the critical importance of ongoing treatment monitoring and are a major reason why this practice is included in AAP (2011) and AACAP (2007) practice parameters.

The MTA identified response patterns based on gender and socioeconomic status (SES). Gender was significant for core externalizing symptoms, as combined treatment and MTA-managed medication were superior to traditional care in males, while only combined treatment was superior to traditional care in females. Low-SES (e.g., poverty-threshold, Medicaid beneficiaries) was significant for peripheral externalizing symptoms as this group demonstrated superior symptom improvement from combined treatment compared with other treatment conditions. Examinations of the moderating influences in low-SES families are an important consideration given the barriers to behavioral

treatment adherence encountered by this group (Dumas & Wahler, 1983). Arnold and colleagues (2003) found ethnic minorities (i.e., African Americans, Latinos) in general improved from combined treatment while Caucasian youth did not. Nonetheless, there is a general paucity of treatment literature pertaining to ethnicity-related differences (Weisz, Huey, & Weersing, 1998).

Despite wide dissemination of the AAP practice parameters, an alarming amount of literature indicates poor adherence to these guidelines by PCPs (e.g., Epstein, Langberg, Lichtenstein, Kolb, & Simon, 2013). Although 92% of pediatricians report familiarity with AAP guidelines, only 78% implement the recommendations (Rushton et al., 2004). Regarding other primary care specialties, Rushton and colleagues found only 60% of family physicians were familiar with AAP guidelines, and only 39% incorporate the guidelines. Specifically, only 53% of PCPs report they conduct follow-up visits (Rushton et al., 2004). Moreover, PCPs who conduct these critical follow-up visits may actually be lower than PCP self-report. In a study using chart review audits, Epstein and colleagues (2008) found only 27% of patients/families had telephone contact with their PCP within two weeks of medication initiation, and only 52% had in-office follow-up visits within six weeks. Further, only 9% of PCPs had obtained follow-up ratings to evaluate treatment response and adverse effects.

Survey results of 1,374 PCPs indicated that difficulty in making external mental health referrals and limited insurance reimbursement for time-rendered services were primary barriers to adhering to AAP guidelines (Rushton et al., 2004). Epstein and colleagues (2013) posit that these logistical concerns severely limit PCPs' ability to use rating scale data for treatment monitoring. Additionally, PCPs may not have the appropriate knowledge of rating scales that are sensitive to change for the prescribed treatment. Also, due to short appointment times, it is not always feasible to explain the purpose of the scales, distribute the scales, establish a protocol for efficient return of the scales, and score and interpret results.

#### **Pharmacological treatment: Limitations.**

Results of large-scale randomized clinical trials, such as the MTA (1999a), have affected how PCPs view stimulant use given the benefits of these medications

at one year and two years post treatment. However, there are limitations, as much research conducted over the past 25 years has indicated that stimulants do not produce long-term (i.e., two or more years) benefit once medication has been discontinued. As part of the eight-year MTA follow-up, Molina and colleagues (2007) demonstrated that no beneficial effects of stimulant medication were apparent. Additionally, the APA (2006) Task Force Report found that stimulant medication lacks evidence for effectiveness beyond one year. Finally, some research has indicated that stimulants may not address peripheral areas of functional impairment such as academic, social, and behavioral skill-building (Pelham & Smith, 2000), with some studies finding minimal improvement in academic performance ( $ES = 0.12$ ; MTA, 1999a, 2004). These findings are difficult to reconcile with improvements found by Faraone and Buitelaar (2010), Crenshaw and colleagues (1999), and Pliszka and colleagues (2006) cited previously.

Data also suggest between 20 and 30% of children may not respond to stimulant medication and/or may not be able to tolerate side effects (e.g., sleep disturbance, appetite suppression, emotional lability, exacerbation of tic disorders; Schachter, Pham, King, Langford, & Moher, 2001). Given the impairments related to sleep and appetite, research has long cited suppression in height and weight as being linked with stimulants (Safer & Allen, 1973). For example, the MTA (1999a) revealed a two cm reduction in height gain and a two kg reduction in weight gain for children between ages 10 and 12 (Swanson & Volkow, 2009).

#### **Pharmacological treatment: Conclusions.**

Numerous double-blind, placebo-controlled trials have demonstrated the short-term efficacy of stimulant medication. In the MTA (1999a), stimulant medication was more effective than behavioral treatment alone and comparable to the combined treatment. However, limitations were present in the design of the MTA trial, which may have predisposed a favorable outcome for medications (Pelham, 1999). Stimulant medications may be more accessible than behavioral treatment, which often requires considerable allocation of financial and family resources. Research also shows that children (70-80%; Schachter et al., 2001) respond to medications with rapid symptom improvement.

While stimulant medication is clearly the PCP standard of care, this treatment's inability to demonstrate long-term effectiveness after discontinuation, coupled with the demonstrated side-effects, highlights the need for a thorough risk to benefit analysis. Additionally, while outcome efficacy from stimulant use is contingent on close monitoring of treatment response, the current PCP standard of care is not consistent with the level of ongoing monitoring recommended by AAP.

Of note, in developing treatment recommendations to be included as part of the *Texas Children's Medication Algorithm for ADHD Project*, a panel of researchers (Pliszka et al., 2006) reaffirmed from a literature review the well-established efficacy of stimulant medication for treating core symptomatology (i.e., inattention, hyperactivity, impulsivity). However, in a significant departure from previous guidelines, the panel in their 2005 revisions to the Texas Algorithm recommended using conjunctive behavioral intervention along with stimulant medication for treating peripheral externalizing ADHD symptoms such as aggression (p. 644).

**Behavioral treatment.** While there are fewer published behavioral intervention studies than medication studies, behavioral intervention is an effective treatment for children with ADHD and the externalizing behavioral problems demonstrating effect sizes in the moderate to large range (APA, 2006; O'Leary, Pelham, Rosenbaum, & Prince, 1976; Patterson et al., 1993). Organizations in the fields of education (U.S. Department of Education, 2004), mental health (APA Division of Clinical Child and Adolescent Psychology, 2012; National Association of School Psychologists, 2002; AACAP, 2007), and medicine (American Academy of Pediatrics, 2011) have recommended these treatments as evidence-based interventions. Further, some (e.g., Pelham, 2012; Pelham et al., 1998) recommend that behavioral treatments consisting of parent and teacher training and child intervention be used as a first-line therapy, with adjunctive medication use if needed, given that the effects of several behavioral treatments are comparable to those obtained with low to moderate dose stimulant.

The following sections review these behavioral treatment components, while emphasizing those

most extensively studied and included in empirical evaluations of multimodal treatment programs (e.g., MTA, 1999a) and algorithms (Pelham, 2007). One evidence-based treatment model is the *Buffalo Treatment Algorithm for ADHD* (Pelham et al., 2000; Pelham, 2007), which uses multimodal behavioral treatment across the home, clinic and school. This treatment approach is consistent with modalities used in the Summer Treatment Program for ADHD (STP; Pelham et al., 2000), which has been recognized by APA Divisions 53 and 37 as a model program and is listed by the Substance Abuse and Mental Health Services Administration (SAMHSA) as a *well-established* program for ADHD. This program along with the behavioral package in the MTA (1999a) consists of parent and teacher training and child intervention.

Poor parenting practices are a strong predictor of negative long-term outcomes in children with externalizing ADHD symptoms (Chamberlain & Patterson, 1995). Therefore, parent training should be included as a component of any comprehensive treatment plan (Pelham et al., 1998). During parent training, parents are taught skills to remediate behavior problems (Reiff & Stein, 2011). Parent training programs (e.g., *Living with Children*; Patterson, 1976; *The Incredible Years*; Webster-Stratton & Hammond, 1997; *Helping the Noncompliant Child*; McMahon & Forehand, 2003; *Parent-Child Interaction Therapy*; McNeil & Hembree-Kigin, 2011) are effective for treating disruptive behaviors (Pelham et al., 1998). More specifically, programs that include psychoeducation and specific strategies for delivering and following through on commands, shaping behaviors through principles of reinforcement, contingency-based programs using daily behavior charts, positive attending, planned ignoring, and time-out protocols, have shown positive effects on primary outcomes including externalizing ADHD symptoms (e.g., hyperactivity, impulsivity) as well as peripheral symptoms, such as aggression, oppositionality, conduct problems, child-parent interactions, and parental stress (McCleary & Ridley, 1999; Patterson et al., 1993; Reiff & Stein, 2011). Although parent programs are effective in isolation, their treatment effects may be enhanced across settings, when combined with teacher skills training and child intervention (MTA, 1999a).



Despite the effectiveness of parent behavior management training programs (e.g., Patterson et al., 1993), research has highlighted a lack of generalization into other settings such as the classroom (Webster-Stratton & Hammond, 1997). Principles of behavioral therapy often included in parent training have shown useful for teacher-led implementation in the classroom (e.g., labeled praise, positive attending, reflections, behavioral descriptions, establishing instructional control, response cost, contingent reinforcement; McIntosh, Rizza, & Bliss, 2000). Initial teacher training typically includes recommendations on limiting classroom distractions and providing accountability by using daily classroom incentive programs for on-task behavior, work completion, and following directions and also includes consultation regarding the implementation of home/school interventions (e.g., DRC). Corkum, McKinnon, and Mullane (2005) evaluated a 10-week training program for teachers of children with ADHD symptoms (no subtype differentiation), and found that children displayed greater improvement compared with those whose teachers did not participate in the training. The training topics in the teacher program included psychoeducation, effective communication, and behavioral management techniques including ignoring, natural consequences, time-out, and token systems.

Another effectiveness study used Parent-Child Interaction Therapy, and adapted its components for the preschool setting (i.e., Teacher-Child Interaction Therapy). Lyon and colleagues (2009) found that this adapted classroom intervention yielded systematic increases in teachers' use of the trained skills (e.g., positive attending, contingent reinforcement), which have shown to be effective in the parent version. Most research pertaining to teacher training for treating externalizing core and peripheral symptoms has examined programs included as one component of comprehensive behavioral packages (e.g., MTA, 1999a). For example, the behavioral treatment component of the MTA (1999a) consisted of up to 24 sessions of biweekly teacher behavior management consultation via face-to-face visits and telephone contacts (Wells et al., 2000). Although these types of treatments have demonstrated effectiveness, the research literature has consistently demonstrated maximum improvements in the generalization and

maintenance of child behavioral improvements occur when teacher training is used in combination with parent training and child intervention.

Child intervention often focuses on addressing observable maladaptive or inappropriate responses to environmental stimuli through behavior modification techniques (e.g., shaping, extinction, contingency management, token reinforcement). Given the multiple impairments that children with core and peripheral externalizing symptoms exhibit at home and school, interventions should target behavior change across settings. Daily report cards (DRCs) for behavior are effective instruments for treating externalizing ADHD symptoms (Christ, Riley-Tillman, & Chafouleas, 2009; Fabiano et al., 2010; Pelham et al., 2000;). Four decades ago, O'Leary and colleagues (1976) conducted an experimental evaluation of children with externalizing symptoms of ADHD using a DRC intervention. This intervention yielded statistically significant improved scores compared with control group on the Conners' (1969) Teacher Rating Scale and the Problem Behavior Rating scale. While DRCs are effective in isolation (Fabiano et al., 2010), research has looked at DRCs as a component of behavioral treatment packages (e.g., Owens et al., 2008; Wells et al., 2000). DRCs also offer utility and empirical support for monitoring progress over time and in varying treatment conditions (e.g., Sibley et al., 2011).

**Behavioral treatment: Limitations.** Despite the empirical support that behavioral treatments have garnered for treating core and externalizing ADHD symptoms, there are limitations to their use. Evidence from comprehensive behavioral treatment programs (e.g., MTA, 1999a; Pelham's STP, 2007) indicate that their intensive multimodal therapies must be implemented simultaneously across settings in order to conjointly affect functional domains given the lack of generalization when used in isolation. Also, compared with medication, behavioral treatments do not show the same rapidity or magnitude of effects (MTA, 1999a). Several studies have also found that the effects of behavioral interventions are limited to the window in which the treatment occurs (Pelham et al., 2000). Even the APA (2006) Task Force Report found behavioral treatments to be ineffective one year after treatment discontinuation.

**Behavioral treatment: Conclusions.** A review of the research including multisite longitudinal studies and extant meta-analyses indicates that behavioral interventions can reasonably be offered as first-line care for core and peripheral externalizing ADHD symptoms with moderate to large effect sizes (APA, 2006). In fact, the effects of many behavioral treatments for externalizing behaviors are comparable to the effects obtained with low to moderate dose stimulant medication (e.g., Pelham et al., 1998). One important distinction is the focus in medication research on reducing the core symptoms of ADHD (i.e., inattention, hyperactivity, impulsivity), while behavioral treatment studies have focused more extensively on peripheral domains of functional impairment (e.g., aggression, oppositionality, conduct problems) which are thought to mediate secondary outcomes (e.g., parent-child/peer relationships, academic/school functioning; APA, 2006; Fabiano et al., 2010).

Guidelines of the APA (2006) Task Force Report state that selecting a treatment must be guided by the balance between benefits and harms. Since there is short-term effectiveness for both medication and behavioral treatment, the report recommends that behavioral treatment be used as first-line care due to its absence of side-effects. However, given results of Jensen and colleagues' (2005) cost-effectiveness analysis of behavioral treatments used in the MTA, implementing these interventions to the intensity established in the research literature (e.g., weekly group and individual training sessions lasting for several months in the MTA) and recommended by professional organizations (AAP, 2011; APA, 2006) is not feasible within the fast-paced and insurance driven context of most primary care settings.

**Combined treatment.** Due to the limitations in using medication or behavioral treatment as a single modality, much research (e.g., AAP, 2011; Klein & Abikoff, 1997; Majewicz-Hefley & Carlson, 2007; MTA, 1999a; Pelham et al., 1988) suggests using a combination may be most effective. A consideration in using combined treatment is how the intensity level of each modality (e.g., medication dosage, frequency of behavioral intervention) influences the cumulative treatment effect. Much of the literature (e.g., Klein & Abikoff, 1997; Pelham et al., 1988)

has demonstrated the additive effects of combining medication with behavioral treatment in yielding larger treatment gains. Further, much of this literature has evaluated the relative effectiveness of "dose" (i.e., can similar or enhanced therapeutic outcomes be obtained with lower intensities of combined treatments than when either is used alone?).

In a meta-analysis of combined treatment efficacy studies, Majewicz-Hefley and Carlson (2007) found large effect sizes for inattention (1.27), hyperactivity (1.27), and impulsivity (0.91). These effect sizes are higher than those reported from meta-analytic medication monotherapy efficacy studies (e.g., Conner, Fletcher, & Swanson, 1999; Crenshaw et al., 1999). The MTA (1999a) found that while teachers reported greater improvements in hyperactive-impulsive symptoms, parents reported greater improvements in inattentive symptoms in children receiving combined treatment. This study also found combined treatment and medication alone to be comparable in treating both sub-types of ADHD symptoms. However, Swanson and colleagues (2001) found the combined treatment in the MTA study yielded a higher percentage (68%) of participants who obtained "excellent" normalized responses than medication alone (56%). Though not a statistically significant difference, it is up to individual interpretation whether this difference is clinically meaningful.

Another study conducted by Klein and Abikoff (1997) randomly assigned 89 children to eight weeks of behavioral therapy, medication, or a combination (no subtype differentiation). Behavior ratings from teachers, parents, and psychiatrists found global improvement rates of 93%, 93%, and 97%, respectively, for children receiving combined treatment; considerably higher than the 50-79% improvement rates reported for the behavioral treatment and medication groups. The behavior therapy involved function-based treatment planning for home and school consisting of parent and teacher training on operant methods (e.g., reinforcers, punishment, time out, consequences, rewards). In a case study evaluating the isolated and interactive effects between the use of common classroom contingencies for behavior (contingent teacher reprimands, time-out, no contingency) and methylphenidate at different dosage levels on disruptive and off-task behaviors, Northrup and

colleagues (1999) found the effects of medication to be moderated by the behavioral contingency conditions in place. Other studies (e.g., Blum, Mauk, McComas, & Mace, 1996; Schell, Pelham, Bender, & Andree, 1986) have confirmed these additive effects between combinations of medication and behavioral treatment.

Some studies (e.g., Abramowitz, Eckstrand, O'Leary, & Dulcan, 1992; Hoza, Pelham, Sams, & Carlson, 1992) have examined the combined effects of treatment as a function of dosage or intensity (i.e., dosage effects). Results demonstrate that a low dose medication and high dose behavioral treatment have approximately equivalent effects as high dose medication and high dose behavioral treatment. With the exception of the MTA (Vitiello et al., 2001) which showed that children receiving combined treatment ultimately received endpoint doses that were 20% less (31.1 mg/day) than those receiving medication only (38.1 mg/day), no research was found which has examined whether high dose behavioral treatment can be effectively improved by adding medication or whether comparable therapeutic effects can be obtained with a lower medication dosage if behavioral treatment is implemented previously or conjunctively. This is an important consideration with health and safety implications, especially if the same outcome can be obtained with fewer side-effects. In fact, Pelham (2012) posits that physicians can effectively prescribe lower doses of stimulant medication when starting with behavioral treatment. This option may improve parental views of acceptability and increase treatment adherence.

**Externalizing symptoms of ADHD treatment: Acceptability and adherence.** Several researchers (e.g., Eckert & Hintze, 2000; Kazdin, 1980; Witt & Elliot, 1985) have proposed that treatment *acceptability* is a significant determinant in treatment *adherence*. Further, treatments with high acceptability correlate with greater clinical outcomes, compliance and motivation, treatment satisfaction, and lower drop-out rates (Cross-Calvert & Johnston, 1990). The following sections review literature pertaining to parent and teacher treatment preferences for externalizing ADHD symptoms and their likelihood for implementing treatment as intended by the therapist.

Acceptability is defined as the overall evaluation by nonprofessionals, lay persons, and consumers of treatment regarding whether the treatment procedures are fair, appropriate, reasonable, unobtrusive, and consistent with conventional notions about treatment procedures (Kazdin, 1980). Using treatments with a high-degree of acceptability is an important consideration for providers given the sequential and reciprocal relationship between perceived treatment acceptability and treatment use, treatment integrity, and treatment effectiveness (Witt & Elliot, 1985). Research evaluating acceptability for externalizing ADHD treatments tends to be overlooked, as this information is not typically included in outcome studies (Springer & Reddy, 2004). The studies that do report this data are predominantly focused on assessing parental views toward treatment, not teachers (Vereb & DiPerna, 2004).

Research has demonstrated that behavioral treatments are more acceptable than medication for externalizing ADHD symptoms (Johnston, Hommersen, & Seipp, 2008; Wilson & Jennings, 1996). However, it is noteworthy that studies evaluating parent acceptability have revealed patterns indicating that higher acceptability ratings of proposed treatments tend to be correlated positively with a greater severity of behavior problems (Gage & Wilson, 2000; Kazdin, 1980). For example, Gage and Wilson (2000) found that while parents of children with ADHD rated medication as a more acceptable treatment than parents of children without ADHD, those parents of children without ADHD rated the behavioral treatments higher than parents of children with ADHD.

Another survey of parents ( $N = 96$ ) of children with ADHD (no subtype differentiation) found behavioral treatment alone was rated more acceptable than behavioral treatment combined with medication (Wilson & Jennings, 1996). In fact, no parent reported medication alone would be their preferred treatment. Additionally, behavior therapies in this study (e.g., self-management therapy, parent training, parent-child interaction training) involving both the child and parent were rated higher than therapies involving either the child or parent alone.

In a study of mothers ( $N = 109$ ) randomly assigned to read descriptions of parent training or

stimulant medication as treatments for a child with ADHD in a vignette, mothers rated behavioral parent training as more acceptable (Johnston, Hommersen, & Seipp, 2008). Similarly, parents in the MTA (1999a) reported they would recommend behavioral treatment more often. In fact, Pelham and colleagues (submitted for publication) found that 34% of parents assigned to the medication treatment in the MTA reported being disappointed in their assignment, compared to 7% assigned to behavioral treatment, and 4% assigned to combined treatment. In another survey of parents ( $N = 55$ ) of children with ADHD, behavior therapy was rated more acceptable than medication (Krain, Kendall, & Power, 2005; no subtype differentiation). In addition to assessing parent acceptability ratings, Krain and colleagues (2005) examined the role of ethnicity in treatment preferences, as Caucasian parents' ratings of medication were significantly higher than those of non-Caucasian parents. These results are similar to another study which found that ethnic group differences may indicate preferences for treatment, as ethnic minority (e.g., African American, Latino) parents were found to rate medication treatments less positively than Caucasians (Pham, Carlson, & Kosciulek, 2009).

Using the Treatment Evaluation Inventory (TEI; score range = 0 to 115; Kazdin, 1980), Gage (2002) found parents ( $n = 126$ ) and teachers ( $n = 45$ ) differed in how acceptable they found treatments for ADHD (no subtype differentiation). While teachers ( $M = 79.60$ ;  $SD = 11.29$ ) rated the combined treatment as more acceptable ( $M = 70.07$ ;  $SD = 19.44$ ), parents ( $M = 84.38$ ;  $SD = 14.50$ ) rated the behavioral intervention as more acceptable ( $M = 80.27$ ;  $SD = 10.82$ ). The implications of these results are unclear; however, it is important to note both parents and teachers found behavioral and combined treatments more acceptable than medication use alone.

In a study assessing teacher acceptability ratings ( $N = 147$ ) of two behavior interventions (i.e., daily report card, response-cost procedure) and medication in elementary and middle school students diagnosed with ADHD (no subtype differentiation), results indicated the DRC was more acceptable than response cost or medication (Power, Hess, & Bennett, 1995). Further, teachers rated medication as more acceptable when used in combination with behavioral intervention.

This study, along with Witt and Robbins (1985), evaluated teacher acceptability ratings of ADHD treatment (no subtype differentiation) as a function of teaching experience. For example, Power and colleagues (1995) found a moderately negative relationship between elementary and middle school teachers' experience and their acceptability ratings of medication, while Witt and Robbins (1985) found a positive relationship between elementary, middle, and high school teachers' experience and their acceptability ratings of behavioral intervention. Much teacher acceptability research on ADHD treatment is focused on the use of DRCs. DRCs are consistently found to be acceptable presumably due to time-effectiveness and ease of use (Chafouleas, Riley-Tillman, & Sassu, 2006; Waschbusch et al., 2011). Chafouleas, Riley-Tillman, and Sassu (2006) found over 60% of a teacher sample had used a home-school note program to some degree. In another study in which elementary school teachers ( $N = 79$ ) were asked to rate the acceptability of various evidence-based treatments (daily report card, medication) for use in a vignette case study of a child with ADHD (no subtype differentiation), the DRC received the highest mean ratings on the *Intervention Rating Profile-10* ( $M = 46.5$ ;  $SD = 8.4$ ) and was rated significantly more acceptable than other treatments, including medication ( $M = 37.78$ ;  $SD = 9.53$ ; Girio & Owens, 2009). Treatment adherence has an influence on the effectiveness of behavioral and psychoeducational interventions (Schmidt & Woolaway-Bickel, 2000). Treatment adherence refers to the client's level of participation in the treatment process, as well as their utilization of the treatment techniques as intended by the therapist (Springer & Reddy, 2010). Research has consistently found that treatment acceptability correlates positively with treatment adherence. In fact, Kazdin, Holland, and Crowley (1997) found parental beliefs regarding the relevance of the treatment for addressing a problem evidenced the largest magnitude in effect size differences for distinguishing who completed ( $ES = 1.70$ ) the training program and who dropped out prematurely ( $ES = 1.07$ ).

Given parental views regarding the acceptability of medication, it is not surprising that many parents demonstrate reluctance to begin or persist with a

medication trial. In a parent survey of children with ADHD (no subtype differentiation) who recently initiated a medication trial, 42% of parents reported their child discontinued medication within one month, 33% within 2 to 3 months, 21% within 4 to 6 months, and only 4% continued the regimen longer than 6 months (Toomey et al., 2012). Parents cited medication side-effects (62%) and inadequate effectiveness (34%) as primary discontinuation reasons. This study also inquired about parental attitudes regarding ADHD medication use, and found over half of participants reported to worry about long-term effects (55%) and believed ADHD medications were over-prescribed (58%). In another parental attitudes study of children ( $n = 48$ ) diagnosed with ADHD from outpatient primary care and mental health clinics affiliated with a large university hospital, only 29% of parents believed medication was necessary (dosReis et al., 2009).

In addition to parent surveys, several large insurance claims analyses have found strong indirect indicators of parental reluctance for long-term medication use. For example, Winterstein and colleagues (2008) found using a Medicaid insurance claims analysis of over 40,000 children with ADHD (no subtype differentiation) that most (51%) children who received a prescription were non-adherent one year later. In another study investigating utilization trends in a sample of 11,698 children prescribed with at least one stimulant, roughly one fourth (24%) of those children never filled a second prescription (Habel, Schaefer, Levine, Bhat, & Elliot, 2005).

Several studies have explored demographic differences in treatment adherence. For example, in an MTA study exploring the moderating effects of ethnicity on the use of behavioral parent management training strategies for ADHD (no subtype differentiation), Jones and colleagues (2010) found ethnicity to be a moderating factor for treatment adherence, though not for treatment effects. In a similar study of ADHD treatment acceptability (no subtype differentiation), Krain, Kendall, and Power (2005) found Caucasian parents' ratings of medication were higher than those of non-Caucasian parents and that Caucasian families were more likely to pursue pharmacological treatment than non-Caucasian families.

Few studies exist which assess teacher adherence to classroom-based treatments for ADHD. Many of the teacher-implemented intervention studies have found that intervention implementation is poor and adherence tends to decline after the initial intervention phase (Noell et al., 2005). However, these adherence rates vary substantially depending on the type of treatment and the teachers' level of engagement (Power et al., 2009). For example, Power and colleagues (2009) found teacher engagement (i.e., teacher support, cooperation, involvement in intervention process) in treatment was positively correlated with parent ratings of the family-school relationship in two efficacious family-school interventions for ADHD (e.g., *Family-School Success* [FSS]; *Coping with ADHD through Relationships and Education* [CARE]).

Given the high acceptability reported by teachers on a DRC intervention, it is no surprise that Vujnovic (2009) found in a study assessing the effectiveness of a classroom-based behavioral intervention using a DRC for students ( $N = 33$ ) with ADHD (no subtype differentiation), that teachers implemented adequate levels of integrity (80%) on the intervention which was found to be effective. Also, in a teacher survey ( $N = 1000$ ; Chafouleas et al., 2006) tracking the relationship between teacher-reported acceptability and the reported use of DRC intervention for behavior concerns, teachers indicated high adherence levels. Nearly half (43%) of teachers reported conducting DRC ratings, with almost a third (27%) reporting they conduct ratings more than once daily. Given DRCs are used frequently by teachers (64%), the DRC appears to be a practical intervention for implementation in schools.

**Externalizing symptoms of ADHD treatment: Conclusions.** Key conclusions can be gleaned from the literature with implications for how current treatment methods align with recommendations from practice parameters of professional organizations, which treatments have empirical support, which treatments are likely to demonstrate high acceptability and adherence, and how current treatment methods might be improved. It is noteworthy that the practice parameters from each of the three major organizations that influence child-focused providers all differ in their recommendations.

First, the AACAP (2007) recommends stimulant medication as first-line treatment for ADHD, with behavioral treatment as second-line. This is due to the compelling evidence for stimulant medications as effective treatments for ADHD. However, this recommendation was based on data obtained from the MTA-prescribed medication group in the MTA (1999a) trial, which does not align with the current PCP standard of care in community-based clinics. Next, the AAP (2011) recommends a combination of stimulant medication with behavioral treatment across settings as first-line treatment for school-age children with ADHD. Finally, in its review of pharmacological, behavioral, and combined interventions, APA (2006) concluded that all three treatments are well-established as acute interventions. However, because only medication causes side-effects, based on a risk-to-benefit analysis, the report recommended that behavioral treatment be used as first-line care, with adjunctive medication if needed.

If children are unresponsive to behavioral treatment, evidence clearly supports using combined treatment due to effectiveness and safety considerations, especially when considering the dosage effects reviewed herein possibly allowing for lower stimulant dosages when behavioral treatment is used previously and concurrently. While the practice parameters do not offer specific guidelines for treatment of externalizing ADHD symptoms, given the emphasis they provide on implementing behavioral treatments in the context of a more globally-defined construct of ADHD, it stands to reason there would only be more impetus to deliver behavioral treatment, particularly when treating core and peripheral externalizing symptoms. In sum, coupled with the evidenced parental reluctance to initiate and adhere to medication long-term, it is clear that systems of ADHD care that include behavioral treatment options may provide a viable alternative to the current standard of care for this common concern.

### **Implications for School Psychologists**

School psychologists can play a significant role in providing behavioral intervention for children displaying core and peripheral externalizing behavior problems associated with ADHD. After a thorough multi-factor assessment, school

psychologists can implement one of the several evidence-based treatment packages or techniques to address parenting skills, teacher strategies, as well as other evidence-based strategies, such as shaping, contingency management, and time out to directly intervene with the child. Improvements in ADHD symptomology obtained from behavioral interventions are often limited to the setting in which they occur (e.g., home, school; Webster-Stratton & Hammond, 1997). However, school psychologists are well positioned to coordinate efforts between the home and school and monitor treatment effectiveness, acceptability, and adherence. For example, DRCs, which are rated as highly acceptable by teachers (Girio & Owens, 2009), are a time efficient evidence-based intervention that can synchronize efforts.

School psychologists can also work to improve the effectiveness of medication treatments by collaborating with PCPs. One barrier to effective stimulant use in children and adolescents with ADHD is a lack of progress monitoring data collected by the PCP (Epstein et al., 2008). School psychologists have the assessment knowledge and progress-monitoring skills necessary to monitor the effectiveness of stimulant medications through the administration of standardized rating scales, direct observations, and other assessment techniques to maximize treatment effects. This level of data collection goes above and beyond the typical PCP standard of care, and may assist PCPs in efficiently finding an optimal dose. For more information regarding the monitoring of treatment effects of psychopharmacological medications in the school setting see Carlson and Shahidullah (2014).

### **References**

- Abramowitz, A. J., Eckstrand, D., O'Leary, S. G., & Dulcan, M. K. (1992). ADHD children's responses to stimulant medication and two intensities of a behavioral intervention. *Behavior Modification, 16*, 193–203. doi: 10.1177/01454455920162003
- American Academy of Child and Adolescent Psychiatry. (2007). Practice parameters for the

- assessment and treatment of children and adolescents with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46, 894-921. doi: 10.1097/chi.0b013e318054e724
- American Academy of Pediatrics. (2011). ADHD: Clinical practice guideline for the diagnosis, evaluation, and treatment of ADHD in children and adolescents. *Pediatrics*, 128, 1007-1022. doi:10.1542/peds.2011-2654
- American Psychological Association. (2006). *Report of the Working Group on Psychotropic Medications for Children and Adolescents: Psychopharmacological, Psychosocial, and Combined Interventions for Childhood Disorders: Evidence Base, Contextual Factors, and Future Directions*. Washington, DC: Author. Retrieved from <http://www.apa.org/pi/families/resources/child-medications.pdf>
- American Psychological Association, Division of Clinical Child and Adolescent Psychology. (2012). *Evidence-based psychosocial treatments for Attention-Deficit/Hyperactivity Disorder*. Retrieved from <http://effectivechildtherapy.com/content/attention-deficithyperactivity-disorder-adhd-0>
- Arnold, L. E., Elliott, M., Sachs, L., Bird, H., Kraemer, H., Wells, K.,... Wigal, T. (2003). Effects of ethnicity on treatment attendance, stimulant response/dose, and 14-month outcome in ADHD. *Journal of Consulting and Clinical Psychology*, 71, 713-727. doi: 10.1037/0022-006X.71.4.713
- Barkley, R. A. (2014). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment* (4<sup>th</sup> Ed.). New York: Guilford Press.
- Blum, N. J., Mauk, J. E., McComas, J. J., & Mace, F. C. (1996). Separate and combined effects of methylphenidate and a behavioral intervention on disruptive behavior in children with mental retardation. *Journal of Applied Behavior Analysis*, 29, 305-319. doi: 10.1901/jaba.1996.29-305
- Carlson, J. S., & Shahidullah, J. D. (2014). Best practices in assessing the effects of psychotropic medications on student performance. In A. Thomas & P. Harrison (Eds.), *Best practices in school psychology* (6th ed., pp. 361-374). Bethesda, MD: National Association of School Psychologists.
- Chafouleas, S. M., Riley-Tillman, C. T., & Sassu, K. A. (2006). Acceptability and reported use of daily report cards among teachers. *Journal of Positive Behavior Interventions*, 8, 174-182. doi: 10.1177/10983007060080030601
- Chamberlain, P., & Patterson, G. R. (1995). Discipline and child compliance in parenting. In M.H. Bornstein. (Eds.), *Handbook of parenting, Vol. 4: Applied and practical parenting* (pp. 205-225). Mahwah, NJ: Lawrence Erlbaum.
- Christ, T. J., Riley-Tillman, T. C., & Chafouleas, S. M. (2009). Foundation for the development and use of direct behavior rating to assess and evaluate child behavior. *Assessment for Effective Intervention*, 34, 201-213. doi: 10.1177/1534508409340390
- Conner, D. F., Fletcher, K. E., & Swanson, J. M. (1999). A meta-analysis of clonidine for symptoms of ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38, 1551-1559. doi: 0890-8567/99/3812-1551
- Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, 126, 884-888. doi: 10.1176/ajp.126.6.884
- Corkum, P. V., Mckinnon, M. M., & Mullane, J. C. (2005). The effect of involving classroom teachers in a parent training program for families of children with ADHD. *Child and Family Behavior Therapy*, 27, 29-49. doi: 10.1300/J019v27n04\_02
- Crenshaw, T. M., Kavale, K. A., Forness, S. R., & Reeve, R. E. (1999). Attention Deficit-Hyperactivity Disorder and the efficacy of stimulant medication: A meta-analysis. In Scruggs, T. & Mastropieri, M. (Eds.), *Advances in learning and behavioral disabilities*, (pp. 135-165). Greenwich, CT: JAI Press.
- Cross-Calvert, S. & Johnston, C. (1990). Acceptability of treatments for child behavior problems: Issues and implications for future research. *Journal of Clinical Child psychology*, 19, 61-74. doi: 10.1207/s15374424jccp1901\_8
- dosReis, S., Mychailyszyn, M. P., Evans-Lacko, S. E., Beltran, A., Riley, A. W., & Myers, M. A. (2009). The meaning of attention-deficit/hyperactivity disorder medication and parents' initiation and continuity of treatment for their child. *Journal of Child and Adolescent Psychopharmacology*, 19, 377-383. doi: 10.1089=cap.2008.0118

- Dumas, J. E., & Wahler, R. G. (1983). Predictors of treatment outcome in parent training: Mother insularity and socioeconomic disadvantage. *Behavioral Assessment, 5*, 301-313.
- Eckert, T. L., & Hintze, J. M. (2000). Behavioral conceptions and applications of acceptability: Issues related to service delivery and research methodology. *School Psychology Quarterly, 15*, 123-148. doi: 10.1037/h0088853
- Epstein, J. N., Langberg, J. M., Lichtenstein, P. K., Kolb, R. C., & Simon, J. O. (2013). The myADHDportal.com improvement program: An innovative quality improvement intervention for improving the quality of ADHD care among community-based pediatricians. *Clinical Practice in Pediatric Psychology, 1*, 55-67. doi: 10.1037/cpp0000004
- Epstein, J. N., Langberg, J. M., Lichtenstein, P. K., Mainwaring, B. A., Luzader, C. P., & Stark, L. J. (2008). A community-wide intervention to improve the ADHD assessment and treatment practices of community physicians. *Pediatrics, 122*, 19-27. doi: 10.1542/peds.2007-2704
- Fabiano, G. A., Vujnovic, R. K., Pelham, W. E., Waschbusch, D. A., Massetti, G. M., Pariseau, M. E.,...Volker, M. (2010). Enhancing the effectiveness of special education programming for children with ADHD using a daily report card. *School Psychology Review, 39*, 219-239. Retrieved from <http://eric.ed.gov/?id=EJ891850>
- Faraone, S. V., & Buitelaar, J. (2010). Comparing the efficacy of stimulants for ADHD in children and adolescents using meta-analysis. *European Child and Adolescent Psychiatry, 19*, 353-364. doi: 10.1007/s00787-009-0054-3
- Gage, J. D. (2002). *Parents' and teachers' acceptability of treatment for attention-deficit/hyperactivity disorder: The effects of presentation and information delivery* (Doctoral dissertation). Retrieved from <http://sunzi.lib.hku.hk/ER/detail/hkul/2694321>
- Gage, J. D., & Wilson, L. J. (2000). Acceptability of attention-deficit/hyperactivity disorder interventions: A comparison of parents. *Journal of Attention Disorders, 4*, 174-182. doi: 10.1177/108705470000400304
- Girio, E. L., & Owens, J. S. (2009). Teacher acceptability of evidence-based and promising treatment for children with attention-deficit/hyperactivity disorder. *School Mental Health, 1*, 16-25. doi:10.1007/s12310-008-9001-6
- Habel, L., Schaefer, C., Levine, P., Bhat, A., & Elliott, G. (2005). Treatment with stimulants among youths in a large California health plan. *Journal of Child and Adolescent Psychopharmacology, 15*, 62-67. doi:10.1089/cap.2005.15.62.
- Hoza, B., Pelham, W. E., Sams, S. E., & Carlson, C. L. (1992). An examination of the "dosage" effects of both behavior therapy and methylphenidate on the classroom performance of two ADHD children. *Behavior Modification, 16*, 164-192. doi:10.1177/01454455920162002
- Jensen, P. S., Garcia, J. A., Glied, S., Crowe, M., Foster, M., Schlander, M.,...Wells, K. (2005). Cost-effectiveness of ADHD treatments: Findings from the multimodal treatment study of children with ADHD. *American Journal of Psychiatry, 162*, 1628-1636. doi: 10.1176/appi.ajp.162.9.1628
- Jensen, P. S., Hinshaw, S. P., Kraemer, H. C., Lenora, N., Newcorn, J. H., Abikoff, H. B., Vitiello, B. (2001). ADHD comorbidity findings from the MTA study: Comparing comorbid subgroups. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 147-158. doi: 10.1097/00004703-200102000-00008
- Johnston, C. Hommersen, P., & Seipp, C. (2008). Acceptability of behavioral and pharmacological treatments for attention-deficit/hyperactivity disorder: Relations to child and parent characteristics. *Behavior Therapy, 39*, 22-32. doi: 10.1016/j.beth.2007.04.002
- Jones, H. A., Epstein, J. N., Hinshaw, S. P., Owens, E. B., Chi, T. C., Arnold, E. L.,...Wells, K.C. (2010). Ethnicity as a moderator of treatment effects on parent-child interaction for children with ADHD. *Journal of Attention Disorders, 13*, 592-600. doi: 10.1177/1087054709332158
- Kazdin, A. E. (1980). Acceptability of alternative treatments for deviant child behavior. *Journal of Applied Behavior Analysis, 13*, 259-273. doi: 10.1901/jaba.1980.13-259
- Kazdin, A. E., Holland, L., & Crowley, M. (1997). Family experience of barriers to treatment and premature termination from child therapy. *Journal of Consulting and Clinical Psychology, 65*, 453-463. doi: 10.1037/0022-006X.65.3.453



- Klein, R. G., & Abikoff, H. (1997). Behavior therapy and methylphenidate in the treatment of children with ADHD. *Journal of Attention Disorders*, 2, 89–114. doi:10.1177/108705479700200203
- Krain, A. L., Kendall, P. C., & Power, T. J. (2005). The role of treatment acceptability in the initiation of treatment for ADHD. *Journal of Attention Disorders*, 9, 425–434. doi:10.1177/1087054705279996
- Lyon, A. R., Gershenson, R. A., Farahmand, F. K., Thaxter, P. J., Behling, S., & Budd, K. S. (2009). Effectiveness of teacher-child interaction training in a preschool setting. *Behavior Modification*, 33, 855–884. doi: 10.1177/0145445509344215
- Majewicz-Hefley, A., & Carlson, J. S. (2007). A meta-analysis of combined treatments for children diagnosed with ADHD. *Journal of Attention Disorders*, 10, 239–250. doi: 10.1177/1087054706289934
- McCleary, L., & Ridley, T. (1999). Parenting adolescents with ADHD: Evaluation of a psychoeducation group. *Patient Education and Counseling*, 38, 3–10. doi: 10.1016/s0738-3991(98)00110-4
- McIntosh, D. E., Rizza, M. G., & Bliss, L. (2000). Implementing empirically supported interventions: Teacher-child interaction therapy. *Psychology in the Schools*, 37, 453–462. doi: 10.1002/1520-6807(200009)37:5
- McMahon, R. J., & Forehand, R. (2003). *Helping the noncompliant child: Family-based treatment for oppositional behavior* (2<sup>nd</sup> ed.). New York: Guilford Press.
- McNeil, C. B., & Hembree-Kigin, T. L. (2011). *Parent-child interaction therapy*, 2nd edition. New York: Springer.
- Mohammadi, M. R., & Akhondzadeh, S. (2007). Pharmacotherapy of ADHD: Nonstimulant medication approaches. *Expert Review of Neurotherapeutics*, 7, 191–205. doi: 10.1586/14737175.7.2.195
- Molina, B. S., Flory, K., Hinshaw, S. P., Greiner, A. R., Arnold, L. E., Swanson, J. M., Hechtman, L. (2007). Delinquent behavior and emerging substance use in the MTA at 36 months: Prevalence, course, and treatment effects. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46, 1028–1040. doi: 10.1097/chi.0b013e3180686d96
- MTA Cooperative Group. (1999a). A 14-month randomized clinical trial of treatment strategies for ADHD. The MTA Cooperative Group. Multimodal Treatment Study of Children with ADHD. *Archives of General Psychiatry*, 56, 1073–1086. doi: 10.1001/archpsyc.56.12.1073
- MTA Cooperative Group. (1999b). Moderators and mediators of treatment response for children with ADHD. *Archives of General Psychiatry*, 56, 1088–1096. doi: 10.1001/archpsyc.56.12.1088
- National Association of School Psychologists (2002). *Helping students with ADHD in the classroom: Information for teachers*. Retrieved from <http://www.nasponline.org/resources/handouts/revisedPDFs/adhd.pdf>
- National Institutes of Mental Health. (2013). *Attention Deficit Hyperactivity Disorder in Children and Adolescents Fact Sheet*. Retrieved from <http://www.nimh.nih.gov/health/publications/attention-deficit-hyperactivity-disorder-in-children-and-adolescents/index.shtml>
- Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L.,...Duhon, G. J. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review*, 34, 87–106 Retrieved from <http://eric.ed.gov/?id=EJ683507>
- Northup, J., Fusilier, I., Swanson, V., Huete, J., Bruce, T., Freeland, J.,...Edwards, S. (1999). Further analysis of the separate and interactive effects of methylphenidate and common classroom contingencies. *Journal of Applied Behavior Analysis*, 32, 35–50. doi: 10.1901/jaba.1999.32-35
- O’Leary, K. D., Pelham, W. E., Rosenbaum, A., & Price, G. H. (1976). Behavioral treatment of hyperkinetic children. *Clinical Pediatrics*, 15, 510–515. doi: 0.1177/000992287601500603
- Owens, J. S., Murphy, C. E., Richerson, L., Girio, E. L., & Himawan, L. K. (2008). Science to practice in underserved communities: The effectiveness of school mental health programming. *Journal of Clinical Child and Adolescent Psychology*, 37, 434–447. doi: 10.1080/15374410801955912
- Patterson, G. R. (1976). *Living with children: New methods for parents and teachers*. (Rev. ed.). Champaign, IL: Research Press.
- Patterson, G. R., Dishion, T. J., & Chamberlain, P. (1993). Outcomes and methodological issues

- relating to treatment of antisocial children. In T. R. Giles (Ed.), *Handbook of effective psychotherapy* (pp. 43-88). New York: Plenum Press.
- Pelham, W. E. (1999). The NIMH Multimodal Treatment Study for ADHD: Just say yes to drugs alone? *Canadian Journal of Psychiatry*, *44*, 981-990. doi: 10.1177/070674379904401004
- Pelham, W. E. (2007). Against the grain: A proposal for a psychosocial-first approach to treating ADHD – the Buffalo treatment algorithm. In K. McBurnett, & L.J. Pfiffner (Eds.), *Attention Deficit/Hyperactivity Disorder: Concepts, Controversies, New Directions* (pp. 301-316). New York: Informa Healthcare.
- Pelham, W. E. (2012). Overmedicating America's children: Medication and alternatives to treating ADHD. In N. A. Cummings & W. O'Donohue (Eds.), *Restoring Psychotherapy as the First Line Intervention in Behavioral Care* (pp. 297-325). New York: Ithaca Press.
- Pelham, W. E., Erhardt, D., Gnagy, E. M., Greiner, A. R., Arnold, L. E., Abikoff, H. B.,...& Wigal, T. (2016). *Parent and teacher evaluation of treatment in the MTA: Consumer satisfaction and perceived effectiveness*. Manuscript submitted for publication.
- Pelham, W. E., Gnagy, E. M., Greiner, A. R., Hoza, B., Hinshaw, S. P., Swanson, J. M.,...& McBurnett, K. (2000). Behavioral versus behavioral and pharmacological treatment in ADHD children attending a summer treatment program. *Journal of Abnormal Child Psychology*, *28*, 507-525. doi: 10.1023/A:1005127030251
- Pelham, W. E., Schnedler, R.W., Bender, M.E., Miller, J., Nilsson, D., Budrow, M., Ronnei, M., Paluchowski, C., & Marks, D. (1988). The combination of behavior therapy and methylphenidate in the treatment of attention deficit disorders: A therapy outcome study. In L. Bloomingtondale (Ed.) *Attention deficit disorders III: New research in attention, treatment and psychopharmacology* (pp. 29-48). London: Pergamon.
- Pelham, W. E., & Smith, B. H. (2000). Prediction and measurement of individual responses to Ritalin by children and adolescents with ADHD. In L. Greenhill & B. P. Osman (Eds.), *Ritalin: Theory and practice* (2<sup>nd</sup> ed., pp. 193-217). New York: Mary Ann Liebert.
- Pelham, W. E., Wheeler, T., & Chronis, A. (1998). Empirically supported psychosocial treatments for attention deficit hyperactivity disorder. *Journal of Clinical Child Psychology*, *27*, 190-205. doi: 10.1207/s15374424jccp2702\_6
- Pham, A. V., Carlson, J. S., & Kosciulek, J. F. (2009). Ethnic differences in parental beliefs of attention-deficit/hyperactivity disorder and treatment. *Journal of Attention Disorders*, *13*, 584-591. doi: 10.1177/1087054709332391
- Pliszka, S. R., Crismon, M. L., Hughes, C. W., Conners, K. C., Emslie, G. J., Jensen, P. S., Lopez, M. (2006). The Texas Children's Medication Algorithm Project: Revision of the Algorithm for Pharmacotherapy of ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 642-657. doi: 10.1097/01.chi.0000215326.51175.eb
- Power, T. J., Hess, L. E., & Bennett, D. S. (1995). The acceptability of interventions for ADHD among elementary and middle school teachers. *Developmental and Behavioral Pediatrics*, *16*, 238-243. doi:10.1097/00004703-199508000-00005
- Power, T. J., Soffer, S. L., Mautone, J. A., Costigan, T. E., Jones, T. E., Clarke, A. T., & Marshall, S. A. (2009). An analysis of teacher investment in the context of a family-school intervention for children with ADHD. *School Mental Health*, *1*, 107-117. doi: 10.1007/s12310-009-9005-x
- Reiff, M. I., & Stein, M. T. (2011). Attention-deficit/hyperactivity disorder. In R. R. Voight & S. M. Myers (Eds.) *Developmental and Behavioral Pediatrics*, (pp. 327-348). American Academy of Pediatrics.
- Rushton, J. L., Fant, K. E., & Clark, S. J. (2004). Use of practice guidelines in the primary care of children with ADHD. *Pediatrics*, *114*, 23-28. doi: 10.1542/peds.114.1.e23
- Safer, D. J., & Allen, R. P. (1973). Factors influencing the suppressant effects of two stimulant drugs on the growth of hyperactive children. *Pediatrics*, *51*, 660. Retrieved from <http://pediatrics.aappublications.org/content/51/4/660>
- Schachter, H. M., Pham, B., King, J., Langford, S., & Moher, D. (2001). How efficacious and safe is short acting methylphenidate for the treatment of attention-deficit disorder in children and

- adolescents: A meta-analysis. *Canadian Medical Association Journal*, *165*, 1475-1488.
- Schell, R. M., Pelham, W. E., Bender, M. E., & Andree, J. A. (1986). The concurrent assessment of behavioral and psychostimulant interventions: A controlled case study. *Behavioral Assessment*, *8*, 373-384.
- Schmidt, N., & Woolaway-Bickel, K. (2000). The effects of treatment compliance on outcome in cognitive-behavioral therapy for panic disorder: Quality versus quantity. *Journal of Consulting and Clinical Psychology*, *68*, 13-18. doi: 10.1037/0022-006X.68.1.13
- Sibley, M. H., Pelham, W. E., Evans, S. W., Gnagy, E. M., Ross, M., & Greiner, A. R. (2011). An evaluation of a summer treatment program for adolescents with ADHD. *Cognitive and Behavioral Practice*, *18*, 530-544. doi: 10.1016/j.cbpra.2010.09.002
- Springer, C., & Reddy, L. A. (2004). Measuring adherence in behavior therapy: Opportunities for research and practice. *The Behavior Therapist*, *27*, 1-9.
- Springer, C., & Reddy, L. A. (2010). Measuring parental treatment adherence in a multimodal treatment program for children with ADHD: A preliminary investigation. *Child and Family Behavior Therapy*, *32*, 272-290. doi: 10.1080/07317107.2010.515522
- Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, E. L., Conners, C. K., Abikoff, H. B., Wu, M. (2001). Clinical relevance of the primary findings of the MTA: Success rates based on severity of ADHD and ODD symptoms at the end of treatment. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 168-179. doi:10.1097/00004583-200102000-00011
- Swanson, J. M., & Volkow, N. D. (2009). Psychopharmacology: Concepts and opinions about the use of stimulant medications. *Journal of Child Psychology and Psychiatry*, *50*, 180-193. doi: 10.1111/j.1469-7610.2008.02062.x.
- Toomey, S. L., Sox, C. M., Rusinak, D., & Finkelstein, J. A. (2012). Why do children with ADHD discontinue their medication. *Clinical Pediatrics*, *51*, 763-769. doi:10.1177/0009922812446744
- U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs. (2004). *Teaching children with ADHD: Instructional strategies and practices*. Washington, DC: U.S. Department of Education. Retrieved from <http://www2.ed.gov/rschstat/research/pubs/adhd/adhd-teaching-2006.pdf>
- Vereb, R. L., & DiPerna, J. C. (2004). Teachers' knowledge of ADHD, treatments for ADHD, and treatment acceptability: An initial investigation. *School Psychology Review*, *33*, 421-428. Retrieved from <http://www.nasponline.org/publications/spr/abstract.aspx?ID=1715>
- Vitiello, B., Severe, J. B., Greenhill, L. L., Arnold, E. L., Abikoff, H. B., Bukstein, O. G., Cantwell, D. P. (2001). Methylphenidate dosage for children with ADHD over time under controlled conditions: Lessons from the MTA. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 188-196. doi: 10.1097/00004583-200102000-00013
- Vujnovic, R. K. (2009). *Examining the influence of treatment integrity: Accuracy of daily report card intervention implementation and student outcome*. Dissertation.
- Waschbusch, D. A., Cunningham, C. E., Pelham Jr., W. E., Rimas, H. L., Greiner, A. R., Gnagy, E. M., Scime, M. (2011). A discrete choice conjoint experiment to evaluate parent preferences for treatment of young, medication naïve children with ADHD. *Journal of Clinical Child and Adolescent Psychology*, *40*, 546-561. doi: 10.1080/15374416.2011.581617
- Webster-Stratton, C., & Hammond, M. (1997). Treating children with early-onset conduct problems: A comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology*, *65*, 93-109. doi: 10.1037/0022-006X.65.1.93
- Weisz, J. R., Huey, S. J., & Weersing, V. R. (1998). Psychotherapy outcome research with children and adolescents: The state of the art. In T. H. Ollendick & R. J., Prinz (Eds.), *Advances in Clinical Child Psychology*, *2*, 49-91. New York: Plenum.
- Wells, K. C., Pelham, W. E., Kotkin, R. A., Hoza, B., Abikoff, H. B., Abramowitz, A., Schiller, E. (2000). Psychosocial treatment strategies in the MTA study: Rationale, methods, and critical issues in design and implementation. *Journal of*

- Abnormal Child Psychology*, 28, 483–505. doi: 10.1023/A:1005174913412
- Wilson, L. J., & Jennings, J. N. (1996). Parents' acceptability of alternative treatments for ADHD. *Journal of Attention Disorders*, 1, 114-121. doi: 10.1177/108705479600100204
- Winterstein, A. G., Gerhard, T., Shuster, J., Zito, J., Johnson, M., Liu, H., & Saidi, A. (2008). Utilization of pharmacologic treatment in youth with attention deficit/hyperactivity disorder in Medicaid database. *Annals of Pharmacotherapy*, 42, 24-31. doi: 10.1345/aph.1K143
- Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), *Advances in school psychology* (Vol. 4, pp. 251-288). Hillsdale, NJ: Lawrence Erlbaum.
- Witt, J. C., & Robbins, J. R. (1985). Acceptability of reductive interventions for the control of inappropriate child behavior. *Journal of Abnormal Child Psychology*, 13, 59-67. doi: 10.1007/BF00918372