ISSN 2159-2020 (print)
ISSN 2161-1505 (online)

2011 Volume 15

CASP
CALIFORNIA ASSOCIATION OF SCHOOL PSYCHOLOGISTS

Contemporary School Psychology

Formerly The California School Psychologist

Published by:

California Association of School Psychologists

Special Topic Section:

Innovative Applications of a Public Health Approach to School Psychology Practice

Design, Layout and Formatting of Contemporary School Psychology, Volume 15
Completed by Heidi Holmblad and Dean L. Fregoso

ISSN 2159-2020 (print), ISSN 2161-1505 (online)

California Association of School Psychologists
Suzanne Fisher, Executive Director
1020 12th Street, Suite 200, Sacramento, CA 95814
Phone: 916.444-1595 Fax: 916.444.1597 www.casponline.org

Contemporary School Psychology

2011, Volume 15

TOC

EDITOR'S INTRODUCTION

TOC

The California School Psychologist becomes Contemporary School Psychology: CSP Continues to Link Theory and Practice

Michael Hass, PhD.,
College of Educational Studies, Chapman University

This first volume of *Contemporary School Psychology* marks yet another important milestone in the efforts of the California Association of School Psychologists (CASP) to provide its members and the profession with timely and practical scholarship. As we celebrate this accomplishment, it is important to acknowledge that this inaugural issue builds on the significant accomplishments of *The California School Psychologist* and the work of past editors and associate editors Shane Jimerson, Michael Furlong, Brent Duncan, Stephen Brock, Kristin Powers and Marilyn Wilson. Their efforts have provided the foundation on which this inaugural issue of *Contemporary School Psychology* builds.

In June of 2010, the current editorial team, Kelly Kennedy, Brian Leung and I, proposed that CASP change the name of *The California School Psychologist* to *Contemporary School Psychology*. The goal was to give the journal a name that would both reflect its breadth of scholarship – many authors outside of California have often contributed articles – and enhance the ability of the journal to attract high quality scholarship from a wider pool of authors. Ultimately, this is in service of the journal's prime goal of continuing to be, as Shane Jimerson put it in his introduction to the 2009 issue of the *California School Psychologist*, an "…invaluable resource for faculty, students, and practitioners in school psychology across the state of California." (Jimerson, 2009, p.3). Our new name communicates our commitment to continue this tradition and our aspiration to become a valuable resource for faculty, students, and practitioners nationwide.

Contemporary School Psychology is devoted to current issues in school psychology with the goal of publishing high-quality articles that make explicit the connection between theory and practice. The journal seeks to provide a venue for articles that: (1) critically review research on topics of general interest to school psychologists nationwide, (2) report research relevant to practicing school psychologists, (3) present promising practices or programs that address the needs of children and youth, and (4) critically reflect on the profession of school psychology and the challenges faced by the profession. As before, the intent of the journal is to highlight the diversity of viewpoints in the field and of the students, parents, and communities we serve.

This first issue of *Contemporary School Psychology* begins with a special topic section that includes five articles that address *innovative applications of a public health approach to school psychology practice*. These innovations include how directors of special education view Response to Intervention (RtI) at the secondary level, applications of RtI to promoting mental health and positive behavior, prevention and treatment of selective mutism, and applications of a problem-solving model to struggling readers who need intense interventions. Each of these articles explores applications of the public health model, thus expanding our thinking about school psychologists' role in prevention, Response to Intervention (RtI) and related approaches to meeting the needs of all children.

In addition, this issue of CSP contains five articles that address a wide range of topics important to school psychologists who work with diverse pupils and families, including student engagement, prevention of substance abuse, school readiness, school psychologists working with Native American youth, and psychoeducational assessment of Spanish-speaking pupils. These articles provide both practitioners and academics with timely and useful scholarship that promotes best practice.

Address correspondence and reprint requests to Michael R. Hass; Chapman University; College of Educational Studies; One University Drive, Orange, CA, 92668 or mhass@chapman.edu

This issue also introduces two new features; *Tools for Practice* and a long form *Book Review*. These sections seek to further the aim of the journal to link the world of ideas and theory to the world of practice. *Tools for Practice* will focus on specific strategies that school psychologists can use in their daily practice. Our book review section seeks to introduce readers to relevant literature, often outside typical professional texts, that will promote dialogue and inform practice.

In ending, I would like to express my gratitude to Kelly Kennedy, Brian Leung, and Stephanie Domzalski for their efforts in getting this first issue of *Contemporary School Psychology* to press. Speaking on behalf of the editorial team, I would like to thank the reviewers of manuscripts and the Board and professional staff of CASP for their continuing support of the journal. A special thanks goes to Heidi Holmblad for her efforts in ensuring that CSP is of the highest quality.

REFERENCES

Jimerson, S. (2009). The California School Psychologist Brings Science to Practice: 3 University and School Collaboration to Promote Student Success. *The California School Psychologist, 14*, 3-8.

SPECIAL FOCUS SECTION

TOC

Perspectives of Special Education Directors on Response to Intervention in Secondary Schools

Frank J. Sansosti, PhD., NCSP,
Kent State University

Shannon Goss, M.Ed.,
Olmsted Falls City School District, Ohio

Amity Noltemeyer, PhD., NCSP,
Miami University in Oxford, Ohio

Despite intensified interest in secondary school applications of Response-to-Intervention (RtI), research in this area remains sparse. This study utilized a qualitative focus group methodology to explore special education directors' perceptions of current barriers, facilitators, roles, and practices related to RtI implementation in secondary settings. Based on their unique potential to affect change and promote collaboration between general and special educators, special education directors were selected as participants. Across two focus groups, four themes emerged: systems structures, roles and attitudes, evidence-based practices, and training and professional development needs. Each theme is explored in depth, followed by practical implications, limitations, and recommendations for practice. Although numerous barriers emerged, they should be viewed not as limitations to RtI in secondary schools but rather as serving to identify the systemic factors needed to support the complexity of an RtI initiative beyond the elementary school years.

KEYWORDS: Response to Intervention, Educational Reform, Special Education, Alternative Service Delivery

To achieve the stated goals of both the *No Child Left Behind Act* (NCLB, 2001) and the *Individuals with Disability Education Act* (IDEA, 2004), Response to Intervention (RtI), a service delivery approach for providing services and interventions to students at increasing levels of intensity based on progress monitoring and data analysis (Batsche, Elliott et al., 2006), has been endorsed by educational professionals and policymakers. Moreover, RtI has been recognized as a framework that can address the academic and behavioral needs of all students, with the goal of achieving positive student outcomes within less-restrictive environments (Fuchs & Fuchs, 1998). Generally, RtI methodologies encompass high-quality, research-based classroom instruction; continuous and frequent progress monitoring; implementation of research-based academic and behavioral interventions at multiple levels; and assessment of intervention integrity (Batsche, Elliott et al., 2006).

At its inception, RtI was designed to address the academic difficulties of children suspected of having high-incidence disabilities, namely a specific learning disability, within primary grades (Bender & Shores, 2007; Mellard, Byrd, Johnson, Tollefson, & Boesche, 2004; Vaughn & Fuchs, 2003) and typically in the area of reading (Fuchs & Fuchs, 2006; Marston, 2005). More recently, practitioners have broadened the scope of RtI to include systemic approaches for the identification of and the development of interventions for behavioral difficulties (Malecki & Demaray, 2007; Sandomeirski, Kincaid, & Algozzine, 2007). From this perspective, RtI implementation serves all students, with the goal of achieving positive academic and behavioral outcomes through prevention, early identification, and intervention matched to their specific needs.

Correspondence may be sent to Frank J. Sansosti, PhD., Kent State University, College of Education, Health, and Human Services, 405 White Hall, Kent, OH 44242 or email: fsansost@kent.edu

Given that extant RtI practices largely have been applied within primary grades and typically in the area of reading, there is a growing interest among educational professionals and researchers about the degree to which RtI can be used in secondary settings. Batsche, Kavale, and Kovaleski (2006) contend that RtI can be applied to all grades, as long as there is the presence of (a) clear performance expectations and (b) methods to measure positive changes within these expectations. As school districts across the country consider the ways in which RtI can enhance student learning in a secondary setting, it is essential to view RtI as an educational change initiative rather than as an educational program or curriculum that is in vogue. Such a perspective necessitates that schools foster a structure that builds the capacity of the educational professionals and the system in which they work to sustain effective practices (Schaughency & Ervin, 2006). This concept of building capacity is not new to education, as similar change initiatives, such as Positive Behavior Support, have demonstrated that common features of successful implementation include (a) staff buy-in, (b) shared vision for change, (c) professional development/ongoing technical assistance, (d) organizational restructuring, and (e) committed administrative leadership (George, White, & Schlaffer, 2007; Kincaid, Childs, Blaise, & Wallace, 2007).

The new challenge since the passage and subsequent union of NCLB and IDEA is the manner in which administrative leaders appropriately merge general and special education to attain both schoolwide and district-wide achievement. Typically, the responsibilities of special education directors encompass overseeing programming for those students who require specialized services or accommodations within academic settings. Within this more traditional role, these administrators may operate more as supervisors or managers rather than agents of change. The practice and, more importantly, sustainability of RtI likely will require an emerging role for special education directors who increasingly "promote collaboration between general and special education teachers . . . to assure that high quality education programs are accessible to all students" (Lashley & Boscardin, 2003, p. 3). Therefore, it will be necessary for special education directors to (a) possess knowledge of the variables that influence change, (b) transform their traditional role of manager to an effective instructional leader, and (c) hone essential leadership strategies for establishing best practices for educators and positive educational outcomes (Boscardin, 2005). Unfortunately, the extant literature does not adequately address how special education directors can help build and sustain educational practices that promote a merger between general and special education practices within the context of RtI.

The purpose of this study was to obtain a better understanding of the perceptions of special education directors in regard to current barriers, facilitators, roles, and practices related to RtI implementation in secondary settings. Specifically, the researchers were interested in determining the factors that impede or facilitate RtI implementation at the secondary level. Although a variety of stakeholders could contribute to such an exploration, special education directors were chosen based on the impact that administrative leadership has on student achievement (Berends et al., 2002) as well as the essential role of administrative leadership in a successful system-wide change initiative (George et al., 2007; Kincaid et al., 2007). Based on these factors, it was anticipated that special education directors would provide a unique and informed perspective on the systemic issues influencing RtI implementation at the secondary level.

METHOD

Design

Focus group (qualitative) methodology was used to investigate special education directors' perceptions of RtI implementation in secondary schools. Due to the exploratory nature of the study and the desire to determine perspectives and emerging themes, a focus group approach was selected (Hesse-Biber & Leavy, 2006). Specifically, this study utilized a single-category approach (Krueger & Casey, 2000). In this type of design, several focus groups are conducted with a single audience (i.e., special education directors), and this approach is ideally suited for conducting this exploratory investigation (Morgan, 1998).

Setting and Participants

A total of 20 public secondary schools (grades 9-12) were randomly selected from the 85 public secondary schools within four counties in a Midwestern state. The special education director from the school district in which each school was located was invited via email to participate in one of three focus group sessions. Of the 20 invited special education directors, 19 initially indicated they would attend one of the three focus group sessions. Due to low participation (n = 1), one of three focus groups was canceled. In addition, one participant who had previously committed canceled prior to the session. In all, 17 participants attended two focus group sessions, with eight attendees at the first session and nine attendees at the second session. Both focus groups were conducted in a conference room at a professional development center for educators.

Following the focus group session, demographic information about each secondary school represented by the participants was obtained from the state department of education website. Table 1 presents the relevant comparisons to state data.

Table 1 *Characteristics of Participating Schools versus State Average*

Descriptor	Participant Mean	State Mean
Students (N)	1,655	N/A
Student Characteristics		
% Economically disadvantaged	11.6	35.5
% with Disabilities	11.7	14.5
% English Language Learners	0.6	5.5
% Caucasian	89.0	79.8
% African American	6.2	27.0
% Hispanic	0.7	5.4
% Asian/Pacific Islander	2.3	3.5
% Multiracial	1.3	4.0
Students-per-teacher ratio	17.3	15.6
% Graduation rate	94.6	88.8

In addition, to obtain further information about their professional background, involvement with RtI, and opinions on RtI, participants were asked to complete a five-question electronic survey, of which 12 were completed. These 12 respondents had a mean of 17 years of experience as a special education director, with a range of three to 25 years, and represented a variety of backgrounds, including school psychologists and regular and special education teachers. Table 2 presents a summary of their experiences and beliefs regarding RtI.

Table 2 *Participant Self-Reported Experiences and Beliefs during 2006-2007 School Year (N = 12)*

Item	*n* (Percentage)
District was involved in RTI	10 (83)
Secondary school was involved in RTI	8 (67)
Knowledge about RtI topics	
Significantly below average	0 (0)
Below average	2 (17)
Average	6 (50)
Above average	2 (17)
Significantly above average	2 (17)
General experiences with RtI	
Strongly negative	0 (0)
Negative	2 (17)
Neutral	0 (0)
Positive	10 (83)
Strongly positive	8 (67)

Data Collection

The authors developed a questioning route (Appendix A) to examine participants' perceptions of RtI practice in secondary schools. Based on the purpose of the study, pilot discussions with several practicing special education directors were used to develop the questioning route development. Minor revisions were made to the questioning route based upon expert review from two nationally recognized RtI researchers. The final questioning route included nine open-ended questions. Of the nine questions, five were "key questions" (Krueger & Casey, 2000) that drove the study and that required the greatest attention in the analysis phase.

Two digital audio recorders, a primary and backup recorder, were used to record each focus group discussion. Following each focus group, the audio file on the primary recorder was downloaded onto a computer and used for transcription. In addition to the audio files, relevant comments from participants were summarized on a flip chart by a moderator and on a notepad by an assistant.

Procedure

The author and another school psychology faculty member affiliated with the author's institution each moderated one focus group. Prior to conducting the first focus group, they met to review the questioning route and general procedures for conducting focus groups (Krueger & Casey, 2000). Each moderator was accompanied by one of two graduate students in school psychology who volunteered to be assistants during the focus groups.

Both the moderator and assistant arrived at each session 20 minutes early to prepare for the focus group session. The duties of the moderator included introducing the topic, leading the focus group discussion according to the questioning route, keeping participants on topic, prompting follow-up questions, and writing notes on a large flip chart. The assistant's responsibilities included greeting participants, securing informed consent, setting up snacks and drinks, operating the audio recorder, taking

notes, completing the integrity checklist, providing incentives at the end of the session, and addressing any unexpected problems that occurred during the focus group (e.g., lighting, noise). Each focus group lasted approximately 90 minutes. At the end of each group, participants were thanked and given a $20 gift card for a national discount retailer.

The assistant completed an integrity checklist (available upon request from the first author) to assess the degree to which the focus group procedures were implemented as intended. The procedural checklist included procedures prior to the focus group (e.g., testing audio recorders, arranging seating), during the focus group (e.g., ensuring that the moderators asked questions in sequence, avoiding the sharing of personal opinions, limiting "why" questions, writing key points on an easel pad, and providing a summary at the end of the session), and after the focus group (e.g., checking the audio recording). Procedural reliability was computed as a percentage by dividing the number of items checked on the list by the total number of focus groups conducted and then multiplying by 100. Across both focus group sessions, 100% of the procedural steps were completed.

Data Analysis

Staff from a university Bureau of Educational Research transcribed each audio file into a document. In the infrequent event that portions of the transcript were inaudible or unclear, notes taken by the moderators and the assistants were consulted to provide clarification. The three authors then used a "long table" approach (Krueger & Casey, 2000) to analyze the data contained within the document. Specifically, responses to each question were read aloud and assigned a categorical descriptor, which was arrived at through discussion and consensus. As statements were coded, they were cut from the written transcripts and sorted into labeled categories underneath that question on a "long table." Data from both focus groups were combined, resulting in five to 20 descriptive categories for each question. A summarization document that provided a written record of these categories (sorted by question), as well as direct quotations supporting each, was prepared. Next, all authors reviewed the summarization document, identifying overarching themes that were suggested by the comments across the entire questioning route. Themes were constructed based on a comparison of codes assigned to segments of text that identified unique and shared characteristics. The comments associated with each of the initial descriptive categories were then recoded into four themes through discussion and consensus: (a) *systems structures*, (b) *roles and attitudes*, (c) *evidence-based practices*, and (d) *training and professional development*.

RESULTS

Systems Structures

A strong theme that emerged from the focus group data was the importance of systems structures in RtI implementation. Unfortunately, most of the current structures were perceived as barriers to RtI implementation within the participants' secondary schools. Most participants noted that, due to inflexibility of student schedules, finding time to provide interventions to students within a secondary schedule is difficult. In addition, a few participants indicated their concern that providing interventions may require sacrificing other content, which takes away from the ability to focus on teaching the humanities or nurturing artistic talents. Related to limited time to implement interventions and the sacrificing of content, participants felt that teachers are not afforded adequate time to engage in RtI-related activities (i.e., planning time, collaborative problem-solving meetings, data collection).

Several participants also reported struggling with unanswered questions regarding how RtI should be implemented locally within their secondary schools. Specifically, the participants agreed that more clarification and guidance in translating theory and research related to elementary-based applications of RtI into procedures and structures within secondary schools was needed. Overall, participants indicated that a lack of site-specific implementation strategies translates into inconsistencies across grade levels, especially within secondary schools.

In addition to inconsistencies at the local level, special education directors also viewed state and federal policy as potential barriers to RtI at the secondary level. For example, a few participants expressed

concern and uncertainty regarding the relationship between RtI and procedural safeguards. These participants commented that they "are very used to the old way in terms of the legal ramifications and due process and all of that" and shared apprehension of how such legalities "fit into this new way of identifying kids." Moreover, several participants indicated unease regarding the lack of leadership and clarity at the state department of education concerning the combined roles of general education and special education in RtI.

Finally, the availability and management of funds to support RtI was discussed as a primary barrier to the implementation of RtI within secondary schools. Although a few participants acknowledged that 15% of IDEA funds can be used for early intervention, these funds were not viewed as sufficient for supporting the personnel and intervention programs necessary for successful implementation of RtI. As a result of a perceived disproportionate allotment of such funds to kindergarten through third grade initiatives, these concerns were thought to be particularly salient at the secondary level.

Despite the several barriers that were discussed, a few systems structures were described as facilitating RtI within the participants' schools. Participants stated that the presence of a district leadership team and the possibility of small school size would facilitate the implementation of RtI. In addition, the special education directors indicated that increased options for co-teaching at secondary levels could have the potential to assist students who are struggling academically.

Roles and Attitudes

A second theme that emerged from the focus groups, and generated a great deal of discussion, was the need for changes in the roles and attitudes of educational professionals, parents, and community members. Most often, current roles and attitudes were viewed as barriers to RtI implementation. Participants indicated that, to implement RtI effectively, their own role as special education directors would have to change. The participants discussed the need for better collaboration with other district-level administrators, particularly curriculum directors. Depending on the district, such relationships were noted both as facilitative and inhibitory. In addition to collaborating with others, the majority of participants envisioned their roles as leading and managing change. Specifically, they viewed their role in the implementation of RtI within secondary schools as agents of accountability as well as supporters to student support personnel (e.g., school psychologists).

Second, the participants perceived that principals would need to change their roles and attitudes to better support RtI practices. A few participants discussed the need for principals to be good instructional leaders who have knowledge of content areas and grade-level expectations. The consensus among the participants was that principals who have limited understanding of systemic approaches likely would inhibit the growth of RtI within secondary schools.

Third, several participants indicated that the content-oriented beliefs of secondary teachers conflict with the student-oriented focus of RtI. The perception was that secondary school teachers focus much of their time on limited blocks of instruction and often share limited concern for students who are struggling academically in all areas. Participants noted that secondary teachers should take more ownership of their students, even if the concerns raised extend beyond the subject area that they teach.

Fourth, the attitudes of many parents and outside agencies frequently were mentioned as inhibitory to the RtI process. Participants noted that many parents and advocacy groups still operate under the assumption that the only goal is to obtain an Individualized Education Program (IEP). Consequently, parents and/or advocates may resist RtI procedures that they perceive are wasting time toward this end. To remedy this situation, participants suggested that parents and outside agencies may need more information and education on RtI.

Finally, there were numerous comments indicating the need for a shared commitment among all of the above-noted stakeholders. Although this shared commitment may have been a facilitator for a few of the schools, it appeared to be a barrier for most participants. In general, comments from the special education directors reflected the notion that RtI needs to be a long-term commitment and that general education and special education must work together to reap maximum benefits.

Evidence-Based Practices

Analysis of the data indicated that evidence-based practices were viewed as essential to RtI success. In a few instances, evidence-based practices were mentioned as facilitators. For example, one participant indicated that her district has common assessments for high school, middle school, and elementary school. More often, however, the current lack of evidence-based practices at the secondary level was noted as a barrier. For example, when asked what would be needed to bring RtI to the next level in their secondary schools, participants mentioned data-driven decision making, universal screening, intervention integrity, effective interventions, and better data collection methods.

Training and Professional Development

Participants indicated that a great deal of professional development will be necessary for RtI to be implemented effectively within secondary environments. Specifically, participants focused their discussion on both proximal and distal training opportunities/requirements necessary for RtI to be brought up to scale. With regard to proximal training, special education directors indicated that in-service trainings aim at increasing the knowledge and skills of administrators, teachers, school psychologists, and other student support personnel (e.g., guidance counselors, speech language pathologists). In addition to in-service trainings, participants indicated that the implementation of a coaching model would ensure greater success for implementing RtI.

Distally, participants overwhelmingly agreed that adequate college preparation and teacher selection were necessary for RtI success. Participants discussed a need for college training programs to emphasize components of RtI as well as for districts to select teachers who can provide quality instruction.

DISCUSSION

The present investigation provides a preliminary examination of special education directors' perceptions of the facilitators of and barriers to RtI implementation at the secondary level. Based on the impact of administrative leadership on student achievement as well as the essential role of administrative leadership in a successful system-wide change initiative, special education directors were chosen as the target audience. Because little is known about the implementation of RtI within middle and high schools, this study is viewed as a first step toward understanding how special education directors perceive the feasibility of RtI in these settings. Moreover, the complexity of today's school issues requires that research methods capture the rich, nuanced, varied, and sophisticated feedback that administrators are capable of sharing. By utilizing focus groups, this study provided practical data (i.e., directors' perspectives) that can be used by school-based practitioners assisting with RtI implementation in secondary settings.

During the focus groups, special education directors indicated several components that may inhibit successful implementation of RtI in secondary schools. Most often, the inherent features of a secondary setting (i.e., structured class periods, rigid schedules, requirements for graduation, high population of students in a high school) were described as major barriers to bringing RtI to fruition. The majority of participants endorsed the view that inflexible teachers and student schedules leave virtually no time for RtI practices. For educators, this translates into inadequate collaborative opportunities for problem solving, implementing, and monitoring student interventions. For students, this suggests that benefiting from interventions could lead to the potential of foregoing required credits in a content-oriented environment.

Coupled with the complex configuration of most secondary schools were additional inhibitory factors cited by the participants as germane to this setting. The participants agreed that RtI's vitality is dependent on connecting RtI research and theories to school-based practices; the presence of administrative support through district-wide and school-based leadership teams; collaboration between general and special education; and the availability of funding, state level support, and evidence-based practices (e.g., universal screening measures, high-quality instruction, interventions, progress monitoring tools). Most special education directors remarked, however, on the scarcity of these beyond elementary grades.

Despite the overall perception that many essential RtI components are lacking in the secondary setting, participants noted that, as these emerge, their strength relies on a major shift in all key stakeholders' roles and attitudes. The extant literature confirms that, without possessing the attitudes and beliefs that an

educational change initiative such as RtI requires for improving student outcomes, fidelity of implementation and sustainability over time are less likely to occur (Sarason, 1996). Participants also stated that, for educational professionals to be more apt to support RtI, they must be highly skilled in RtI practices (e.g., high-quality instruction, data collection, intervention implementation and integrity), yet there was concern that college preparation programs are not aligned with RtI practices. Moreover, concerns that professional training must promote positive beliefs and attitudes and foster knowledge and skills beyond the pre-service years emerged.

These concerns are in concert with the participants' beliefs that a shared commitment among stakeholders, including parents and the community, is critical to establishing and maintaining the essential RtI components. This also parallels previous research efforts that, for positive and systemic outcomes, receptive beliefs and attitudes about the change and sharing a common vision must be endorsed (George et al., 2007; Kincaid et al., 2007). Nevertheless, the group conceded that this shared commitment was more often absent than present in their districts. Taken together, the participants' described views of key stakeholders' roles, attitudes, skills, and knowledge substantiated how existing educational, psychological, and administrative practices need to change (and as early as the pre-service level) to meet the diverse needs of students within an integrated educational system such as RtI.

The participants in this study also noted that the role of the special education director as a leader and change agent is critical to successful implementation of RtI in secondary settings. Specifically, participants believed that they should work with district-level administrators and student support personnel (e.g., school psychologists) to foster a shared interest in improving student outcomes. Nevertheless, careful examination of the participants' comments demonstrated uncertainty regarding their receptiveness to implementing RtI. For example, concerns raised regarding the legal bounds of RtI, the lack of clarity on how to merge general and special education, and the lack of guidance from the state department of education on how to troubleshoot these issues suggest discomfort regarding RtI implementation in the secondary setting. These concerns hint at the ancillary role the special education directors may perceive for themselves in an RtI model, which is incompatible with current recommendations for leading and managing educational reform (Boscardin, 2005).

Despite numerous negative comments (barriers) in regard to RtI implementation, participants also identified facilitating factors within their district and their secondary setting. First, some participants mentioned that they are reaching out to create a shared belief that RtI is a systemic approach that will assist in helping students. Second, a few participants described the presence of leadership teams and co-teaching between general and special education in their secondary setting as promising because it has the potential to ease a transition into RtI implementation and channels the message that all personnel must work within an integrated model. Third, one participant indicated that common assessments and student database systems presently in use were viewed as essential to RtI implementation and would be sustained in their district within RtI practices. Finally, another participant proposed that a smaller school size would lend itself to a less complicated implementation. The descriptions of these emerging and facilitating RtI components illustrate a recognition that districts will need to examine their own strengths and weaknesses to structure a system that fits their unique needs.

The results of this investigation align with previous research demonstrating the inhibitory factors often present when implementing systemic change within schools (e.g., IAT process, PBS). For example, Kincaid et al. (2007) discovered that barriers to both IAT and PBS processes include a lack of support from administrators, a lack of knowledge (e.g., misperceptions, misunderstanding), and the absence of school-wide systems structures (e.g., data collection systems, bank of interventions from which to pull). In the Kincaid study and within the findings of the current investigation, systems issues remained the most significant barriers to systemic change. Further, Kincaid et al. (2007) identified administrative support, ongoing professional development, and cohesive team membership as significant facilitators of the change process. The results of the current study offer further support for such findings and suggest that the success of RtI within secondary schools rests on the future support of and advocacy within schools as well as the additional training and professional development, including ongoing technical assistance, of various educators.

Limitations of Current Research

The results of the current analysis of the facilitators of and barriers to RtI implementation are limited by several factors. First, special education directors who participated in the focus groups were those willing to talk about their experiences with, or perceptions of, RtI implementation within secondary settings. Therefore, the extent to which the participants' reports are representative of all special education directors in other locales or to other special education directors not willing to speak about their experiences is unknown. Second, the number of participants was relatively small and may have reflected only those who were actively participating in RtI implementation at the time of this investigation. Third, the sample schools that were recruited did not demonstrate significant racial/ethnic diversity. Such a limitation may minimize the generalizability of the results to other geographic locations. Fourth, there was limited validation of the focus group questions following the authors' revisions. While this is a minor limitation, such validation would add empirical validity to the questioning route. Fifth, resultant themes identified by the authors were not validated by other experts or the participants following the analysis of data. Finally, the use of a qualitative analysis only measures perceptions of participants and does not allow for more direct measures of the actual presence or absence of factors within secondary school environments related to RtI implementation.

Recommendations for Future Research

Additional research is required to provide a more detailed picture of RtI implementation within secondary schools. Future researchers may wish to expand beyond focus groups and employ methods that provide more details regarding factors that contribute to the success or failure of RtI efforts. Additional qualitative methods such as nominal group process (Delbecq, Van de Ven, & Gustafson, 1975) and Concept Mapping™ (Novak, 1998) may be particularly useful. Specifically, the nominal group process may provide a higher level of detail by delineating which facilitators of and barriers to RtI implementation are perceived as having the highest importance. Further, Concept Mapping™ may be useful for understanding the relationships between identified themes. In addition to various qualitative methods, future researchers should utilize more direct observation or systematic analyses that measure the actual presence or absence of factors (e.g., tiered interventions, administrative support, data collection systems) within school environments. Future research also should attempt to employ larger and more diverse samples of special education directors. It will be imperative to gather data from a wide range of teachers (i.e., general and special education), services providers (e.g., school psychologists, speech-language pathologists), and administrators (i.e., school-based and district-based). Finally, future research, using longitudinal analyses of RtI implementation, should identify the variables within a variety of secondary schools that have contributed to successful systemic change. Given that significant change likely will be needed for secondary schools to sustain RtI implementation efforts, it will be helpful to examine the enabling factors within successful schools that have supported full-scale adoption.

Implications for Practice

Although many barriers were discussed, these results should not be viewed as justification for abandoning RtI initiatives in secondary schools nor for accepting these barriers as insurmountable. Rather, it is anticipated that knowledge of such barriers can serve as a starting point for positive systems change. School psychologists can play an active role in helping schools avoid barriers and establish facilitators of RtI in secondary schools in several ways. First, school psychologists should conduct a needs assessment within the building to determine the setting-specific barriers and facilitators to RtI that exist. They should work collaboratively with a building leadership team to develop a specific action plan to address any identified obstacles to RtI. Second, school psychologists should maintain an active role in professional development efforts designed to promote understanding and knowledge of RtI. This may include formalized activities (e.g., workshops, consulting) as well as more informally distributed information (e.g., an RtI bulletin board, teacher-friendly articles). Third, school psychologists should help staff to recognize the broader importance of RtI for improving the outcomes of students (as opposed to eligibility determination) and refocus discussions on the preventive nature of RtI so that everyone is approaching RtI

from a similar theoretical framework. Fourth, school psychologists could create a binder/library of resources on secondary assessment and intervention resources that teachers can access. They should begin by consulting reputable websites and professional journals to collect and create user-friendly handouts on different research-based strategies. Then, they should consider having teachers contribute ideas and resources that they have used and allow shared access to the collection of materials. Finally, school psychologists should collect outcome data on the results of RtI in the school. It is important to understand that systems change can be a difficult process with many frustrations in the first few years of change. To maintain morale, avoid burnout, and increase the likelihood of sustainability, it is necessary to have data to share that demonstrate successes along the way.

CONCLUSIONS

Although projected to be a central group for reform (Boscardin, 2005) and essential to the success of RtI, as seen in the results of this study, special education directors experience frustration and uncertainty regarding the application of RtI in secondary schools. It is hoped that this analysis will not be interpreted as evidence against RtI implementation in secondary schools but rather as an identification of the major systemic factors that need to be adapted to build the capacity necessary to support the complexity of an RtI initiative in secondary settings. For example, as a result of the analysis of this, albeit small, sample, it is evident that there is a need for pre-service training programs and ongoing professional development to address the attitudes/beliefs, knowledge, and skills of general educators, special educators, and administrators to embrace an integrated systems model to meet the needs of secondary students. In addition, there is a need for practitioners to provide examples of how systems can be modified (e.g., block scheduling, students earning credits toward graduation by participating in RtI interventions) to ensure the sustainability of RtI approaches in secondary schools. Future research and demonstration projects likely will provide continued expansion of RtI into various aspects of educational service delivery.

- - -

Frank J. Sansosti, PhD., NCSP, is an assistant professor in the School Psychology Program at Kent State University. His primary research and professional interests focus on the development and implementation of behavioral and social skills interventions for young children with autism spectrum disorders, best practice approaches for the inclusion of students with low-incidence disabilities, and the effective application of RtI practices within schools.

Shannon Goss, M.Ed., is a practicing school psychologist at Olmsted Falls City School District located in northeastern Ohio, and currently is a doctoral student at Kent State University in the School Psychology Program. Her research and professional interests include RtI, curriculum and assessment approaches for preschool and school-age children, and acquisition of cognitive and meta-cognitive study skill strategies for intermediate and middles school students.

Amity Noltemeyer, PhD., NCSP, is an assistant professor in School Psychology at Miami University in Oxford, Ohio. Her research and professional interests include RtI, systems change, disproportionality, educational equity, and literacy assessment/intervention

REFERENCES

Batsche, G., Elliott, J., Graden, J.L., Grimes, J., Kovaleski, J.F., Prasse, D., Reschly, D.J., Schrag, J., & Tilly, D.W., III. (2006). *Response to intervention: Policy considerations and implementation*. Alexandria, VA: National Association of State Directors of Special Education.

Batsche, G.M., Kavale, K.A., Kovaleski, J.F. (2006). Competing views: A dialogue on response to intervention. *Assessment for Effective Intervention*, 32(1), 6-19.

Bender, W.N., & Shores, C. (2007). *Response to intervention: A practical guide for every teacher*. Thousand Oaks, CA: Corwin Press.

Berends, M., Kirby, S.N., Naftel, S., McKelvey, C., Stockley, S., Briggs, R.J., Chun, J., Schuyler, G., Gill, B., & Heilbrunn, initial(s). (2002). NAS designs and academic achievement. In M. Berends, S. Bodilly, & S.N. Kirby (Eds.), *Facing challenges of whole-school reform* (pp. 123-142). Santa Monica, CA: Rand Corp.

Boscardin, M .L. (2005). The administrative role in transforming schools to support problem-solving based assessment for educational intervention. In R. Brown-Chidsey (Ed.), *Assessment for intervention: A problem-solving approach* (pp. 291-303). New York, NY: Guilford Press.

Delbecq, A.L., Van de Ven, A.H., & Gustafson, D.H. (1975). Group techniques for program planning: *A guide to nominal group and Delphi processes*. Glenview, IL: Scott-Foresman.

Fuchs, D., & Fuchs, L. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41(1), 93-99.

Fuchs, L., & Fuchs, D. (1998). Treatment validity: A unifying construct for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research & Practice*, 13, 204-219.

George, M.P., White, G.P., & Schlaffer, J.J. (2007). Implementing school-wide behavior change: Lessons from the field. *Psychology in the Schools*, 44(1), 41-51.

Hesse-Biber, S.N., & Leavy, P. (2006). *The practice of qualitative research*. Thousand Oaks, CA: Sage.

Kincaid, D., Childs, K., Blaise, K.A., & Wallace, F. (2007). Identifying barriers and facilitators in implementing schoolwide positive behavior support. *Journal of Positive Behavior Interventions*, 9, 174-184.

Krueger, R.A., & Casey, M.A. (2000). *Focus groups: A practical guide for applied research* (3rd ed.). Thousand Oaks, CA: Sage.

Lashley, C., & Boscardin, M.L. (2003). Special education administration at crossroads. *Journal of Special Education Leadership*, 16, 63-75.

Malecki, C.K., & Demaray, M.K. (2007). Social behavior assessment and response to intervention. In S.R. Jimerson, M.K. Burns & A.M. Vanderheyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 161-171). New York, NY: Springer.

Marston, D. (2005). Tiers of intervention in responsiveness to intervention. *Journal of Learning Disabilities*, 38(6), 539-544.

Mellard, D.F., Byrd, S.E., Johnson, E., Tollefson, J.M., & Boesche, L. (2004). Foundations and research on identifying model responsiveness to intervention sites. *Learning Disabilities Quarterly*, 27(4), 243-256.

Morgan, D.L. (1998). *Qualitative research methods* (Vol. 16). Thousand Oaks, CA: Sage.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 stat. 1425 (2001).

Novak, J.D. (1998). *Learning, creating, and using knowledge: Concept Maps™ as facilitative tools in schools and corporations*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sandomeirski, T., Kincaid, D., & Algozzine, B. (2007). Response to intervention and positive behavior support: Brothers from different mothers or sisters with different misters? Retrieved from http://www.pbis.org/news/New/Newsletters/Newsletter4-2.aspx

Sarason, S. (1996). *Revisiting the culture of school and the problem of change*. New York, NY: Teachers College Press.

Schaughency, E., & Ervin, R. (2006). Building capacity to implement and sustain effective practices to better serve children. *School Psychology Review*, 35(2), 155-166.

U.S. Department of Education. (2004). Individuals with Disabilities Improvement Act of 2004, Pub. L. 108-466. *Federal Register*, 70, 35802-35803.

Vaughn, S., & Fuchs, L.S. (2003). Redefining learning disabilities as adequate response to instruction: The promise and potential problems. *Learning Disability Research & Practice*, 18, 137-146.

APPENDIX A

Focus Group Questioning Route

Question Type	Question
Introductory Comments	Say something like: "I want to thank you for attending to today's focus group session on Response-to-Intervention (RtI). My name is (insert name), and I am a (insert profession). As you may recall from the initial email you received, the purpose of tonight's focus group is to learn more about (insert profession)'s thoughts about RtI at the secondary level." "I want to let you know that there are no right or wrong answers to tonight's questions. It is expected that you may have different points of view on the issues discussed, and I encourage you to share your viewpoint whether or not it differs from what others have said. Feel free to engage with one another in dialogue about the questions. I am here to ask questions, listen, and make sure that everyone has a chance to share. I am interested in hearing from each of you, although you shouldn't feel obligated to respond to every question. Feel free to eat or drink during the focus group session."
Opening	To begin, please tell us your name and type of setting(s) you work in.
Introductory	What is the first thing that comes to mind when you hear the phrase "Response to Intervention"?
Transition	What messages are you hearing about RtI from the national, state, and district level?
Key	Thinking about your secondary school, what are the structures or beliefs in place that facilitate RtI implementation?
	Thinking about your secondary buildings, what are the barriers to RtI implementation?
	If your school were to move toward an RtI model (or if it already has), what would you envision your role (or what was your role) in the process?
	If resources were unlimited, describe what RtI would look like if were implemented within a secondary school setting.
	Consider where your school is right now. What do you think it would take to build the capacity of your school to move RtI implementation to the next level?
Ending	Are there any concerns or hopeful comments about RtI at the secondary level that have not already been mentioned?

Tier III Assessments, Data-Based Decision Making, and Interventions

Kristin Powers and Arpita Mandal,
California State University, Long Beach

Within the Response-to-Intervention framework, students who fail to profit from high-quality general education instruction, accommodations, and supplemental instruction progress to a more intensive intervention program, sometimes referred to as "Tier III." This article describes a problem-solving approach to designing such intensive, data-based, and scientifically supported interventions for students with pervasive reading problems who have failed to respond to less rigorous services. The application of well-established (i.e., progress monitoring) and emerging methods (i.e., brief experimental analysis) for optimizing interventions are described. Two case studies are provided to illustrate how these techniques may be combined to implement Tier III interventions.

Over the past 15 years, the way in which students with learning disabilities (LD) are identified and served has received considerable scrutiny, resulting in a movement away from a traditional, psychometrically driven IQ/achievement discrepancy model toward one focused on intervention and measurable outcomes – termed Response-to-Intervention (RtI). The reauthorization of IDEA in 2004 codified this shift by requiring state educational agencies (SEA) to consider and even support local educational agencies (LEA) in their use of RtI-based eligibility criteria. At the same time, SEAs were also prohibited from requiring LEAs to use IQ/achievement discrepancy criterion in determining LD eligibility (Fuchs & Fuchs, 2006; Kame'enui, 2007).

Two major axioms of RtI are (a) learning difficulties are not inherently due to child deficits and (b) most students will experience significant educational gains from targeted, empirically based, and closely monitored interventions (Torgesen, 2007). For many students RtI represents an opportunity to avoid or exit the cycle of failure. For a few, namely those who fail to respond to sustained, empirically validated interventions implemented with fidelity, RtI results in continued individualized intervention offered by special education programs to students with learning disabilities (Gresham, 2001). Many researchers and policy makers, such as the National Reading Panel (National Institute of Child Health and Human Development, 2000), National Joint Committee on Learning Disabilities (2005), and President's Commission on Excellence in Special Education (2002), suggest that a shift to RtI instructional and diagnostic practices is preferable to former practices because students who underperform will be identified and serviced earlier, before underachievement becomes entrenched; students will qualify for the intense, individualized support of special education based on their need for that support rather than arbitrary test scores; and the assessment and intervention processes that lead to LD qualification will inform subsequent educational programming decisions.

In practice, RtI has a number of distinguishing components, including (a) universal screening to identify students at-risk for learning difficulties early, (b) progress monitoring of students' response to interventions, (c) scientifically based multi-tiered interventions to ensure intervention intensity is proportional to individual students' learning needs, and (d) problem-solving to select and improve interventions (Fuchs & Fuchs, 2006). Most students make adequate progress when provided high-quality classroom instruction (i.e., Tier I). Those who do not, receive increasingly intense and targeted interventions (i.e., Tier II) until they either make adequate improvement or are referred for special education services (i.e., Tier III). When less resource-intensive interventions such as those provided in classrooms and small groups are unsuccessful, the student progresses to "Tier III" interventions, which are more individual-

Correspondence may be sent to Kristin Powers, College of Education, ED-1, California State University, Long Beach, 1250 Bellflower Blvd, Long Beach, CA 90840-2201. Correspondence may be sent to Arpita Mandal, California State University, Long Beach. E-mail: arpita.k.mandal@gmail.com

ized, frequent, and intensive and may be provided by professionals with greater expertise (Gersten et al., 2008). In some service delivery models, Tier III represents a last push to improve academic performance before qualifying the student for special education services; however, in other models Tier III is special education (Vaughn, 2006). Regardless of the model, the purpose is to provide students with some intervention that is evidence-based and more individualized from those previously attempted.

Researchers have begun to identify assessment and intervention approaches for these so called "non-responders." For example, brief experimental analysis (BEA) has received considerable attention in school psychology literature lately because of its potential to quickly identify the most effective intervention from among a pre-set number of interventions for an individual child (Burns & Wagner, 2008). BEA examines whether a student's underachievement is due to a skill or performance deficit and involves testing a number of interventions in a short period of time to determine which method(s) best improves academic performance. The interventions all target one sub-skill and are designed to either increase academic performance by offering a reward for achieving an increased performance goal (motivation), directly teaching the skill (skill acquisition), or providing an opportunity to practice the skill (skill fluency). The intervention method in which there is the greatest improvement in academic performance may identify both the cause of underachievement and the best way to raise performance to grade level standards. Furthermore, BEA individually administered interventions appear to generalize to effective classroom-based interventions (Duhon et al, 2004).

Continuous data collection in order to modify interventions to maximize student progress has is associated with improved academic performance (Stecker, Fuchs, & Fuchs, 2005). In addition, one-on-one targeted instruction may be necessary to produce the frequent and accurate response rate associated with improved results among struggling learners (McMasters, Fuchs, Fuchs, & Compton, 2005). The purpose of this article is to present how these advancements in assessment and intervention can be integrated and applied to assisting students with pervasive reading problems. Two case studies, which are composites of actual intervention cases, will be used to illustrate the practical application of these strategies to providing Tier III interventions.

SETTING AND PARTICIPANTS

The assessment and intervention practices described in this article are based on the services provided by school psychology graduate students in a university clinic while under the direct supervision of a faculty member. Faculty observe the academic assessment and intervention sessions through one-way mirrors, assist by listening devices, and provide continual consultation to ensure scientifically based assessment and intervention practices are employed. Students are referred to the clinic by their parents and accepted for treatment based on a history of underachievement. Approximately half of the students are identified by their school districts as having a learning disability (LD) and all students exhibit at least a two year delay in one or more content areas. Teacher reports indicate most had received some type of intervention prior to being referred to the clinic. Academic difficulties in the areas of reading, writing, and/or math are targeted within this clinic; however, for the purpose of this article, students with academic concerns in the area of reading will remain the focus. The students meet individually with the clinician (i.e., graduate student) for one hour, twice a week for eight weeks. Assessments are conducted during the first two weeks and interventions are implemented during the remaining six weeks. At the end of the treatment, the clinicians meet with the students and their parents to review the results.

Problem Solving

Based on the work of Bergan and Kratochwill (as cited in Shapiro, 2004), problem solving is an iterative process in which interventions are selected based on individual need and the results of the interventions are considered in developing subsequent interventions. Problem solving consists of the following steps: (a) *problem identification:* student achievement deficits are operationalized in terms of discrepancy from benchmark , standard or norm group, (b) *problem analysis:* various hypotheses about why the student has the deficit are generated, (c) *intervention development and implementation:* a treatment plan is developed and implemented based on the hypothesized causes of the deficits and sci-

entific research on intervention effectiveness, and (d) *intervention evaluation:* fidelity of treatment (i.e., whether the treatment plan was implemented as designed) and intervention effectiveness are evaluated to determine whether to continue, modify, or discontinue the intervention. Continuous progress monitoring prior to (i.e., baseline) and throughout the intervention is critical to maximizing the impact of the intervention (Stecker et al., 2005). Problem solving has been found to improve the effectiveness of student study teams (also known as building assistance teams and pre-referral evaluation teams; Kovaleski, Gickling, Morrow, & Swank, 1999; Rosenfield & Gravois, 1996). Problem solving at Tier III to guide one-on-one instruction can be highly effective because student progress is closely monitored, interventions are highly responsive to student learning, and students have a high number of learning trials, which can lead to rapid mastery of skills. For these reasons, problem solving guided the advanced academic assessment and intervention services provided to the students who sought the services of the clinic.

Problem Identification

The clinicians first identified their students' content area or skill-set in greatest need of intervention by reviewing their application (i.e., report card, statewide assessment results, teacher checklist, etc.) and interviewing the student and his/her parent. The Woodcock-Johnson 3rd Edition (WJ-III) (Woodcock, Mather, & McGrew, 2001), an individually administered norm-referenced standardized test of achievement[1], was administered to confirm that the appropriate skill was targeted for intervention. Oral Reading Fluency[2] (ORF), as measured by a reading Curriculum -Based Measurement (CBM), was assessed in each case to establish the child's current level of functioning, identify an intervention goal based on expected rates of growth and pre-established benchmarks, and monitor the child's progress toward the intervention goal. There is considerable evidence that CBM-Reading provides reliable, valid, and sensitive data for making instructional decisions (Fuchs & Fuchs, 1999). Furthermore, instructional placement standards and goal setting guidelines necessary for establishing intervention goals exist for ORF (Shapiro, 2004).

As presented in Table 1, both students profiled had the greatest area of need in reading. A Survey Level Assessment (SLA) of ORF consisting of administering a grade-level CBM probe and subsequent lower grade-level probes until the student met the benchmark for the grade level of the probe, was administered to identify each student's current reading instructional level. Common ORF benchmarks suggested by Fuchs and Deno are: 40 – 60 words correct per minute (WCPM) for 1st and 2nd grade text, 70 – 100 WCPM for 3rd through 6th grade text, and > 120 WCPM for 7th grade text and greater (as cited in Shapiro, 2004).

Thomas, a 9th grade student who receives special education services for a learning disability, met the minimum benchmark for 4th grade material. Thus, 4th grade reading material is not overly frustrating for him but there is ample room for improvement in his ability to fluently read 4th grade material. Britney, who is repeating 3rd grade, easily passed the 2nd grade benchmark but did not meet the 4th grade benchmark; therefore, 3rd grade material will be used to monitor the effects of her intervention. Identifying students' instructional level is critical to designing effective interventions because if the text is too difficult the student will become frustrated from a lack of success, may potentially rehearse errors, and will certainly have a lower and less accurate response rate compared to reading text appropriately matched to his skills. Similarly, providing direct instruction on text that is too easy will produce minimal benefits due to a limited capacity for growth (i.e., ceiling effect) and lack of opportunity to benefit from corrective feedback.

---

[1]The Woodcock-Johnson Test of Achievement 3rd Edition (WJ-III) was administered during the first three sessions to identify intra-individual areas of strength and weakness and to compare the student's performance to a norm-group. In most cases, the value added by these data is limited because similar information may be gleaned from past performance on annual state tests. However, administering the test provided a non-intervention activity to engage in during the first three sessions while baseline data were collected. It is much easier to arrange short sessions to collect base-line data and lengthen the sessions once the intervention is added in a school-based application of a Tier 3 intervention.

[2]ORF probes are available at Dibels.uoregon.edu and Aimesweb.edformation.com

To ensure the Survey Level Assessment adequately measured each child's ORF skills, three probes on three separate days, for a total of nine probes, were administered to each child. The median of each group of three probes was identified, and the median of these three scores was used as the baseline (see Table 1; Good, Simmons, & Kame'enui, 2001). Ideally, baseline data should be collected until they are stable in order to infer changes in the data are due to the intervention and not some pre-existing trend toward improvement; however, most educators and clinicians do not wish to delay intervention long enough to follow this procedure, so it is common practice to collect three data points per day, over the course of three different days to determine a baseline (Shinn, 2002). In the clinic setting, these data points were collected over three separate days spanning a two-week period; however, in a school setting where educators have more frequent access to students, it is possible to obtain a baseline over three consecutive days.

To identify the intervention goals, the guidelines offered by Fuchs, Fuchs, Hamlett, Walz and German (1993) for setting ambitious goals were applied. Accordingly, Thomas' ORF should increase by 1.1 WCPM per week to achieve a post-intervention goal of 78 WCPM (7 WCPM increase from the 71 WCPM baseline). An intervention goal of 61 WCPM was set for Britney based on a baseline score of 52WCPM and an expect growth rate of 1.5 WCPM per week.

Table 1. *Problem Identification Data*

	Thomas	Britney
Demographic Data	Age 14, 9th grade, identified as having a learning disability, receives special education services.	Age 9, 3rd grade, general education, repeating 3rd grade.
WJ-III SS between 85 and 115 are average o Broad Reading o Broad Math o Broad Writing	o 54 SS o 85 SS o 70 SS	o 82 SS o 87 SS o 109 SS
Oral Reading Fluency (ORF); WCPM = words correct per minute	9th Grade Probe: 64 WCPM 8th Grade Probe: 61 WCPM 7th Grade Probe: 60 WCPM 6th Grade Probe: 64 WCPM 5th Grade Probe: 68 WCPM 4th Grade Probe: 70 WCPM	4th Grade Probe: 42 WCPM 3rd Grade Probe: 40 WCPM 2nd Grade Probe: 54 WCPM
Instructional Level	Fourth Grade	Third Grade
Baseline	71 WCPM	52 WCPM
Goal	78 WCPM	61 WCPM

Problem Analysis

Functional assessment of academic behavior (FAAB). Semi-structured parent and student interviews contained in the FAAB were used to collect information on each student's learning ecology, including past and present achievement, motivation, opportunity to learn at home and at school, and home-school communication (Ysseldyke & Christenson, 2002). This information is useful for understanding how long the student has struggled with reading and some possible ecologically based interventions. The FAAB confirmed that the main concern for Thomas was his reading. Thomas reported that he had struggled with reading his whole life, and he tried to avoid reading when he could. His mother reported that he did very little reading at home. She described Thomas as a very motivated and positive student, noting she often received notes from school indicating Thomas was excelling in their citizenship program.

Both Thomas and his mother attended Thomas's last Individualized Education Program (IEP) meeting; however, they were unable to describe the nature of Thomas's disability and the type of interventions and accommodations he was provided. When utilizing the FAAB with Britney, her mother reported that Britney's main concern was reading, because she was retained in 3rd grade for failing to pass the local school district's reading assessment. Her mother also reported that Britney often did not know how to do her homework and made a lot of mistakes. Britney's mother indicated it was hard for her to help her daughter, because she did not always understand the assignment since English is her second language. She reported that she had very little contact with Britney's teacher. Britney reported that she liked school, particularly math, and she knew how to get help on her class work as needed.

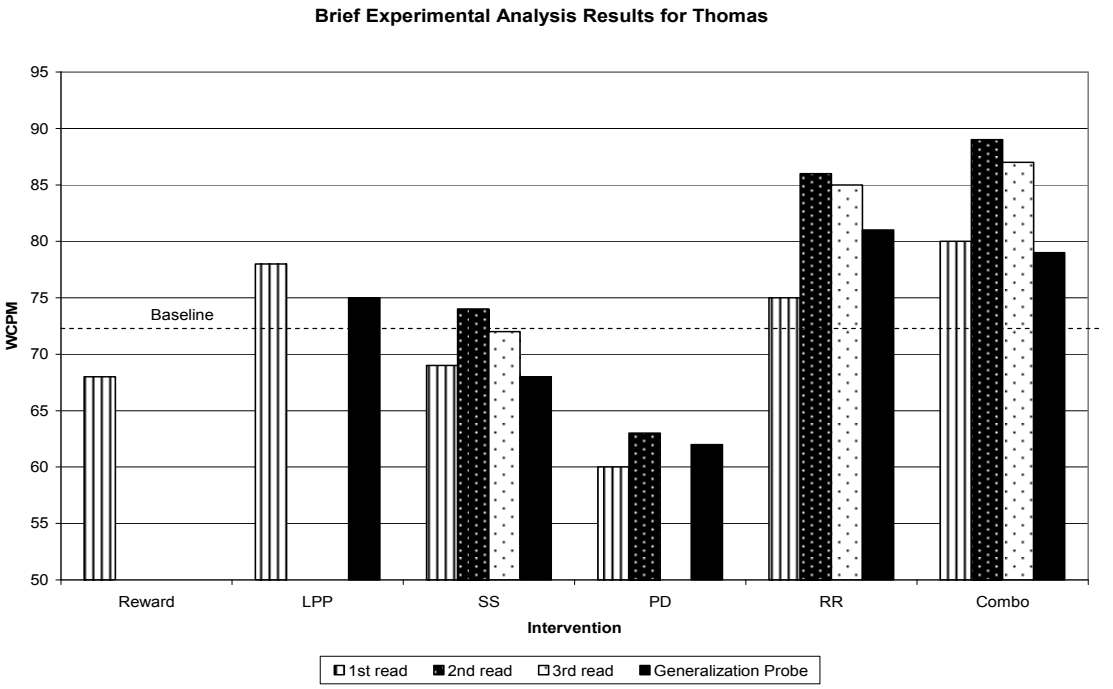
Error analysis. An error analysis of the students' responses to the WJ-III reading subtests and ORF probes was conducted by examining patterns of additions, omissions, and substitutions. An analysis of Thomas' errors indicated that he frequently left off the ending of verbs, occasionally skipped entire words and phrases without noticing, and misread content vocabulary (e.g., photosynthesis, chlorophyll). Britney was observed to read most consonant-vowel-consonant (CVC) words accurately but to have difficulty decoding longer, multi-syllable words. She also misread some high frequency words (e.g., their, said, one), read through punctuation, and lost her place easily if not tracking with her finger.

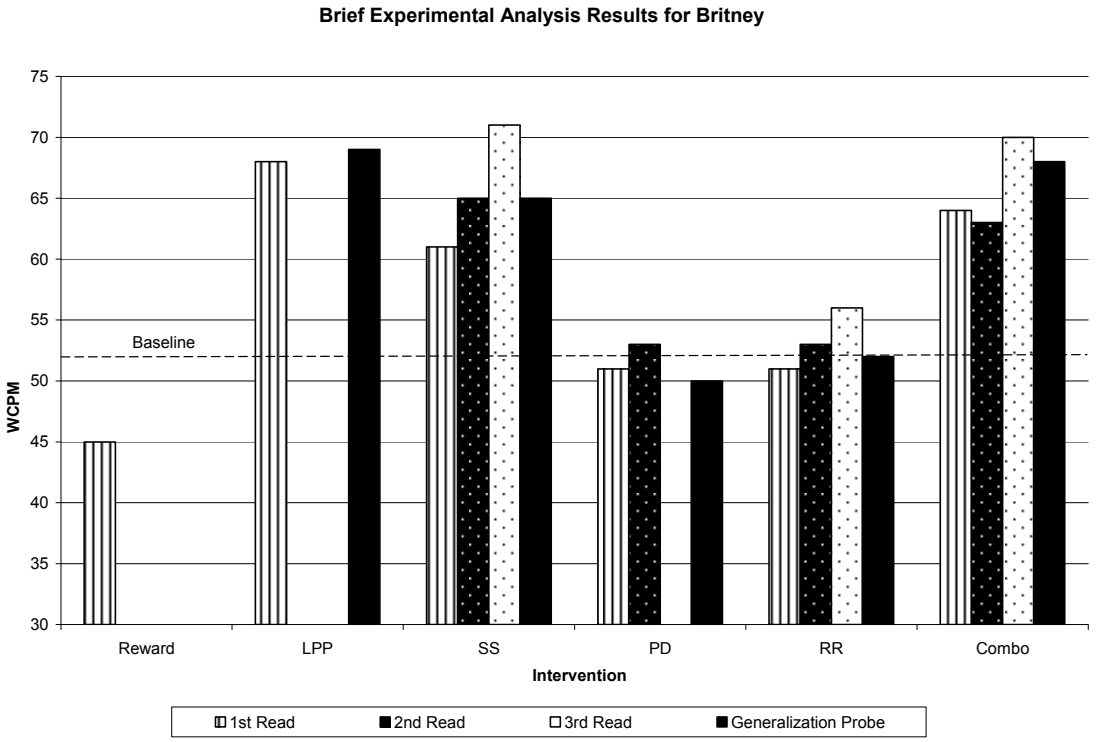
Table 2. *Problem Analysis Data*

Assessment	Thomas	Britney
Functional Assessment of Academic Behavior (FAAB) - Parent & Child Report	The main concern is reading; he has struggled with reading his whole life; he tries to avoid reading and does very little reading at home; he is a motivated and positive student; and his mother reported that his teachers keep her informed.	The main concern is reading; she was retained in 3rd grade because she didn't pass the reading benchmark; her mother also reported Britney often does not know how to do her homework and makes a lot of mistakes.
Error Analysis	Frequently leaves off the ending of verbs, skips words and phrases while reading, misreads content vocabulary (photosynthesis, chlorophyll), slowly decodes compound words.	Over-relies on alphabetic-principle to decode, difficulty blending, misses some high frequency sight words, loses place easily if not tracking with finger.

Brief experimental analysis of behavior (BEA). Next, the clinicians conducted a BEA in order to identify the type of intervention likely to produce the best results. In BEA, a single subject design is employed to observe the student's response to various interventions (Daly, Andersen, Gortmaker, & Turner, 2006). For example, to determine whether oral reading fluency can be improved by employing motivational strategies, the clinician offers the student a highly desirable reward if she improves upon her last ORF rate by 30%. If the student's performance is improved in this condition, the clinician may conclude that underachievement is due to lack of motivation rather than a skill deficit and subsequent interventions would focus on increasing the student's motivation to read. By comparing Britney's and Thomas' performance in the 'reward' condition to their baseline, one can conclude that offering incentives to increase their motivation to read was not effective (see Figure 1); therefore, their reading problems were likely due to a skill rather than motivation deficit. Subsequent analyses focused on determining whether modeling, corrective feedback, or practice lead to improved oral reading fluency rates (Daly, Witt, Marten, & Dool, 1997).

Figure 1. *Results of brief experimental analysis for Thomas and Britney*





Based on the work of Daly et al. (2006), four interventions for reading skill deficits were attempted independently and then in one combination, with each student. To examine the effects of modeling on the children's ORF, listening passage preview (LPP) was attempted. In LPP, the clinician reads a passage aloud while monitoring that the student follows along with her finger. Next, the student reads the same instructional passage aloud. Both Britney and Thomas read more fluently compared to their baseline performance after the passage had been read aloud to them (see Figure 1). In order to assess whether they generalized the knowledge they gain through LPP, they read a novel probe which contained many of the same words as the instructional probe (i.e., high word overlap). Britney's performance on the generalization probe remained much better than her baseline performance suggesting that she not only learned to read the instructional probe more accurately after hearing it read aloud, but she was able to generalize the knowledge. Thomas, however, did not perform as well on the generalization probe compared to his baseline performance suggesting he might profit more from some other type of intervention.

The next two interventions, syllable segmentation (SS) and phrase drill (PD), provide corrective feedback and accurate practice to increase ORF. SS involves providing corrective feedback and direct instruction in alphabetic principle (i.e., letter-sound correspondence and blending). The student reads an instructional passage twice while the clinician notes errors (misread or omitted words). The clinician then provides direct instruction on decoding each syllable of the error word. The student practices reading each syllable and blending the syllables to form the word. After receiving error correction on each misread word, the student re-reads the entire passage. Britney made dramatic improvement over her baseline performance in this condition, and she maintained this improvement on the generalization probe. Thomas' reading, however, did not appear to be improved much by SS.

Phase Drill (PD) combines corrective feedback and practice to improve ORF. In this condition, the student reads the entire passage while the clinician identifies any error words. Next, the clinician reads the first misread word aloud and then prompts the student to read the word aloud in isolation. If he pronounces it correctly, he reads the entire phrase or sentence containing the error word three times. After this process is completed for each misread word, the student reads the entire instructional passage and the generalization passage. Neither Thomas' nor Britney's ORF increased much past baseline in this condition, which suggests that PD is not a very effective intervention for these two students.

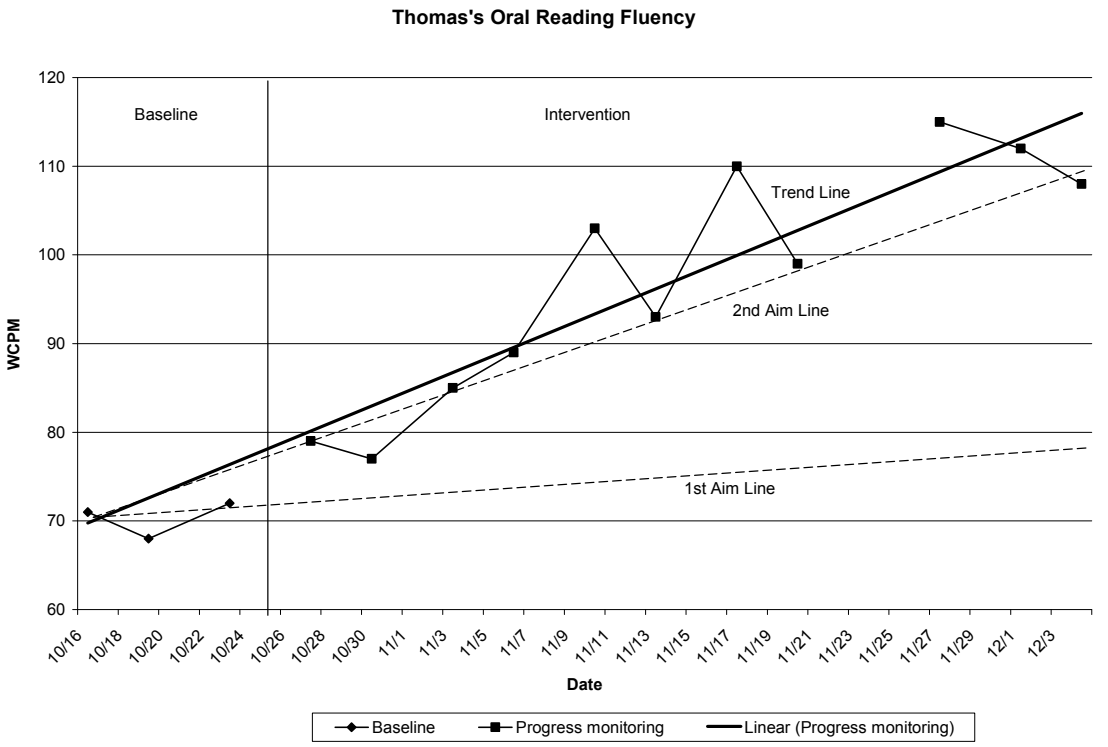
Another intervention, Repeated Reading (RR), involves having a student read a passage three times and providing feedback about her fluency after each read. Students like Thomas benefit from the multiple opportunities to respond as his performance on both the third read of the instructional probe and the generalization probe (which is read only once) was much improved over baseline. Britney did not seem to benefit from RR probably because she repeated many of her errors.

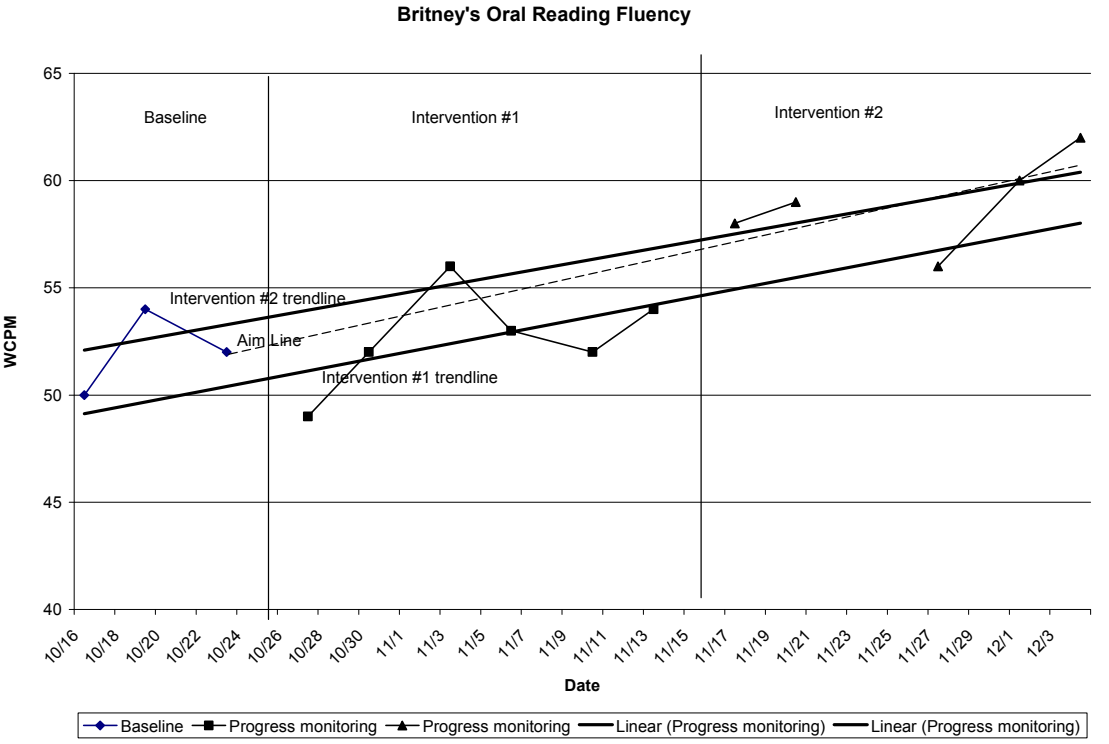
The final intervention investigated in BEA is a combination of the two most effective conditions for each student to determine if two interventions are better than one. Accordingly, Britney received a brief intervention based on combining LPP with SS and Thomas received a brief intervention based on combining LPP with RR. Both students ORF was much greater than their baseline ORF in the combination condition, and Thomas' was even greater than his performance in the most effective single intervention condition (i.e., LPP).

Intervention Development and Implementation

Individualized interventions based on direct assessment of each student's reading skills were developed. For Thomas, the results of the BEA, error analysis, and the FAAB suggest he would profit from more practice in reading. Therefore, a portion of the intervention was devoted to Guided Repeated Oral Reading of text at his instructional level (4th grade) to improve his fluency and accuracy. The effectiveness of guided oral reading is well established (National Reading Panel, 2000) and repeated reading with corrective feedback appears to be one of the most effective types of guided oral reading (Otaiba & Rivera, 2006). Thus, Thomas repeatedly read a 4th grade passage until he met the pre-established criterion of 95 WCPM. When Thomas made an error (such as dropping the suffix of a word, substituting a word, or omitting a word) the clinician tapped her pencil, which signaled Thomas to re-read the word. If Thomas failed to read the word correctly, the examiner provided the word. The error analysis indicated

Figure 2. *Progress monitoring data for Thomas and Britney*





his content vocabulary may have been underdeveloped. Accordingly, the examiner pre-viewed vocabulary and modeled reading passages from Thomas' science and social studies textbooks prior to Thomas reading the passage (i.e., Listening Passage Preview). A list of misread words was created and Thomas and the clinician defined each word using a dictionary, examined the Greek and Latin roots to the words (when applicable), and wrote sentences containing the words (Bromely, 2007). Thomas then entered each word onto a page in his vocabulary notebook, which included listing synonyms and antonyms, writing a paragraph that contained the word at least twice, and drawing a picture to represent the word (Bromely, 2007). Words selected at random from his vocabulary notebook were written on flashcards and reviewed throughout the intervention. Finally, Thomas and his mother, in consultation with the clinician, established a household rule that Thomas would read 35 minutes every day before watching television or playing videogames. In addition, Thomas was involved in charting his ORF data, which encouraged him to consider his progress toward the intervention goal.

Britney's difficulty decoding multi-syllable words and favorable response to the BEA Syllable Segmentation interventions suggest she would benefit from direct instruction on reading the parts of the multi-syllable word and then blending the parts to read the whole word (Bursuck & Damer, 2007). Britney read a list of multi-syllable words from a third grade text by segmenting the parts (i.e., 'im-ple-ment') and then blending the parts to read the whole word. If she made an error on a part, she was instructed to say the sounds of each letter in the part (i.e., /m/ /e/ /n/ /t/) and then re-read each part and finally read the whole word. Once she read the list three times with no errors, she read the passage that contained the words. The Fold-in Technique was employed to improve Britney's sight word vocabulary. This technique provides "a high level of repetition with guaranteed success" (Shapiro, 2004, p. 203) by drilling the student on seven "known" and three "unknown" sight words. Britney is bilingual, speaking primarily Spanish at home and English at school, so the clinician confirmed that Britney knew the definition of each word before adding it to the fold. In addition, Britney's mother was encouraged to read to Britney in Spanish at home on a nightly basis because many literacy skills developed in one language transfer to another (Goldenberg, 2008). Finally, Britney was directed to use a mask when reading which involves covering the text she had previously read in order to focus her gaze as she read.

Intervention Evaluation

Faculty observations of the graduate student clinicians followed by consultation sessions ensured high fidelity to the intervention plans. The parents were also asked to complete a reading log with their children to determine how often the home interventions were implemented in order to increase accountability all around.

Thomas responded very favorably to the intervention. His reading log indicated he spent at least 20 minutes per day pre-viewing and reading his text books. He reported that his general education science and social studies grades had improved from pre-viewing the text and vocabulary. Thomas also began to enjoy quiet reading time at home. Finally, Thomas' performance on three ORF probes (the median of which are graphed in Figure 2) administered at the end of each session indicated dramatic improvement. By the second week of the intervention, it became apparent that Thomas would exceed the intervention goal so a more ambitious goal was set. By the end of the six-week intervention, Thomas had made remarkable gains in ORF. The trend line trajectory indicates his ORF of 4th grade passages increased by 44 WCPM. A Survey Level Assessment found sixth grade material to be within his instructional range. Finally, Thomas correctly matched 22 vocabulary words (selected at random from his vocabulary notebook) to their definitions on a five-minute probe (Espin, Shin, & Busch, 2005).

By the third week, it was apparent that Britney would not meet her intervention goal unless the intervention was modified. Based on a review of the data collected in the problem analysis phase, observations, interviews, and error analysis conducted during the first three weeks of the intervention phase, the following modifications were made: (a) discontinued the mask as it appeared to be more of a distraction than an asset; (b) added Listening Passage Preview (LPP) of 3rd grade high-interest story books; (c) provided direct instruction on prefixes and suffixes; (d) taught Britney to recognize cognates (i.e., words that are similar in English and Spanish such as "telephone' and 'teléfono') to encourage transfer of her

literacy skills across her two languages; and (e) loaned Spanish "We Both Read" shared reading books (published by Treasure Bay) to Britney's mother to read at home after she had reported uncertainty about how to promote her daughter's literacy skills. The data graphed in Figure 2 clearly indicate the revised intervention produced better results.

Since each intervention plan consisted of a number of different scientifically based interventions selected for their likelihood of success based on BEA, FAAB, and error analysis data bundled together, it is impossible to discern which intervention or combination of interventions caused the positive responses; however, improving the trajectories of struggling readers is more important than identifying which intervention is most effective, especially considering that what is effective in one time and place with a particular student may not be effective at a later time or different place. What is critically important is that students' responses are monitored using formative assessments in order to make mid-course corrections if the intervention is failing to produce the desired outcome. Students who require Tier III interventions have little time to lose on ineffectual interventions because reading problems persist and tend to become more severe with the passage of time (Good, Simmons, & Smith, 1998).

APPLICATION TO SCHOOL SETTING

How can clinical training in Tier III assessments and intervention apply to psychologists working in schools? In our work, we have found it takes considerable effort and determination to carry these activities into professional school-based practices. Certain limitations and adjustments need to be addressed, including the availability of resources, the need for informed parent consent, and fidelity of treatment.

A relevant issue in implementing a three-tiered model successfully is the availability of resources in schools. Considering the case studies presented in this article focused on one clinician working with one student at a time, it is necessary to address whether it is possible to replicate a similar student to practitioner ratio within a school setting. Successful implementation is more a matter of utilizing available resources efficiently and appropriately, rather than attempting to acquire additional resources. Tier I requires high-quality instruction within the general education curriculum, not requiring any additional supports in the form of educators, materials, or time. Students who do not make adequate progress within the general curriculum receive additional academic support in Tier II. The critical components of Tier II include identifying these students who require more targeted interventions and implementing these interventions within small groups. To make this process as efficient as possible within the classroom, Tier II interventions can include small groups targeting specific skill deficits matched to each of the five main areas of reading instruction (phonemic awareness, alphabetic principle, fluency with text, vocabulary, and comprehension) identified by the National Reading Panel (2000; Burns & Coolong-Chaffin, 2006). Each teacher can be matched up with one group of students focusing on one of these five areas, allowing for highly focused interventions and grouping across grades (Burns & Coolong-Chaffin, 2006).

With a system such as this put into place, resources are maximally utilized and all students receive differentiated instruction matched to their needs. It also allows for school psychologists, reading specialists, and resource specialists, as well as other support staff, to be available for both consultation throughout the first two tiers and implementation of individualized instruction for those students now in Tier III who did not make adequate progress in Tiers I and II. The case studies presented in this article employ individualized interventions implemented in a one-on-one setting. Tier III can include this type of one-on-one support or small groups (i.e., two or three students) which employ the same strategies described throughout this article.

In regard to treatment fidelity, clinicians in these two case studies were under direct supervision by a faculty member who ensured that intervention practices were employed the way they were intended. In a school setting, such practices may include having a consultant observe the intervention and provide performance feedback or developing a checklist the interventionist completes after each session. In schools, it is important to utilize all available teachers, specialists, and other support staff in implementing and monitoring interventions; therefore, it may be most beneficial and time efficient for interventionists to complete a self-monitoring tool assessing intervention integrity specific to the components included in the intervention plan. In order to evaluate integrity, the components of the intervention must be opera-

tionally defined and then a checklist or rating scale including those components can be developed (Roach & Elliot, 2008). Intervention integrity monitoring tools may provide practitioners the opportunity to address both interventionists' difficulties and successes in implementation, and lead to more meaningful intervention plans (Roach & Elliot, 2008).

Not all of the services provided by school psychologists require informed parent consent. According to the National Association of School Psychologists (2010), parent consent is not required for "a school-based school psychologist to review a student's educational records, conduct classroom observations, assist in within-classroom interventions and progress monitoring, or to participate in educational screenings conducted as part of a regular program of instruction" (p. 4). Parent consent is required, however, if the consultation about a particular student is likely to be extensive and ongoing (NASP, 2010). Depending on LEA guidelines, an assessment plan signed by a parent or guardian with educational rights may or may not be required to move through Tier III. If there is a suspected disability and intervention data will be used to make a special education eligibility decision, then it is necessary to get informed parent consent (i.e., signed assessment plan). If students are receiving differentiated instruction and individualized interventions as a part of their regular program of instruction within the curriculum, informed parent consent may not be necessary. It should be noted, however, that a signed assessment plan is required when administering an individually administered, norm-referenced, standardized assessment, such as the WJ-III, demonstrated in the case studies.

CONCLUSION

The dawning of RtI does not eliminate the need for professionals with expertise in reading delays and learning disabilities; rather, it presents an opportunity for retooling by adopting some of the promising assessment and intervention methodologies that have emerged in the past decade. Many of the activities required in a data-driven system such as RtI are already in a school psychologist's skill-set (i.e., assessment, consultation, data analysis, and intervention design), making for a very natural shift in their roles in schools (Allison & Upah, 2006). There remains considerable debate about whether tests of processing and cognition, essential to traditional refer-test-diagnosis practices, produce information that enhances students' learning outcomes (Flanagan, Ortiz, & Alfonso, 2008; Gresham, Restori, & Cook, 2008). Parents often want to know why their child is struggling to learn to read when the skill comes more easily to other children. One day science may provide these answers but until then, clinicians and specialists who collect data to design and evaluate interventions and provide high-quality services proportional to student need are providing the utmost scientifically based practices.

- - -

Kristin Powers is a professor of school psychology and director of the Community Clinic for Counseling and Educational Services at California State University, Long Beach. She is co-author of the fourth edition of the book Collaborative Consultation in the Schools with Thomas Kampwirth and co-project director of an OSEP grant to train school psychologists to be effective instructional consultants.

Arpita Mandal is a school psychology graduate student at California State University, Long Beach (CSULB). She is currently completing her internship year and participating in an OSEP Personnel Preparation training grant for school psychology students to become instructional consultants in a response-to-intervention (RtI) model. Arpita serves as president of CSULB's School Psychology Student Association and is on the Political Action Committee in the Orange County Association of School Psychologists (OCASP).

REFERENCES

Allison, R., & Upah, K. (2006). The danger of response to intervention to school psychology: Why we shouldn't be afraid. *NASP Communiqué, 34*(5).

Bromley, K. (2007). Nine things every teacher should know about words and vocabulary instruction. *Journal of Adolescent and Adult Literacy, 50*(7), 528-537.

Burns, M.K., & Coolong-Chaffin, M. (2006). Response to intervention: The role of and effect on school psychology. *School Psychology Forum: Research in Practice, 1*(1), 3-15.

Burns, M.K., & Wagner, D. (2008). Determining an effective intervention within a brief experimental analysis of reading: A meta-analytic review. *School Psychology Review, 37*, 126 -136.

Bursuck, W.D., & Damer, M. (2007). Reading instruction for students who are at risk or have disabilities. Boston: Pearson.

Daly, E.J., Andersen, M., Gortmaker, V., & Turner, A. (2006). Using experimental analysis to identify reading fluency interventions: connecting the dots. *The Behavior Analyst Today, 7*(1), 133-150.

Daly, E.J., Witt, J. D., Martens, B.K., & Dool, E.J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review, 26*, 554-574.

Duhon,G.J., Noell, G.H., Witt, J.C., Freeland, J.T., Dufrene, B.A., & Gilbertson, D.N. (2004). Identifying academic skills and performance deficits: The experimental analysis of brief assessments of academic skills. *School Psychology Review, 33*, 429-443.

Espin, C.A., Shin, J., & Busch, T.W. (2005). Curriculum-based measurement in the content areas: Vocabulary-matching as an indicator of social studies learning. *Journal of Learning Disabilities, 38*(4), 353-363.

Flanagan, D.P., Ortiz, S.O., & Alfonso, V.C. (2008) Response to intervention (RTI) and cognitive testing approaches provide different but complementary data sources that inform SLD identification, *NASP Communiqué, 36*(5), 16-17.

Fuchs, D., & Fuchs, L.S. (2006). Introduction to response to intervention: What, why and how valid is it? *Reading Research Quarterly, 41*, 93-99.

Fuchs, L.S., & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom-based assessments. *School Psychology Review, 28*(4), 659-671.

Fuchs, L.S., Fuchs, D., Hamlett, C.L., Walz, L., & German, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*, 34.

Gersten, R., Compton, D., Connor, C.M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W.D. (2008). *Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide*. (NCEE 2009-4045). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved October 15, 2009 from http://ies.ed.gov/ncee/wwc/publications/practiceguides/.

Goldenberg, C. (2008). Teaching English language learners: What the research does – and does not – say. *American Educator, 32*(2), 8-23, 42-45.

Good, R.H., Simmons, D.C., & Kame'enui, E.J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading, 5*(30), 257-288, Lawrence Erlbaum Associates, Inc.

Gresham, F. (2001). *Responsiveness to intervention: An alternative approach to the identification of learning disabilities*. Paper presented at the Learning Disabilities Summit. Washington, DC. Retrieved May 8, 2002 from ERIC database.

Gresham, F., Restori, A.F., & Cook, C.R. (2008). To test or not to test: Issues pertaining to response to intervention and cognitive testing. *NASP Communiqué, 37*(1), 16-17.

Kame'enui, E.J. (2007). Responsiveness to intervention. *Teaching exceptional children, 39*, 6 -7.

Kovaleski, J.F., Gickling, E.E., Morrow, H., & Swank, P.R. (1999). High versus low implementation of instructional support teams. *Remedial and Special Education, 20*(3), 170-183.

McMaster, K., Fuchs, D., Fuchs, L.S., & Compton, D.L. (2005). Responding to nonresponders: An experimental field trial of identification and intervention methods. *Exceptional Children, 71*, 445-463.

National Association of School Psychologist. (2010). *Principles for professional ethics*. Bethesda, MD.

National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

National Joint Committee on Learning Disabilities (2005). Responsiveness to intervention and learning disabilities. Retrieved February 20, 2008 from http://www.ncld.org/images/stories/downloads/njcldreports/njcld_rti2005.pdf.

National Reading Panel (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Retrieved October 18, 2010 from http://www.nichd.nih.gov/publications/nrp/report.cfm.

Otaiba, S.A., & Rivera, M.O. (2006). Individualized guided oral reading fluency instruction for students with emotional and behavioral disorders. *Interventions in School and Clinic, 41*(3), 169-174.

President's Commission on Excellence in Special Education (2002). A new era: Revitalizing special education for children and their families. Washington, DC: US Department of Education.

Roach, A.T., & Elliot, S.N. (2008). Best practices in facilitating and evaluating intervention integrity. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V*. (pp. 195-208). Bethesda, MD: National Association of School Psychologists.

Rosenfield, S., & Gravois, T.A. (1996). *Instructional consultation teams: Collaborating for change*. New York: Guilford Press.

Shapiro, E.S. (2004). *Academic skills problems: Direct assessment and interventions, 3ed*.. New York: Guilford Press.

Shinn, M.R. (2002). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV*. (pp. 671–698). Bethesda, MD: National Association of School Psychologists.

Stecker, P.M., Fuchs, L.S., & Fuchs, D. (2005). Using CBM to improve student achievement: A review of research. *Psychology in the Schools, 42*, 795-818.

Torgesen, J.K. (2007). *Using an RtI model to guide early reading instruction: Effects on identification rates for students with learning disabilities* (Florida Center for Reading Research Technical Report No. 7). Retrieved October 15, 2007, from http://www.fcrr.org/TechnicalReports/Response_to_intervention_Florida.pdf.

Woodcock, R.W., Mather, N., & McGrew, K.S. (2001). Woodcock Johnson Psychoeducational Battery, 3rd ed. Chicago: Riverside

Vaughn, S. (2006). The state of reading research in instruction for struggling readers. *Interventions in School and Clinic*, 41(3), 169-174.

Ysseldyke, J., & Christenson, S. (2002). Functional assessment of academic behavior. Sopris West: Longmount, CO.

TOC

Response to Intervention as a Vehicle for Powerful Mental Health Interventions in the Schools

John Mark Froiland,
University of Northern Colorado

School psychologists can work within a Response to Intervention (RtI) framework to increasingly promote the mental health of students. This article shares the unfolding of two composite case studies that exemplify how a practicing school psychologist can use a problem-solving framework to deliver effective mental health interventions to individual students. The first case involved an elementary school student who markedly improved in terms of on-task behavior, and the second case concerned a high school student who overcame school refusal and negative affect. The advantages of providing brief counseling and psychological consultation prior to a possible full evaluation or special education placement are discussed.

KEYWORDS: mental health, intervention, problem-solving, elementary school students, high school students

The reauthorization of the *Individuals with Disabilities Education Act* (IDEA) emphasizes Response to Intervention (RtI) as an alternative to IQ-achievement discrepancies in the identification of learning disabilities (U.S. Department of Education, 2004). This is likely the reason that the RtI literature has emphasized academic skills. As an example of the emphasis on academics, consider that the software for keeping track of Tier I, Tier II, and Tier III curriculum-based measurement (CBM) data (e.g., AIMSWEB) have been around for quite some time, whereas behavioral progress monitoring systems and related software (e.g., the Behavior Intervention Monitoring and Assessment System; Bardos, McDougal, & Meier, 2010) are still being developed. As Gresham (2004) noted, the use of RtI logic to measure intervention response is better established in the academic domain than it is in the mental health domain, but RtI is well suited for addressing behavioral problems. Despite its potential usefulness, relatively few articles have discussed the application of RtI to behavioral problems (Hawken, Vincent, & Schumann, 2008), and RtI has been underutilized by school systems for addressing social and emotional challenges (Cheney, Flower, & Templeton, 2008).

One may view RtI as a promising framework within which school psychologists can prevent children from being unnecessarily labeled, placed, or medicated for attention and emotional disorders. In fact, in a recent study of RtI and behavior, the majority of students at risk for emotional disturbance across nine elementary schools responded well to a thorough intervention (Cheney et al., 2008). Because approximately 20% of children in the U.S. have significant emotional and behavioral difficulties, and most of these children do not receive mental health services, utilizing RtI to better serve children at risk for emotional and behavioral disorders presents an extremely important option for school psychologists (Cheney et al., 2008; Gresham, 2005; U.S. Surgeon General, 1999).

At Tier III, or the tertiary intervention level, in RtI, there is the expectation of the delivery and formative assessment of brief (i.e., 10 weeks), individualized and intense interventions prior to a full evaluation or special education placement (VanDerHeyden, Witt, & Gilbertson, 2007). This article presents two powerful case studies (composites to protect confidentiality) to educate school psychologists about how Tier III intensive interventions can meaningfully improve the mental health of children and youth in the schools. Both cases occurred in districts that were accustomed to the traditional role of the psychologist, but, in these cases, the psychologist gradually increased the implementation of a problem-solving

Correspondence concerning this article should be addressed to John Mark Froiland, University of Northern Colorado, Department of School Psychology, School of Applied Psychology and Counselor Education, McKee 298, Box 131, Greeley, CO, 80639. E-mail: john.froiland@unco.edu

model of RtI. Problem solving involves the application of the scientific method through clear problem identification, in-depth problem analysis, intervention development, ongoing measurement of students' intervention response, and refinement of the intervention based on data (Marston, 2005).

Although advanced single-subject designs for treating significant behavior problems are found frequently in school psychology journals (e.g., Lieberman, Yoder, Reichow, & Wolery, 2010), this article will present the unfolding of two cases that involved the implementation of designs and methods that are suitable to practitioners with heavy caseloads and that were implemented by a full-time practicing school psychologist serving various school districts. Although this sort of case study has been reported in the context of academic concerns (e.g., Bolt, 2005; Barnett et al., 2006), very few such cases in the context of emotional or behavioral concerns were found in school psychology journals. Although Barnett, VanDerHeyden, and Witt (2007) provide a clear, but very brief, example of how the same metric (rate of aggressive acts/minute) can be used across Tiers I, II, and III to demonstrate a positive response in Tier III for a Head Start student, the current article focuses solely on Tier III, so that readers can see how one might approach the different phases of an intense intervention case.

Barnett, Daly, Jones, and Lenz (2004) describe a case study of a Tier III intervention in which a student's disruptive behavior was decreased and the two key components of the intervention were identified through multiple treatment phases that were graphed. While the design of the case studies below is not as advanced, more case details are provided about what the school psychologist did at different phases of the problem-solving process (i.e., problem identification, problem analysis, intervention development and implementation, and intervention evaluation).

CASE STUDIES

Case 1: Problem Solving for Attention Difficulties

The referral concerns for this case involved educators stating the likelihood of a diagnosis of both ADHD and a reading disorder. There was no recorded behavioral observation data for Tier I or II; rather, the team had completed forms indicating that previous interventions (i.e., sitting toward the front of the class in Tier I and filling out a reward-based contract to pay attention more in Tier II) led to no improvement. The teacher felt overwhelmed, not knowing how to help the student. During the school-wide assistance team meeting, he said that the student was always off-task. All too often, at this point in the problem-solving process, school-wide assistance teams or individual professionals engage in ineffective practices, rapidly distributing short forms of ADHD scales and then either making a quick referral to the primary care physician or rushing into a full psychoeducational evaluation.

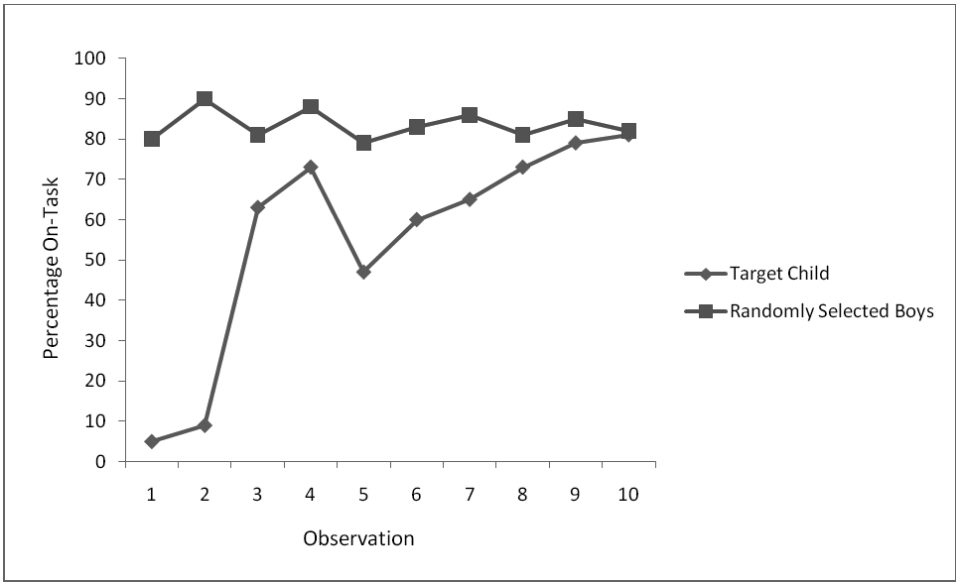
Problem identification

The teacher originally reported that the child was "always off-task" (e.g., staring out the window or at the ceiling, talking out of turn). During the team meeting, further probing elicited an estimate of the child's being on-task for a few minutes every hour. When asked for specifics about what it would look like if the child were much better, after interventions were implemented (the miracle question adapted from solution-focused therapy; Walter & Peller, 1992), the teacher explained that the student would listen attentively when the teacher was talking, talk when it was his or her turn or after being called upon, and read more quickly and accurately. The psychologist explained the importance of collecting baseline data (i.e., a period of recording observations prior to the intervention) on the child's on-task behavior that could be used to analyze the problem and provide reference points for later progress monitoring. The teacher's and the administrator's concerns about wasting time (by not beginning a full evaluation immediately) were addressed by explaining that the baseline observation data and later intervention response data could be useful even if the child did not significantly improve. In fact, by using these measurements of on-task behavior, the effects of medication or special education could be tested and compared to psychosocial interventions. The problem identification meeting concluded with the teacher, school psychologist, and other team members agreeing to a written plan for collecting the data.

Problem analysis

The student averaged a 7% on-task rate over the two weeks of baseline recording (Figure 1). While a longer baseline would have been preferable, the intervention team members were eager to see whether an intervention would work. The observations were conducted by an independent observer (another teacher), who was trained and supervised in the use of interval recording for time-on-task by the school psychologist. Simultaneous observations of randomly selected boys in the classroom indicated that the average boy was on-task approximately 84% of the time (Figure 1). After considering the data during an individual consultation with the teacher, the agreed-upon goal was for the student to reach an average of 80% on-task behavior within roughly eight weeks.

Figure 1. *Case 1: Percentage of observation intervals on-task during baseline and intervention.*



The psychologist explained that an evidence-based, self-monitoring intervention would likely lead to significant improvement in on-task behavior. Because on-task behavior was extremely low, the psychologist suggested that a sustained elevation in on-task behavior would likely lead to spillover improvements in oral reading fluency.

Intervention development and implementation

During the intervention phase, the school psychologist provided the student with approximately 30 minutes of training in self-monitoring procedures each week. The training included cognitive modeling, positive self talk, role playing with the use of the student's self-monitoring sheet, and perspective taking. Perspective taking involved the psychologist's inviting the student to play the role of the teacher by talking about his favorite hobby while standing at a chalk board and making illustrations. The student enjoyed the role play; however, what was notable was that he liked it better when the psychologist paid close attention to his presentation as opposed to when the psychologist acted bored or distracted. When the psychologist asked, "Now do you understand how your teacher feels when you do not pay attention to what he is interested in?" he appeared to have an epiphany and replied, "Yes, I do!"

At randomly selected times throughout the day, the teacher prompted the student to self-monitor by recording whether he was on-task. The student was trained to compliment himself when he was on-task and encourage himself when he rated himself as off-task. The teacher also used his own copy of the sheet to record whether the student was on-task. To build the student's self-monitoring accuracy and understanding of the teacher's view of on-task vs. off-task behavior, at the end of each day, the student

and teacher reviewed each self-monitoring instance. The teacher's ratings were used for determining whether rewards were earned, and daily rewards at school and weekly rewards at home were provided whenever the student met appropriately challenging self-set goals for attentiveness. Over the course of the intervention, the level of agreement between the student and the teacher increased from 40% to 95%, suggesting that the student was becoming a more accurate self-monitor.

The independent observer continued to collect behavioral observation data. One week after the treatment began, the student's behavior improved remarkably. However, during the third week of treatment, the psychologist noticed a significant dip in on-task level (observation 5 in Figure 1). To determine whether this reduction in on-task behavior was related to a shift in treatment integrity, the psychologist asked the teacher whether he did anything differently that week. The teacher explained that he had discontinued the intervention for a few days because he wanted to see whether the student could maintain his remarkable progress without it. Through analyzing the data together, the psychologist and teacher determined that the student needed the intervention to continue; however, the number of self-monitoring prompts was reduced so as to interfere with instruction less frequently.

Intervention evaluation

After an increase in treatment integrity, the student again increased his on-task behavior significantly. At the end of week nine, the teacher reported that the student was on-task as much as the average male student, which is in accordance with Figure 1. This favorable comparison with peers provides social validation for the student's behavioral improvement (Gresham, 2005). The teacher explained that he would adapt the intervention so that all children in the class could self-monitor simultaneously a few times a day. This enabled a successful fading of the intensity of the intervention (e.g., reduced self-monitoring frequency and a simplified self-monitoring sheet) for the targeted student, while also promoting the increased on-task behavior of children who chatted or otherwise engaged in mild levels of off-task behavior during instructional time. This class-wide adaptation of the intervention by the teacher is an example of a mental health and problem-solving consultation ideal (Gresham, 2004), promoting attentiveness and self-regulation throughout the class. This illustrates how psychologists can model a calm, confident, scientific approach to difficult problems and can help teachers become more effective problem solvers. The teacher, the reading specialist, and the parents were encouraged that the child's reading fluency gradually rose from the low-average to the average range during the course of this intervention.

Case 2: RtI Framework Made Room for the Effective Treatment of School Refusal

Problem identification

A high school guidance counselor noticed that a 10th grade girl often appeared sad, tired, and irritable. Her grades were dropping, and she often vehemently refused to attend school. Although she was perennially popular among classmates, she was starting to withdraw from them. Upon further probing, the guidance counselor explained that the student had 20 school refusals in the first 20 weeks of the school year and that her grades dropped from Bs in previous years to Cs at the time of the meeting. She also cried at school often and no longer engaged in conversation with her peers during free time (e.g., during lunch). After sharing this, the guidance counselor asked the school psychologist, "Can you evaluate her for depression and refer her to a physician, if necessary?"

The school had explained to all students at the beginning of the school year that regular attendance was expected, a physician's note was required for an absence, and that students would serve one detention for each unexcused absence after the second unexcused absence. Records showed that the student averaged approximately one unexcused absence per week for the first 10 weeks of the year. During Tier II, the student signed a contract with the guidance counselor, agreeing to increase her attendance. The guidance counselor was to write positive notes to the family for her improving her attendance rate for any given week, and the family and student would be reminded of the school discipline policy and the consequences associated with continued absenteeism. During 10 weeks of Tier II, the student continued to average approximately one unexcused absence per week. Toward the end of Tier II, the guidance counselor suspected that depression, rather than defiance, was the underlying reason for school refusal.

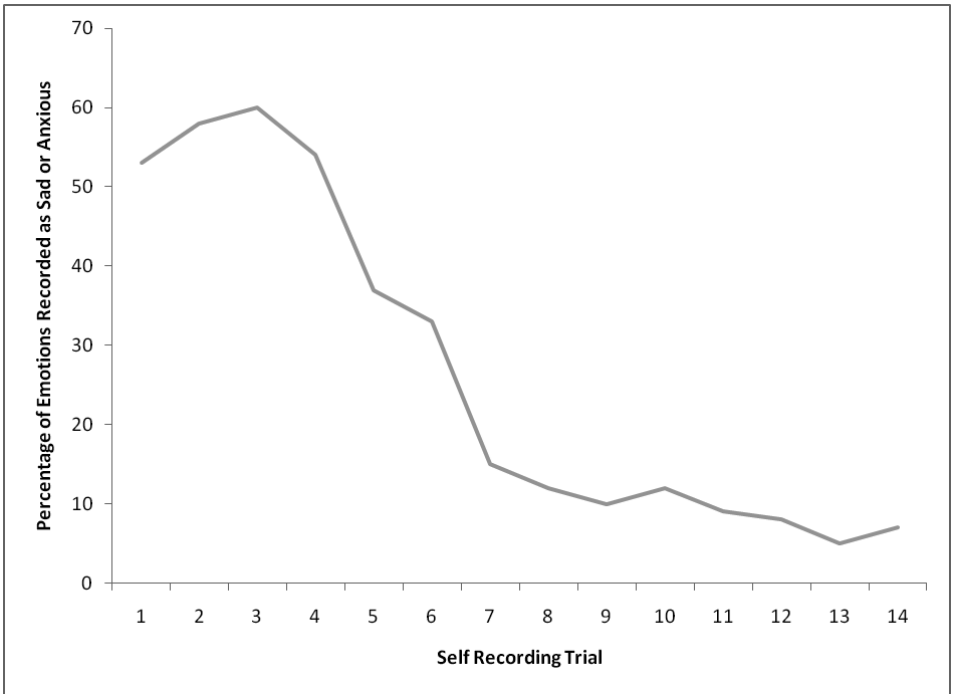
Typically, many school mental health professionals in similar circumstances screen for depression, make sure that such students are not suicidal, and make an outside referral if necessary. If the school psychologist would have done the last item without first developing an intervention, the student would have likely ended up on medication or would have accumulated so many days of absence that she would not have completed her courses that year.

Problem analysis

Pre-intervention self-rating scales provided by the school psychologist revealed that the student struggled with clinically significant levels of depression, including feeling sad, loss of interest in many daily activities, hyposomnia, fatigue, self-disparaging statements, and body image distortion as well as some symptoms of anxiety.

The school psychologist loaned the student a beeper that vibrates at random times, which he set to an average of once an hour. He taught her how to record on a handout whether she felt happy, calm, angry, sad, nervous, or neutral each time that the beeper vibrated. Because this was a school-based intervention, she had the beeper on only during the school day. During a week of baseline recordings, she felt sad or nervous an average of 55% of the time (Figure 2).

Figure 2. *Case 2: Self-recorded sad or anxious feelings during baseline and intervention.*



Intervention

The psychologist approached the case from a combination of a problem-solving and a traditional model. The psychologist suggested that he provide counseling from a cognitive-behavioral perspective, track intervention response data (i.e., attendance and student self-monitoring of emotions), and measure levels of depression again in eight weeks using in-depth behavioral rating scales. The psychologist stated that, if the student was still in the clinically significant range on the rating scales and did not improve on the RtI measures, he would provide a psychological report that the mother could bring to a physician or clinical psychologist to obtain further treatment. Although this approach involved a paradigm shift in how the school psychologist would approach a situation like this, the guidance counselor agreed to the

plan, contingent on the mother's acceptance of the plan. The mother agreed to the plan at the end of a one-hour phone consultation session.

The student was interested in the interview process as well as learning about the interpretation of the depression and anxiety inventory results. Due to the student's initially being somewhat guarded during counseling, the psychologist decided to first develop a more collaborative relationship through non-directive counseling. The student made progress during these first two weeks, as seen in her perfect attendance and in a decrease in the percentage of times that she recorded feeling sad or nervous (Figure 2). She decreased her sad/nervous feelings, however, from an average of 55% of the time to an average of only 35%. Thus, the school psychologist decided to determine whether she could make greater improvement by directly teaching her how to replace negative or irrational thoughts (e.g., "I can't stand being at school") with more adaptive ones (e.g., "I'd rather be shopping, but I can learn at least one thing in this class today"). She was instructed to use the beeper not only as a self-monitoring tool but also as a trigger to identify what she was thinking and then to replace any irrational or negative thoughts with positive or more rational thoughts. This led to a stronger decline in self-reported sadness/anxiousness (Figure 2).

Intervention evaluation

The guidance counselor tracked the student's weekly attendance rates. Initially, the student had refused to attend school three times in the week prior to intervention. Improvement was noticed early on, in that there were no instances of refusal to attend school during the first few weeks of intervention. By sharing the progress with the family and school personnel, momentum was gained, and greater expectation for further improvement was developed.

At the end of the semester, a school administrator reported to the psychologist that the student had only one absence in the 15 weeks after the intervention commenced. This is a remarkable improvement on a social impact measure (i.e., an indicator that is considered to be critical for functioning in society), particularly because social impact measures are considered more difficult to change than more proximal measures such as anxiety (Gresham, 2005). The student also came to the psychologist's office to share that she obtained straight A's for the first time in her life. More importantly, she viewed herself as capable of succeeding in school if she applied herself, overcoming a previously low academic self-esteem. In accordance with improvement on daily ratings of her emotions (Figure 2), depression *t*-scores scores dropped from within the clinically significant range ($t = 75$; 50, with a standard deviation of 10, is average) to the average range ($t = 52$) at the end of eight weeks and then reached the below average range ($t = 40$) at 15 weeks. The maintenance of gains at 15 weeks boded well, particularly because the frequency of the counseling sessions had been reduced to twice a month after the first eight weeks.

The student's self-reported social anxiety and physiological symptoms of anxiety also were reduced from the at-risk to the normal range. In addition, the student reported falling to sleep easily, getting plenty of sleep, and waking up feeling refreshed. The student also noted that she felt like "a brand new person with a bright outlook on life." Her mother reported that she observed her daughter as smiling more, crying less, and speaking more positively than prior to the onset of the school refusal. The parent's report provided further social validation for the effectiveness of the intervention (Gresham, 2004) and indicated that skills gained in school-based counseling were being transferred to the home.

DISCUSSION

The case study involving self-monitoring provides an example of successful behavioral consultation (in Tier III) provided by a school psychologist. While the intervention and its measurement were similar to the single-subject design experiments published in school psychology journals (see Briesch & Chafouleas, 2009, for a review), the methods were implemented by one school psychologist and available school staff, rather than by a team of university researchers. This complements the work of university researchers by showing how a school psychologist, who is inundated with numerous cases and the exigencies of everyday practice, may be able to successfully apply a field-friendly, single-subject design and a self-monitoring intervention. This stands in contrast to the belief of single-subject design experts that practitioners have to negotiate the design and methods with problem-solving team members, often sacrificing internal validity for field expedience (Barnett et al., 2004).

In case study 1, a short baseline reduced the internal validity, but the comparison with randomly selected boys indicates that those not receiving the intervention had stable levels of on-task behavior. Case study 2, involving the student who overcame school refusal and negative affect, provides an example of utilizing both progress monitoring and behavioral rating-scale data in achieving and evaluating the success of an intervention provided by a school psychologist. Whereas many Tier III interventions are delivered by consultees, there are situations in which brief counseling provided by the school psychologist may be invaluable as a component of the intervention plan.

A group of school psychologists once asked me how to overcome resistance to implementing a problem-solving model of RtI in their buildings. Examples such as the case studies above were shared, along with an explanation that the easiest way to jump start problem-solving implementation in the face of skepticism is to develop powerful demonstration cases, which can provide momentum and credibility as well as elevate expectations. Many of these school psychologists lamented that they never had one case like those mentioned above, but many also began to develop hope for a better practice.

Clinical psychologists, school psychologists, family therapists, and mental health counselors have been encouraged with many such case examples during an American Psychological Association accredited professional development course using the course workbook provided (Froiland, 2006). Psychologists and counselors of diverse experience levels were happy to gain a fresh perspective and to realize that they were capable of doing even finer mental health work than they had been doing.

The current article was developed in the hope that school psychologists will increase their expectation that similar case successes can happen (or can happen more often) in their practices or in the practices of those whom they train. In addition to having the necessary psychological techniques and knowledge, it is vital that school psychologists believe (and convey an expectation) that they can help children and consultees better mobilize the curative resources available to them (Wampold, 2007). However, one also needs a service delivery framework for psychological interventions that are informed by ongoing assessment and, when indicated, summative assessment. RtI, especially with the emphasis on intense individualized interventions at Tier III, affords such potential.

Children and youth deserve an opportunity to thrive in the least restrictive environment. Because behavioral consultation (Noell, Duhon, Gatti, & Connell, 2002) and a variety of psychological interventions are effective (Wampold, 2007), school psychologists have the potential to help far more children thrive socially and emotionally. Special education placement, as well as the diagnostic labels required for placement, may lead to teasing and exacerbate self-esteem issues. Further, one does not know whether special education is actually the least restrictive environment unless effective consultation and interventions have first been implemented. Brief psychosocial interventions are likely to help children thrive in their general environment and are often worth implementing before a referral for a full psychological evaluation is made.

School psychologists can utilize RtI to further consult with schools and families to develop potent interventions that foster students' mental health, social success, and resilience. The case studies presented here show how a school psychologist might approach cases in the context of a school-based practice in schools that understand that RtI can serve as a vehicle for powerful mental health interventions.

- - -

John Mark Froiland, PhD., is an assistant professor in the Department of School Psychology at the University of Northern Colorado. Previously, Dr. Froiland practiced as a school psychologist for five years and then served as an Institute of Education Sciences Postdoctoral Fellow for two years in the Department of Child Development and Family Studies at Purdue University. His research interests include community supports for parent involvement and childhood development, as well as interventions to support children's autonomous motivation. Dr. Froiland has provided several full day problem-solving seminars for psychologists and is an Advisor for SmartNow.com in San Francisco. Dr. Froiland is a member of the National Association of School Psychologists and the American Psychological Association (APA), including membership in the Developmental Psychology and School Psychology Divisions of APA.

REFERENCES

Bardos, A., McDougal, J.L., & Meier, S. (2010). *Empirically based progress monitoring for behavioral and socio-emotional interventions*. Paper presented at the Annual Convention of the National Association of School Psychologists, Chicago.

Barnett, D.W., Daly, E.J., III, Jones, K.M., & Lentz, F.E., Jr. (2004). Response to intervention: Empirically based special service decisions from single-case designs of increasing and decreasing intensity. *The Journal of Special Education, 38*, 66-79.

Barnett, D.W., Elliot, N., Graden, J., Ihlo, T., Nantais, M., & Prasse, D. (2006). Technical adequacy for response to intervention practices. *Assessment for Effective Intervention, 32*, 20-31.

Barnett, D.W., VanDerHeyden, A.M., & Witt, J.C. (2007). Achieving science-based practice through response to intervention: What it might look like in preschools. *Journal of Educational and Psychological Consultation, 17*, 31-54.

Bolt, S.E. (2005). Reflections on practice within the Heartland problem-solving model: The perceived value of direct assessment of student needs. *The California School Psychologist, 10*, 65-79.

Briesch, A.M., & Chafouleas, S.M. (2009). Review and analysis of literature on self-management interventions to promote appropriate classroom behaviors (1988-2008). *School Psychology Quarterly, 24*, 106-118.

Cheney, D., Flower, A., & Templeton, T. (2008). Applying response to intervention metrics in the social domain for students at risk of developing emotional or behavioral disorders. *The Journal of Special Education, 42*, 108-126.

Froiland, J.M. (2006). *Problem-Solving for Success: The latest applications in learning, attention, and behavior disorders* [Course workbook]. Nashville, TN: Cross Country Education.

Gresham, F.M. (2004). Current status and future directions of school-based behavioral interventions. *School Psychology Review, 33*, 326-343.

Gresham, F.M. (2005). Response to intervention: An alternative means of identifying students as emotionally disturbed. *Education and Treatment of Children, 28*, 328-344.

Hawken, L.S., Vincent, C.G., & Schumann, J. (2008). Response to intervention for social behavior: Challenges and opportunities. *Journal of Emotional and Behavioral Disorders, 16*, 213-225.

Lieberman, R.G., Yoder, P.J., Reichow, B., & Wolery, M. (2010). Visual analysis of multiple baseline across participants graphs when change is delayed. *School Psychology Quarterly, 25*, 28-44.

Marston, D. (2005). Tiers of intervention in responsiveness to intervention: Prevention outcomes and learning disabilities identification patterns. *Journal of Learning Disabilities, 38*, 539-544.

Noell, G.H., Duhon, G.J., Gatti, S.L., & Connell, J.E. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education. *School Psychology Review, 31*, 217-234.

U.S. Department of Education. (2004). Individuals with Disabilities Improvement Act of 2004, Pub. L. 108-466. *Federal Register, 70*, 35802-35803.

U.S. Surgeon General. (1999). *Mental health: A report of the surgeon general*. Retrieved June 23, 2010, from http://www.surgeongeneral.gov/library/mentalhealth/home.html

VanDerHeyden, A.M., Witt, J.C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a Response to Intervention (RTI) model on identification of children for special education. *Journal of School Psychology, 45*, 225-256.

Walter, J.L., & Peller, J.E. (1992). Becoming solution-focused in brief therapy. Levittown, PA: Brunner/Mazel.

Wampold, B.E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist, 62*, 857-873.

Response to Intervention (RtI) in the Social, Emotional, and Behavioral Domains: Current Challenges and Emerging Possibilities

Elina Saeki, Shane R. Jimerson, James Earhart, Shelley R. Hart,
Tyler Renshaw, Renee D. Singh, & Kaitlyn Stewart,
University of California, Santa Barbara

As many schools move toward a three-tier model that incorporates a Response to Intervention (RtI) service delivery model in the social, emotional, and behavioral domains, school psychologists may provide leadership. The decision-making process for filtering students through multiple tiers of support and intervention and examining change is an area where school psychologists are encouraged to apply their expertise regarding assessment and evidence-based interventions. This paper describes an implementation of behavioral and social-emotional RtI in an elementary school setting. Issues and challenges related to measurement of change (i.e., responsiveness to intervention) and identification of students for additional supports as well as emerging possibilities of incorporating qualitative information in the process are discussed.

Education professionals continue to struggle to address the needs of an increasing number of students who have social, emotional, or behavioral difficulties (U.S. Department of Education, 2008). Social-emotional and behavioral problems among students are associated with a variety of poor school-related outcomes (Greenberg et al., 2003; Zins et al., 2004). Too frequently, schools address student behavior problems by employing consequences such as detention, suspension, and expulsion (Skiba & Rausch, 2006). A disproportionate amount of time and energy is spent on serving a small number of students with social-emotional and behavior problems. One potential remedy for this issue is early identification of and intervention with students at risk for such problems. This approach, known as Response to Intervention (RtI), has been recommended as an alternative to the "wait-to-fail" approach that some schools have historically used (Gresham, 2005). Until recently, RtI has been primarily utilized in the academic domain to identify students with specific learning disabilities (Jimerson, Burns, & VanDerHeyden, 2007). However, RtI may also serve as an effective approach for preventing and remedying the social, emotional, and behavioral problems of students who respond to behavioral interventions and therefore do not need more intensive services in special education. On the other hand, students who continue to display behavior problems despite early interventions should continue to receive increasingly targeted services.

The intent of this discussion is to explore RtI as a service delivery model for social, emotional, and behavioral problems in schools. Specifically, the aim of this work is to describe the decision-making process to effectively and efficiently provide appropriate services to students in need of social, emotional, or behavioral support. To illustrate the social-emotional-behavioral RtI methodology presented herein, a case example from an action research collaborative project will be included. Finally, issues and challenges related to measurement of responsiveness to intervention and emerging possibilities will be discussed.

OVERVIEW OF RESPONSE TO INTERVENTION IN THE SOCIAL, EMOTIONAL AND BEHAVIORAL DOMAIN

Response to Intervention (RtI) has been characterized as "the science and practice of assessment and intervention" (Jimerson, Burns, & VanDerHeyden, 2007). RtI may also be described as the change in behavior as a function of intervention (Gresham, 1991, 2002). RtI is typically comprised of five core

Correspondence may be sent to Elina Saeki, UCSB, GGSE, CCSP, Santa Barbara, CA 93106-9490 or e-mail: esaeki@education.ucsb.edu or jimerson@education.ucsb.edu

components: (a) a continuum of evidence-based services available to all students (Martson, Muyskens, Lau, & Canter, 2003); (b) ongoing monitoring of student progress (Gresham et al., 2005); (c) a systematic decision-making process of determining student progress in the academic or behavioral domain (Vaughn, Linan-Thompson, & Hickman, 2003); (d) implementation of increasingly intensive interventions when students do not demonstrate improvements in response to other interventions (Fairbanks, Sugai, Guardino, & Lathrop, 2007); and (e) evaluation of special education services for students who do not demonstrate improvements despite implementation of increasingly intensive interventions (Fuchs, Mock, Morgan, & Young, 2003).

One of the fundamental principles of RtI, underlying all of the core components noted above, is the importance of matching the severity of student problems with appropriate intervention intensities (Gresham, 2004). Toward this end, the United States Public Health Service delineates three levels of prevention outcomes: primary prevention, secondary prevention, and tertiary prevention. Primary prevention seeks to prevent harm and secondary prevention seeks to reverse harm for those students at-risk for school problems (Gresham, 2004). Tertiary prevention also seeks to reduce harm, but is aimed at students with the most severe difficulties (Sugai, Horner, & Gresham, 2002).

All levels of prevention outcomes in the social, emotional, and behavioral RtI framework call for effective interventions utilizing evidence-based strategies that prevent problems rather than react to problems by employing aversive consequences. Early identification and intervention can prevent the escalation of problems into more debilitating forms of social-emotional and behavioral functioning.

Burns, Deno, and Jimerson (2007) discuss RtI in terms of the basic steps involved in problem-solving to unify efforts to operationalize RtI and to test its efficacy in identifying students for service. Assorted problem-solving models exist in the literature, perhaps the most specific details the steps as: (a) Identify the problem, (b) Define the problem, (c) Explore alternative solutions to the problem, (d) Apply a solution, and (e) Look at the effects of the application (IDEAL; Bransford & Stein, 1984). Figure 1 illustrates the Three-Tier Response to Intervention model and the IDEAL Problem Solving Model.

Figure 1. *Matrix Representing the Three-Tier Response to Intervention model and the IDEAL Problem Solving Model (Bransford & Stein, 1984).*

		IDEAL Problem Solving Model				
		Identify the problem	Define the problem	Explore alternative solutions to the problem	Apply a solution	Look at the effects of the application
Three-Tier Response to Intervention Model	**Tier I** Provided to all students	Implement core curriculum and universal screening to identify problems	Collect data to rule out classwide or curricular problems	Generate potential classwide interventions if necessary	Implement classwide remedial interventions or make instructional modifications	Continue benchmark assessment to determine if the class progresses
	Tier II Addressing needs of 10-15% of students		Collect data that are essential to understanding and clearly defining the basis for the problem	Generate a list of evidence-based strategies to intervene classwide or small group level	Implement explicit instructional strategies to address the problem area for a small-group of children	Outcome assessment to examine progress on at least monthly basis
	Tier III Addressing needs of 5-10% of students		Collect additional data that are essential to understanding and clearly defining the basis for the problem	Generate a list of evidence-based intensive individualized interventions	Implement evidence-based intensive individualized interventions to address the problem area	Frequent (twice weekly) outcome assessment to examine progress

Figure reprinted with permission of authors, original source: Burns, M. K, Deno, S., & Jimerson, S. R. (2007). Toward a unified model of Response to Intervention. In Jimerson, S. R., Burns, M. K, & VanDerHeyden, A. M. (Eds.), *The Handbook of Response to Intervention: The Science and Practice of Assessment and Intervention* (pp. 428-440). New York: Springer.

Screening

A variety of strategies for screening have been proposed to identify students who may qualify for selected interventions. Behavior rating scales can provide global estimates of student behavior across various domains (Riley-Tillman, Kalberer, & Chafouleas, 2005). For example, the Behavioral and Emotional Screening System (BESS; Kamphaus & Reynolds, 2007) assesses a range of behaviors that can be used by parents, teachers, and the student.

Whereas the BESS is a screening *instrument*, the Systematic Screening for Behavior Disorders (SSBD; Cheney, Flower, & Templeton, 2008) is a general screening *process*. It identifies students who may be at risk for developing externalizing and internalizing behavior disorders (U.S. Department of Education, 1995). The SSBD uses a multiple gating procedure that utilizes progressively more precise and specific screening instruments to identify youth who need help. The three stages in the SSBD screening process utilize teacher nominations, ratings, and observations. Within a behavioral RtI framework, these screening procedures may occur after implementation of a school-wide or class-wide universal intervention to identify students who warrant further attention.

Conducting systematic and direct observations is another strategy for identifying students who may benefit from selected interventions. This requires a trained observer to operationalize the target behavior, observe the student in a systematic way, and score the data in a consistent manner (Salvia & Ysseldyke, 2004). For example, the Behavior Observation of Students in Schools (BOSS; Shapiro, 2003) uses momentary time sampling to determine the frequency of a specific behavior, such as how often the student is out of his or her seat.

Intervention and Evaluation

After the identification of students for selected intervention at the secondary prevention level, the next important consideration stems from the purpose of, or the answer sought from, implementation of selected interventions. Within the three-tiered RtI model, the question at the primary prevention level is whether students are responding to a systematic, evidence-based universal intervention. At the tertiary prevention level, the concern is whether students are responding to targeted and individualized interventions. In this way, the questions that the primary and tertiary prevention levels seek to answer are fairly straightforward. The secondary prevention level, between the primary and tertiary levels, follows up on the universal prevention efforts and seeks to answer whether the lack of response to the universal intervention was due to an insufficient intensity of intervention or a poor match to the student's needs (McIntosh, Campbell, Carter, & Dickey, 2009). Answering this question incorrectly may lead to inappropriate special education eligibility decisions (Fuchs, Mock, Morgan, & Young, 2003). Malecki and Demaray (2007) offer further discussion and guidelines relevant to the implementation of assessment for social behaviors using an RtI framework (see table 1).

This leads to the challenge of establishing decision rules for determining the extent of effectiveness or student's responsiveness to the intervention. This is particularly salient when implementing social and behavioral interventions, in which the criteria for improvement are less clear in comparison to academic interventions targeted to improve specific skills based on a preset criteria (e.g., the student is able to read 90 words per minute). In fact, in academic domains, the RtI model has extensive amounts of research, particularly in the area of reading (National Institute of Child Health and Human Development, 2000). National benchmarks provide educators with specific criteria whether students are on target, advanced, or falling behind the normative sample. However, such a metric does not currently exist in the domain of social behavior. Gresham (2005) adds that metrics for determining the effectiveness of RtI in the social and behavioral domain are still in their infancy.

Cheney and colleagues (2008) examined the utility of metrics recommended by Gresham (2005) to evaluate the progress of students who were at risk of developing emotional disturbance (ED) and enrolled in a selected group intervention. They found that percentage of change and effect sizes were the most useful metrics to identify students who responded to the intervention versus those who did not respond. This suggests that a systematic and effective process of evaluating responsiveness to intervention may include continuous, ongoing progress monitoring and calculation of percentage of change and

Table 1. *Summary Outline to Guide Implementation of Assessment of Social Behaviors within an RtI Framework.*

Question to Answer	Assessments to Use	Considerations	Purpose of Assessment
What is the target behavior?	Choose appropriate target social behavior(s)	The target behavior must be clearly and operationally defined.	To choose an appropriate target behavior for prevention/intervention based on school need or goals
What are the current levels of this behavior? Are 80 to 90% of students succeeding in this area at Tier I?	Conduct Tier I Assessment with Review, Interview, Observe, and/or Test	Tier I assessments should be easy to collect on entire school population	To understand the current levels of the target social behavior in the school, to create normative data or benchmark criteria
Implement Tier I universal interventions school-wide and continue data collection. Assess at-risk students for potential Tier II interventions.			
Do some children need more intensive intervention (Tier II)?	Conduct Tier II Assessment with Review, Interview, Observe, and/or Test	Tier II assessments should provide information to aid problem analysis and intervention development.	To determine how to develop interventions for children that are not responding to Tier I interventions.
Are the interventions being implemented effective (at Tier II)?	Conduct Tier II Assessment with primarily Observe & Test	Data collected to monitor Tier II interventions should be able to be gathered repeatedly and reliably.	To determine students' response to intervention at Tier II.
Are the Tier I interventions being implemented effective?	Conduct Tier I screening Assessment with Review, Interview, Observe, and/or Test	Tier I assessments should be easy to collect on entire school population	To monitor the levels of the target social behavior in the school and compare against previously identified benchmark criteria
Continue Tier I universal interventions school-wide with necessary changes and continue data collection. Implement and progress monitor Tier II interventions. Identify and develop interventions for Tier III			
Do a few children need more intensive intervention?	Conduct Tier III Assessment with Review, Interview, Observe, and/or Test	Tier III assessments should provide information to aid problem analysis and intervention development.	To determine children that did not respond to Tier II interventions and are in need of more intensive interventions
Are the interventions being implemented effective (at Tier III)?	Conduct Tier III Assessment with primarily Observe & Test	Data collected to monitor Tier III interventions should be able to be gathered repeatedly and reliably.	To determine students' response to intervention at Tier II.

Table reprinted with permission of authors, original source: Malecki, C. K., & Demaray, M. K. (2007). Social behavior assessment and response to intervention. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *The Handbook of Response to Intervention: The Science and Practice of Assessment and Intervention*. (pp. 161-171). New York: Springer.

effect size based on the data gathered through progress monitoring.

The current action research project took place at a public elementary school within a suburban school district in a midsized city in Southern California. The principal and school psychologist identified concerns regarding increasing numbers of discipline referrals and overall disruptive behavior among third-grade students. A collaborative problem-solving team consisting of the principal, school psychologist, university faculty supervisors, and school psychologist interns met and discussed concerns and possible solutions to this issue.

METHOD

Participants

Participants for this study included 55 third-grade students from three classrooms (2 third-grade classes and 1 third- and fourth-grade combination class). Of the participating students, 49% were female, 71% identified as Hispanic/Latino, 9% Caucasian, 6% Asian, 6% African American, and 9% other.

Procedures

Universal Intervention (Tier I). Considering the presenting challenges, the problem-solving team determined that all students could benefit from lessons on impulse control. Two school psychologist interns implemented lessons from the *Second Step: A Violence Prevention Curriculum* in to all third grade students[1]. Specifically, the Impulse Control lessons were implemented. These interventions were delivered twice per week for a total of five sessions. One intern facilitated the lessons while the other provided classroom support, ensuring that the students were engaged and participating. The classroom teacher was also present during the lessons and helped to manage disruptive behavior when necessary. The students were administered the Knowledge Assessment both before and after the intervention. Each session was approximately 45 minutes long.

Selected Intervention (Tier II). After the universal intervention was implemented, the teachers completed a student nomination form, listing students they believed might benefit from additional support through selected interventions in a small group setting. Two groups were formed: six boys were selected to focus on an additional impulse control lessons and three girls were selected to work on empathy and friendship skills. *The Second Step Curriculum* was utilized for these additional lessons; however, other materials were used to supplement emerging issues within the groups. For example, as the sessions progressed, it became apparent that the students from both the girls and the boys group could benefit from assertiveness training. Role-plays and games were used to demonstrate the importance of being assertive (e.g., telling a peer to stop distracting you; telling a friend that your feelings are hurt). In addition, behavioral modification methods (i.e., star charts in which students earn stars and redeem rewards) were used to encourage meaningful participation within the small groups. The two groups met separately, once per week. Each intern was in charge of one group and facilitated the lessons. Each of the groups met for 15 sessions.

Table 2. *Pre- and Post-Assessment Mean Scores for Each Domain (Tier II)*

	Boys (n = 6)		Girls (n = 3)	
	Pre	Post	Pre	Post
Cooperation	11	14.5	14.6	15.3
Assertion	12.5	14.3	14	14.6
Empathy	12.7	13.1	16.6	15.6
Self-Control	12.7	13.3	11	12.6

Note. Higher raw scores indicate better adjustment in the given domain.

[1]*Second Step: A Violence Prevention Curriculum* (Committee for Children, 1992) is an intervention that promotes social skill development. The curriculum teaches specific skills that are important for healthy social-emotional development, including empathy, impulse control and problem solving, and anger management. Evaluations of *Second Step* suggest that the program yields sustained improvements in students' actual behaviors, knowledge, attitudes, and motivation (Committee for Children, 2002). Due to the presenting concerns regarding the behaviors of third grade students, lessons on impulse control were selected and implemented as the universal intervention.

Targeted Intervention (Tier III). While the students were receiving selected interventions, the interns engaged in periodic check-ins with the teachers. One student in particular was identified as struggling with emotional issues, and it was suggested that he receive intensive, targeted intervention. This provided transition from the delivery of selected interventions to targeted interventions. Given that the presenting concern was primarily emotional (e.g., fluctuating moods) rather than a skill deficit (e.g., impulse control, empathy), the school psychologist implemented strength based, cognitive-behavioral therapy. This student continued to receive targeted intervention for the remainder of the school year (i.e., 2 months).

Measures

The following measures were used to examine changes in student knowledge and behaviors, and to inform decisions regarding the relative need for additional support.

Knowledge Assessment for Second Step (KASS; Committee for Children, 2004). The *KASS* is a self-report measure developed by the authors of the *Second Step* curriculum to assess knowledge in social-emotional skills. The *KASS* consists of several problem situations and related social-emotional skills knowledge questions presented to students that they respond to in writing. It is designed to be utilized in a pre- and post-test format. Administration, scoring, and interpretation are standardized with directions provided in the *KASS* manual. Instructions, problem situations, and questions are provided both orally and in writing for students, who are allowed as much time as needed to complete the assessment. The authors indicate that a pilot test and a field test were conducted in 2002-2003, followed by revisions and further field testing in 2003-2004; however, no data is available (Committee for Children, 2004).

Social Skills Rating System (SSRS; Gresham & Elliot, 1990). The SSRS is a multi-rater assessment used to observe changes in attitudes, beliefs, and behaviors. It may be used to identify students at risk for academic and social difficulties as well as their strengths, to inform follow-up assessments, and to guide the selection of intervention strategies. Parent, teacher, and child reports are available. The child self-report version of the *SSRS* includes four factors: cooperation, assertion, empathy, and self-control. The internal consistency of the SSRS subscales ranged from .51 to .91, with a mean internal consistency of .75 (Gresham & Elliot, 1990).

RESULTS

Universal Intervention

Descriptive statistics on the *KASS* raw scores indicate that the mean scores for the pre- and post-assessment at this level were 6.24 (SD = 3.62) and 8.70 (SD = 4.04), respectively (the raw scores were used because the *KASS* does not provide conversions to standardized scores). Of the 43 students who took both the pre- and post-assessment, 81% of the students improved their score (n = 35), 14% of the students' scores stayed the same (n = 6), and 5% of the students' scores decreased (n = 2). A calculation of effect size based on mean scores and standard deviations of the pre- and post-assessment data reveal a large effect (d = .64). Additionally, the percentage of change from pre- to post-assessment was 39.4%.

Selected Intervention

For the students identified for Tier II interventions, mixed results were demonstrated on the *KASS*. Several students' scores improved (n = 4), one decreased (n = 1), several revealed no change (n = 2), and others were absent during either the pre- or post-assessment administration (n =2). Essentially, over half of the students that the teachers identified for selected intervention improved their scores from pre- to post-assessment. Additionally, given this discrepancy, only the information gathered from the teacher nomination process (i.e., the list of suggested students to receive selected interventions) was included as part of the decision-making process of filtering students from Tier I to Tier II interventions. A discussion will follow regarding the implications of this.

The SSRS was also administered to the children during the first and last session.

Boys' group. The mean raw scores on the *SSRS* pre-assessment suggest higher functioning and adjustment in the domain of empathy and self-control compared to cooperation and assertion (See Table 2).

This was contrary to teacher reports of these students having the most difficulties with impulse control. The teachers suggested that the social skills group focused on self-control, given that this was an area they had observed the students having the most difficulty with in the classroom. With this feedback from the teachers, the decision was made to provide lessons on self-control despite the pre-assessment suggesting that they may be fairly well adjusted in this particular skill set.

The pre- and post-assessment data indicates that of the six students, four students' scores improved overall while two students' scores decreased (See Table 3). An increase in standard scores was observed from pre- to post-assessment (M = 81.9, SD = 9.1; M = 86.7, SD = 9.9, respectively; d = .51). The percentage of change from pre- to post-assessment was 5.91%.

Girls' group. The mean raw scores on the *SSRS* pre-assessment for the girls suggested higher functioning and adjustment in the domain of empathy (See Table 2). Similar to the findings for the boys' pre-assessment, these results were contrary to teacher reports, who communicated that the students could benefit the most from friendship and empathy training. However, their mean score for empathy was the highest among the four domains. In addition, their mean score was lowest for self-control, suggesting this to be an area demonstrating the most need for additional support. However, based on teacher reports, the decision was made to have empathy training as the focal point of the selected small group intervention.

The pre- and post-assessment data indicate that of the three students in the group, two of their scores improved and one student's score decreased (See Table 3). An increase in standard scores was observed from pre- to post-assessment (M = 89.3, SD = 3.5; M = 95, SD = 7, respectively; d = .81). The percentage of change from pre- to post-assessment was 6.35%.

Table 3. *Pre- and Post-Assessment Standard Scores for the Social Skills Rating System (Tier II)*

Boys		Girls	
Pre	Post	Pre	Post
72	94	86	103
69	71	93	90
87	91	89	92
84	80	-	-
91	86	-	-
88	98	-	-

Targeted Intervention

The student identified for targeted intervention received ongoing, intensive individual counseling. The school psychologist met with him weekly and used clinical judgment to detect the student's progress in counseling. In the present context, no pre- or post-test measures were utilized at this stage given the school psychologist's use of clinical judgment in lieu of quantitative data collection. However, measures addressing specific target areas and sensitivity to change are optimal in this context.

DISCUSSION

A three-tiered, RtI framework for social, emotional, and behavioral issues affords an opportunity to provide additional, meaningful supports for students who are at-risk but may not qualify for special education services. The current action research collaborative project illustrates how a social, emotional, behavioral-oriented RtI system may be implemented in an elementary school setting.

One of the major challenges faced was in the evaluation of students' responsiveness to the interventions and the identification of students requiring additional supports (see Malecki & Demaray, 2007 for further discussion). In the identification of students to receive both Tier II and Tier III interventions, pre- and post-assessment data and teacher nominations were used. The challenge rested on the cases in which the pre- and post-data did not match teacher reports. For example, of the nine students who received selected interventions in Tier II, four of the students' scores from the pre- and post-assessment in Tier I improved, while several of the students' scores either remained the same or went down (several were also absent either during the pre- or post-assessment date).

There may be a variety of reasons for each student's performance that is not directly related to their responsiveness to intervention. For instance, students may not have scored higher for the post-assessment because their pre-assessment scores were already high to begin with. Other students may have had difficulties comprehending the items and thus scored low on both the pre- and post-assessment. In addition, knowledge, or a score indicating knowledge, may not necessarily translate to the desired behavior. Students may know how to respond to questions about controlling their impulses, but may not be able to act out the skills that they articulated in their responses. These are considerations to be made when interpreting student's pre- and post-assessment performance.

Regardless of students' pre- and post-assessment performance, had the teachers nominated students to receive selected interventions (Tier II) *before* the implementation of the universal intervention (Tier I), it is possible that they would have nominated the same students that they actually identified *after* the students received the universal intervention. That is, the teachers would have chosen the students they observe to be in need of additional support, regardless of the pre- and post-data.

For RtI in the academic domain, the use of quantitative measures (e.g., words read correctly per minute as a proxy for reading fluency) may be sufficient to determine student's current skills and performance. However, for RtI in the behavioral domain, the use of pre- and post-assessment data by itself has been elusive in fully detailing a student's social, emotional, and behavioral functioning. Metrics to evaluate the progress of at-risk students have been developed, with research supporting its usefulness to gauge responsiveness to intervention (Cheney et al., 2008; Gresham, 2005). However, as Gresham (2005) states, metrics for determining the effectiveness of RtI in the social and behavioral domain are still in their infancy.

In a time of increasing accountability, school psychologists likely feel pressure to provide evidence to demonstrate the effectiveness of various interventions that they deliver. Pre- and post-assessments are administered to detect changes in students' attitudes, knowledge, and behavior and can be used as a tool to demonstrate the effectiveness of an intervention. In addition, it can be used as a proxy of students' responsiveness to intervention, such that students who made minimal improvements may be identified for increasingly intensive and targeted interventions.

In addition to these quantitative assessments (e.g., self-, teacher-, and parent-report measures), qualitative information may provide value in the decision-making process providing additional information relevant to which students filter through the RtI service delivery model. Systematic observations of students in a variety of settings (e.g., classroom, playground) and teacher interviews can provide information that may not be apparent when examining only self-report questionnaire data. In fact, in this instance, teacher input was the primary influence in the identification of students to receive selected (Tier II) and targeted interventions (Tier III). This points to the importance of a multifaceted evaluation process by which information is gathered from a variety of sources.

However, without a systematic procedure for incorporating qualitative information, school psychologists run the risk of compromising objectivity in the decision-making process. The collection and interpretation of both quantitative and qualitative data must be objective and psychometrically sound. For instance, systematic direct observation systems such as the Student Observation System (SOS; Reynolds & Kamphaus, 2005) and the Behavioral Observation of Students in Schools (BOSS; Shapiro, 2003) could be used to gather a random sampling of student's behaviors. In addition, teacher interviews should be coded systematically. Unfortunately, collecting systematic direct observation data for even a handful of students suspected to be at-risk for behavioral and social-emotional difficulties and coding information gathered through interviews is beyond the scope of resources for most school-based practitioners.

Limitations and Future Directions

Manipulation of methodologies employed in this action research project was limited by school-level factors (e.g., requests by principal and school psychologist to focus interventions on third grade). However, this real world context offers a discussion of challenges and future directions in the implementation of social, emotional, and behavioral RtI.

The inclusion of a school-wide systematic screening to identify areas of risk was discouraged by the

principal, given her preference for utilizing qualitative and professional judgment rather than quantitative information. The limited time and resources available may create resistance from the administration in implementing systematic screening to the entire school population. School psychologists should emphasize the importance of objectivity in assessment, and qualitative information should be considered in conjunction with quantitative data.

CONCLUSION

The roles and responsibilities of some school psychologists may change considerably as schools move toward an RtI service delivery model, not only in the academic domain but in the behavioral and social-emotional domains as well. In some instances, the role may shift from an emphasis on assessment and evaluation for special education services to early identification, intervention, and progress monitoring. Given their unique training and knowledge, many school psychologists are in a position to provide leadership in the development and implementation of processes pertaining to RtI. Applying principles of data-based decision making, school psychologists should corroborate information from various sources to meet the needs of students within an RtI system.

Elina Saeki, M.Ed., is a doctoral student at the University of California, Santa Barbara.

Shane R. Jimerson, PhD., is a Professor at the University of California, Santa Barbara.

James Earhart, M.Ed., is a doctoral student at the University of California, Santa Barbara.

Shelley R. Hart, M.S., is a doctoral student at the University of California, Santa Barbara.

Tyler L. Renshaw, M.Ed., is a doctoral student at the University of California, Santa Barbara.

Renee D. Singh, M.Ed., is a doctoral student at the University of California, Santa Barbara.

Kaitlyn Stewart, M.Ed., is a graduate student at the University of California, Santa Barbara.

REFERENCES

Achenbach, T.M. (1991). *Manual for the child behavior checklist and 1991 profile*. Burlington, VT: University Associates in Psychiatry.

Bransford, J., & Stein, B. (1984). The ideal problem solver: *A guide for improving thinking, learning and creativity*. San Francisco: W.H. Freeman.

Burns, M.K, Deno, S., & Jimerson, S.R. (2007). Toward a unified model of Response to Intervention. In Jimerson, S.R., Burns, M.K, & VanDerHeyden, A.M. (Eds.), *The Handbook of Response to Intervention: The Science and Practice of Assessment and Intervention* (pp. 428-440). New York: Springer.

Cheney, D., Flower, A., & Templeton, T. (2008). Applying response to intervention metrics in the social domain for students at risk for developing emotional or behavioral disorders. *The Journal of Special Education, 42*, 108-126.

Committee for Children (1992). *Second Step: A violence prevention curriculum, grades 1-3* (2nd ed.). Seattle, WA: Author.

Committee for Children (2002). Second Step Grades 1-3: Teacher's guide. Retrieved from http://www.cfchildren.org/programs/ssp/research/

Conners, C.K. (1997). *Conners' Rating Scales-revised: Technical manual*. North Tonawanda, NY: Multi-Health Systems.

Daly, E.J., Martens, B.K., Barnett, D., Witt, J. C., & Olson, S.C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement, instruction, and intensity. *School Psychology Review, 36*, 562-581.

Fabiano, G.A., Vujnovic, R., Naylor, J., Pariseau, M., & Robins, M. (2009). An investigation of the technical adequacy of a daily behavior report card (DBRC) for monitoring progress of students with attention-deficit/hyperactivity disorder in special education placements. *Assessment for Effective Intervention, 34*, 231-241.

Fairbanks, S., Sugai, G., Guardino, D., & Lathrop, M. (2007). Response to intervention: Examining classroom behavior support in second grade. *Exceptional Children, 73*, 288-310.

Fuchs, D., Mock, D., Morgan, P.L., & Young, C.L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18*, 157-171.

Furlong, M.J., Morrison, G.M., Chung, A., Bates, M., & Morrison, R. (1997). School violence: A multicomponent reduction strategy. In G. Bear, K. Minke, & A. Thomas (Eds.), *Children's needs II: Development, problems, and alternatives* (pp. 245-256). Bethesda, MD: National Association of School Psychologists.

Greenberg, M.T., Weissberg, R.P., O'Brien, M.U., Zins, J.E., Fredericks, L., Resnik, H., & Elias, M.J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*, 466-474.

Gresham, F.M. (1991). Conceptualizing behavior disorders in terms of resistance to intervention. *School Psychology Review, 20*, 23-26.

Gresham, F.M. (2002). Responsiveness to intervention: An alternative approach to the identification of learning disabilities. In R. Bradley, L. Danielson, & D. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 467-519). Mahwah, NJ: Lawrence Erlbaum.

Gresham, F.M. (2004). Current status and future directions of school-based behavioral interventions. *School Psychology Review, 33*, 326-343.

Gresham, F.M. (2005). Response to intervention: An alternative means of identifying students as emotionally disturbed. *Education and Treatment of Children, 28,* 328-344.

Gresham, F.M. (2007). Response to intervention and emotional and behavioral disorders: Best practices in assessment for intervention. *Assessment for Effective Intervention, 32*, 214-222.

Gresham, F.M., & Elliot, S.N. (1990). *Social skills rating system manual*. Circle Pines, MN: American Guidance Service.

Hawken, L.S., Vincent, C.G, & Schumann, J. (2008). Response to intervention for social behavior: Challenges and opportunities. *Journal of Emotional and Behavioral Disorders, 16*, 213-225.

Hintze, J.M., & Matthews, W.J. (2004). The generalizability of systematic direct observations across time and setting: A preliminary investigation of the psychometrics of behavioral observation. *School Psychology Review, 33*, 258-270.

Hsieh, H., & Shannon, S. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*, 1277-1288.

Jimerson, S.R., Burns, M.K, & VanDerHeyden, A.M. (2007). Response to intervention at school: The science and practice of assessment and intervention. In Jimerson, S.R., Burns, M.K, & VanDerHeyden, A.M. (Eds.), *The Handbook of Response to Intervention: The Science and Practice of Assessment and Intervention.* (pp. 3-9). New York: Springer.

Malecki, C.K., & Demaray, M.K. (2007). Social behavior assessment and response to intervention. In S.R. Jimerson, M.K. Burns, & A.M. VanDerHeyden (Eds.), *The Handbook of Response to Intervention: The Science and Practice of Assessment and Intervention.* (pp. 161-171). New York: Springer.

Martson, D., Muyskens, P.L., Lau, M., & Canter, A. (2003). Problem-solving model for decision making with high-incidence disabilities: The Minneapolis experience. *Learning Disabilities Research and Practice, 18*, 187-200.

McIntosh, K., Campbell, A.L., Carter, D.R., Dickey, C.R. (2009). Differential effects of a tier two behavior intervention based on function of problem behavior. *Journal of Positive Behavior Interventions, 11*, 82-93.

National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction* (NIH Publication No. 00-4769). Washington, DC: US Government Printing Office.

Reynolds, C.R., & Kamphaus, R.W. (2005). *Behavior Assessment System for Children* (2nd ed.). Circle Pines, MN: American Guidance Service.

Riley-Tillman, T.C., Chafouleas, S.M. (2009). Foundation for the development and use of direct behavior rating (DBR) to assess and evaluate student behavior. *Assessment for Effective Intervention, 34*, 201-213.

Riley-Tillman, T.C., Kalberer, S.M., & Chafouleas, S.M. (2005). Selecting the right tool for the job: A review of behavior monitoring tools used to assess student response-to-intervention. *The California School Psychologist, 10*, 81-91.

Rose, L.C., & Gallup, A M. (2005). The 37th annual Phi Delta Kappa/Gallup Poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 87*, 41-57.

Salvia, J., & Ysseldyke, J.E. (2004). *Assessment* (9th ed.). Princeton, NJ: Houghton Mifflin.

Shapiro, E.S. (2003). *Behavior observation of students in schools*. San Antonio, TX: Harcourt Assessment.

Skiba, R.J., & Rausch, M.K. (2006). School disciplinary systems: Alternatives to suspension and expulsion. In G. Bear & K. Minke (Eds.), *Children's Needs III: Understanding and Addressing the Developmental Needs of Children*. Washington, D. C.: National Association of School Psychologists.

Sugai, G., Horner, R.H., & Gresham, F.M. (2002). Behaviorally effective school environments. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II* (pp. 315-350). Bethesda, MD: National Association of School Psychologists.

Sugai, G., Lewis-Palmer, T., & Hagan, S. (1998). Using functional assessments to develop behavior support plans. *Preventing School Failure, 43*, 6-13.

United States Department of Education (1995). Education programs that work. Retrieved September 5, 2008, from www.ed.gov/pubs/EPTW/eptw12/eptw12h.html.

U.S. Department of Education, National Center for Education Statistics (2009). *Digest of Education Statistics, 2008.* Washington, DC.

Vaughn, S., Linan-Thompson, S., & Hickman, P. (2003). Response to instruction as a means of identifying students with reading/learning disabilities. *Exceptional Children, 69*, 391-409.

Zins, J.E., Weissberg, R.P., Wang, M.C., & Walberg. H. J. (Eds.). (2004). *Building academic success on social and emotional learning: What does the research say?* New York: Teachers College Press.

Selective Mutism: A Three-Tiered Approach to Prevention and Intervention

R.T. Busse and Jenna Downey,
Chapman University

Selective mutism is a rare anxiety disorder that prevents a child from speaking at school or other community settings, and can be detrimental to a child's social development. School psychologists can play an important role in the prevention and treatment of selective mutism. As an advocate for students, school psychologists can work with teachers, parent caregivers, speech pathologists, and other support staff toward helping children who may develop or have selective mutism. The purpose of this article is to present school-based prevention and intervention approaches within a three-tiered approach that may reduce the incidence and severity of selective mutism. We present theories and research on the etiology and prevalence of the disorder, followed by a review of intervention methods and research at each tier. Based on the theoretical and research literature base, we conclude that early intervention may result in the prevention and amelioration of many occurrences of selective mutism.

KEYWORDS: Selective Mutism, Childhood Anxiety Disorders, Social Phobia, Prevention, Treatment

The purpose of this article is to present school-based prevention and intervention approaches within a three-tiered approach that may reduce the prevalence and severity of selective mutism. Children with selective mutism (SM) experience a "consistent failure to speak in specific social situations (in which there is an expectation for speaking, e.g., at school) despite speaking in other situations" (American Psychiatric Association [APA], 2000, p. 78). To be diagnosed with SM, a child's lack of speech: a) must last for at least one month, excluding the first month of school; b) must interfere with educational or occupational achievement or with social communication; c) cannot be due to any lack of knowledge or discomfort with the spoken language; and d) cannot solely be due to a communication disorder, pervasive developmental disorder, schizophrenia, or any other psychotic disorder (APA, 2000).

Selective mutism is widely characterized as a disorder primarily linked with social anxiety (Bergman, Piacentini, & McCracken, 2002; Chavira, Shipon-Blum, Hitchcock, Cohan, & Stein, 2007; Ford, Sladeczek, Carlson, & Kratochwill, 1998; Kratochwill, 1981; Stone, Kratochwill, Sladeczek, & Serlin, 2002). SM often can be confused with other speech issues, such as the silent period some children experience when learning a second language, the absence of speech due to aphasia or deafness, or the absence of speech sometimes associated with autism (Cline & Baldwin, 2004). The primary characteristic that differentiates this disorder from related conditions is that children who experience SM usually speak freely in other environments, and their failure to speak usually occurs at school (Leonard & Dow, 1995).

Not speaking in school may hinder a child's academic performance and social development in particular, although more research needs to be conducted on the short and long term negative consequences of SM. Not surprisingly, the short-term effects have been found to include heightened anxiety and social skills deficits (e.g., Bergman, et al., 2002; Cunningham, McColm, & Boyle, 2006; Ford, et al., 1998). The long-term effects of SM have been infrequently studied, with two well-controlled studies indicat-

Correspondence concerning this article should be addressed to R.T. Busse, Counseling and School Psychology Program, College of Educational Studies, Chapman University, One University Drive, Orange, CA 92866. Email: busse@chapman.edu

ing that the majority of cases remitted without intervention, however young adults with former selective mutism described themselves as less independent and having more social problems than controls (Remschmidt, Poller, Herpetz-Dahlmann, Hennighausen, & Gutenbrunner, 2001; Steinhausen, Wachter, Laimbock, & Metzke, 2006). Furthermore, many cases of SM persist if not treated (Crundwell, 2006; Ford, et al. 1998; Stone, et al., 2002), which indicates the need for intervention. Interventions with SM are especially important in elementary schools, because the majority of cases are first identified in preschool or kindergarten (Leonard & Dow, 1995; Stone, et al, 2002). School psychologists can play an important role in implementing prevention at the universal level, and providing more focused interventions for children who may be at-risk for SM.

In this article, we present interventions for SM, following a three-tiered approach (see Table 1). The first tier, or primary prevention, focuses on prevention methods that may be implemented school-wide to reduce the development of SM. The second tier, or early onset interventions, involves interventions that can be implemented with groups or in the classroom for children showing signs of SM. The third tier focuses on individual treatment methods implemented both in and outside the classroom for children who have developed the disorder.

Table 1 *Three Tiered Interventions for Selective Mutism*

Tiers	Examples of Intervention Methods
Tier I	• Parent/Caregiver newsletters and trainings on the identification and prevention of potential anxiety problems • School-wide oral communication strategies: Maintaining expectancies for speaking, providing opportunities to respond, wait-time for responses, minimizing reinforcement of nonverbal communication • Preparation of preschoolers and families for the transition to kindergarten
Tier II	• Early identification of children who are at-risk for or have selective mutism • Child-focused oral communication strategies: Maintaining expectancies for speaking, providing opportunities to respond, wait-time for responses, minimize reinforcement of nonverbal communication • Contingency management • Shaping • Group therapy
Tier III	• Family and play therapy • Contingency management • Shaping • Social skills training • Stimulus fading • Systematic desensitization/relaxation training • Self-modeling • Psychopharmacological therapy

CONCEPTUALIZATON AND ETIOLOGY OF SELECTIVE MUTISM

A disorder like SM was first described in the late 1800s by Adolf Kussmaul, who called the disorder *aphasia voluntaria*, which stemmed from the interpretation that the disorder involved a voluntary decision not to speak (Cohan, Chavira, & Stein, 2006; Krysanski, 2003; Standart & Le Couteur, 2003). In the early 1930s, the disorder was referred to as *elective mutism* which came to be called *selective mutism* in the 1970s and 80s (Krysanski, 2003), the term that is used in the most current version of the DSM (APA, 2000). The change in terminology reflects an emphasis on a child's "consistent failure" to speak in select environments. This emphasis represents an adjustment in the criteria for SM from former definitions which described the disorder as a "refusal to speak." The word "refusal" was changed because it indicated that children with the disorder simply were being oppositional or defiant in choosing not to speak (Cline & Baldwin, 2004).

Early theories on the causes of SM often focused on the family and experiences with trauma, such as a hostile home environment, physical or sexual abuse, or tragic events such as the death of a loved one (Leonard & Dow, 1995). Although trauma may still be believed to be the cause for some cases of SM, there is limited evidence to support this theory. Psychodynamic theorists often conceptualize SM as a child's reaction to an unresolved conflict with parents or caregivers to gain control over some aspect of the child's life (Krysanski, 2003). Similarly, family systems theorists often view SM as a product of conflicting familial relationships (Anstendig, 1998). Behaviorists typically view SM as a result of negatively reinforced learning patterns that teach the child to use silence as a method of reducing or controlling their anxiety in reaction to specific stimuli (Krysanski, 2003).

Selective mutism is most commonly found to be co-morbid with social anxiety. The majority of the research and literature base over the past 30 years supports this relationship (Bergman, et al., 2002; Chavira et al., 2007; Ford, et al., 1998; Kratochwill, 1981; Leonard & Dow, 1995; Krysanski, 2003; Manassis, et al., 2007; Morris & Kratochwill, 1985; Standart & Couteur, 2003; Steinhausen, et al., 2006; Yeganeh, Beidel, & Turner, 2006). Researchers have found that the majority of children diagnosed with SM also matched the criteria for social phobia (Black & Udhe, 1995; Dummit, et al., 1997, as cited in Chavira, et al. 2007; Yeganeh, et al., 2006), a specific type of social anxiety that includes "a marked and persistent fear of one or more social or performance situations in which the person is exposed to unfamiliar people or to possible scrutiny by others" (APA, 2000, p.456).

Bergman et al. (2002) surveyed 125 teachers who reported on 2256 kindergarten through second grade students and found that children diagnosed with SM were rated higher on levels of internalizing, withdrawn, and anxious/depressed characteristics than comparison children. Steinhausen et al. (2006) focused on personality traits in a longitudinal study on 33 children with SM and matched controls. After 13 years of study, they concluded that "…SM and child anxiety disorders share similarities in their temperamental, environmental and biological etiologies, and that SM also co-occurs with various specific anxiety disorders such as social phobia, separation anxiety, and posttraumatic stress disorder" (p. 754). Temperamental characteristics that correlate with anxiety, and those that are prevalent in children with SM, include shyness and behavioral inhibition, or a slow-to-warm temperament (Cline & Baldwin, 2004; Ford, et al., 1998). Although there is a clear link between SM and anxiety, more research needs to be conducted to determine the factors involved.

Familial factors also may play a role in SM, although the research is mixed. For example, Chavira et al. (2007) found generalized social phobia occurred in 37% of a sample of 140 parents/caregivers with children diagnosed with SM, versus 14.1% of 62 control group parents/caregivers. Similarly, Schwartz, Freedy, and Sheridan (2006) surveyed the parents/caregivers of 33 children with SM and found that "33% reported a relative with social anxiety disorder and 12.1% reported a relative with SM" (p. 46). These findings indicated that a genetic component, or indirect familial factor, may influence the development of SM, although the results also indicated that most of the parents/caregivers under study did not evidence anxiety or social withdrawal. Cunningham, McHolm, Boyle, and Patel (2004) conducted a study that compared 52 children with SM and their families to a control group and found no differences between the groups on measures of family functioning.

Some researchers have posited that a biological factor may be related to shyness (Kagan, 1997; Marshall & Stevenson-Hinde, 2001), which occurs in many, but not all, children with SM. In a review of 100 cases of SM, 85% of children with SM were rated as shy (Steinhausen & Juzi, 1996), although this finding probably is an overestimate due to the 'shy-like' behaviors associated with SM. A biological correlate for shyness may be indicative of a related biological factor for SM but the data are not strong and are in need of further empirical research, such as twin and adoption studies to examine genetic and biological influences.

Incidence, Prevalence, and Demographics

There is a limited amount of research on the prevalence (the proportion of the population with a disorder) of SM, and virtually nothing is known about the incidence (rate of new occurrences) of the disorder. SM typically is estimated to occur in less than one percent of the population (APA, 2000), although the DSM refers only to "individuals seen in mental health settings" (p.126). Some prevalence data have indicated that approximately seven per 1000 children are affected in the United States (Bergman, et al., 2002) and Israel (Elizur & Perednik, 2003). Due to relatively low prevalence, it is difficult to conduct large controlled studies to ascertain accurate estimates (Krysanski, 2003; Leonard & Dow, 1995; Standart & Couteur, 2003). Establishing accurate incidence and prevalence rates is hampered by the different levels of severity and common misdiagnoses of children with SM. The apparent prevalence rate also could increase in areas with higher immigrant populations (Cline & Baldwin, 2004). Thus, the prevalence estimates may change with more awareness of the disorder and more accurate diagnosis.

Research on the demographics of SM has mainly focused on sex and age. The majority of the data indicate that SM is more prevalent in females than in males, with the ratio ranging from 1.6–3:1 (Kolvin & Fundudis, 1981; Krysanski, 2003; Leonard and Dow, 1995; Standart & Couteur, 2003). The higher prevalence of SM in females builds another potential link between SM and anxiety disorders, which are also more often diagnosed in females (Leonard & Dow, 1995). The variation in the data on sex ratios likely is due to the limited amount of sample sizes and inability to control for the selection of participants with SM. The onset of SM appears to range from three to six years of age, with a majority of referrals occurring during the first years of school (Cohan, et al., 2006; Leonard & Dow, 1995).

Children from immigrant backgrounds have been found to be more likely than non-immigrant children to be diagnosed with SM (Elizur & Perednik, 2003; Steinhausen & Juzi, 1996; Toppelberg, Tabors, Coggins, Lum, & Burger, 2005). These findings likely are due to misdiagnoses. Studies on the prevalence of SM within diverse samples can be difficult to control, due to the overlap of characteristics between SM and what is known as 'the silent period' for English language learners. Children from immigrant backgrounds may be more prone to be misdiagnosed with SM if they are experiencing an initial nonverbal stage before becoming comfortable speaking the language of their adopted country (Toppelberg, et al., 2005). More research is warranted on children from immigrant backgrounds, with a focus on differentiating children who have SM and those who are experiencing a silent period. Related variables that warrant further investigation are cultural influences that may contribute to the diagnosis, or lack thereof, of SM. Virtually nothing is known about the incidence or prevalence of SM across different races/ethnicities or the cultural variables that may influence the rate of occurrence or diagnosis. Cross-cultural research has shown that SM occurs in different countries (e.g., Elizur & Perednik, 2003; Remschmidt, et al., 2001). Overall, the literature has shown that SM is a rare disorder, with unclear statistics on its incidence, prevalence, and demographics.

PRIMARY PREVENTION OF SELECTIVE MUTISM (TIER I)

We are not aware of any research that has focused on prevention efforts for SM, perhaps because there is no consensus regarding the causes of SM. We believe that the most plausible explanation of the etiology of SM is within a behavioral model. If one considers SM as a learned behavior akin to social phobia, then prevention methods may be directed at minimizing antecedent and consequent events that may lead to SM. Thus, although research is needed on the prevention of SM, we offer potential methods based on behavioral theory, and on Tier II and III interventions.

Most cases of SM are not identified until a child begins attending school, where the child's teacher is usually the first to bring concerns to parents/caregivers and other school personnel (Crundwell, 2006; Leonard & Dow, 1995; Schwartz, et al., 2006; Standart & Couteur, 2003). Primary prevention methods may reduce the frequency of severe cases and may save the school and parents/caregivers cost. Primary prevention can be focused on reducing the number and severity of cases by expanding awareness of SM, training teachers on communication strategies that may be used in the classroom, and minimizing the anxiety associated with entry into the school environment.

Because SM is rare, most school personnel and parent caregivers likely are unaware of the condition. Awareness about internalizing behaviors like SM and anxiety is important because externalizing behaviors, such as attention deficit hyperactivity disorder, often may overshadow the more 'quiet' internalizing disorders. Informing teachers and caregivers about SM and other forms of anxiety in children who are entering school may increase the chances of addressing SM early. Early diagnosis and intervention is important because many cases of SM worsen with time, the disorder often interferes with a child's academic and social development (Crundwell, 2006), and, if not treated, SM may become an accepted part of the child's identity (Omdal, 2008). Caregivers and teachers of children entering school can be informed with a letter addressing the early signs of SM and other anxiety related problems, or can be invited to a training program about early warning signs (Cline & Baldwin, 2004).

Research is lacking on the prevention of anxiety problems in school. Given the relative rarity of SM (and other anxiety related disorders), it may be best to address all types of anxiety in a prevention model. Dadds and Roth (2008) conducted a study to examine the effectiveness of a parent/caregiver-training program toward preventing anxiety associated with school. The study included families from 12 intervention and 13 comparison preschools. In the experimental condition the caregivers were trained on building social competence in their children, and improvement was monitored in first grade through caregiver and teacher reports. The intervention consisted of six sessions across a 12 week period. The sessions were organized around responding to stress, behavior management, cognitive-behavioral intervention for challenging self-talk and attributional styles, and using a problem-solving model to cope with anxiety. The results were mixed, but provided some indication that parent training may be effective toward alleviating some school-based anxiety. Although the results were weak and the study was not directed at SM per se, it does provide a potential rough framework for the prevention of anxiety.

In addition to caregiver training, teacher awareness of SM may be beneficial toward prevention of the disorder, or to ameliorating its effects. School psychologists and other support staff can meet with teachers to discuss the characteristics of SM or supply information about the disorder and how to differentiate it from other issues, such as the silent period of English language learners or another disorder, such as autism.

Another potential way to prevent or lessen the occurrence of SM is to train all teachers on oral communication strategies in the classroom. From a prevention perspective, maintaining expectancies for speaking and providing opportunities to respond may be helpful toward facilitating oral communication with all children. For SM, opportunities to respond may include providing situations that allow for speaking, such as avoiding closed yes/no questions, calling on children rather than waiting for them to volunteer, providing a 'wait time' (perhaps 3-5 seconds) for responses, and creating small group classroom activities that include verbal responding. To maintain the expectation for speaking, it also may be beneficial not to reinforce nonverbal responding such as head nodding, pointing, or note writing in lieu of speaking (see Porjes, 1992; Watson & Kramer, 1992). To prevent potential selective mutism, these simple tactics would be used from the very first day of school, rather than waiting for children to begin speaking. It is important, however, to avoid creating excessive anxiety (although a little anxiety may be therapeutic) by pressuring or forcing a child to speak. (These methods also may be adapted by caregivers to use in the home and community before a child's entry into school.)

A third intervention that may reduce the chance of a child developing SM is preparing them for the transition to school, and establishing connections between neighboring preschools, elementary schools, and related service professionals. Research has not been focused on effective methods of preparation for children entering the early school grades. Crundwell (2006) highlighted empathizing with a child's

anxious feelings about entering school. It may be useful to examine schools that provide a visiting day(s) to caregivers and children prior to the beginning of the school year to measure the effects on reducing children's anxiety. Although some preschool children may not develop SM until kindergarten, characteristics related to SM should be acknowledged and monitored (Ullrich, Carroll, Prigot, & Fagen, 2002), and intervention should occur quickly before the problem becomes more entrenched. Although the prevention methods described may prevent SM from occurring or may lessen its severity, research is needed on the effectiveness of prevention strategies.

EARLY ONSET INTERVENTION FOR SELECTIVE MUTISM (TIER II)

Early onset, or Tier II, interventions can focus on helping children who are at-risk for SM. The goal is to minimize the need for individual treatment and resolve issues before they become more serious. Interventions at this level may include anxiety screening, classroom-wide techniques, and group therapy. There have only been a handful of Tier II intervention studies for SM, therefore although empirically little is known, we offer suggestions for Tier II methods based on theory and the available research base.

If school personnel decide to implement a transition program for incoming kindergartners, they have the advantage of identifying in advance who may develop SM or other potential behavioral and academic issues. Key signs of a child at-risk for SM are behaviors related to anxiety, such as amplified separation anxiety, shy behaviors, and slow-to-warm temperament (Bergman, et al., 2002; Ford, et al., 1998). Assessment tools such as the Selective Mutism Questionnaire (SMQ; Letamendi, et al., 2008) can be completed by caregivers to garner information about their child's level of communication and anxiety. Training for caregivers may be focused on teaching them how to cope with their child's anxiety and how to reduce anxiety by altering their own behavior (Sharkey, Nicholas, Barry, Bogley, & Ahern, 2008), such as providing opportunities for response and not reinforcing non-speaking behaviors at home or in the community. Teachers also can be notified of any incoming children at-risk for SM and accordingly prepare their classroom structure to identify SM related behaviors early, and to get a head start on ameliorative classroom based intervention.

Before engaging in more intrusive procedures, support personnel, such as school psychologists, can consult with teachers and caregivers to attempt the implementation of classroom and home-based strategies, including: opportunities to respond, contingency management, shaping, successive approximation, and monitoring. Opportunities to respond in this tier is more involved than Tier I because the method is focused on a specific child(ren). In contingency management, the teacher (and/or caregiver) attempts to positively reinforce all signs of verbal behavior while ignoring nonverbal behavior (Cohan, et al., 2006; Watson & Kramer, 1991; Wulbert, Nyman, Snow, & Owen, 1973). Pairing contingency management with shaping, adjustments can be made on which target behaviors are reinforced. For example, a child who does not exhibit verbal behavior may need lower, more attainable goals, such as participation with a group or any form of verbal responses in the classroom (Giddan, Ross, Sechler, & Becker, 1997).

Drawing from Tier III interventions, the most effective strategies with SM typically follow a gradual, systematic approach to ease the child into speaking, also known as successive approximation (Crundwell, 2006; Omdal, 2008). It is important for teachers and caregivers to be patient and to recognize small steps towards improvement. Lastly, data should be collected on the student's verbal and nonverbal behaviors to assess improvements that have occurred and what strategies were effective (Kearney & Vecchio, 2007). This process also may provide useful data if the child requires additional supports, such as Tier III intervention by the school psychologist.

Group therapy is a third possible intervention that fits within Tier II. This type of therapy allows the school to enhance efficiency by treating a group of children instead of each individual with SM. This mode of treatment appears to be seldom used, probably due to the low number of cases that occur each year. Therefore, group therapy may require the inclusion of students with shyness or other anxiety related problems, or other issues such as limited social skills. Group therapy often is focused on nonverbal and verbal goals (e.g., making eye contact, saying hello when greeted, responding yes or no, initiating conversation) to increase communication for use in the school and community. Bozigar and Hansen (1984) were successful using group therapy to treat three Hispanic-American girls and one African-American

girl (ages 6-9). Sharkey et al. (2008) conducted a study in Ireland using an 8-week group therapy program for five children with SM and their caregivers. The therapy yielded mixed results, with two of the five children no longer meeting the criteria for SM at post-treatment and at a six-month follow-up. This study indicated the need for more research on group therapy, and on the effects of involving caregivers in the therapy process. The research on group therapy has indicated that it may be an effective approach for SM, but the method requires more empirical support.

TERTIARY TREATMENT OF SELECTIVE MUTISM (TIER III)

Children with SM have received a variety of individual treatments, including psychoanalysis, behavior therapy, cognitive-behavioral therapy, psychopharmacology, and various combinations of these therapies, typified as multimethod interventions. Tier III interventions focus on the individual child, are the most common type of interventions for children with SM, and are the most researched and tangible approaches for the disorder.

Psychoanalysis

Psychodynamic theorists focus on understanding past events that may have influenced a child to stop speaking, and usually search for a conflict between the child and family members (Leonard & Dow, 1995). Psychoanalysts often use strategies such as family therapy and play therapy in an attempt to unfold certain feelings the child and family members might be having (Cohan, et al., 2006). There has been little documented success using psychoanalysis to treat SM, perhaps because often it involves a protracted treatment process (Giddan, et al., 1997). There also has been a lack of generalization for studies that have shown initial success in treating SM in clinical settings (Cline & Baldwin, 2004). Whereas psychoanalysis may be less effective on its own for the treatment of SM, psychoanalytic strategies such as play therapy and family therapy might be useful when combined with behavioral interventions.

Behavioral and Cognitive-Behavioral Therapy

Behavioral and cognitive-behavioral therapies have been shown to be more successful in treating SM (Beare, Torgerson, & Creviston 2008; Cohan, et al., 2006; Ford, et al, 1998; Giddan, et al., 1997; Kratochwill, 1981; Labbe & Williamson, 1984; Porjes, 1992; Stone, et al., 2002). In a review of the literature on SM from 1990-2005, Cohan et al. (2006) concluded that among all the interventions used to treat SM, behavioral and cognitive-behavioral interventions were the most effective and had the strongest research support. Behavioral and cognitive-behavioral strategies include: contingency management, shaping, social skills training, stimulus fading, systematic desensitization, relaxation training, and self-modeling.

Contingency management, shaping, and social skills training. Contingency management techniques involve the use of operant conditioning methods such as positive reinforcement in a strategic way to modify behavior (Giddan, et al., 1997). For example, teachers who implement a token economy are practicing contingency management. Porjes (1992) successfully incorporated contingency management in the treatment of two children with SM. The contingency consisted of receiving a desired reinforcer for verbal behavior. Other studies have been conducted that were successful in treating SM with contingency management methods, typically as part of an intervention package (see Cohan et al., 2006). Whereas contingency management has been shown to be an effective approach which often is recommended for use in the classroom, it may be more effective when paired with other behavioral strategies (Labbe & Williamson, 1984). Shaping is a common strategy to combine with contingency management, and involves small steps of successive approximation that serve as target behaviors to be reinforced. Children with SM may be sensitive to verbal communications and often may need smaller steps that decrease anxiety toward the ultimate goal of speaking to someone (Crundwell, 2006; Omdal, 2008). An additional technique which may be paired with contingency management is social skills training (SST), however, the research is limited. It is difficult to determine how effective SST is individually because it is often combined with other strategies such as contingency management (Cohan, et al., 2006; Fisak, Oliveros, & Ehrenreich., 2008).

Stimulus fading. One of the most effective behavioral strategies for children with SM is stimulus fading (Beare, et al., 2008; Kratochwill, 1981; Labbe & Williamson, 1984; Morris & Kratochwill, 1985; Watson & Kramer, 1992; Wulbert, et al., 1973). Stimulus fading is a Tier III intervention because it requires more focus on an individual and outside involvement during the process. Commonly paired with contingency management, stimulus fading is the process of reducing the control of a stimulus on a certain behavior. For a child with SM, the stimulus is usually a parent caregiver or other family member with whom the child is comfortable speaking. The process starts with the child in a controlled environment with that stimulus (Cohan, et al., 2006). Once the child has shown the willingness or ability to speak, a person with whom the child is less comfortable (e.g., the teacher) will gradually be added to the situation. Simultaneously, the comfortable stimulus (e.g., caregiver) will gradually become distant, and "fade" out. Shaping and reinforcement appear to be important in the process of stimulus fading because small steps need to be established and the child may require help to overcome the anxiety associated with each step (Cohan, et al., 2006). Stimulus fading should include classroom and teacher involvement, with the ultimate goal of generalizing speaking beyond the initial fading procedure into the daily school environment and community. Several studies have been conducted to examine the effectiveness of stimulus fading on SM (see Cohan et al., 2006). For example, Wulbert et al. (1973) paired stimulus fading with contingency management to successfully treat a 6-year-old girl with SM. This article stands out due to the specificity offered on the treatment steps of the fading process.

Systematic desensitization and relaxation therapy. Cognitive-behavioral therapy (CBT) has been shown to be an effective approach for increasing speech in children with SM (Cohan, et al., 2006; Grover, Hughes, Bergman, & Kingery, 2006; Schwartz, et al., 2006). CBT strategies often are paired with stimulus fading and contingency management (Cohan, et al., 2006). Systematic desensitization is the process of gradually exposing a person to a hierarchy of anxiety-provoking stimuli, with the goal of reducing the person's level of anxiety in each situation. This technique often is paired with relaxation training to manage the anxiety symptoms and improve outcomes. A hierarchy for a child with SM may include different levels of speaking situations, such as starting from whispering, to speaking in a small group, and ultimately to speaking audibly in front of an entire class. Suveg, Comer, Furr, and Kendall (2006) treated an 8-year-old girl identified with SM, social phobia, generalized anxiety disorder, and cognitive delays by training her in relaxation exercises, and then exposing her to a nine-step hierarchy that dealt with conversing with others (e.g., calling a friend on a phone or having a conversation with an unfamiliar adult). Relaxation exercises included writing assignments about situations in which she stated she was anxious, and conducting progressive muscle relaxation in anticipation of the anxiety provoking situation. The child was able to reach the higher steps in her hierarchy before the end of treatment. Overall, the data have indicated that systematic desensitization and relaxation training can be effective in the treatment of SM.

Self-modeling. Self-modeling involves using a video or audio device that records the child speaking, and then inserting the recording into an environment in which the child typically does not speak (Cohan, et al., 2006). The child is exposed to the recording with the goal of increasing their familiarity with the anxiety provoking stimulus (i.e., the classroom environs) which may result in more comfort that allows them to speak. Pigott and Gonzales (1987) used video self-modeling with a child who only spoke when his mother or brother were present in the classroom. Therapists in the study made a videotape of the classroom when the child's mother and brother were present, and edited the video to show the child answering direct questions and other tasks without his mother or brother present. The student watched the short videos before school for two weeks and received reinforcement for verbal behaviors in class. The intervention resulted in an overall increase in answering and asking questions in class. In some cases, video self-modeling has influenced children with SM to initiate conversations and speak freely within the classroom (Kehle, Madaus, Baratta, & Bray, 1998; Kehle & Owen, 1990). A related method may be to use virtual reality methods to place the child in a computer generated scenario. Although research has yet to be conducted, this technique may have the advantage of increasing child involvement due to the novelty of the method.

Psychopharmacological Therapy

Research findings supporting the link between SM and anxiety have influenced the use of psychopharmacological treatment (Carlson, Mitchell, & Segool, 2008; Leonard & Dow, 1995). For example, Golwyn and Weinstock (1990) conducted a study on the use of phenelzine with a 7-year-old girl with SM (phenelzine is a monoamine oxidase inhibitor (MAOI) that has been used for the treatment of depression and anxiety disorders). After 16 weeks, the child was speaking freely in school and in the community. The medication was gradually decreased and a five-month follow-up showed that the student remained talkative without the medication.

Carlson, Kratochwill, and Johnston (1994) surveyed 308 psychiatrists on how they might treat a child with SM. Of the sample, 199 reported having treated a child with SM, and 36% reported that they prescribed medication. Interestingly, only 14% of the psychiatrists who had treated SM reported that therapy that included medication was the most effective method of treatment. The method with the highest endorsement rate was psychotherapy (24%), followed by a combination of psychotherapy and family therapy (17%), and behavior therapy (14%).

It is the school psychologist's role to consult with the a child's caregivers and to collaborate on a decision whether a child should be referred to a physician, and to be aware that choosing a more invasive intervention should be done with caution (Schwartz et al., 2006). Phenelzine has produced undesirable side effects and is now rarely considered for treatment (Cline & Baldwin, 2004). Whereas other psychotherapeutic drugs such as selective serotonin reuptake inhibitors (SSRIs) have been shown in some cases to be effective in reducing symptoms of SM (see Carlson, et al., 2008 for a review of research on pharmacotherapy for SM), medication probably is best used as a last resort because of the potential side effects, and the unknown effects of psychotropic medication on brain development. However, medications may be required for more intractable or long-term cases or for children who also exhibit extreme depression or other anxiety disorders (Carlson, et al. 2008).

Multimethod Treatment

All the treatments described in this article have some level of empirical support for their use, which makes it confusing to know which treatment is the best to implement. However, a trend appears among most of these treatment studies: a combination of multiple strategies can be an effective way to treat SM. For example, stimulus fading often has been combined with contingency management, shaping, and relaxation training. Kehle et al. (1998) implemented a treatment with three students using combinations of self-modeling, reinforcement, stimulus fading, shaping, and medication. Each case showed improvement toward the end of treatment. Similarly, psychodynamic strategies such as play therapy and family therapy have been combined with behavioral strategies (Cline & Baldwin, 2004). The literature indicates that the successful treatment of SM is typified by the use of multiple methods of intervention (Astendig, 1998; Cohan, et al. 2006).

CONCLUSION

As an advocate for children, school psychologists can work with teachers, parent caregivers, speech pathologists, and other support staff toward helping children who may develop or have selective mutism. School psychologists also are responsible for updating their knowledge on the research regarding etiology and intervention strategies for SM.

A review of the literature on SM revealed that more research is needed in all the areas presented in this article: incidence, prevalence, demographics, primary prevention, early onset interventions, and tertiary interventions. Although SM may be a rare and complex disorder, many treatments have been shown to be successful in treating SM. A growing body of research has supported Tier III interventions, with the preponderance of research supporting behavioral and cognitive-behavioral therapies. More research particularly is needed to strengthen the theoretical, research, and practice base of Tier I and Tier II interventions. Research with children showing symptoms of anxiety and SM supports the possibility that primary and secondary interventions, such as caregiver and teacher training, kindergarten preparation,

anxiety screening, classroom intervention, and group therapy may be successful in schools. Based on the available theoretical and research base, we believe that Tier I and Tier II intervention may result in the prevention and amelioration of many occurrences of selective mutism.

- - -

R.T. Busse, PhD., is an associate professor in the Counseling and School Psychology Program at Chapman University. Dr. Busse has worked with children with selective mutism for over 15 years. Dr. Busse's other interests include methods for assessing response to intervention outcomes, curriculum-based assessment, and child and adolescent psychopharmacology.

Jenna Downey is an EdS student in the School Psychology Program at Chapman University. Her professional interests include selective mutism, autism, applied behavior analysis, and mental health disorders.

REFERENCES

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

Anstendig, K. (1998). Selective mutism: A review of the treatment literature by modality from 1980-1996. *Psychotherapy, 35*, 381-391.

Beare, P., Torgerson, C., & Creviston, C. (2008). Increasing verbal behavior of a student who is selectively mute. *Journal of Emotional and Behavioral Disorders, 16*, 248-255.

Bergman, L.R., Piacentini, J., & McCracken, J.T. (2002). Prevalence and description of selective mutism in a school-based sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*. 938-946.

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

Bozigar, J.A., & Hansen, R.A. (1984). Group treatment for elective mute children. *Social Work, 29*, 478-480.

Carlson, J., Kratochwill, T., & Johnston, H. (1994). Prevalence and treatment of selective mutism in clinical practice: A survey of child and adolescent psychiatrists. *Journal of Child and Adolescent Psychopharmacology, 4*, 281-291.

Carlson, J.S., Mitchell, A.D., & Segool, N. (2008). The current state of empirical support for the pharmacological treatment of selective mutism. *School Psychology Quarterly, 23*, 354-372

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

Cline, T., & Baldwin, S. (2004) *Selective mutism in children*. London, England: Whurr Publishers Ltd.

Cohan, S.L., Chavira, D.A., & Stein, M. B. (2006). Practitioner review: Psychosocial interventions for children with selective mutism: A critical evaluation of the literature from 1990-2005. *Journal of Child Psychology and Psychiatry, 47*, 1085-1097.

Crundwell, R.M.A. (2006). Identifying and teaching children with selective mutism. *Teaching Exceptional Children, 38*, 48-54.

Cunningham, C E., McHolm, A.E., Boyle, M.H. (2006). Social phobia, anxiety, oppositional behavior, social skills, and self-concept with specific selective mutism, generalized selective mutism, and community controls. *European Journal of Child and Adolescent Psychiatry, 17*, 245-255.

Cunningham, C.E., McHolm, A., Boyle, M.H., & Patel, S. (2004). Behavioral and emotional adjustment, family functioning, academic performance, and social relationships in children with selective mutism. *Journal of Child Psychology and Psychiatry, 45*, 1363-1372.

Dadds, M.R., & Roth, J.H. (2008). Prevention of anxiety disorders: Results of a universal trial with young children. *Journal of Child and Family Studies, 17*, 320-335.

Elizur, Y., & Perednik, R. (2003) Prevalence and description of selective mutism in immigrant and native families: A controlled study. *Journal of American Academy of Child and Adolescent Psychiatry, 42*, 1452-1459.

Fisak, B.J., Oliveros, A., & Ehrenreich, J.T. (2006). Assessment and behavioral treatment of selective mutism. *Clinical Case Studies, 5*, 382-402.

Ford, M., Sladeczek, I.E., Carlson, J., & Kratochwill, T.R. (1998). Selective mutism: Phenomenological characteristics. *School Psychology Quarterly, 13*, 192-227.

Giddan, J.J., Ross, G.J., Sechler, L.L., & Becker, B.R. (1997). Selective mutism in elementary school: Multidisciplinary interventions. *Language, Speech & Hearing Services in Schools, 28*, 127-133.

Goldwyn, D.H., & Weinstock, R.C. (1990). Phenelzine treatment of elective mutism: A case report. *Journal of Clinical Psychiatry, 51*, 384-385.

Grover, R.L., Hughes, A.A., Bergman, R.L., & Kingery, J.N. (2006). Treatment modifications based on childhood anxiety diagnosis: Demonstrating the flexibility in manualized treatment. *Journal of Cognitive Psychotherapy: An International Quarterly, 20*, 275-286.

Kagan, J. (1997). Temperament and the reactions to the unfamiliar. *Child Development, 68*, 139-143.

Kearney, C.A., & Vecchio, J.L. (2007) When a child won't speak. *The Journal of Family Practice, 56*, 917-921.
Kehle, T.J., Madaus, M.R., Baratta, V.S., & Bray, M.A. (1998). Augmented self-modeling as a treatment for children with selective mutism. *Journal of School Psychology, 36*, 247-260.
Kehle, T., & Owen, S.V. (1990). The use of self-modeling as an intervention in school psychology: A case study of an elective mute. *School Psychology Review, 19*, 115-121.
Kolvin, I., & Fundudis, T. (1981) Elective mute children: Psychological development and background factors. *Journal of Child Psychology and Psychiatry, 22*, 219-232.
Kratochwill, T.R. (1981). *Selective mutism: Implications for research and practice*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Kratochwill, T., & Morris, R. (1985). Conceptual and methodological issues in the behavioral assessment and treatment of children's fears and phobias. *School Psychology Review, 14*, 94-107.
Krysanski, V. (2003). A brief review of selective mutism literature. *The Journal of Psychology, 137*, 29-40.
Labbe, E.E., & Williamson, D.A. (1984). Behavioral treatment of elective mutism: A review of the literature. *Clinical Psychology Review, 4*(3), 273-292.
Leonard, H., & Dow, S. (1995). Selective mutism. In S. March (Ed.), *Anxiety Disorders in Children and Adolescents* (pp. 235-250). New York: Greenwood.
Letamendi, A.M., Chavira, D.A., Hitchcock, C.A., Roesch, S.C., Shipon-Blum, E., & Stein, M.B. (2008). Selective mutism questionnaire: Measurement structure and validity. *Journal of American Academy of Child and Adolescent Psychiatry, 47*, 1197-1204.
Manassis, K., Tannock, R., Garland, E.J., Minde, K, McInnes, A., & Clark, S. (2007). The sounds of silence: Language, cognition, and anxiety in selective mutism. *Journal of American Academy of Child and Adolescent Psychiatry, 46*, 1187-1195.
Marshall, P.J., & Stevenson-Hinde, J. (2001). Behavioral inhibition: Physiological correlates. In W.R. Crozier & L.E. Alden (Eds.), *International Handbook of Social Anxiety: Concepts, research and interventions relating to the self and shyness* (pp. 53-76). New York: John Wiley & Sons.
Morris, R.J., & Kratochwill, T.R. (1985). Behavioral treatment of children's fears and phobias: A review. *School Psychology Review, 14*, 84-93.
Omdal, H. (2008). Including children with selective mutism in mainstream schools and kindergartens: Problems and possibilities. *International Journal of Inclusive Education, 12*, 301-315.
Pigott, H.E., & Gonzales, F.P. (1987). Efficacy of videotape self-modeling in treating an electively mute child. *Journal of Clinical Child Psychiatry, 12*, 106-110.
Porjes, M.D. (1992). Intervention with the selectively mute child. *Psychology in the Schools, 29*, 367-376.
Remschmidt, H., Poller, M., Herpertz-Dahlmann, B., Hennighausen, K., & Gutenbrunner, C. (2001). A follow-up study of 45 patients with selective mutism. *European Archives of Psychiatry and Neuroscience, 251*, 284-296.
Schwartz, R., Freedy, A., & Sheridan, M. (2006). Selective mutism: Are primary care physicians missing the silence? *Clinical Pediatrics, 45*, 43-48.
Sharkey, L., McNicholas, F., Berry, E., Begley, M., & Ahern, S. (2008). Group therapy for selective mutism – A parents' and children's treatment group. *Journal of Behavior Therapy and Experimental Psychiatry, 39*, 538-545.
Standart, S., & Couteur, A.L. (2003). The quiet child: A literature review of selective mutism. *Child and Adolescent Mental Health, 8*, 154-160.
Steinhausen, H., & Juzi, C. (1996). Elective mutism: an analysis of 100 cases. *Journal of American Academy of Child and Adolescent Psychiatry, 35*, 606-614..
Steinhausen, H., Wachter, M., Laimbock, K., & Metzke, C.W. (2006). A long-term outcome study of selective mutism in childhood. *Journal of Child Psychology and Psychiatry, 47*, 751-756.
Stone, B.P., Kratochwill, T.R., Sladezcek, I., & Serlin, R.C. (2002). Treatment of selective mutism: A best-evidence synthesis. *School Psychology Quarterly, 17*, 168-190.
Suveg, C., Comer, J.S., Furr, J.M., & Kendall, P.G. (2006). Adapting Manualized CBT for a cognitively delayed child with multiple anxiety disorders. *Clinical Case Studies, 5*, 488-510.
Toppelberg, C.O., Tabors, P., Coggins, A., Lum, K., & Burger, C. (2005). Differential diagnosis of selective mutism in bilingual children. *Journal of American Academy of Child and Adolescent Psychiatry, 44*, 592-595.
Ullrich, A., Carroll, M., Prigot, J., & Fagen, J. (2002). Preschoolers' inhibition in their home: Relation to temperament. *The Journal of Genetic Psychology, 163*, 340-359.
Watson, T.S., & Kramer, J.J. (1992). Multimethod behavioral treatment of long-term selective mutism. *Psychology in the Schools, 29*, 359-366.
Wulbert, M., Nyman, B.A., Snow, D., & Owen, Y. (1973). The efficacy of stimulus fading and contingency management in the treatment of elective mutism: A case study. *Journal of Applied Behavior Analysis, 6*, 435-441.
Yeganeh, R., Beidel, D.C., & Turner, S.M. (2006). Selective mutism: More than social anxiety? *Depression and Anxiety, 23*, 117-123.

TOC

GENERAL ARTICLES

TOC

The Student Engagement in Schools Questionnaire (SESQ) and the Teacher Engagement Report Form-New (TERF-N): Examining the Preliminary Evidence

Shelley R. Hart, Kaitlyn Stewart & Shane R. Jimerson,
University of California Santa Barbara

Student engagement in school is an important construct that has been associated with student success. For the current study, researchers examined the psychometrics of the Student Engagement in Schools Questionnaire (SESQ) and the Teacher Engagement Report Form (TERF-N) of student engagement. The results revealed that both the SESQ and the TERF-N have good internal consistency. The exploratory factor analysis results for the SESQ demonstrated alignment with the theoretically driven development (five factors: Affective Engagement-Liking for Learning, Affective Engagement-Liking for School, Behavioral Engagement-Effort & Persistence, Behavioral Engagement-Extracurricular, and Cognitive Engagement) whereas the results for the TERF-N were more complicated. The items did not load as conceptualized in a 3-factor model, but instead loaded on one, General Engagement factor. Finally, while it may be that teachers viewed a student's level of engagement as a global construct, the correlations between the measures indicated that they might be used to provide helpful, convergent information obtained from a variety of sources regarding a student's levels of engagement. Future directions and implications for school psychologists are discussed.

Engagement is a growth-producing activity through which an individual allocates attention in active response to the environment (Csikszentmihalyi, 1990). Engagement related to school activity (or student engagement) has become an important concept related to multiple educational outcomes (e.g., achievement, attendance, behavior, dropout/completion; e.g., Finn, 1989; Jimerson, Campos, & Greif, 2003; Jimerson, Renshaw, Stewart, Hart, & O'Malley, 2009). Student engagement has been identified as a primary variable in understanding dropout, particularly as a gradual process operating in a student's life and influencing that final decision to withdraw (Jimerson et al., 2009). Numerous studies have linked student engagement with improved academic performance and it has repeatedly demonstrated to be a robust predictor of achievement and behavior in the schools (Appleton, Christenson, & Furlong, 2008; Shernoff & Schmidt, 2008). It has also been correlated with both health compromising (e.g., substance abuse, depression, suicidality, aggression, early sexual activity) and health promoting (e.g., exercise, nutrition, safe sex activities) behaviors (Carter, McGee, Taylor, & Williams, 2007).

As a result of its demonstrated relationships with a variety of outcomes, it is postulated that an understanding of student engagement might help educators prevent deleterious outcomes and promote positive ones for at-risk students. Student engagement is a construct that resonates with most consumers of education, including students and parents (Appleton, Christenson, & Furlong, 2008) and presents an attractive focus for researchers and educators, in that compared to other predictors of academic success that are static (e.g., socioeconomic status [SES], ethnicity), it is believed to be a malleable characteristic and therefore a more appropriate focus for interventions (e.g., Christenson, Sinclair, Lehr, & Godber, 2001). In addition, both the individual *and* the environment shape a student's level of engagement, thus, there are many factors in the school environment (e.g. interpersonal relationships, recognition) that may enhance it (Fredricks, Blumenfeld, & Paris, 2004). Indeed, researchers have shown that effective interventions to promote student engagement and motivation also enhance the probability of high school completion (Reschly, Appleton, & Christenson, 2007). For these reasons it can be viewed as an asset associated with positive student outcomes (Furlong et al., 2003).

Correspondence may be sent to Shelley R. Hart, UCSB, GGSE, CCSP, Santa Barbara, CA 93106-9490 or e-mail: shart@education.ucsb.edu or jimerson@education.ucsb.edu

DEFINING AND MEASURING STUDENT ENGAGEMENT IN SCHOOL

Despite its apparent utility, student engagement remains a nebulous construct with researchers using ambiguous or inconsistent definitions resulting in equally nebulous measures. Several recent reviews have focused on defining this meta-construct and setting the stage for future scholarship (see Appleton et al., 2008; Fredricks et al., 2004; Jimerson et al., 2003). These scholars (i.e., Appleton et al., 2008; Jimerson et al., 2003) suggest that student engagement in school is multi-dimensional and appears to overlap with several similar constructs (e.g., school connectedness, school bonding). The proposed definition includes both *indicators* (i.e., affective, behavioral, and cognitive) and *facilitators* (i.e., both personal and contextual factors that influence engagement) of engagement (Appleton et al., 2008). Each component is vital to a complete understanding of student engagement. Appleton and colleagues (2008) have suggested that indicators are proposed to "…convey a student's degree or level of connection with learning"; while facilitators are "…factors [that] influence the strength of the connection" (p. 382).

The current study is focused primarily on the indicators of student engagement, and therefore, each indicator will be discussed further. *Affective* engagement refers to a student's feelings toward his school, learning, teachers, and peers (e.g., the student has positive feelings toward his teachers; Jimerson et al., 2003). The terms psychological and emotional engagement have also been used in the current literature to describe this construct (Appleton, Christenson, Kim, & Reschly, 2006; Reschly et al., 2007). *Behavioral* engagement includes observable student actions or participation while at school and is investigated through a student's positive conduct, effort, and participation (e.g., participation in extracurricular activities, attendance, and work habits; (Fredricks et al., 2004). Historically, research has been focused primarily on this aspect of student engagement. *Cognitive* engagement includes a student's perceptions and beliefs associated to school and learning (e.g., I will do well in this class if I try). It refers to the cognitive processing a student brings to academic tasks as well as the amount and type of strategies a student utilizes (Walker, Greene, & Mansell, 2006).

Some researchers propose the notion of academic engagement as a fourth indicator of student engagement (e.g., Reschly & Christenson, 2006). Academic engagement has been defined as time spent in academic learning. We contend that academic engagement can be better explained as an aspect of one of the three more commonly identified indicators (e.g., time-on-task is more accurately described as a behavioral indicator) or as an *outcome* of student engagement (e.g., Grade Point Average [GPA]).

Whereas there seems to be a general consensus that three indicators of engagement exist, there still remain differences in precisely *how* these indicators are defined and measured. For example, Jimerson et al. (2003) locate motivation within the affective engagement indicator, while Fredricks et al. (2004) define this construct as a cognitive indicator of engagement, and Patrick, Ryan, and Kaplan (2007) describe it as a cognitive precursor to engagement. Therefore, an obvious challenge remains for researchers of student engagement in parsing out the characteristics of each component.

PURPOSE OF THE PRESENT STUDY

It appears to be the multidimensional nature of student engagement that has created confusion in the field. For example, researchers may focus on only one component (unidimensional approach) or mix elements of several components (mixed approach), nonetheless operationalizing it as "student engagement." We argue that part of the reason for this confusion is the lack of a comprehensive measure to examine the meta-construct of student engagement. Thus, a psychometrically sound, universal measure of student engagement would advance scholarship in this area.

In order to fill this need for a comprehensive instrument, researchers from more than 19 countries collaborated to develop such a measure and to study student engagement internationally (Lam & Jimerson, 2008). *The Student Engagement in Schools Questionnaire* (SESQ) was the product of this collaboration. As a self-report measure, the SESQ obtains information from the student's perspective related to both the indicators and facilitators of engagement. Optimally, in an assessment of any construct, information would be collected from a variety of sources, in a variety of contexts, through a variety of methods, and over a period of time. To further this aim and provide a complement to the self-report SESQ and to evaluate the three indicators of engagement from the teacher perspective, the *Teacher Engagement*

Report Form - New (TERF-N)[1] was also examined.

The main purpose of this study is to establish the psychometric properties of the SESQ and TERF-N. Specifically, reliability and validity evidence will be evaluated through internal consistency estimates, exploratory factor analysis and correlations between measures.

METHOD

Participants

The present study utilized a sample drawn from one junior high and one high school located in the central coast area of California. For analyses of the SESQ, a sample of N = 428 seventh- through ninth-grade students was obtained. There were very few eighth-grade students included in the sample (5%), while ninth-graders composed the majority of the sample (59%), followed by seventh-graders (36%). Fifty-four percent of the sample was male, 42% Hispanic, 25% African American, 6% White (non-Hispanic), and 2% other. Due to the return rate of the TERF-N by the teachers (N = 4), for these analyses, a subsample (N = 129 seventh-grade students; 48% male) of the larger sample was utilized. The classrooms are considered to be representative of the schools, as well as the community because the demographics of the classrooms from which the same was drawn are similar to the both the schools and the communities where they are situated.

Procedure

Participation was requested through direct contact with school administrators. Two schools agreed to participate. Next the teachers at the two schools were contacted to determine their interest in and availability for the project. Ten of the teachers contacted agreed to participate. Finally, researchers obtained consent from the students and the survey was completed during one class period in spring 2008 and spring 2009. The researchers, school psychology graduate students from the University of California, Santa Barbara, presented the surveys to students, provided directions, fielded questions, and collected completed surveys. Students were provided with an alternative to opt-out of the data collection procedure. Teachers completed their ratings of students, while students were completing the self-report.

Measures

Student Engagement in Schools Questionnaire (SESQ). Scholars from more than 19 countries collaborated in the development of the SESQ (see Lam & Jimerson, 2008, for a description of this process and the international scholars who participated). The SESQ is a 109-item paper-and-pencil, Likert-type, self-report questionnaire focused on the comprehensive assessment of the construct of student engagement. After agreeing on the definition of student engagement, scholars developed a questionnaire to encompass this construct. Items were drawn from existing research, increasing the content validity of the measure (see Lam & Jimerson, 2008, for a detailed description of this process and the resulting measure). The SESQ contains four composites (i.e., Student Engagement in the Schools, Motivational Beliefs, Social-Relatedness Contexts, Student Outcomes) within which are 13 domains and 15 sub-domains. Students respond according to a Likert-type scale of 1-5 (e.g., *1 = never*, *5 = always*) and typically require approximately 35-minutes to complete. Due to the sampling restrictions associated with a survey of 109-items (i.e., a very large sample would be required for evaluation of the entire survey), for purposes of this study the items representing only the indicators of engagement (i.e., Affective, Behavioral, and Cognitive) are examined in the exploratory factor analysis; therefore, only the composite of Student Engagement in Schools (ENG; 33 items) was explored. However, reliability estimates are examined for the entire survey. As proposed by researchers, the SESQ-ENG is composed of five factors (Affective Engagement: Liking for Learning; Affective Engagement: Liking for School; Behavioral Engagement: Effort and Persistence; Behavioral Engagement: Extracurricular Activities; and Cognitive Engagement; Lam & Jimerson, 2008).

[1]An original teacher report (*Teacher Engagement Report Form-Original* [TERF-O]; Lam & Jimerson, 2008) was developed as part of the international collaboration. However, we anticipated that there were indicators of engagement that were not measured by the original 6 questions requested; therefore we developed the TERF-N to expressly access teachers' impressions of all three indicators (i.e., affective, behavioral, and cognitive engagement) for each student.

Teacher Engagement Report Form (TERF-N). The TERF-N is a 10-item, paper-and-pencil chart, where the teacher fills in 10 boxes, one for each item, per student. Each item is completed using a Likert-type scale of 1-5 (e.g., *1 = strongly disagree*, *5 = strongly agree*). The TERF-N requires approximately 45 minutes to complete for 30 students. The questionnaire items address aspects of affective, behavioral, and cognitive engagement.

Data Analyses

Analyses for this study were selected for scale development and validation. Establishing the reliability of a measure is a crucial first step in scale development. Therefore, the analyses for each measure began with internal consistency estimates to examine reliability. Next, as these measures have not been analyzed prior, exploratory factor analyses (EFA) were conducted on each to examine the latent factor structure of each measure. Finally, correlations were conducted to examine the relationship of the scales and to examine external validity. Internal consistency estimates and correlations were conducted utilizing the SPSS package (version 16.0; SPSS, 2007), and the EFA's were conducted using Mplus software (version 5.21; Muthén & Muthén, 2009).

RESULTS

Preliminary Analysis

Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multi-collinearity, with no serious violations noted. Skewness and kurtosis for each item within both the SESQ-ENG and the TERF-N were evaluated to be within acceptable (±2) limits.

Internal Consistency

Cronbach's coefficient alpha (α) is the most common coefficient of reliability (Cronbach & Shavelson, 2004). Alpha is the ratio of the variance of the true scores to the observed scores; therefore, the higher the reliability, the closer the true scores will be to the observed scores (Gliner & Morgan, 2000). This measure of internal consistency is used to demonstrate how well a set of items measures a unidimensional latent construct (e.g., affective engagement). For this reason, separate coefficient analyses were run for each domain of the SESQ and TERF-N. The literature demonstrates a range of acceptable alpha levels from $.60 \leq \alpha \leq .90$ (Gliner & Morgan, 2000). Acceptable coefficients for this study were set at $\alpha \geq .70$.

Table 1 lists the internal consistency estimates for both the domains of the SESQ and the overall TERF-N. In general, both measures demonstrate good reliability. Estimates for the SESQ range from $.65 \leq \alpha \leq .95$. Only one domain (Attributions) did not demonstrate the acceptable level of $\alpha \geq .70$. The data for the TERF-N indicates good internal consistency ($\alpha = .83$) between the 10 items.

Table 1 *Cronbach's Alpha (α) Estimates for the Domains of the Student Engagement in Schools Questionnaire (SESQ) and the Overall Teacher Engagement Report Form (TERF-N)*

Domain	α	Domain	α
SESQ: Affective Engagement	.88[a]	*SESQ*: Teacher Support	.83[a]
SESQ: Behavioral Engagement	.85[a]	*SESQ*: Peer Support	.84[a]
SESQ: Cognitive Engagement	.93[a]	*SESQ*: Peer Aggression	.84[a]
SESQ: Goal Orientations	.85[a]	*SESQ*: Peer Victimization	.78[a]
SESQ: Attributions	.65	*SESQ*: Parental Support	.82[a]
SESQ: Learning Self-Efficacy	.84[a]		
SESQ: Motivating Instructional Contexts	.95[a]	*TERF-N*	.83[a]

Note. [a] Domain meets or exceeds the acceptable $\alpha \geq .70$.

Factor Structure

Establishing the internal (factor) structure of a measure is an important step in psychometrics. Factor analysis is a common way to do this. This process allows researchers to a) clarify the number of factors within a set of items, b) reveal the associations among the factors, and c) link items to factors (Furr & Bacharach, 2008). It accomplishes these tasks through a number of statistical techniques that aim to simplify complex data sets (Kahn, 2006; Kline, 2002). As described by Brown (2006), factors (or latent, unobservable, variables/constructs) account for the variation and covariation among a set of observed indicators (or items). That is, the indicators are intercorrelated because they are influenced by a common construct/factor/latent variable and if that latent variable were partialed out, the correlations between the indicators would be zero. As such, factor analysis provides for a more parsimonious understanding of a construct, as there are fewer factors than indicators. Exploratory factor analysis (EFA) is the first step in establishing the factors underlying a measure and in conducting data reduction. Therefore, EFA is the focus of the current study.

EFA requires several decisions as a result of a variety of rotation and estimation options. For any analysis with two or more factors, there exist an infinite number of equally good-fitting solutions, with each solution represented by a different factor-loading matrix (Brown, 2006). This means that any multiple-factor EFA model does not have a unique solution and researchers must make decisions about the solution interpreted among those infinite equally fitting solutions (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Pursuit of the property of simple structure is the criterion most typically used for the selection of solutions (Thurstone, 1947). Simple structure is defined as a solution where a) each factor is defined by a subset of indicators that load highly (primary loading) on the factor[2], and b) each indicator has a high loading on one factor and a trivial (i.e., close to zero) loading on any remaining factors (i.e., secondary loadings). In order to obtain simple structure and increase interpretability, rotation of the factors in multidimensional space is recommended.

Two basic types of rotation exist: orthogonal and oblique (Brown, 2006). In an orthogonal rotation, the factors are constrained to be uncorrelated (i.e., the axes of the factors remain at 90° angles), whereas with an oblique rotation, the factors are allowed to intercorrelate (i.e., the axes are allowed to be more or less than 90° angles). An oblique rotation is believed to provide a more realistic representation of the interrelated nature of the factors (or underlying constructs; Fabrigar et al., 1999), provides solutions that are more likely to generalize to CFA, and if the factors are truly not correlated, an oblique rotation will provide a solution virtually the same as an orthogonal rotation (Brown, 2006). Therefore, oblique rotation was selected in this study. Several oblique rotation methods have demonstrated satisfactory solutions, without one approach clearly dominating research (Fabrigar et al., 1999; Kline, 2002); therefore, Geomin rotation (Yates, 1987) was selected, as it is the default rotation used in the Mplus software for EFA.

Additionally, an estimation procedure (i.e., finding parameter values of a model that best fit the data) needs to be indicated. Maximum Likelihood (ML) estimation was chosen due to its ability to provide goodness-of-fit statistics. Goodness-of-fit statistics provide an overall summary of the model's ability to reproduce the observed covariance matrix. There are three different types of fit statistics: absolute, parsimony corrected, and comparative/incremental. Within these types exist a variety of indices; at least one index should be considered from each type when evaluating the overall fit of the model (Brown, 2006). Absolute fit statistics assess the magnitude of discrepancy between the sample and fitted or predicted variance-covariance matrices (Hu & Bentler, 1999) and include the chi-square (χ^2) statistic and the standardized root mean square residual (SRMR). Parsimony correction fit statistics incorporate a penalty for poor model parsimony (i.e., more than needed freely estimated model parameters). The most commonly used statistic from this category is the root mean square error of approximation (RMSEA). Finally, comparative fit statistics assess the fit of the model specified with a null model -- typically one where the covariances among all indicators are fixed to zero (variances are not constrained; Brown, 2006), with the most popular being the comparative fit index (CFI). The best fitting model would have a non-significant

[2]Factor loadings are indicated by lambda (λ) and are defined as completely standardized estimates of the regression slopes predicting the indicators by the latent variable, or factor (Brown, 2006).

chi-square (χ^2) statistic (although this is very sensitive to sample size; Brown, 2006), a Comparative Fit Index of > .95 (CFI), a Root Mean Square Error of Approximation of < .06 (RMSEA), a Standardized Root Mean-Squared Residual of < .08 (SRMR; Hu & Bentler, 1999).

Recommendations for the number of factors to extract, an optimal sample size, and factor loading cut-offs vary in the literature. One common method for determining the number of factors to retain is Kaiser's criterion (retain factors with eigenvalues over 1.00); however, after reviewing numerous studies that indicated this procedure tended to overestimate the number of factors, Fabrigar et al. (1999) stated knowledge of "…no study of this rule that shows it to work well" (p. 278). Additionally, it has been suggested that this criterion is only appropriate with principal components analysis (PCA; Kahn, 2006). Another method utilizing eigenvalues is Cattell's "scree test" (Cattell, 1966). The scree plot is examined to identify the last substantial drop in magnitude of the eigenvalues. A non-horizontal line can be drawn through the point where the eigenvalues "flatten out," and the number of factors before that flattening out are retained. This criterion has been criticized due to its subjectivity (Fabrigar et al., 1999). Both methods are provided in this study as they are still frequently referenced in the literature.

Sample size is another issue debated in the literature. The ratio of item-to-subject for the SESQ was 428:33 (or 13:1), while recommendations typically state at least a 5:1 ratio (Kahn, 2006). The sample size for the TERF-N was smaller (N=87, still resulting in a 8:1 ratio). A recent study exploring best practices in EFA (Costello & Osbourne, 2005) suggests that 60-70% of the time, the SESQ ratio would likely result in the correct factor structure, while 40-60% of the time, the TERF-N ratio would result in the correct factor structure. However, it has also been discussed that the importance in sample size is related to the communalities (not simply the size of the population). An item's communality is the variance explained in that item by the common factors. For communalities in the moderate range (such as demonstrated by the SESQ and TERF-N), a sample size of at least 200 is recommended (Fabrigar et al., 1999).

While EFA is a data-driven approach (i.e., the number of factors is not specified a priori), it is important to note that in order for an EFA to be useful it must make sense. Theory may play a large part in the determination of the number of factors extracted. Additionally, a balance between parsimony and plausibility must be struck (i.e., the fewest number of factors used to account for the correlations among the variables; Fabrigar et al., 1999). After the number of factors was determined, factor-loading cut-offs of $\lambda \geq .32$ were used in the current study per recommendations in the literature (Tabachnick & Fidell, 2007).

Student Engagement in the Schools Questionnaire-Engagement Composite (SESQ-ENG)

Goodness-of-fit statistics for the Engagement items of the SESQ (SESQ-ENG) EFA are located in Table 2. Eigenvalues and the scree plot can be found in Figure 1. According to the Kaiser criterion, a six-factor model is suggested; while the SESQ scree plot indicated a two-factor model and the chi-square statistic remains significant. Again, these methods for determining the number of factors to extract are problematic. The CFI indicated a six-factor model, while the RMSEA and the SRMR might indicate either the five- or six-factor model. The improvements in fit for the CFI and the RMSEA are incremental from the five- to the six-factor model; therefore, in the interest of parsimony and theory (a five-factor model was hypothesized) the five-factor model was evaluated.

Table 3 displays the factor loadings and communalities of the SESQ-ENG items. The factors were all well-determined (i.e., at least three items loaded on each factor) and relatively simple structure (i.e., each item loaded significantly on only one factor) was established. Results indicated that items loaded significantly on hypothesized and theoretically determined factors. The cumulative percentage of variance explained by the five factors was 61.45% (a breakdown by factor can be found in Table 4). The correlations between the factors indicated medium to large correlations (i.e., between .37 - .66), but not too large to indicate multicollinearity; therefore, it appears that each factor represents a unique construct within the concept of student engagement (Table 5 displays the factor correlations). Two items cross-loaded (i.e., item loaded on more than one factor at a significant level). These two items loaded on both factors one (Affective - Liking for Learning) and factor three (Behavioral - Effort and Persistence). However, these items loaded more strongly on the theoretically-driven factor. The majority of the correlations were significant at $p < .01$. There is a wide range of values from trivial-to-large ($r = .01$ - $.74$). The average inter-item correlation is in the moderate range ($r = .34$).

Table 2 Goodness-of-Fit Statistics for the SESQ-ENG EFA

Model	χ^2	DF	CFI	RMSEA	SRMR
1	3400.96***	495	.61	.12 (.11-.12)	.10
2	2124.26***	463	.78	.09 (.09-.10)	.06
3	1488.41***	432	.86	.08 (.07-.08)	.05
4	1160.08***	402	.90	.07 (.06-.07)	.04
5	900.98***	373	.93	.06 (.05-.06)[a]	.03[a]
6	739.17***	345	.95[a]	.05 (.05-.06)[a]	.03[a]

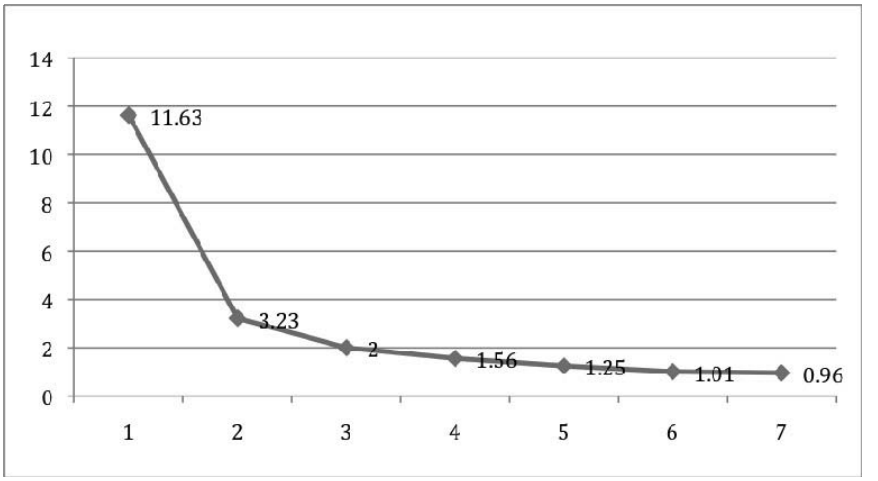
Note. DF=Degrees of Freedom. CFI = Comparative Fit Index. RMSEA = Root Mean Square of Error of Approximation. SRMR = Standard Root Mean-Squared Residual. AIC = Aikake Information Criterion. BIC = Bayesian Information Criterion.
*** $p < .001$.
[a] Model meets standards for that index established in the literature.

Table 3 *Factor Loadings for the SESQ-ENG Items*

Factor	Item	λ	Communality
	I am very interested in learning. (a1)***	.44 (.38)	.57
1 Affective: Liking for Learning	I think what we are learning in school is interesting. (a3)	.85	.74
	I like what I am learning in school. (a5)	.79	.70
	I enjoy learning new things in class. (a7)	.52	.61
	I think learning is boring. (a9R)***	.39 (.33)	.41
2 Affective: Liking for School	I like my school. (a11)	.92	.79
	I am proud to be at this school. (a13)	.87	.78
	Most mornings, I look forward to going to school. (a15)	.34	.36
	I am happy to be at this school. (a17)	.77	.77
3 Behav.: Effort & Persist.	I try hard to do well in school. (a2)	.45	.50
	In class, I work as hard as I can. (a4)	.45	.53
	When I'm in class, I participate in class activities. (a6)	.32	.40
	I pay attention in class. (a8)	.63	.55
	When I'm in class, I just act like I'm working. (a10R)	.66	.45
	In school, I do just enough to get by. (a12R)	.62	.41
	When I'm in class, my mind wanders. (a14R)	.42	.19
	If I have trouble understanding a problem, I go over it again until I understand it. (a16)	.50	.36
	When I run into a difficult homework problem, I keep working at it until I think I've solved it. (a18)	.52	.52
4 Behav.: Extracurr	I am an active participant of school activities such as sport day and school picnic. (a20)	.72	.59
	I volunteer to help with school activities such as sport day and parent day. (a22)	.82	.71
	I take an active role in extracurricular activities in my school. (a24)	.59	.59
5 Cognitive	When I study, I try to understand the material better by relating it to things I already know. (b1)	.59	.47
	When I study, I figure out how the information might be useful in the real world. (b2)	.58	.38
	When learning new information, I try to put the ideas in my own words. (b3)	.62	.52
	When I study, I try to connect what I am learning with my own experiences. (b4)	.76	.57
	I make up my own examples to help me understand the important concepts I learn from school. (b5)	.62	.51
	When learning things for school, I try to see how they fit together with other things I already know. (b6)	.70	.53
	When learning things for school, I often try to associate them with what I learnt in other classes about the same or similar things. (b7)	.78	.64
	I try to see the similarities and differences between things I am learning for school and things I know already. (b8)	.79	.66
	I try to understand how the things I learn in school fit together with each other. (b9)	.73	.62
	I try to match what I already know with things I am trying to learn for school. (b10)	.81	.67
	I try to think through topics and decide what I'm supposed to learn from them, rather than studying topics by just reading them over. (b11)	.64	.50
	When studying, I try to combine different pieces of information from course material in new ways. (b12)	.70	.65

Note.*** Items that cross-loaded. Both items loaded on Factors 1 & 3. Values on the conceptualized factors were higher (values in parentheses are the factor loadings on the second factor). The average communality for the SESQ = .55.

Figure 1. *Scree Plot of the Eigenvalues for the SESQ*



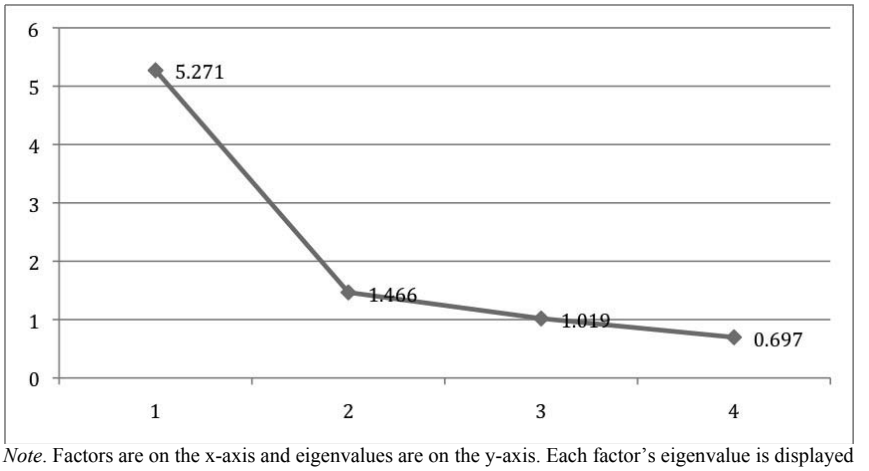
Note. Factors are on the x-axis and eigenvalues are on the y-axis. Each factor's eigenvalue is displayed in the chart (e.g., factor 1 had an eigenvalue of 11.63).

Table 4 *Variance Explained in the SESQ-ENG*

Factor	Percent of variance explained
Cognitive Engagement	36.82%
Behavioral Engagement (Effort & Persistence)	9.81%
Affective Engagement (Liking for Learning)	6.21%
Affective Engagement (Liking for School)	4.73%
Behavioral Engagement (Extracurricular Activities)	3.88%

Note. Cumulative percent of variance explained = 61.45%

Figure 2. *Scree Plot of Eigenvalues for the TERF-N*



Note. Factors are on the x-axis and eigenvalues are on the y-axis. Each factor's eigenvalue is displayed in the chart (e.g., factor 1 had an eigenvalue of 5.271).

Table 5 *Factor Correlations for the SESQ-ENG*

	AFF: Lkg Lrg	AFF: Lkg Sch	BEH: Eff&P	BEH: ECA
AFF: Lkg Lrg				
AFF: Lkg Sch	.51**			
BEH: Eff&P	.66**	.41**		
BEH: ECA	.38**	.37**	.42**	
COG	.44**	.38**	.47**	.45**

Note. Correlation is significant at the $p < .01$ level.

Teacher Engagement Report Form - New (TERF-N). The scree plot with eigenvalues is displayed in Figure 2. Utilizing Kaiser's criterion, a three-factor model would be optimal; while Cattell's scree test would indicate a one-factor model. The fit indices (Table 6) provide divergent indications of the best-fitting model, although it seems the four-factor model is the best fit. An ultra-Heywood case[3] was evident in the four-factor model rendering this model un-interpretable. The five-, three-, and two-factor models resulted in several undetermined factors and complex structure. Additionally, in each of these models, most of the items loaded on one factor. Therefore, the one-factor model is interpreted.

Table 7 lists the factor loadings for the factor (titled General Engagement). With the exception of two items, all items significantly loaded on the factor. The percentage of variance explained by the factor was 52.88%. The correlations for the TERF-N items range from trivial ($r = -.07$) to large ($r = .87$). With the exception of item 6, the majority of correlations are significant. The average inter-item correlation is moderate ($r = .43$).

Correlations between the SESQ and the TERF-N

After establishing internal consistency and structure of the measures, establishing that these measures are correlated is important. Bivariate correlations between the SESQ and the TERF-N were conducted. This analysis demonstrated positive, significant, moderate correlations between the TERF-N (total score) and the SESQ for the Affective ($r = .33$), Behavioral ($r = .35$), and Total ($r = .33$) engagement composites, each statistically significant at $p < .01$. The Cognitive engagement composite correlation was not significant ($r = .06$).

Table 6 *Goodness-of-Fit Statistics for the TERF-N EFA*

Model	χ^2	df	CFI	RMSEA	SRMR
1	169.58***	35	.85	.17 (.15-.20)	.09
2	96.28***	26	.92	.15 (.11-.18)	.04
3	59.65***	18	.95[a]	.13 (.10-.17)	.03
4	19.16[a]	11	.992[a]	.08 (.00-.13)	.02
5	7.70[a]	5	.997[a]	.07 (.00-.15)	.01[a]

Note. DF=Degrees of Freedom. CFI = Comparative Fit Index. RMSEA = Root Mean Square of Error of Approximation. SRMR = Standard Root Mean-Squared Residual. AIC = Aikake Information Criterion. BIC = Bayesian Information Criterion.
*** $p < .001$.
[a] Model meets standards for that index established in the literature.

[3]When parameter estimates have out-of-range values, they are referred to as Heywood cases. In the case of the TERF-N, the communality parameter estimate was more than one, which is referred to as an ultra-Heywood case and, which renders the solution invalid and un-interpretable (Brown, 2006).

Table 7 *Factor Loadings for the TERF-N Items*

ITEM & HYPOTHESIZED DOMAIN	General Engagement λ	Communality
AFFECTIVE		
Seems interested in school	.89	.92
Gets along with peers	.56	.51
Seems to care about grades	.94	.90
BEHAVIORAL		
Has good attendance	.50	.97
Participates in class discussions/activities	.74	.70
Is referred for out-of-class disciplinary procedures	-.29	.54
Is respectful to staff	.60	.62
COGNITIVE		
Persists on more challenging tasks	.23	.87
Demonstrates appropriate effort for task	.91	.89
Is self-motivated	.92	.91

Note. The average communality for the TERF-N = .60.

DISCUSSION

The primary purpose of this study was to begin to examine the psychometric properties of the Student Engagement in Schools Questionnaire (SESQ) and the Teacher Engagement Report Form (TERF-N). Analyses included; a) internal consistency estimates to examine reliability, b) exploratory factor analysis to examine the latent factor structure of each measure, and c) correlations to examine the relationship of the scales and begin to examine external validity.

Internal Consistency

Both measures yielded good internal consistency ($\alpha \geq .70$) with the exception of the Attributions domain ($\alpha = .65$) of the SESQ. Examination of the Attributions domain indicates that it may be problematic due to the structure of the questions, in particular the response format appeared to create confusion (as evidenced by questions from students during the administration and also by the number of responses that were not plausible). Based on careful review of responses to this item, it is recommended that this portion of the questionnaire be eliminated or altered to avoid this challenge.

Factor Structure

The factor structure of the SESQ Engagement items (i.e., the items measuring the affective, behavioral, and cognitive indicators of engagement) retained its conceptualized structure through exploratory factor analysis. As hypothesized, a five-factor model consisting of Affective (i.e., Liking for Learning and Liking for School), Behavioral (i.e., Effort & Persistence, Extracurricular Activities), and Cognitive Engagement factors was demonstrated. Two items (a1 & a9R) from the Affective: Liking for Learning factor cross-loaded on the Behavioral: Effort & Persistence factor. While these items loaded more strongly on the conceptualized factors, they demonstrated the weakest loadings on the factor. These items were also among the lowest communalities of the factor. Next steps for the SESQ will include reducing the length of the survey. These two cross-loading items will likely be candidates for elimination.

The EFA for the TERF-N produced interesting results. While a three-factor structure was hypoth-

esized, a one-factor model of General Engagement was the most viable, interpretable model. It may be that a student's levels of engagement are viewed through less differentiated eyes by a teacher (i.e., teachers see students who are engaged and those who are not rather than students who are affectively engaged and not behaviorally engaged) in comparison to students self-reporting (i.e., students may feel more variation than is actively displayed). Another consideration is that the teacher relies on easily observable characteristics and creates an overall, rather stable impression of a student, while the student relies on internal characteristics and is able to see more variation in his attitudes and beliefs over time and situation. A final explanation regarding the unidimensionality of the TERF-N may be related to the size of the questionnaire (i.e., 10 items).

Two items did not load significantly on the General Engagement factor (i.e., "Is referred for out-of-class disciplinary procedures" and "Persists on more challenging tasks"). These items also had very low communalities. While these items did not load as expected on the factor, it is believed by the researchers that these items provide important information to the evaluation of student engagement. According to the conceptualization of the construct, out-of-class discipline and persistence in academic tasks are considered related to a student's level of engagement at school. Additionally, the sample size of the TERF-N study was relatively small. It will be important to further evaluate the TERF-N on larger sample sizes before making any decisions about eliminating or changing items.

External Validity

Finally, moderate, significant correlations between the SESQ Affective and Behavioral Engagement composites and the overall TERF-N scores were evident. These correlations imply a relationship between how teachers and students view engagement; however, the moderate size of the correlations suggests that important information is gathered from both informants. The Cognitive Engagement composite score was not statistically significant and close to zero, implying no relationship between the two. One possible reason for this might be that cognitive engagement is an internal characteristic of an individual. Behavioral engagement is typically easily observable by the classroom teacher in a way that cognitive engagement is not. These results suggest that information regarding a student's cognitive engagement in school is best recognized by the individual themselves. Following further exploration, it may be that the cognitive dimension of the TERF-N is not feasible.

Limitations and Future Directions

The interpretation of the results of this exploratory study warrant caution, related to the limitations of the study. In particular, the small size of the TERF-N sample and the geographically limited sample of both analyses are notable limitations. Future studies can be conducted to further examine the results of the current study by utilizing a larger sample size, including more classroom teachers, with a variety of grade levels. Second, the length of the SESQ is problematic. The self-report form, consisting of 109 items, may cause fatigue in the students. Future efforts will include creating a psychometrically sound brief version of the SESQ. Additionally, a computer-based version of both forms would be beneficial. With a computer-based questionnaire students and teachers could complete the form outside of class time and over more than one session if necessary. A computer-based questionnaire would also increase accessibility and widen the population from which to sample.

Additionally, researchers should focus on gathering more reliability and validity data. Information regarding outcomes (e.g., grades, attendance, dropout) should be collected in order to demonstrate correlations between the SESQ and TERF-N and these outcomes. Ideally, longitudinal data will be gathered to demonstrate the importance of these measures. Also, measures accessing both similar and different constructs (e.g., Psychological Sense of School Membership, Goodenow, 1993; Children's Depression Inventory; Kovacs, 1992) should be utilized in order to confirm convergent and divergent validity. Internal consistency and confirmatory factor analyses should continue to be investigated in order to provide statistical bases for the measures. Future studies should gather gender and ethnicity information in more reliable ways (e.g., from student records), and analyze gender and ethnicity differences in student engagement.

Finally, the goal of assessment is to drive interventions. The SESQ and TERF-N will provide important information to the understanding of a student's experience at school. Once measures are created and tested, the next logical step is to use the information from those measures to inform interventions; future research should focus on creating and evaluating school-based engagement programs for prevention and intervention. There are many opportunities for research in student engagement in the schools.

CONCLUSIONS

Student engagement is a complex construct that continues to be important in promoting positive outcomes for students. Challenges remain in the conceptualization and measurement of this construct. One question that persists is: Should engagement be measured by its apparent components or should all types of engagement be lumped together as one overarching construct? Researchers should focus on the areas of examining and measuring engagement in order to further refine measures (e.g., short forms), and create interventions directly related to an individual student's needs. Further efforts are warranted in the development of a more comprehensive perspective of assessing student engagement, for example, using self-, teacher-, and parent-reports. Information from the multiple sources may provide a better understanding of students. Additionally, student engagement data collected on the large scale can provide administrators with additional school climate information, and may direct interventions at the universal, school-wide level.

Ongoing efforts related to the conceptualization and measurement of student engagement also need to seek out helpful information about how the construct relates directly to positive student outcomes. The current models of student engagement assume that there are multiple components, thus, outcomes should be present and measureable in, thus, defined components.

It is important for school psychologists to be aware of the literature and the ongoing research efforts in the area of student engagement in the schools. Interventions suggested by researchers (e.g., Reschly et al., 2007) can be targeted for specific students who are at-risk, or used for prevention efforts at the school-wide level. School psychologists can provide a context that is consultation- and collaboration-friendly; they can use their knowledge to provide psychoeducation to teachers on the importance of engagement in the classroom, in addition to strategies to enhance student engagement in school.

This study contributes to continued efforts by school staff, school psychologists, and educational researchers to investigate student engagement. Through collaboration among practitioners and academics, student engagement interventions may be part of the key to promoting school completion and academic outcomes. It is hoped that further work in this area will result in better measures of student engagement in school and increased positive outcomes for all students.

- - -

Shelley R. Hart, PhD., is a recent graduate of the University of California, Santa Barbara.

Kaitlyn Stewart, M.Ed., is a recent graduate of the University of California, Santa Barbara.

Shane R. Jimerson, PhD., is a professor at the University of California, Santa Barbara.

REFERENCES

Appleton, J.J., Christenson, S.L., & Furlong, M.J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45*, 369-386. doi: 10.1002/pits.20303

Appleton, J.J., Christenson, S.L, Kim, D., & Reschly, A.L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology, 44*, 427-445. doi: 10.1016/j.jsp.2006.04.002

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Carter, M., McGee, R., Taylor, B., & Williams, S. (2007). Health outcomes in adolescence: Associations with family, friends and school engagement. *Journal of Adolescence, 30*, 51-62. doi: 10.1016/j.adolescence.2005.04.002

Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Christenson, S.L., Sinclair, M.F., Lehr, C.A., & Godber, Y. (2001). Promoting successful school completion: Critical conceptual and methodological guidelines. *School Psychology Quarterly, 16*, 468-484. doi: 10.1521/scpq.16.4.468.19898

Costello, A.B., & Osborne, J.W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9.
Cronbach, L.J., & Shavelson, R.J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement, 64*, 391-418. doi: 10.1177/0013164404266386
Csikszentmihalyi, M., (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.
Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299. doi: 10.1037/1082-989X.4.3.272
Finn, J.D. (1989). Withdrawing from school. *Review of Educational Research, 59*, 117-142.
Fredricks, J.A., Blumenfeld, P.C., & Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research 74*(1), 59-109. doi: 10.3102/00346543074001059
Furr, R.M., & Bacarach, V.R. (2008). *Psychometrics: An introduction*. Thousand Oaks, CA: Sage.
Furlong, M.J., Whipple, A.D., St. Jean, G., Simental, J., Soliz, A., & Punthuna, S. (2003). Multiple contexts of school engagement: Moving toward a unifying framework for educational research and practice. *The California School Psychologist, 8*, 99-114.
Gliner, J.A., & Morgan, G.A. (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Mahweh, NJ: Lawrence Erlbaum & Associates.
Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in the Schools, 30*, 79-90. doi: 10.1002/1520-6807(199301)30:1<79::AID-PITS231030011 3>3.0.CO;2-X
Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. doi: 10.1080/10705519909540118
Jimerson, S.R., Campos, E., & Greif, J.L. (2003). Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist, 8*, 7-27.
Jimerson, S.R., Renshaw, T.L., Stewart, K., Hart, S., & O'Malley, M. (2009). Promoting school completion through understanding school failure: A multi-factorial model of dropping out as a developmental process. *Romanian Journal of School Psychology, 2*, 12-29.
Kahn, J.H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *The Counseling Psychologist, 34*, 684-718. doi: 10.1177/0011000006286347
Kline, P. (2002). *An easy guide to factor analysis*. New York: Routledge.
Kovacs, M. (1992). *The Children's Depression Inventory*. New York: Multi-Health Systems.
Lam, S.F., & Jimerson, S.R. (2008). *Exploring student engagement in schools internationally: Consultation paper*. Chicago, IL: International School Psychologist Association.
Muthén, L.K., & Muthén, B.O. (2009). *Mplus Statistical Analysis with Latent Variables: User's Guide* (5th Ed.). Los Angeles, CA: Author.
Patrick, H., Ryan, A.M., & Kaplan, A. (2007). Early adolescents' percpetions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology, 99*, 83-98. doi: 10.1037/0022-0663.99.1.83
Reschly, A.L., Appleton, J.J., & Christenson, S.L. (2007). Student engagement at school and with learning: Theory and interventions. *NASP Communiqué, 35*(8).
Reschly, A.L., & Christenson, S.L. (2006). Prediction of dropout among students with mild disabilities: A case for inclusion of student engagement variables. *Remedial and Special Education, 27*, 276-292. doi: 10.1177/07419325060270050301
Shernoff, D.J., & Schmidt, J.A. (2008). Further evidence of an engagement-achievement paradox among U.S. high school students. *Journal of Youth and Adolescence, 37*(5), 564-580. doi: 10.1007/s10964-007-9241-z
SPSS, Inc. (2007). *SPSS Base 16.0 for Mac User's Guide*. Chicago, IL: Author.
Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics*. Boston, MA: Allyn & Bacon.
Thurstone, L.L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
Walker, C.O., Greene, B. A., & Mansell, R.A. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learning and Individual Differences, 16*, 1-12. doi: 10.1016/j.lindif.2005.06.004
Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. Albany, NY: State University of New York Press.

TOC

The Relative Effects of Chronological Age on Hispanic Students' School Readiness and Grade 2 Academic Achievement

Michael Furlong,
Matthew Quirk,
Counseling, Clinical, and School Psychology Department,
University of California, Santa Barbara

This study examined the relations of age, preschool experience, and gender with children's school readiness levels at kindergarten entry. The sample included 5,512 children of predominantly Hispanic heritage and from families experiencing low socioeconomic circumstances. A series of between-subjects ANOVAs indicated that age (Eta^2 .019 to .043), preschool experience (Eta^2 .104 to .204), and gender (Eta^2 .015 to .022) were significantly related to children's school readiness as measured by the Kindergarten Student Entrance Profile (KSEP). Logistic regression examined the unique contribution of these variables to predict students' academic achievement at the end of Grade 2 with a subsample of 980 students. The strongest achievement predictor was school readiness – the odds of students rated in the top 25% on the KSEP having proficient or advanced scores on the English Language Arts portion of the California Standards Test at the end of Grade 2 were 4.51 greater than the odds of students rated in the bottom 75% on the KSEP having proficient or advanced scores. Similar results were found for students' achievement in mathematics. The key findings of this study showed that formal preschool experiences play an important role in preparing children of Hispanic descent and who live in households experiencing low income for kindergarten entry. However, children's readiness at entry into kindergarten was more strongly related to later academic achievement than age at kindergarten entry and preschool experience.

KEYWORDS: school readiness, Kindergarten Student Entrance Profile, age of school entry, academic achievement, Hispanic students, gender

As a child nears five years of age, parents often consider the skills he or she needs to be successful in school. Similarly, schools may use developmental screening assessments to evaluate if entering kindergarteners are prepared to be responsive learners (Saluja, Scott-Little, & Clifford, 2000). In response to these concerns, the California legislature passed a law moving the cutoff date for entry into kindergarten from December 2 to September 1 of the year in which the child turns five years old (Kindergarten Entry Age, CA SB. 1381, 2010). This change in policy was based on the belief that children entering kindergarten before five years of age are too young and that beginning school at an older age would benefit their academic and social development. In most U.S. states, children must be 5 years old by August 31-September 1 to be eligible to enroll in kindergarten, and in some states, the entry date is June 1 (Stipek, 2002). California's policy change is reflective of a number of states changing the school entry date and requiring younger children to wait an additional year before entering kindergarten (Stipek, 2002).

Despite the laudable aims of this educational policy, which is backed by some research evidence (Bisanz, Dunn, & Morrison, 1995; Diamond, 1983; Langer, Kalk, & Searls, 1984; Stipek, 2003; Uphoff & Gilmore, 1985), there is still debate regarding the long-term effects of later entry into kindergarten,

Address correspondence to Michael Furlong, University of California, Santa Barbara, Gevirtz Graduate School of Education, Counseling, Clinical, and School Psychology Department, Santa Barbara, CA 93106. E-mail: mfurlong@education.ucsb.edu

Partial funding for this article was provided by the Santa Barbara County First 5 Commission. The authors thank Karin Dominquez and Mona Anderson for their support in the completion of this study.

particularly for students who come from disadvantaged socioeconomic backgrounds (Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007).

Research over the past decade has demonstrated that children's school readiness is related to a variety of positive school-related outcomes (Duncan et al., 2007; Ladd, Herald, & Kochel, 2006; Le, Kirby, Barney, Setodji, & Gershwin, 2006; Quirk, Furlong, Lilles, Felix, & Chin, 2011). However, few studies have specifically examined the combinatorial influences of chronological age, gender, and preschool experience on children's readiness at entry into kindergarten (Gullo & Burton, 1992), and none have examined these issues with at-risk student populations. Thus, the current study examined the unique and interactive effects of age, preschool experience, and gender on children's school readiness at kindergarten entry with a sample of over 5,500 Hispanic children from a predominantly low income, agricultural community in central California. In addition, longitudinal data on a subsample of 980 students were used to explore the relations between these factors and students' performance on high stakes mathematics and English assessments at the end of Grade 2.

Age and School Readiness

Current policies and practices regarding kindergarten eligibility are primarily derived from a maturationist perspective (Snow, 2006). From this perspective, children are viewed as ready for kindergarten entry when they have achieved a certain level of physical and psychological maturity. Therefore, it is often assumed that, in terms of school readiness, older is better (Meisels, 1996; Shepard & Smith, 1986).

Concerns regarding the readiness of children entering kindergarten are not limited to school administrators and policymakers. An increasing number of parents whose children born at or near existing kindergarten cutoff dates are deciding to hold their children back from kindergarten entry (NCES, 1997; Stipek, 2002), a process sometimes referred to as *redshirting* (see Frey, 2005). In addition, teachers often cited age as a significant factor when describing children who struggle in kindergarten (Heaviside & Farris, 1993; NCES, 1993) and reported that they consider age when deciding if a child who is held back in kindergarten will fit well in the subsequent cohort of students (Shepard & Smith, 1986).

Taken together, public sentiment appears to support the notion that older children are more ready for schooling than younger children. Embedded within this assumption is the belief that relatively older children will benefit more from kindergarten instruction than younger children; however, the evidence from research examining this assertion is mixed. Some research evidence indicates that children who enroll in kindergarten at an older age tend to outperform children who enter kindergarten at a younger age on standardized tests during the elementary school years (Carter, 1956; Cascio & Schanzenbach, 2007; Crosser, 1991; Dickinson & Larson, 1963). In addition, Gullo and Barton (1992) found that age, years of preschool experience, and the interaction between age and years of preschool all contributed significantly to children's school readiness at the end of kindergarten. However, they also found that when preschool experience was controlled, the effects of age were not nearly as strong, indicating that age is not as influential for children who have had extensive preschool experiences. Other researchers have found that younger kindergarteners made progress similar to their older peers during the kindergarten year (Loeb et al., 2007; Mayer & Knutson, 1999; Morrison et al., 1997) and that any short-term deficits observed during the early school years tended to dissolve by the end of elementary school (Stipek & Byler, 2001).

Preschool Experience and School Readiness

In contrast to the maturationist model, the constructivist model is based on readiness emerging from children's interactions with skilled peers and adults. The constructivist model is used to support efforts to provide high-quality preschool experiences (Andrews & Slate, 2001). Numerous studies have examined the impact of preschool experiences on various aspects of children's school readiness. For example, beginning childcare at an earlier age has been shown to have a significant positive effect on readiness development of many children (Gullo & Burton, 1992; Howes, 1988). In addition, the positive impact of preschool on academic and social development is particularly strong for children of minority or low-income backgrounds (Connell & Prinz, 2002; Lee, Brooks-Gunn, Schnur, & Liaw, 1991; Loeb et

al., 2007; Turney & Kao, 2009). However, a large scale study by the National Institute of Child Health and Development (NICHD) was conducted to examine the effects of childcare on a number of different developmental outcomes and found no relation between hours of childcare enrollment and cognitive or language skills (NICHD Early Child Care Research Network, 2000). Most of these studies did not examine these relations with large numbers of Hispanic children, or families where English was not the primary language spoken. Hence, more research is needed to better understand the influence of preschool experiences for these populations of children.

Study Purpose

The primary purpose of this study was to examine the relation between chronological age at enrollment into kindergarten with later academic advantages. We tested the hypothesis that age at kindergarten enrollment is positively associated with school readiness and later school outcomes by examining the following research questions:

1. Is age at entry to kindergarten related to children's school readiness? To answer this question, we compared the school readiness ratings of children by age and hypothesize that those children born in the winter and spring months (December 2 to April 30) would have significantly higher school readiness ratings than younger children born in the summer and fall months (May to December 1). In addition, we examined the association of age at kindergarten entry with gender and preschool experience, variables that others studies have indicated are associated with an adaptive transition into formal education.
2. Do those students who are held out of kindergarten for a year, even when they are eligible by age to enroll, have higher school readiness ratings than typical kindergartners when enrolled? For this analysis, we examined the school readiness ratings of students who were "redshirted" and compared them to the ratings of other children.
3. School readiness is just one indicator of a child's possible early educational trajectory; hence, a related important question examined in this study was: Does age at kindergarten entry predict academic achievement at the end of Grade 2? And, if it does, is the effect size larger than those of gender, preschool experience, English language skills, and a measure of school readiness taken at entry to kindergarten?

METHOD

Participants

The students in this study were all enrolled in a K-8 school district located in a community with a population of about 100,000, with a distance of more than 60 miles to another similar-sized population area. District enrollment was more than 12,500. The average class size was 23.2 in kindergarten, 20.8 in Grade 1, and 20.0 in Grade 2. In 2007-2008, 31% of the district's students obtained scores of "proficient" or "advanced" on the California Standards Test (CST) English-Language Arts assessment compared to 46% of similar-aged students throughout California. On the CST mathematics standards assessment, the district's students had scores similar to the statewide averages for similar-aged students (44% versus 43% proficient or advanced, respectively). School enrollment questionnaires completed by parents during kindergarten enrollment indicated that a significant proportion of the children were from families experiencing low socioeconomic circumstances, with 79.2% of the entering kindergarten students receiving free or reduced-price lunch services.

Beginning in 2004, all kindergarten students were assessed with the Kindergarten Student Entrance Profile (KSEP), a universal school readiness instrument administered as a standard school practice in the district. The data for this study included all students with school readiness ratings who entered kindergarten in the fall of 2005, 2006, 2007, and 2009. Table 1 shows the characteristics of each cohort. Across all four cohorts, school readiness ratings were available for 5,512 children. These participating students were exactly 50% boys and 50% girls. The majority of these children were of Hispanic heritage (89.2%), followed by children of White (nonHispanic) (4.1%) and Filipino (1.9%) ethnicities. All remaining ethnicities were less than 1.0% of the sample. As an indicator of language background, the

kindergarten teachers recorded the language the children were observed to use as part of the school readiness assessment. Most of the children used English (46.1%), followed by a combination of English and Spanish (36.7%), and Spanish only (17.2%). Data for one or more of the variables used in the analysis were missing for 1.9% of these cases, which were not included in the analysis.

In addition, the 2006 cohort included 980 children who were rated with the KSEP at Kindergarten entry and for whom Grade 2 CST scores were available. This subgroup was comprised of 47.9% boys and 52.1% girls; 91.6% of whom were of Hispanic heritage. Data for this subgroup of students were used to examine the relations between school readiness at kindergarten entry and later academic achievement.

Measures

School readiness. The Kindergarten Student Entrance Profile (KSEP; Santa Maria–Bonita School District, First 5 of Santa Barbara County, and University of California Santa Barbara, 2005) is a universal screening measure of students' school readiness. The KSEP is an observational rating scale completed by teachers after they observe a child for three weeks in the natural classroom environment. All district kindergarten teachers completed a two-hour training session focused on procedures for administering and scoring the KSEP. This training explained the KSEP rubric, which provides operational definitions for each KSEP item (see Lilles et al. [2009] and Quirk et al. [2011] for detailed information about the KSEP).

The KSEP includes 16 items linked to domains of social-emotional, physical, and cognitive elements of children's school readiness. Each item is rated on a 4-point scale: 1 = *not yet,* 2 = *emerging*, 3 = *almost mastered*, and 4 = *mastered*. A rating rubric for each individual item identifies type of behaviors associated with varying levels of readiness. The rubric for the KSEP item assessing impulse control, for example, has the following descriptions: (1) *not yet – unable to delay having wants and needs met;* (2) *emerging – distracted by getting wants and needs met, yet able to be redirected by others;* (3) *almost mastered – distracted by getting wants and needs met but redirects self;* and (4) *mastered – able to delay wants and needs until appropriate time*. The 16 items are combined into a total composite score (range 16-64); with a score of 64 indicating that the student demonstrated mastery on all items. The KSEP does not assess language proficiency; therefore, children can demonstrate mastery using any language or mode of communication. For this study, the internal consistency (Cronbach's alpha) of the 16-item KSEP ratings for the four cohorts were .91, .92, .92, and .93, respectively, with total scores ranging from 16-64 (M = 48.71, SD = 9.8, across all 4 cohorts; Table 1 shows means for each cohort.). KSEP ratings have been shown to predict Grade 1 and 2 curriculum-based reading fluency probes (Quirk et al., 2011). A copy of the KSEP and related information is available online (http://web.me.com/michaelfurlong/KSEP).

Table 1 *Estimated Marginal Mean Kindergarten Student Entrance Profile Ratings Across Four Cohorts of Kindergartners by Gender, Season of Child's Birth, and Child's Preschool Experience*

	Kindergarten Cohort											
	August 2005			August 2006			August 2007			August 2009		
Study variables	*N*	*M*	*SE*	*N*	*M*	*SE*	*N*	*M*	*SE*	*N*	*M*	*SE*
Gender												
Male	519	50.08	0.47	637	47.79	0.42	708	48.15	0.40	838	48.16	0.31
Female	509	52.64	0.45	680	51.31	0.39	697	51.01	0.38	821	50.84	0.32
Age at Grade K entry [a]												
Winter (oldest)	253	53.06	0.69	351	51.86	0.58	362	51.97	0.54	395	51.89	0.44
Spring	259	52.23	0.65	335	50.76	0.56	337	51.16	0.57	410	50.42	0.45
Summer	278	50.95	0.62	313	49.08	0.58	354	48.73	0.54	422	48.84	0.44
Fall (youngest)	238	42.20	0.65	318	46.50	0.58	352	46.46	0.56	432	46.85	0.44
Preschool experience												
Head Start	151	56.94	0.70	201	52.93	0.63	227	54.45	0.60	281	53.71	0.51
State preschool	183	53.50	0.63	278	51.08	0.53	262	51.36	0.55	397	53.39	0.43
Summer camp	108	48.49	0.83	154	48.67	0.73	193	46.57	0.67	325	46.75	0.48
No known preschool	586	46.11	0.35	684	45.51	0.34	723	45.93	0.33	656	44.15	0.33
Mean KSEP Score	1,028	49.32	9.49 [b]	1,317	48.31	9.65 [b]	1,405	48.34	9.90 [b]	1,659	48.43	9.80 [b]

Note. See Table 2 for analysis of variance results. [a] Denotes time of year when child was born: winter (December, January, February), spring (March, April, May), summer (June, July, August), and fall (September, October, November). [b] = standard deviation

California Standards Test (CST). The CST is used to monitor student academic progress from Grades 2-12. The Grade 2 English-Language Arts (E-LA) portion of the exam contains 65 items matched to the California's curriculum blueprint by multiple independent item review teams covering the areas of word analysis, reading comprehension, literary response and analysis, writing strategies, and written conventions. The Grade 2 mathematics portion of the exam has 64 items covering the content areas of number sense, algebra functions, measurement and geometry, statistics, data analysis, and probability.

The internal consistency coefficient for both E-LA and mathematics in the 2008 administration was .93 for Hispanic second graders (California Department of Education Standards an Assessment Division, 2009). Both the E-LA and mathematics scores produce a standard score that is then expressed as one of five general performance levels: far below basic, below basic, basic, proficient, and advanced. The desired level of achievement is proficient or advanced. For the 2008 statewide administration, it was estimated that 92% of the Grade 2 students who scored as proficient or advanced in E-LA and mathematics were correctly classified (California Department of Education Standards an Assessment Division, 2009).

Procedures

The district school readiness kindergarten transition coordinator trained all teachers on the use of the KSEP prior to the beginning of each school year. KSEP ratings were recorded for each student on a standard form and scores were recorded in an Excel database with an associated student identifier (state assigned identification number). All data were collected by the district as part of general education practices and shared with researchers as part of a collaborative effort to better understand the psychometric properties of the district-developed KSEP and to better determine the readiness of students at school entry. California education law encourages school districts and universities to collaborate on data sharing for research purposes. The database used for this study was stripped of any unique student identifiers, per the requirements of the university's institutional review board.

Analysis Plan

To address the first research question, the associations between kindergarten readiness and gender, age, and preschool experience and their possible interaction, four separate 2 (boy vs. girl) x 4 (winter, spring, summer, fall birth) x 4 (Head Start, state preschool, summer orientation camp, and no known preschool) between-subjects analyses of variance (ANOVA) were conducted. Tukey's post-hoc comparisons were used to examine group differences for significant main and interaction effects. Using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) for the design and analyses used in this study (setting the apriori effect size to .25, *p*-level to .05, power to .90, and for 32 total subgroups), a sample of 400 was needed for each of the four cohorts included in the analysis; hence, the sample size for these analyses was sufficient. Each of these four ANOVAs is an independent replication of the analysis with a unique sample. In addition, to synthesize the results across all four cohorts, we combined the cohorts and ran another 4 (age: season of birth) x 4 (preschool experience) ANOVA with KSEP school readiness ratings as the dependent variable. Because this was a second analysis using the same samples, we adjusted the *p*-level to .025 for all ANOVAs.

Another planned analysis examined the relative contributions of gender, age at Grade K entry (season of birth), preschool experience, KSEP language of response, and kindergarten readiness to predict Grade 2 academic achievement. For this analysis, we used logistic regression. The predictor variables were all recoded to be binary in the following fashion: gender (0 = male, 1 = female); age at Grade K entry (season of birth: 0 = summer/fall birth, 1 = winter/spring birth); preschool experience (0 = no known preschool/summer orientation camp, 1 = Head Start/state preschool); KSEP language of response (0 = Spanish/English–Spanish, 1 = English), and school readiness (0 = KSEP percentile rank of 1-74, 1 = KSEP percentile rating of 75-99). The dependent variables for these two analyses were the Grade 2 CST English-Language Arts and Mathematics scores (0 = far below basic/below basic/basic, 1 = proficient/advanced). Using G*Power 3 (Faul et al., 2007) for this design and analysis (setting the apriori effect size to .10, *p*-level to .05, power to .90) with 4 predictors variables, a sample of 159 was needed; hence the sample size for these analyses was sufficient.

RESULTS

Variables Associated with School Readiness

Tables 1 (means and standard deviations) and 2 (ANOVA statistics) show the results of the 2 (gender: boy vs. girl) x 4 (season of birth: winter, spring, summer, fall) x 4 (preschool: Head Start, state preschool, summer orientation camp, and no known preschool) between-subjects analysis of variance for each of the four cohorts. The same pattern of results was replicated across all four cohorts. For each cohort, the main effects for gender (small effect size = Eta^2 .015 to .022), season of birth (small effect size = Eta^2 .019 to .043), and preschool experience (medium to large effect size = Eta^2 .104 to .204) were all significant. All two- and the three-way interaction terms were nonsignificant.

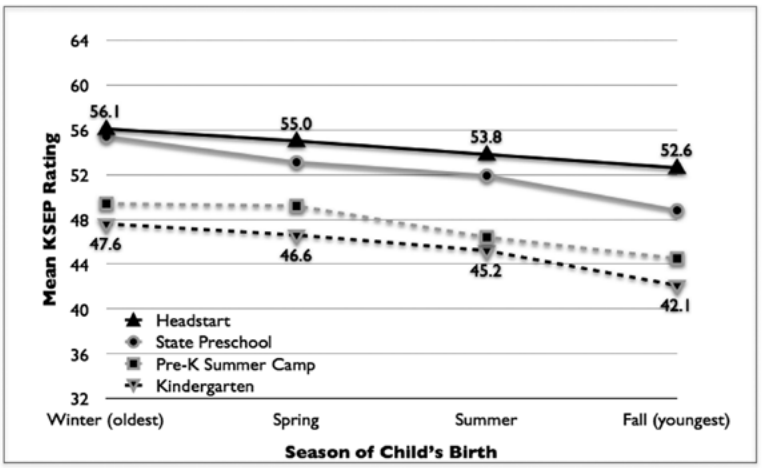
Table 2 *Univariate Analysis of Variance Results for the Kindergarten Student Entrance Profile Ratings Across Four Cohorts of Kindergartners by Gender, Season of Child's Birth (Chronological Age), and Child's Preschool Experience*

	Kindergarten Entry															
	August 2005				August 2006				August 2007				August 2009			
Study variables	*F*	*df*	*p*	*Eta2*	*F*	*df*	*p*	*Eta2*	*F*	*df*	*p*	*Eta2*	*F*	*df*	*p*	*Eta2*
Gender	15.51	1, 996	< .001	.015	37.35	1, 1285	< .001	.028	26.91	1, 1373	< .001	.019	36.71	1, 1627	< .001	.022
Age at Grade K entry [a]	6.45	3, 996	< .001	.019	16.29	3, 1285	< .001	.037	20.31	3, 1373	< .001	.043	23.93	3, 1627	< .001	.042
Preschool experience	82.36	3,996	< .001	.199	49.67	3, 1285	< .001	.104	65.14	3, 1373	< .001	.125	139.40	3, 1627	< .001	.204

Note. See Table 1 for study variable mean values. [a] Denotes month child was born: winter (December, January, February), spring (March, April, May), summer (June, July, August), and fall (September, October, November).

To summarize these results, another 4 (season of birth: winter, spring, summer, fall) x 4 (preschool: Head Start, state preschool, summer orientation camp, and no known preschool) ANOVA was conducted combining all four cohorts (N = 5,409). Significant main effects were found for season of birth (age), F = 68.46 (3, 5393), p < .000, eta^2 = .037 (small effect size), and preschool experience, F = 311.65 (3, 5393), p < .000, eta^2 = .148 (large effect size), with a nonsignificant interaction term, F = 1.37 (9, 5393), p = .194, eta^2 = .002. Post-hoc comparisons using Tukey's method found the following pattern of mean KSEP rating differences (see Figure 1) for season of birth: winter > spring > summer > fall. Post-hoc comparisons for preschool experience (see Figure 1) found this pattern: Head Start > state preschool > preK summer camp > kindergarten (no known formal preschool experience).

Figure 1. *Mean kindergarten Student Entrance Profile ratings of kindergartners by preschool experience and age at K entry*



Students Held Back (Redshirted) for One Year

As an additional analysis, we examined each cohort and identified all students whose birth date indicated that they entered kindergarten one year later than their same-age peers (these students were not included in the previous analyses). There were 108 students (1.5% of all students in the sample) whose enrollment in kindergarten was held back one year. This group was comprised of 55% boys, 45% girls, and 92% of them had no known formal preschool experience. The mean KSEP rating was 48.30 (*SD* = 10.22) compared to a mean rating of 48.55 (*SD* = 9.73) for all students who entered school at a typical age, *t* = 0.25 (*N* = 5506), *p* = n.s.. Only 27% of these children had KSEP ratings in the top 25% of all children, even though they were all older than the other students in their cohort at kindergarten entry.

Predicting Grade 2 Academic Achievement

Assessing school readiness is important for understanding the effects of kindergarten preschool articulation efforts; in addition, distal effects are also important, such as early scholastic achievement. To examine the relative effects of the variables examined in this study on academic achievement, two simultaneous logistic regressions were performed with the following binary predictor variables: season of birth, gender, KSEP, language of response, preschool experience, and KSEP rating. The dependent variables were Grade 2 California Standards Test Scores in English-Language Arts and Mathematics. Table 3 shows the model fit statistics for each logistic regression, which were acceptable.

Table 3 *Summary of Logistic Regression Analysis Model for Predicting Grade 2 California Standards Test Scores in English-Language Arts and Mathematics (0 = Far Below Basic, Below Basic, Basic; 1 = Proficient, Advanced)*

Predictor variables	*B*	*SE B*	Wald [a]	*OR (95% CI)*	*p*
CST English–Language Arts					
Age at kindergarten entry	0.12	.15	0.72	1.13 (0.85–1.51)	.397
Gender	0.35	.15	5.69	1.41 (1.06-1.88)	.017
KSEP language of response	0.51	.15	11.95	1.67 (1.25–2.23)	.001
Preschool experience	-0.20	.16	1.63	0.82 (0.60–1.16)	.202
KSEP rating	1.51	.17	76.50	4.51 (3.22–6.32)	< .0001
Constant	-2.42	.17	77.29		< .0001
Cox Snell R^2 — Nagelkerke R^2 = .114–.157; Model fit = $\chi 2$ (5, *N* = 979) = 118.12, *p* < .0001, goodness of fit test—Homer and Lameshow Test = $\chi 2$ (8, *N* = 979) = 7.09, *p* = .528; classification accuracy = 71.1%					
CST Mathematics					
Age at kindergarten entry	0.30	.14	4.94	1.35 (1.04–1.76)	.026
Gender	-0.07	.13	0.26	0.93 (0.72–1.21)	.608
KSEP language of response	0.21	.14	2.33	1.24 (0.94–1.62)	.127
Preschool experience	-0.33	.15	4.93	0.72 (0.54–0.96)	.723
KSEP rating	1.30	.18	55.03	3.67 (2.60–5.18)	< .0001
Constant	-0.43	.15	8.73		.003
Cox Snell R^2 — Nagelkerke R^2 = .077–.102; Model fit = $\chi 2$ (5, *N* = 977) = 77.99, *p* < .0001; goodness of fit test—Homer and Lameshow Test = $\chi 2$ (8, *N* = 977) = 4.01, *p* = .934; classification accuracy = 60.7%					

[a] degrees of freedom was 1 for each predictor variable.

The following results can be interpreted using Cohen's (1988) guidelines for odds ratio (OR) effect sizes: 1.49 = small, 3.45 = medium, and 9.00 = large. As shown in Table 3, proficient/advanced CST E-LA status was significantly predicted by gender (Wald chi-square = 5.69, OR = 1.41) and KSEP language of response (Wald chi-square = 1.95, OR = 1.67). The strongest predictor was KSEP rating (Wald chi-square = 76.50, OR = 4.51) – the odds of children with a 75-99 KSEP percentile rank being proficient/

advanced on the Grade 2 E-LA CST was 4.51 times greater than the odds of children with a 1-74 KSEP percentile rating having proficient/advanced status. Fifty-nine percent of the children with higher KSEP ratings had CST scores in the proficient/advanced range compared to 27% of children with lower KSEP scores. Age at kindergarten entry (Wald chi-square = 0.72, OR = 1.13) and preschool experience (Wald chi-square = 1.63, OR = 0.82) did not significantly predict E-LA CST status.

Grade 2 CST Mathematic status was also most strongly predicted by KSEP rating (Wald chi square = 55.03, OR = 3.67). This result indicates that the odds of children with KSEP ratings in the 75-99 percentile rank range being proficient or advanced on the Grade 2 E-LA CST was 3.67 times greater than the odds of children with KSEP ratings with 1-74 percentile ranks having proficient or advanced status. Sixty-nine percent of the children with higher KSEP ratings had CST mathematics scores in the proficient-advanced range compared to 43% of children with lower KSEP ratings. The only other variable to modestly predict Grade 2 CST Mathematics status was age at K entry (Wald chi square = 4.94, OR = 1.35). Gender, KSEP language response, and preschool experience were not significant predictors.

DISCUSSION

Overall, the results of this study indicated that age, and gender (significantly, but slightly) and preschool experience (significantly and moderately) were related to Hispanic children's school readiness upon entry into kindergarten. Specifically, children who had some form of preschool experience were rated significantly higher in terms of their school readiness than children with no preschool experience, which is consistent with previous research examining similar populations of children (Connell & Prinz, 2002; Lee et al., 1991; Loeb et al., 2007; Turney & Kao, 2009). Also, similar to previous research findings, older children were, on average, more ready at kindergarten entry than their younger peers (Meisels, 1996; Shepard & Smith, 1986) and girls were slightly more ready than boys. It should be noted that the youngest children in our sample with Head Start preschool experience were rated as more ready ($M = 52.6$) than the oldest children with no known preschool experience ($M = 47.6$). There was no significant interaction between age and preschool experience on school readiness, indicating that age did not moderate the effectiveness of preschool in preparing children for kindergarten entry. Thus, it appears that for Hispanic children from families experiencing low socioeconomic circumstances, preschool experience was more important to improving school readiness than chronological age alone.

Other research of the long-term achievement patterns of Hispanic students indicated that "…half of the achievement gap in fourth grade exists when students walk through the door in kindergarten" (Rumberger & Arellano, 2007, p. 71). The results of this study lend support to this assertion because school readiness at kindergarten entry was the most significant factor in predicting children's later academic success. The effects of age at kindergarten entry, gender, and preschool experience on later achievement were negligible, suggesting a critical transition at the time of kindergarten entry. In other words, it appears that preschool experience and age were important influences on whether these students were ready for school prior to entry into kindergarten; however, once these children entered the K-12 system, school readiness level was the primary factor that significantly predicted subsequent academic achievement.

Finally, our results do not support the practice of academic redshirting, or holding younger children back a year before enrolling them into kindergarten. This finding appears to contradict a study by Datar (2006) that indicated that delaying kindergarten entry for younger at-risk students significantly improved their test scores in both reading and math.

Issues Related to Age at School Entry

The findings of this study pertaining to the effects of age on educational "success" were based only on the student's Grade 2 CST scores. There is a need to extend the analysis to examine academic trajectories across elementary and into secondary school. Moreover, when assessing the influence of age at school entry, other outcomes can be legitimately considered. For example, Dhuey and Lipscomb (2008) found that older age at school entry is associated with higher rates of leadership in high school, which is associated with increased economic earnings. However, these relations are complex and not completely understood. Cascio and Schanzenbach (2007) reported that socioeconomic factors could interact with

age of school entry on distal outcomes such as leaving school early. These authors reported that youth from families experiencing low income who are older at school entry are less likely to take the Scholastic Aptitude Test. Dobkin and Ferreira (2010) also reported that younger children may have lower academic achievement, but as a group, they actually stayed in school longer than their older classmates. This is something to consider as California's kindergarten enrollment age moves to September 1. This will mean that children with September birthdates, for example, will be 18 years, 9 months old at their normal time for graduation from high school.

Issues Related to Effects of Preschool Experience

This study found that children attending preschool were more likely to have cognitive and social-emotional profiles indicative of being ready for school entry. Preschool accounted for about 15% of the variance in KSEP ratings, with age at school entry and gender accounting for substantially less variance. This finding is consistent with other studies showing that children attending center-based preschools, such as Head Start (all Head Start programs in this study were accredited by the National Association for the Education of Young Children), have significantly higher cognitive and social-emotional skills when entering kindergarten compared to a control sample of similar children (U.S. Department of Health and Human Services, Administration for Children and Families, 2010). Nonetheless, when it came to predicting Grade 2 CST scores, the positive effects of preschool experiences were no longer evident, a pattern that mirrors the findings of the Head Start Impact Study (U.S. Department of Health and Human Services, Administration for Children and Families, January 2010) and other studies (Magnuson, Ruhm, & Waldfogel, 2007). One interpretation of this finding is that preschool experiences on average help to improve the basic cognitive and social-emotional skill sets that children need to enter school ready to learn and to be responsive to high-quality instruction. However, children who have not had preschool experiences may nonetheless have had sufficient care and nurturing during the preschool years to have the requisite cognitive and social-emotional skills at school entry. The possibility of a child having nurturing childcare outside of a preschool setting indicates that the variance in later academic skills, such as at Grade 2 in this study, will be more strongly related to school readiness, and not preschool participation per se. What appears to be important is that children get ready for school, whether it is in a formal preschool or in another nurturing child rearing context.

Alternatively, it should be considered that the sample in the present study was predominately Hispanic and many of the youths were from families with low-income circumstances. These children typically have less access to center-based preschools (Turney & Kao, 2009), which diminishes their relative odds of being ready for school. Other factors need to be considered. For example, in a study using the Early Childhood Longitudinal Study dataset, Loeb et al. (2007) found that other variables should be considered for Hispanic children. In particular, length and intensity of the preschool experience was associated with higher reading and math achievement. We were unable to control for this in the present study, which could have accounted for the diminished effects of preschool on academic achievement.

STUDY LIMITATIONS AND NEED FOR FUTURE RESEARCH

The results of this study generalize only to Hispanic children in California and contribute to the literature by increasing knowledge about their early educational needs. Nonetheless, there is a need to use the KSEP in other schools and communities to further assess its predictive ability. Another limitation is that we were limited to variables that were archived by the district and were available for analysis; hence, we could not examine the influence of other factors on school readiness and later school achievement, such as parental involvement and child rearing practices. Although preschool experience was a robust predictor of school readiness, there was unexplained variance. Future research is needed to examine other key factors that influence the school readiness of Hispanic children.

CONCLUSION AND IMPLICATIONS FOR PRACTICE AND POLICY

An easy and practical step that school psychologists can take to help all children have the best opportunity to be ready for school at entry to kindergarten is to open lines of communication between their school district and their state and regional programs that serve preschool children and their families. Major coordinated early childhood development initiatives can be reviewed, for example, in California (First 5, http://www.first5california.com/default.asp), North Carolina (Smart Start, http://www.smartstart-nc.org/), and Colorado (http://smartstartcolorado.org/). In addition, many communities have Head Start programs that can be a focus of efforts to facilitate the school readiness of children with high needs. Among the ways that school psychologists can support efforts to promote school readiness are to: (a) become informed about their community's 0-5 programs and services; (b) discuss ways to articulate services and information from the preschool years through the transition to kindergarten (for example, opening a district cumulative folder for children in the last year of preschool); (c) disseminate information about how parents can support their children's preparation for school by accessing resources such as the Public Broadcasting Services' *A Place of Our Own* (http://aplaceofourown.org) and *Los Niños en Su Casa* (http://www.losninosensucasa.org/index.php); and (d) implement a universal school readiness assessment process that links with the district's instructional support services.

The results of this study lend support to the argument that merely moving to an earlier kindergarten enrollment date may not substantially affect the educational trajectories of Hispanic children from families experiencing low-income circumstances. The results do lend support to efforts to provide high-quality preschool experiences, such as in nationally accredited Head Start preschools, because this was the strongest predictor of school readiness. School psychologists should become aware of their districts' plans to implement these services because merely being "redshirted" for a year was not associated with school readiness or accelerated achievement in the early school years. Finally, the strong relation between school readiness and Grade 2 CST scores emphasizes the need to enhance preschool opportunities in all communities and to facilitate the transition of children into high-quality kindergarten programs.

- - -

Michael J. Furlong, PhD., is a professor in the Counseling, Clinical, and School Psychology Department, Gevirtz Graduate School of Education, University of California Santa Barbara. He is a member of the Society for the Study of School Psychology and a fellow of the American Psychological Association (Division 16) and the American Education Research Association. He is the Editor of the Journal of School Violence and the Director of the Center for School-Based Youth Development (see http://web.me.com/michaelfurlong).

Matthew Quirk, PhD., is an assistant professor of Counseling, Clinical, and School Psychology at the University of California, Santa Barbara. His current research focuses on issues related to children's school readiness, early reading development, and academic motivation/engagement. He can be contacted at 2145 Education Building, Gevirtz Graduate School of Education, University of California at Santa Barbara, Santa Barbara, CA 93106-9490 or by e-mail at mquirk@education.ucsb.edu.

REFERENCES

Andrews, S.P., & Slate, J.R. (2001). Prekindergarten programs: A review of the literature. *Current Issues in Education, 4*, 1-43. Available Online, from http://www.communityresearchpartners.org/uploads/publications/School_Readiness_Assessment.pdf

Bisanz, J., Dunn, M., & Morrison, F.J. (1995). Effects of age and schooling on the acquisition of elementary quantitative skills. *Developmental Psychology, 31*, 221-236. doi.org/10.1037%2F0012-1649.31.2.221

California Department of Education Standards an Assessment Division. (2009). *California Standards Tests technical report: Spring 2008 administration*. Available Online, from http://www.cde.ca.gov/ta/tg/sr/technicalrpts.asp

Carter, L.B. (1956). The effect of early school entrance on the scholastic achievement of elementary school children in the Austin public schools. *Journal of Educational Research, 50*, 91-103.

Cascio, E., & Schanzenbach, D. (2007). *First in the class? Age and the education production function* (December 2007). NBER Working Paper No. W13663. Available at SSRN: http://ssrn.com/abstract=1073646

Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Connell, C.M., & Prinz, R.J. (2002). The impact of childcare and parent–child interactions on school readiness and social skills development for low-income African American children. *Journal of School Psychology, 40*, 177-193. doi:10.1016/S0022-4405(02)00090-0

Crosser, S. L. (1991). Summer birth date children: Kindergarten entrance age and academic achievement. *Journal of Educational Research, 84*, 140-146.

Datar, A. (2006). Does delaying kindergarten entrance give children a head start? *Economics of Education Review, 25*, 43-62. doi:10.1016/j.econedurev.2004.10.004

Diamond, G.H. (1983). The birthdate effect — A maturational effect. *Journal of Learning Disabilities, 16*, 161-164. doi.org/10.1177%2F002221948301600306

Dickinson, D.J., & Larson, J.D. (1963). The effects of chronological age in months on school achievement. *Journal of Educational Research, 56*, 492-493.

Dobkin, C., & Ferreira, F. (2010). Do school entry laws affect educational attainment and labor market outcomes? *Economics of Education Review, 29*, 40-54. doi.org/10.1016%2Fj.econedurev.2009.04.003

Dhuey, E., & Lipscomb, S. (2008). What makes a leader? Relative age and high school leadership. *Economics of Education Review, 27*, 173-183. doi.org/10.1016%2Fj.econedurev.2006.08.005

Duncan, G.J., Dowsett, C.J., Classens, A., Magnuson, K., Huston, A.C., Klebanov, P., … Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*, 1428-1446. doi:10.1037/0012-1649.43.6.1428

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Frey, N. (2005). Retention, social promotion, and academic redshirting: What do we know and need to know? *Remedial and Special Education, 26*, 332–346. doi:10.1177/07419325050260060401

Gullo, D.F., & Burton, C.B. (1992). Age of entry, preschool experience, and sex as antecedents of academic readiness in kindergarten. *Early Childhood Research Quarterly, 7*, 175-186.

Heaviside, S., & Farris, E. (1993). *Public school kindergarten teachers' views on children's readiness for school*. Washington, DC: U.S. Department of Education, Office for Educational Research and Improvement.

Howes, C. (1988). Relations between early child care and schooling. *Developmental Psychology, 24*, 53-57. doi:10.1037//0012-1649.24.1.53

Kindergarten Entry Age, CA S. 1381 California State Legislature (2010).

Ladd, G.W., Herald, S.L., & Kochel, K. P. (2006). School readiness: Are there social prerequisites? *Early Education and Development, 17*, 115-150. doi:10.1207/s15566935eed1701_6

Langer, P., Kalk, J.M., & Searls, D.T. (1984). Age of admission and trends in achievement: A comparison of Blacks and Caucasians. *American Educational Research Journal, 21*, 61-78. doi.org/10.3102%2F00028312021001061

Le, V., Kirby, S.N., Barney, H., Setodji, C.M., & Gershwin, D. (2006). *School readiness, full-day kindergarten, and student achievement: An empirical investigation*. Santa Monica, CA: RAND Corportation. Retreived March 1, 2010, from http://www.rand.org/pubs/monographs/MG558/

Lee, V.E., Brooks-Gunn, J., Schnur, E., & Liaw, F.R. (1991). Are Head Start effects sustained? A longitudinal follow-up comparison of disadvantaged children attending Head Start, no preschool, and other preschool programs. *Child Development, 61*, 495-507.

Lilles, E., Furlong, M., Quirk, M., Felix, E., Dominguez, K., & Anderson, M. (2009). Preliminary development of the Kindergarten Student Entrance Profile. *The California School Psychologist, 14*, 71-80. Available Online, from http://www.casponline.org/

Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R.W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive. *Economics of Education Review, 26*, 52-66. doi.org/10.1016%2Fj.econedurev.2005.11.005

Magnuson, K.A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review, 26*, 33-51. doi.org/10.1016%2Fj.econedurev.2005.09.008

Mayer, S.E., & Knutson, D. (1999). Does the timing of school affect learning? In S.E. Mayer & P.E. Peterson (Eds.), *Earning and learning: How schools matter* (pp. 79-102). Washington, DC: Brookings Institute.

Meisels, S.J. (1996). Performance in context: Assessing children's achievement at the onset of school. In A. Sameroff & M. Haith (Eds.), *The five to seven year shift: The age of reason and responsibility* (pp. 408-431). Chicago, IL: University of Chicago Press.

Morrison, F.J., Griffith, E.M., & Alberts, D.M. (1997). Nature-nurture in the classroom: Entrance age, school readiness, and learning in children. *Developmental Psychology, 33*, 254-262.

National Center for Education Statistics. (1993). *Public school kindergarten teachers' views on children's readiness for school* (Publication No. 065-000-00596-4). Washington, DC: U.S. Department of Education, Office for Educational Research and Improvement. Available Online, from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=93410

National Center for Education Statistics. (1997). *The elementary school performance and adjustment of children who enter kindergarten late or repeat kindergarten: Findings from national surveys* (Publication No. 98097). U.S. Department of Education: Washington, DC: 1997. Available Online, from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=98097

NICHD Early Child Care Research Network. (2000). The relation of child care to cognitive and language development. *Child Development, 71*, 960-980.

Quirk, M., Furlong, M.J., Lilles, E., Felix, E., & Chin, J. (2011). Preliminary development of Kindergarten school readiness assessment for Latinos. *Journal of Applied School Psychology, 27,* 77-102.

Rumberger, R.W., & Arellano, B.D. (2007). Understanding and addressing the California Latino achievement gap in early elementary school. In R.A. Gutiérrez & P. Zavella (Eds.), *Mexican Americans in California: Transformations and challenges* (pp. 61-76). Champaign, IL, University of Illinois.

Saluja, G., Scott-Little, C., & Clifford, R.M. (2000). Readiness for school: A survey of state policies and definitions. *Early Childhood Research & Practice, 2*, 1-16.

Shepard L., & Smith, M. (1986). Synthesis of research on school readiness and kindergarten retention. *Educational Leadership, 44*, 78-86.

Snow, K. (2006). Measuring school readiness: Conceptual and practical considerations. *Early Education and Development, 17*, 7-41. doi:10.1207/s15566935eed1701_2

Stipek, D. (2002). At what age should children enter kindergarten? A question for policy makers and parents. *Society for Research in Child Development Social Policy Report, 16*, 1-16. Available Online, from http://www.srcd.org/index.php?option=com_content&task=view&id=232&Itemid=550

Stipek, D.J. (2003). School entry age. In R.E. Tremblay, R.G. Barr, & R. Peters (Eds.), *Encyclopedia on early childhood development* [online]. Montreal, Quebec: Centre of Excellence for Early Childhood Development, 1-5. Available Online, from http://www.child-encyclopedia.com/documents/StipekANGxp.pdf. Accessed [November 20, 2010].

Stipek, D., & Byler, P. (2001). Academic achievement and social behaviors associated with age of entry into kindergarten. *Applied Developmental Psychology, 22*, 175-189.

Turney, K., & Kao, G. (2009). Pre-kindergarten child care and behavioral outcomes among children of immigrants. *Early Childhood Research Quarterly, 24*, 432-444. doi.org/10.1016%2Fj.ecresq.2009.07.007

U.S. Department of Health and Human Services, Administration for Children and Families. (January 2010). *Head Start impact study. Final report.* Washington, DC: Author.

Uphoff, J.K., & Gilmore, J. (1985). Pupil age at school entrance: How many are ready for success. *Educational Leadership, 43*, 86-90.

A Summary and Synthesis of Contemporary Empirical Evidence Regarding the Effects of the Drug Abuse Resistance Education Program (D.A.R.E.)

Renee D. Singh, Shane R. Jimerson, Tyler Renshaw, Elina Saeki,
Shelley R. Hart, James Earhart, & Kaitlyn Stewart,
University of California, Santa Barbara

The prevention of drug abuse is an especially salient topic for school psychologists and other educational professionals. Schools are the primary setting for providing education and information aimed at the prevention of drug abuse. Previous meta-analyses (Ennett, et al., 1994; West & O'Neal, 2004) indicate that one of the nation's most popular drug prevention programs, the Drug Abuse Resistance Education program (D.A.R.E.), was not effective in reducing illicit drug use among youths. In 2003, D.A.R.E. was modified in an attempt to make it more effective. The purpose of this review is to summarize and synthesize the contemporary empirical evidence, which includes six studies focusing on the old D.A.R.E. and one study focusing on the new D.A.R.E., regarding outcomes associated with the modified D.A.R.E. program. Recent studies offer mixed evidence regarding the effectiveness of the new D.A.R.E curriculum, thus, further systematic investigation is warranted to better understand student outcomes associated with the new D.A.R.E curriculum. This information is particularly valuable for school psychologists, administrators, and other education professionals responsible for identifying empirically supported programs for use in schools.

Through the Drug Abuse Resistance Education (D.A.R.E.) program, law enforcement personnel contribute their expertise to help teach America's youth to resist peer pressure, and to abstain from drugs, gangs, and violence. We all have a responsibility to join these professionals in enabling youth to choose alternatives to violence and dangerous behavior and to lead the next generation of Americans toward a brighter future. NOW, THEREFORE, I, BARACK OBAMA, President of the United States of America, by virtue of the authority vested in me by the Constitution and the laws of the United States, do hereby proclaim April 8, 2010, as National D.A.R.E. Day. (http://www.dare.com/home/tertiary/default1b34.asp)

We find ourselves amidst a socio-political zeitgeist emphasizing "evidence-based" and "empirically supported" prevention and intervention efforts in our nation's schools. Given the above, and the preceding proclamation by President Barack Obama declaring a National D.A.R.E. Day, it begs the question: What does the research reveal regarding student outcomes associated with D.A.R.E.? The following begins with a brief discussion of the importance of prevention programs to address deleterious outcomes associated with drug use, then provides a summary of one of the nation's most popular prevention programs, D.A.R.E. The results of recent meta-analyses are subsequently discussed and recent studies are reviewed. In the final sections, we explore the widespread popularity of D.A.R.E. and the importance of empirical evidence and data-based decision making in choosing school-wide prevention programs.

DELETERIOUS OUTCOMES OF DRUG USE

The harmful effects of drug abuse can be widespread and irreversible. According to the National Institute on Drug Abuse (NIDA), even short-term drug use may lead to detrimental conditions such as cardiovascular disease, stroke, cancer, HIV/AIDS, hepatitis, and lung disease (NIDA, 2008). Moreover, long-term drug use may lead to debilitating psychiatric conditions such as paranoia, depression, exces-

Correspondence may be sent to Renee D. Singh, UCSB, GGSE, CCSP, Santa Barbara, CA 93106-9490 or e-mail: rsingh@education.ucsb.edu or jimerson@education.ucsb.edu

sive aggression, and hallucinations. In addition, drug use is associated with deleterious academic, social, emotional, behavioral, and mental health indicators among adolescents (Hussong, Curran, Moffitt, Caspi, & Carrig, 2004; Martins & Alexandre, 2009). Finally, both short- and long-term drug use may lead to temporary or permanent neurological damage that impairs various aspects of memory and executive functioning, including emotional and behavioral inhibition (NIDA, 2008). Unfortunately, despite these deleterious consequences, many school-aged youth continue to both use and abuse drugs of many kinds.

DRUG USE AMONG YOUTH IN THE UNITED STATES

The recent National Survey on Drug Use and Health (NSDUH) revealed that, between 2002 and 2009, overall rates of drug use among youths ages 12 through 17 were down (from 11.6% to 9.5%) (Substance Abuse and Mental Health Services Administration, 2009). Although there may be several plausible hypotheses for explaining this observed decline in drug use, Kuehn (2006) suggested that it is due, at least in part, to drug prevention programs, which have proliferated in communities and schools during the last few decades. Currently, some of the most popular drug use and abuse prevention programs targeting youth include the Boys and Girls Club of America community prevention efforts, various Substance Abuse and Mental Health Service Administration (SAMHSA) Model Programs, the Drug Free Communities and Supports Program, the Coaches Playbook Against Drugs, the Weed and Seed program, and the Drug Abuse Resistance Education (D.A.R.E.) program (Office of National Drug Control Policy, 2010; http://www.whitehousedrugpolicy.gov/prevent/programs.html).

Despite the importance of such prevention programming, communities and educational agencies have limited economic resources and are therefore faced with the difficult task of discerning which programs — including those focused on drug prevention — are essential and effective and thus deserve to be adopted, as opposed to which programs are superfluous or ineffective and are thus likely to be a waste of valuable resources. For schools especially, this situation is further compounded by current educational policy, which demands that data-based decisions be made and, ultimately, that empirically supported programs be employed. Given this present context, educational professionals involved in decision-making processes regarding the selection of school programming may substantially benefit from a summary and synthesis of contemporary empirical evidence regarding various popular prevention programs. The intent of this work is to aid educational professionals in their decision-making processes by providing such a summary and synthesis for one of the most popular drug prevention programs: D.A.R.E.

THE D.A.R.E. PROGRAM

Original D.A.R.E. Program

The original D.A.R.E. program was developed in 1983 as part of a joint effort between the Los Angeles Police Department and Los Angeles Unified School District to encourage students to resist both drugs and violence (Hansen & McNeal, 1997). The original curriculum consisted of core elements (e.g., resistance, skill training, self-esteem building) from two prototype versions of Project SMART (Self-Management and Resistance Training), supplemented with additional information on gang-related issues as well as legal issues related to drug use. During its first year of implementation, the program was delivered to approximately 6 million students at a cost of $750 million (costs of approximately $125 per child). Following the initial implementation, with the help of funding from the 1994 Safe and Drug-Free Schools and Communities Act, D.A.R.E. reached approximately 26 million students across the globe. Within the last decade at least 75% of school districts in the United States participated in D.A.R.E. (Griffith, 1999). Outside the U.S., D.A.R.E. has been implemented in 44 countries, including New Zealand, Belgium, Mexico, Canada, Sweden, Brazil, and Colombia. Thus, with approximately 33 million youth participating in the program since its inception, D.A.R.E. is often referred to as one of the most popular drug prevention programs in the world (Griffith, 1999). Further evidence of the program's popularity is found in President Barack Obama's recent declaration of a "National D.A.R.E. Day" (April

8, 2010) to honor D.A.R.E. officers' efforts to reduce drug use, gang participation, and violence among the nation's youth.

Currently, D.A.R.E. curricula are available for elementary-, middle-, and high school- aged students. The original curriculum was designed for use with elementary-aged students only, but middle school and high school components were added in 1986 and 1988, respectively, to broaden the program's reach (Des Jarlais, 2006). Although implementation time and lesson objectives vary by grade level, all D.A.R.E. curricula share certain elements in common. For example, all curricula are led by police officers who must attend and graduate from a two-week training program that includes instruction on drugs, gangs, internet safety, several other topics, and teaching techniques (D.A.R.E. America, 2010). All students participating in D.A.R.E. must complete a student workbook and a D.A.R.E. essay, have good attendance, follow D.A.R.E. and school rules, and be good role models and citizens in order to receive an award and graduate from the program.

Revised D.A.R.E. Program

Between 1994 and 2002, several events prompted D.A.R.E. to revise the original curricula (Des Jarlais, 2006). For example, Ennett et al. (1994) published a meta-analysis of several studies that examined the effectiveness of the D.A.R.E. curricula. Among other things, the authors concluded that the D.A.R.E. curricula were ineffective in reducing illicit drug use among youths, especially in the long-term. Similarly, both the Center for Substance Abuse Prevention and the U.S. Department of Education established programs (e.g., the National Registry of Effective Programs) designed to guide the use of federal funds related to drug prevention that did not include D.A.R.E in its registry, as it was judged to be ineffective (Des Jarlais, 2006). As a result of the lack of compelling evidence supporting the use of D.A.R.E., federal funding to train D.A.R.E. officers was reduced. Then, in 2003, in response to growing criticism regarding the program's lack of empirical support, D.A.R.E. made changes within its leadership hierarchy and organizational mission, which were followed by the development of new, science-based curricular components, training models, instructional methodologies, and partnerships (e.g., with the UCLA Anderson School of Business). This forged a more comprehensive program that met National Health Education Standards and may have boosted the program's research-base (D.A.R.E. America, 2010). Ultimately, what resulted was a new D.A.R.E. program that focused on enhancing the decision-making processes and protective factors (e.g., bonding, family, school, and community) that have been shown to foster resiliency in students at-risk for drug use.

New D.A.R.E. elementary school curriculum. The new D.A.R.E. elementary school curriculum is intended for students in 5th through 6th grades and consists of nine lessons and a culminating activity that involve more participation, active learning, supplemental take-home activities (e.g., "Family Talk"), and optional lessons (e.g., use of inhalants; D.A.R.E. America, 2010). A complete list of lesson objectives for the new elementary school curriculum is presented in Table 1. Although the potential dangers of misusing drugs, medicine, and other substances are discussed in all elementary-level curricula, the lessons for younger students focus more on child safety and prevention issues and are briefer than similar lessons for older students. For example, the D.A.R.E. curriculum for students in kindergarten through 2nd grade consists of four sessions, whereas the curriculum for students in 3rd through 4th grade consists of five sessions.

New D.A.R.E. middle school curriculum. The middle school curriculum, which is called "keepin' it REAL (kiR)," was added to the D.A.R.E. curricula in 2006 (D.A.R.E. America, 2010). In the context of the curriculum, "REAL" is an acronym for "Refuse, Explain, Avoid, and Leave." Originally developed by researchers at Pennsylvania State University and Arizona State University, the curriculum consists of 10 lessons designed to be implemented over the course of 10 weeks. Each lesson lasts approximately 40-45 minutes and involves the use of a series of videos, which depict high school-aged students using "REAL" strategies to resist drugs. Prior to being adopted by the D.A.R.E. program, "keepin' it REAL" was listed as an evidenced-based program on SAMHSA'S National Registry of Evidence-based Programs and Practices (see for instance, Hecht et al., 2003; Kulis et al., 2007).

Table 1. *Objectives for the Current D.A.R.E. Elementary School Curriculum (Available online: http://www.dare.com)*

Lesson 1: Purpose and Overview of the D.A.R.E Program
Students will be able to say in their own words the theme of the D.A.R.E program by participating in discussions. Students will explain the steps in the D.A.R.E Decision-Making Model by practicing with a partner. Students will write a personal reflection in response to the lesson.

Lesson 2: Tobacco and You
Students will review the D.A.R.E. Decision-Making Model by applying it to authentic situations. In a graphic representation, students will compare their estimates of the extent of tobacco use among adolescents with estimates reported in national surveys. Students will compare and contrast "common" beliefs about tobacco use through discussion, and analyze the validity of personal beliefs about tobacco use through a think, pair, share activity. They will recognize and correct personal misconceptions about tobacco use. Students will identify harmful effects of tobacco to the body by creating a warning label. Students will write a personal reflection in response to the lesson.

Lesson 3: Smoke Screen
Students will draw conclusions regarding the impact of advertising on the sale of tobacco by analyzing marketing techniques. Students will apply the D.A.R.E. Decision-Making Model to tobacco situations through group consensus. Students will recognize the harmful effects of marijuana use on the body by competing a worksheet in a think, pair, share format. Students will write a personal reflection in response to the lesson.

Lesson 4: Alcohol and You
In a graphic representation, students will compare their estimates of the extent of alcohol use among adolescents with estimates reported in national surveys. Students will compare and contract "common" beliefs about alcohol use through discussion, and analyze the validity of personal beliefs about alcohol use through a think, pair, share activity. They will recognize and correct personal misconceptions about alcohol use. Students will identify harmful effects of alcohol use to the body and the brain by completing a true/false worksheet. Students analyze risky situations and apply the D.A.R.E. Decision-Making Model in addressing strengths, needs, and health benefits of NOT using alcohol. Students will write a personal reflection in response to the lesson.

Lesson 5: The Real Truth
Students will review the D.A.R.E. Decision-Making Model by applying it to authentic personal situations. Students will recognize and describe the powerful effects of alcohol advertising by viewing a video and discussing the impact of alcohol advertising ploys. Students will identify harmful effects of inhalants on the body through questions and answers. Teams of students will brainstorm and list in a relay race the health risks in using drugs such as tobacco, alcohol, marijuana, and inhalants. Teams of students will brainstorm and list in a relay race and healthy things to do with friends that do not involve the use of tobacco, alcohol, marijuana, or inhalants. Students will write a personal reflection in response to the lesson.

Lesson 6: Friendship Foundations
Students will discuss recently observed advertisements and how the concept of friends was used in the ads. Students will describe qualities of positive friendships. Students will identify personal social support networks by starring specific resources. Students will identify types of peer pressure, and effective refusal responses in a written situation. Students will apply the D.A.R.E. Decision-Making Model in written peer pressure situations. Students will write a personal reflection in response to the lesson.

Lesson 7: Putting it Together
Students will discuss recent personal peer pressure situation, and how they responded. Students will demonstrate confident response styles in using "way to be in charge" through role-play. Students will apply the D.A.R.E. Decision-Making Model process by analyses of confident responses to the role-play situations. Students will write a personal reflection in response to the lesson.

Lesson 8: Personal Action
Students will discuss recent personal examples of how they responded Confidently and Responsibly in situations. Students will identify internal sources of personal pressure by group analyses and discussion. Students will develop a plan of action to make healthy and wise decisions about alcohol, tobacco, and other drugs in the D.A.R.E. Report. Students will write a personal reflection in response to the lesson.

Lesson 9: Practice! Practice! Practice!
Students will practice refusal skills to enhance health by using authentic pressure situations in a game activity. Students will evaluate and reflect their learning by sharing the D.A.R.E. Report and Pledge.

Lesson 10: Special Event
Students will participate in an appropriate D.A.R.E. culminating activity, which recognizes individual achievement of all the participants, and reinforces the knowledge and skills they have learned.

New D.A.R.E. high school curriculum. The high school curriculum is designed to extend and reinforce skills that students have already learned through exposure to D.A.R.E.'s kindergarten through middle school curricula (D.A.R.E. America, 2010). Examples of learning objectives include educating students about the personal and social consequences of substance abuse, reinforcing decision-making processes, and resisting peer pressure. Additionally, the curriculum emphasizes "helping students to recognize and cope with feelings of anger without causing harm to themselves or others and without resorting to violence or the use of alcohol and drugs" (D.A.R.E. America, 2010).

OUTCOMES ASSOCIATED WITH D.A.R.E.

As previously noted, Ennett et al.'s (1994) meta-analysis indicated that D.A.R.E. was not effective in reducing illicit drug use among youths. Similarly, a more recent study examining D.A.R.E. outcome results published between 1991-2002, reported an overall effect size that was small and not statistically significant (West & O'Neal, 2004), thus, providing further evidence of the ineffectiveness of D.A.R.E. The 11 studies included in the 2004 meta-analysis were chosen because they contained a control group and had been published in a peer-reviewed scientific journal. Specifically, the authors reported that D.A.R.E. was ineffective by noting that "the overall weighted effect size for the included D.A.R.E. studies was extremely small (correlation coefficient = 0.011; Cohen d = 0.023; 95% confidence interval = −0.04, 0.08) and nonsignificant (z = 0.73, NS)" (p. 1027). Considering this record of null findings, the purpose of this current review is to explore whether the status of D.A.R.E.'s effectiveness has changed since West and O'Neal's (2004) meta-analysis, which included studies published between 1991 and 2002. Specifically, two key questions are addressed in this current summary and synthesis: (1) Do contemporary evaluations of the effectiveness of the D.A.R.E. program distinguish between the old and new versions of the curricula? (2) Has the status of D.A.R.E.'s empirical effectiveness changed since 2003, when major revisions were initially undertaken with the existing curricula?

METHOD

To answer the above questions, a systematic review of contemporary literature was completed using electronic search engines (i.e., ERIC, Google Scholar, PsycINFO, and PubMed). Because Ennett et al.'s (1994) and West and O'Neal's (2004) meta-analyses provide extensive reviews of D.A.R.E. articles published between 1991 and 2002, and because the present analysis is focused on examining the effectiveness of the new D.A.R.E., the current review includes only articles published between 2002 and 2010. Furthermore, given the interest in exploring D.A.R.E.'s status as an empirically supported program, only empirical articles published in peer-reviewed journals were considered.

Review of the Literature

The literature review revealed seven studies that examine the effectiveness of D.A.R.E. These studies are briefly reviewed in the following section. To begin with, an investigation by Ahmed, Ahmed, Benett, and Hinds (2002) examined the effectiveness of D.A.R.E. in preventing the initiation of cigarette smoking in 236 fifth- and sixth-grade students. The measures included a 25-item questionnaire that featured questions from the Center for Disease Control's (1995) *Youth Risk Behavior Survey*, whereas outcome variables included (a) initiation of cigarette smoking, (b) current smoking status, (c) reasons for smoking, (d) knowledge of health risks, and (e) participant's experience of the effectiveness of D.A.R.E. in preventing smoking behavior. Because the program was implemented during the 1998-1999 school year, it appears that the participants received the old version of D.A.R.E. Overall, the results were in support of D.A.R.E.'s effectiveness in preventing the initiation of cigarette smoking indicating that the D.A.R.E. group had significantly lower rates of smoking at follow-up (May 1999) and was less likely to initiate smoking than the non-D.A.R.E. group (odds 4.9; $p = 0.003$; 95% CI: 1.7-14.0). Students in the D.A.R.E. group also had significantly higher knowledge scores on the health risks of smoking ($p = 0.002$), and students with the highest knowledge scores on the health risks of smoking had significantly lower rates of smoking overall ($p = 0.001$).

Table 2. *Summary of Studies Reviewed.*

Reference	Results	Participants	Effect Size(s)
Ahmed, Ahmed, Benett, & Hinds (2002)	The D.A.R.E. group had significantly lower rates of smoking at follow-up (May 1999) and was less likely to initiate smoking than the non-D.A.R.E. group (odds 4.9; $p = 0.003$; 95% CI: 1.7-14.0). Students in the D.A.R.E. group also had significantly higher knowledge scores on the health risks of smoking ($p = 0.002$), and students with the highest knowledge scores on the health risks of smoking had significantly lower rates of smoking overall ($p = 0.001$).	N = 236 fifth and sixth grade students	Not Stated
Perry et al. (2003)	None of the conditions (D.A.R.E. vs. Control; D.A.R.E. Plus vs. Control; and D.A.R.E. Plus vs. D.A.R.E.) yielded significant differences with regard to any of the outcome variables (self-reported tobacco, alcohol, and marijuana use; multi-drug use; violence; and victimization) at baseline or follow-up. However, boys at D.A.R.E. Plus schools were marginally less likely to increase current smoking ($p \leq 0.08$) and victimization ($p \leq 0.10$).	N= 6,237 seventh grade students	Not Stated
Komro et al. (2004)	At second follow-up, boys in the D.A.R.E. Plus condition exhibited fewer acts of physical violence (ES =.10) and verbal violence (ES = .10) than boys in the control condition. However, Komro et al. (2004) state that such differences were "entirely mediated by a decrease of norms that support violence [$p < 0.05$], an increase in outcome expectancies about being violence-free [$p < 0.05$], and an increase in parental consequences for fighting" [$p < 0.05$] (p. 335). No Significant differences were found with regard to any of the remaining outcome variables (weapon carrying and victimization). Among girls, those in the D.A.R.E. Plus condition had significantly lower scores on the Victimization Scale than those in the D.A.R.E. only condition (ES = .09). However, no significant differences were found with regard to any of the remaining outcome variables (physical violence, verbal violence, and weapon carrying).	N = 6,237 seventh grade students	Boys—Acts of Physical Violence = 0.10 Boys—Acts of Verbal Violence = 0.10 Girls—Victimization = 0.09
Zhong et al. (2005)	The results indicated that the new D.A.R.E. had significant ($p < 0.05$) effects on normative beliefs about tobacco (ES = -0.411), alcohol (ES = -0.381), and marijuana (ES = -0.431) use. Similarly, the new D.A.R.E. also had significant ($p < 0.05$) effects on attitudes toward tobacco (ES = -0.176), alcohol (ES = -0.338), and marijuana use (ES = -0.203). With regard to refusal skills, however, the new D.A.R.E. only had significant effects ($p < 0.10$) on refusal skills related to alcohol use (0.085).	N = 513 8th grade students	Normative Beliefs about Tobacco Use = -0.411 Normative Beliefs About Alcohol Use = -0.381 Normative Beliefs About Marijuana Use = -0.431 Attitudes Toward Tobacco Use = -0.176 Attitudes Toward Alcohol Use = -0.338 Attitudes Toward Marijuana Use = -0.203 Refusal Skills Related To Alcohol Use = 0.085
Merrill, Pinsky, Killeya-Jones, Sloboda, & Dilasciol (2006)	Structural/organizational strengths include a well-defined organizational focus, uniform training and means for rapid dissemination, continuing education mechanisms, mechanisms for program monitoring and fidelity of implementation, branding, and predictable and consistent financing. Structural/organizational weaknesses include unstable funding and an inability to incorporate components for continual upgrading of the curricula that reflect "research evidence and 'principles of prevention'" (p. 1).	D.A.R.E. America leadership, 50 D.A.R.E. state coordinators, two city coordinators (from Washington DC and New York City), and two focus groups held with D.A.R.E. officers	-
Pan & Bai (2009)	The overall effects of the D.A.R.E. program on drug use (-0.08) and psychosocial behavior (0.36) were small.	20 studies published between 1991 and 2003	Drug Abuse = -0.08 Psychosocial Behavior = 0.36
Ennet, Tobler, Ringwalt, & Flewelling (2009)	D.A.R.E.'s largest effect was on knowledge (0.42), whereas its smallest effects were on self-esteem (0.06) and drug use (0.06). Mean effect sizes for attitudes about drug use, social skills, and attitude towards police were 0.11, 0.19, and 0.13 respectively.	Eight studies published between 1986 and 1993	Knowledge about Drugs = 0.42 Attitudes about Drug Use = 0.06 Social Skills = 0.11 Self-Esteem = 0.06 Attitude Toward Police = 0.13 Drug Use = 0.06

Next, Perry et al. (2003) studied the effects of D.A.R.E. on drug use and violence. The participants included 6,237 students drawn from 24 schools that were randomly assigned to one of three conditions: D.A.R.E. only, D.A.R.E. Plus, and delayed program control. D.A.R.E. Plus differs from regular D.A.R.E. in several important ways. Firstly, officers teaching D.A.R.E. Plus receive 2-hours of additional training. Secondly, it includes "On the VERGE," a 4-session, classroom-based program, which includes parental involvement. Thirdly, it includes extra-curricular activities and neighborhood action teams created and facilitated by community organizers. Because all D.A.R.E. programs were implemented during the 1991-2001 school years, it is likely that the participants in both conditions received the old version of D.A.R.E. That is, the participants in the D.A.R.E. only condition received just the old version of D.A.R.E.; whereas in the D.A.R.E. Plus condition received the old version of D.A.R.E. plus additional components. Additionally, based on the authors' descriptions of the D.A.R.E. Plus program, it does not appear to be the same as the new D.A.R.E. program developed after 2003. For example, D.A.R.E. Plus does not include "keepin' it REAL (kiR)," which is one of new D.A.R.E.'s empirically supported components. Overall, the study's results regarding D.A.R.E.'s effectiveness were mixed. For example, none of the conditions (D.A.R.E. vs. Control; D.A.R.E. Plus vs. Control; and D.A.R.E. Plus vs. D.A.R.E.) yielded significant differences (p with regard to any of the outcome variables (self-reported tobacco, alcohol, and marijuana use; multi-drug use; violence; and victimization) at baseline or follow-up. However, boys at D.A.R.E. Plus schools were marginally less likely to increase current smoking ($p \leq 0.08$) and victimization ($p \leq 0.10$).

Komro et al. (2004) studied the same participants featured in Perry et al.'s (2003) study. So, the participants again included 6,237 students drawn from 24 schools that were randomly assigned to one of three conditions: D.A.R.E. only, D.A.R.E. Plus, and delayed program control. Primary outcome variables, which were measured via a self-report questionnaire, included physical violence, verbal violence, weapon carrying, and victimization. Because all D.A.R.E. programs were implemented during the 1991-2001 school years and because the authors did not clearly state whether they were evaluating old vs. new D.A.R.E., it is assumed that the participants in both D.A.R.E. conditions received the old version of D.A.R.E. Although the authors concluded that the "D.A.R.E. Plus program was more effective in preventing violence among boys than among girls," their overall results were not in support of D.A.R.E.'s effectiveness (p. 335). For example, at second follow-up, boys in the D.A.R.E. Plus condition exhibited fewer acts of physical violence (ES =.10) and verbal violence (ES = .10) than boys in the control condition. However, Komro et al. (2004) state that such differences were "entirely mediated by a decrease of norms that support violence [$p < 0.05$], an increase in outcome expectancies about being violence-free [$p < 0.05$], and an increase in parental consequences for fighting" [$p < 0.05$] (p. 335). No significant differences were found with regard to any of the remaining outcome variables (weapon carrying and victimization). Among girls, those in the D.A.R.E. Plus condition had significantly lower scores on the Victimization Scale than those in the D.A.R.E. only condition (ES = .09). However, no significant differences were found with regard to any of the remaining outcome variables (physical violence, verbal violence, and weapon carrying).

Using self-report surveys, Zhong et al. (2005) evaluated the new D.A.R.E.'s effectiveness on influencing normative beliefs, attitudes, and refusal skills related to substance use (i.e. tobacco, alcohol, and marijuana use). Zhong et al. (2005) clearly state that they are evaluating the new version of D.A.R.E, which includes curricula related to the study's outcome variables. The participants included 513 eighth-grade students from a larger study conducted by the Institution for Social Policy at the University of Akron. Overall, the authors' findings were in support of D.A.R.E.'s effectiveness. For example, the new D.A.R.E. had significant ($p < 0.05$) effects on normative beliefs about tobacco (ES = -0.411), alcohol (ES = -0.381), and marijuana (ES = -0.431) use. Similarly, the new D.A.R.E. also had significant ($p < 0.05$) effects on attitudes toward tobacco (ES = -0.176), alcohol (ES = -0.338), and marijuana use (ES = -0.203). With regard to refusal skills, however, the new D.A.R.E. only had significant effects ($p < 0.10$) on refusal skills related to alcohol use (negligible ES = 0.085).

Merrill, Pinsky, Killeya-Jones, Sloboda, and Dilasciol (2006) studied the organization, structure and function of the new D.A.R.E. Data were gathered via interviews with D.A.R.E. America leadership,

50 D.A.R.E. state coordinators, two city coordinators (from Washington, DC, and New York City), and two focus groups held with D.A.R.E. officers. Although no outcome data was presented, the authors' results highlighted several structural/organizational strengths and weaknesses of the new D.A.R.E. program. For example, strengths included a well-defined organizational focus, uniform training and means for rapid dissemination, continuing education mechanisms, mechanisms for program monitoring and fidelity of implementation, branding, and predictable and consistent financing. In contrast, weaknesses included unstable funding and an inability to incorporate components for continual upgrading of the curricula that reflect "research evidence and 'principles of prevention'" (p. 1).

Pan and Bai (2009) reviewed 20 studies that assessed the effectiveness of the D.A.R.E. program in the United States. Studies were published between 1991 and 2003, which indicates that the majority of studies they reviewed examined the old D.A.R.E. Their results indicated that the overall effects of the D.A.R.E. program on drug use (-0.08) and psychosocial behavior (0.36) were small.

Finally, Ennet, Tobler, Ringwalt, and Flewelling (2009) conducted a meta-analytic review of eight studies that assessed the effectiveness of D.A.R.E. with certain outcome classes. Studies were published between 1986 and 1993, which indicates that all studies they reviewed examined the old D.A.R.E. Outcome classes included knowledge about drugs, attitudes about drug use, social skills, self-esteem, attitude toward police, and drug use. The results indicated that D.A.R.E.'s largest effect was on knowledge (0.42), whereas its smallest effects were on self-esteem (0.06) and drug use (0.06). Mean effect sizes for attitudes about drug use, social skills, and attitude towards police were 0.11, 0.19, and 0.13 respectively.

DISCUSSION

Considering that D.A.R.E. is referred to as one of the most popular drug prevention programs in the world (Griffith, 1999), it is prudent for school psychologists, administrators, and other educational professionals to consider the empirical evidence regarding outcomes associated with the program. To date, there appears to be a disconcerting lack of evidence supporting the use of D.A.R.E. In aggregate, recent empirical evidence reveals that the *old* D.A.R.E. is not effective in reducing outcomes related to substance use and violence. However, recent research offers preliminary evidence that the *new* D.A.R.E. may be effective in influencing normative beliefs about substance use, attitudes toward substance use, and adoption of refusal skills to resist substance abuse.

Reasons for D.A.R.E.'s Widespread Use, Despite the Lack of Empirical Support

Considering the research findings presented above, one might ask: Why do school districts continue to participate in D.A.R.E. when its effectiveness is unclear? To answer this question, we reviewed several studies examining people's perceptions of D.A.R.E. For example, Lucas (2008) studied 420 parents' perceptions of D.A.R.E.'s effect on parent/child attitudes and behavior. Findings indicated that parents perceived the program to be useful in helping children understand and resist drugs, increasing both parents'/children's awareness of drug problems, and increasing parent-child conversations about drug problems. However, they also stated that "most parents did not see an impact on their child's school performance or overall attendance" (p. 99). Thus, it appears that parents did not perceive the D.A.R.E. program as having a direct link to academic outcomes.

Similarly, Birkeland, Murphy-Graham, and Weiss (2005) studied reasons why school districts continue to offer D.A.R.E. despite unsupportive evaluations of the program's effectiveness. Because data were collected between 2001 and 2003, it appears that participating school districts were using the old version of D.A.R.E. Their sample consisted of 16 school districts in Colorado, Massachusetts, Kentucky, and Illinois. The authors reported that sampled school districts decided to continue participation in D.A.R.E., despite its questionable effectiveness, because school officials believed the evaluations were not sensitive to improvements in relationships between police and students.

Finally, Donnermeyer and Wurschmidt (1997) examined educator's perceptions of D.A.R.E. using a statewide survey. Because the study was published in 1997, these perspectives were regarding the old version of D.A.R.E. The participants included 286 5th and 6th-grade teachers and principals. In general,

the participants assigned favorable ratings to each of the following: Teacher/officer interactions, role playing exercises, the graduation ceremony, program quality, and program impact on students. This indicated that D.A.R.E was generally well-liked and was valued among the participants.

Empirical Evidence and Data-Based Decisions

Despite the lack of empirical evidence supporting the use of D.A.R.E., it appears that in general, the D.A.R.E. program continues to be implemented in many schools throughout the United States, and in other countries around the world. However, the lack of empirical evidence has influenced some school districts to discontinue D.A.R.E. Weiss, Murphy-Graham, and Birkeland (2005) studied the influence of unsupportive evaluations on school districts' participation in D.A.R.E. The authors' findings indicated that one school district dropped D.A.R.E. independent of the unsupportive evaluations. Five districts dropped D.A.R.E. as a result of the unsupportive evaluations. Two districts dropped D.A.R.E. because of the unsupportive evaluations and because the Safe and Drug Free Schools (SDFS) and Communities Act (SDFSCA) only provided federal funding for research-based programs. Two districts dropped D.A.R.E. as a result of SDFSCA requirements and because D.A.R.E. was not featured on the list of approved, evidenced-based programs published by the U.S. Department of Education. Thus, it appears that in some instances the empirical evidence regarding the ineffectiveness of the D.A.R.E curriculum has resulted in districts abandoning this particular strategy.

CONCLUSIONS

The results of this contemporary literature review on D.A.R.E.'s effectiveness indicate that the old D.A.R.E. was not effective in reducing outcomes related to substance use and violence. Secondly, a recent study of the new D.A.R.E. curricula reveals preliminary evidence influencing normative beliefs about substance use, attitudes toward substance use, and adoption of refusal skills to resist substance abuse. However, additional research is needed before it can be concluded that D.A.R.E. is effective in these areas. Finally, future evaluations need to clearly state whether they are evaluating the old vs. new D.A.R.E., considering that with the recently incorporated evidence-based components (e.g., keepin' it REAL), it is possible that the new D.A.R.E. may be more effective than its predecessor. Whereas the objectives of D.A.R.E. are laudable, the empirical evidence to date does not provide compelling evidence of effectiveness. Thus, given the contemporary socio-political zeitgeist emphasizing "evidence-based" and "empirically supported" prevention and intervention efforts in our nation's schools, it is perplexing that a national D.A.R.E day has been established in the United States.

The importance of data-based decision-making cannot be understated, particularly in light of educational policy. School practitioners, particularly school psychologists, are often called upon to provide expertise in identifying empirically-based intervention programs. Thus, it is imperative that educational professionals attend to the extant empirical evidence and also continue to collect data to examine outcomes associated with new curricula. In the case of the D.A.R.E., previous research yielded evidence illustrating that the old D.A.R.E. is largely ineffective in producing changes in youth's attitudes and behaviors, however, with few studies of the new D.A.R.E. curriculum, further research is necessary to discern both short-term and long-term effects.

- - -

Renee D. Singh, M.Ed., is a doctoral student at the University of California, Santa Barbara; Shane R. Jimerson, PhD., is a professor at the University of California, Santa Barbara; Tyler L. Renshaw, M.Ed., Elina Saeki, M.Ed., Shelley R. Hart, M.S., and James Earhart, M.Ed., are doctoral students at the University of California, Santa Barbara; Kaitlyn Stewart, M.Ed., is a graduate student at the University of California, Santa Barbara.

REFERENCES

Ahmed, N.U., Ahmed, N.S., Bennett, C.R., & Hinds, J.E. (2002). Impact of a drug abuse resistance education (D.A.R.E) program in preventing the initiation of cigarette smoking in fifth- and sixth-grade students. *Journal of the National Medical Association, 94*, 249-256. Retrieved from www.csa.com

Birkeland, S., Murphy-Graham, E., & Weiss, C. (2005). Good reasons for ignoring good evaluation: The case of the drug abuse resistance education (D.A.R.E.) program. *Evaluation and Program Planning, 28*, 247-256. doi:10.1016/j.evalprogplan.2005.04.001

Centers for Disease Control and Prevention. (1995). *Youth risk behavioral surveillance questionnaire*. Georgia: Public Health Service.

D.A.R.E. America (19962010). The official D.A.R.E. website: Drug abuse resistance education. Retrieved December 10, 2009, from http://www.dare.com.

Des Jarlais, D.C., Sloboda, Z., Friedman, S.R., Tempalski, B., McKnight, C., & Braine, N. (2006). Diffusion of the D.A.R.E and syringe exchange programs. *American Journal of Public Health, 96*, 1354-1357. doi:10.2105/AJPH.2004.060152.

Donnermeyer, J.F. (1998). Educator perceptions of the D.A.R.E. officer. *Journal of Alcohol and Drug Education, 44*, 1-17. Retrieved from www.csa.com

Donnermeyer, J.F., & Wurschmidt, T. N. (1997). Educators' perceptions of the D.A.R.E. program. *Journal of Drug Education, 27*, 259-276. Retrieved from www.csa.com

Ennett, S.T., Rosenbaum, D.P., Flewelling, R.L., Bieler, G.S., Ringwalt, C.L., Bailey, S.L. (1994). Long-term evaluation of drug abuse resistance education. *Addictive Behavior, 19*, 113-125.

Faine, J.R., & Bolander, E. *Drug abuse resistance education: An assessment of the 1987-88 Kentucky State Police D.A.R.E. program*. Bowling Green: Western Kentucky University, the Social Research Laboratory.

Griffith, J.S. (1999). Daring to be different? A drug prevention and life skills education programme for primary schools. *Early Child Development and Care, 158*, 95-105. doi:10.1080/0300443991580108

Hecht, M.L., Marsiglia, F.F., Elek-Fisk, E., Wagstaff, D.A., Kulis, S., & Dustman, P. (2003). Culturally-grounded substance use prevention: An evaluation of the keepin' it R.E.A.L. curriculum. *Prevention Science, 4*, 233–248.

Hussong, A.M., Curran, P.J., Moffitt, T.E., Caspi, A., & Carrig, M.M. (2004). Substance abuse hinders desistance in young adults' antisocial behavior. *Development and Psychopathology. Special Issue: Transition from Adolescence to Adulthood, 16*, 1029-1046. doi:10.1017/S095457940404012X

Komro, K.A., Perry, C.L., Veblen-Mortenson, S., Stigler, M.H., Bosma, L.M., Munson, K.A., & Farbakhsh, K. (2004). Violence-related outcomes of the D.A.R.E. plus project. *Health Education & Behavior, 31*, 335-354. doi:10.1177/1090198104263337

Kuehn, B.M. (2006). Shift seen in patterns of drug use among teens. *JAMA: Journal of the American Medical Association, 295*, 612-613. doi:10.1001/jama.295.6.612.

Kulis, S., Nieri, T., Yabiku, S., Stromwall, L., & Marsiglia, F.F. (2007). Promoting reduced and discontinued substance use among adolescent substance users: Effectiveness of a universal prevention program. *Prevention Science, 8*, 35-49.

Lucas, W.L. (2008). Parents' perceptions of the drug abuse resistance education program (DARE). *Journal of Child & Adolescent Substance Abuse, 17*, 99-114. doi:10.1080/15470650802292905

Martins, S.S., & Alexandre, P.K. (2009). The association of ecstasy use and academic achievement among adolescents in two U.S. national surveys. *Addictive Behaviors, 34*, 9-16. doi:10.1016/j.addbeh.2008.07.022

Morris, M.C., Cline, R.J. W., Weiler, R.M., & Broadway, S.C. (2006). Prescription drug abuse information in D.A.R.E. *Journal of Drug Education, 36*, 33-45. doi:10.2190/AT0P-NGT8-2KWY-UY48

National Institute on Drug Abuse (NIDA). Medical consequences of drug abuse. Retrieved April 18th, 2010, from http://www.nida.nih.gov/consequences/index.html.

Substance Abuse and Mental Health Services Administration. (2009). *Results from the 2008 National Survey on Drug Use and Health: National Findings* (Office of Applied Studies, NSDUH Series H-36, HHS Publication No. SMA 09-4434). Rockville, MD.

Weiss, C.H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence: How evaluations affect D.A.R.E. *American Journal of Evaluation, 26*, 12-30. doi:10.1177/1098214004273337

West, S.L., & O'Neal, K.K. (2004). Project D.A.R.E. outcome effectiveness revisited. *American Journal of Public Health, 94*, 1027-1029. doi:10.2105/AJPH.94.6.1027

Zhong, H., Zhao, X., & Deason-Howell, L.M. (2005). Does gender matter? Assessing the new D.A.R.E. program effects on mediators across gender groups. *Journal of Alcohol and Drug Education, 49*, 15-25. Retrieved from www.csa.com

School Psychologists Working with Native American Youth: Training, Competence, and Needs

Carol Robinson-Zañartu, Nola Butler-Byrd and Valerie Cook-Morales,
San Diego State University

Paul Dauphinais,
Turtle Mountain Schools, Belcourt, North Dakota

Elvina Charley,
Ganado Unified School District, Ganado, Arizona

Mike Bonner,
Omaha, Nebraska

Author Note
The authors gratefully acknowledge the contributions of Flora Howe, Anslem Yazzie, Eugene Honanie, Sally Baas, Cheryl Sanchez-Ostapchenko, and Marilyn Robinson in developing the pilot survey; David Shriberg for preliminary data analysis; and Lee Huff for his vision.

Despite growing emphases on multicultural competence, Native American youth remain tremendously underserved by schools: low achievement, high dropout rates, and over-identification for special education persist. The authors analyzed responses of 403 school psychologists to a national survey regarding their competence gained in training, in current practice, and that needed for effective work with Native Americans. Respondents reported significant under-preparation in training and inadequate preparation for competent practice. Both ethnicity and length of experience with the population yielded significant differences in perceived levels of competence.

KEYWORDS: Native American, American Indian, school psychology, multicultural, cultural competence.

Persistent problems in learner outcomes for Native American (NA) youth should compel us as school psychologists to examine our capacity to make a difference. NA achievement is significantly below that of "mainstream" children, dropout rates are unacceptably high, and over-identification for special education is actually on the rise (Faircloth & Tippeconnic, 2010; National Center for Education Statistics [NCES], 2008). Historically, NA children, their families, and even their cultures were blamed for problematic academic outcomes. Increased attention to multicultural competence, including use of systemic approaches in service delivery, was expected to positively impact such achievement gaps, helping school psychologists to more competently assess individual youth, and more effectively facilitate relevant systems changes, from school-based attitudes and expectations to culturally appropriate interventions.

Growing bodies of literature have established relationships between cross-cultural competence and effective outcomes for youth (Rogers & Lopez, 2002; Tarver Behring & Ingraham, 1998). However, despite mandates and resources to support multicultural competence, outcomes for Native youth suggest these efforts are insufficient. Sue (2001) cautioned that using widely inclusive parameters around multicultural training might lead to watered-down approaches, inadequate to address the needs of specific groups. Concerned with similar issues following a meta-analytic review of multicultural education, Smith,

Correspondence concerning this article should be addressed to Carol Robinson-Zañartu, Department of Counseling and School Psychology, San Diego State University, San Diego, CA 92128-1179, USA. E-mail: crobinsn@mail.sdsu.edu

Constantine, Dunn, Dinehart, and Montoya (2006) suggested the need for research regarding work with specific racial/ethnic groups. In light of this issue, this study proposed to examine school psychologists' preparedness to work competently with Native American youth.

THE STATUS OF NATIVE AMERICAN YOUTH

With a 20.6% increase in those under age 18 in one decade (U.S. Census, 2000), Native American (NA) youth represent one of our fastest growing populations. Although NA students reside in virtually every state, unless they are part of a large reservation community, many remain "invisible" to educators. Their percentages in most schools are low, and many carry surnames from histories of colonization that lead to misidentification. For instance, although the largest actual number of NAs in the U.S. live in California, where many have Spanish surnames, most schools enroll low percentages. Thus, NA issues are not differentiated, but subsumed under efforts to serve other "minorities."

The UCLA Civil Rights Project's most recent research on education statistics for Native youth (Faircloth & Tippeconnic, 2010) reports that NA graduation rates now range from 30% to 64%, and that on average, fewer than 50% of NA students from the Pacific and Northwest U.S. graduate high school. Achievement levels fall well below those of other students. At both grades 4 and 8, NA students had lower average scores in reading and mathematics than the average for all other students nationally (NCES 2008; Rampey, Lutkus & Weiner, 2006). Further, Native American youth remain over-represented in special education. The U.S. Department of Education (2003) reported that nationally, 11.9% of NA youth receive special education services. However, others dispute and differentiate these data. Reviewing 2005 national education statistics, Stancavage, Mitchell, Bandeira de Mello, Gaertner, Spain, and Rahal (2006) found that 17% of NA students in 4th grade, and 15% in 8th grade were classified as students with disabilities versus only 9% of non-NA students. They also found differences in identification patterns between "high density" and "low density" schools. Low-density schools (fewer than 25% NA youth) identified significantly more Native youth as having disabilities. This over-identification in low-density schools raises multiple hypotheses; for instance, might mainstream teaching styles be incompatible with indigenous learners; or might cultural inexperience limit school psychologists' skills in differentiating difference from disability.

MULTICULTURAL AND CULTURE-SPECIFIC COMPETENCE IN SCHOOL PSYCHOLOGY

Both the American Psychological Association (APA) (1990; 2003) and the National Association of School Psychologists (NASP, 2000) have provided guidelines requiring multicultural content in training. By the early 1990's, although some school psychology training programs had articulated multicultural training components, (c.f., Cook-Morales & Robinson-Zañartu, 1995; Palmer, Juarez & Hughes, 1991), 40% did not yet offer coursework or integrate culturally relevant content (Rogers, Ponterro, Conoley & Wiese, 1992). Two decades later, Rogers and Molina (2006) recognized only 11 psychology programs as highly successful in minority recruitment and retention, supported in part by culturally relevant content.

The APA School Psychology Division's Task Force on Cross-Cultural School Psychology Competencies developed a framework for cultural competence, summarizing existing knowledge bases, and defining competencies within six domains: (a) legal and ethical issues; (b) school culture, policy and advocacy; (c) psycho-educational assessment; (d) academic, therapeutic and consultative interventions; (e) working with interpreters; and (f) research (Rogers et al., 1999). Their generic multicultural guidelines were designed to prepare candidates to transfer cultural awareness and skills gained in training to specific cultural demands in practice. However, no data indicate whether practicing school psychologists do this.

Multiple scholars assert that specific knowledge of NA cultures, issues, and worldviews is critical to effective school-based work and research with this population (Cleary & Peacock, 1998; Deloria & Wildcat, 2001; Robinson-Zañartu, 1996). Specific suggestions fall into areas such as educational policy, history, intervention, educational programs, and research. The need for educators to understand such culturally specific information as, for example, local tribal histories, intergenerational effects of historical traumas on learning, the Indian Child Welfare Act, Title VII, the Johnson-O'Malley Act, tribal behavioral norms, taboos on discussing certain symbols, events, or stories, the role of extended family in a child's education,

tribal research protocols, and the skills to act respectfully with sovereign nations all are considered critical for effective educational practice.

Most scholars writing about work with NA youth have cultural familiarity or longevity of work within NA communities (e.g., Cleary & Peacock 1998; Klug & Whitfield, 2003). Most school psychologists do not bring these experiences to their practice. No data demonstrate what difference this makes in competence to practice with NAs. Thus, it becomes useful to determine whether the multicultural training school psychologists receive prepares them to work with NA populations, and for those working with Native youth, what influence that has on gaining sufficient additional competence. Further, given the differential special education identification rates in high and low density areas, does familiarity with the cultures, either through personal membership or longevity of work experience, influence levels of competence?

The research questions guiding this study explore the following issues in relation to six domains of culturally competent practice: (a) How do school psychologists rate their training, current competence, and need for competence in each of the six areas of practice in relation to serving NA youth? (b) Are there significant differences between perceived levels needed for effective practice and those received in training and current levels of practice? (c) Does length of time working with NA youth make a difference in perceived training, preparedness or perceived need in each domain? and (d) Do perceived level of preparedness, training or need for competence in the six domains vary with ethnicity?

METHOD

Participants

Two strata of participants were sought: the first represented school psychologists as a group nationally, randomly selecting from NASP's membership database; the second represented school psychologists working with greater numbers of NA youth. Because school psychologists may be unaware of serving Native youth, surveying the entire population was warranted. As NASP's membership is slightly over 20,000, we sought a sample size that would produce a 5% margin of error (p), yielding a 95% confidence level for a population size of 20,000 respondents (Isaac & Michael, 1981). The web-based survey was sent to a random sample of 2500 of NASP's members (Stratum 1), and for Stratum 2, to 51 school psychologists in postal codes with 25% or more NA population, and to 96 Bureau of Indian Affairs (BIA) psychologists. Two weeks following the initial request, a reminder was sent. Four hundred twenty-three responses were received, well beyond the minimum number needed to produce confidence in our results. Of those, 18 were removed from analysis due to not working in the field, and two due to missing data, yielding a sample size of 403, representing a 16% return rate.

Demographics of respondents. The distribution of respondents across ethnic categories broadly approximates that of NASP's membership, with the vast majority of the sample (82.8%) identifying as European American (EA). This compares with 92.6 % of NASP's membership who self-identify as Caucasian (J. Charvat, personal communication, November 2, 2008). Although the full over-sampling group was not separable from the total, those who worked in BIA/BIE schools were also predominantly Euro-Americans (72.5%). Overall, more NA or dual ethnicity including NAs responded to the survey (3.9%) than are represented in NASP (0.8%). The over-sampling procedure produced sufficient numbers of respondents who reported working with NA youth (28%) to inform the research questions without compromising the overall integrity of the sample.

Survey Design

Working from literature reviews of cultural competence in school psychology and competent educational or psychological practice with Native American youth, members of NASP's Native American Task Force designed a survey to determine school psychologists' perceptions of their competence gained during training, competence in practice, and levels of competence needed for effective practice. To determine areas of competence, they began with the six areas of culturally competent practice originally identified by the APA School Psychology Divisions' Task Force on Cultural Competence (Rogers et al., 1999) discussed above. Following a review of the literature on competence with Native American youth in education and psychology, members of the NASP Task Force developed a list of exemplars to describe

each category to reflect attention to Native American education issues. A focus group of Native American psychologists then validated or modified the exemplars in each of the categories. For example, understanding intergenerational effects of historical traumas on learning became an exemplar under the areas of school culture, policy and advocacy; knowledge of the Indian Child Welfare Act, Title VII, and the Johnson-O'Malley Act became exemplars under the area of legal and ethical issues; understanding tribal behavioral norms and taboos on discussions of certain symbols, events and stories were skill sets under academic, therapeutic and consultative interventions; and knowledge about tribal research protocols fell under the area of research. These modified areas of cultural competence became the six major subsets of the survey, establishing content validity through use of previously identified areas of competence, areas emerging from the literature, and experience of Native American psychologists in the focus group helping establish the survey questions.

An initial web-based survey was piloted by members of the NASP Task Force. Based on feedback from the pilot group and from NASP's research office, the survey was revised, mainly shortening the exemplar list. In the final 24-item survey, demographic data described respondents' gender and ethnicities, types of schools served, e.g., BIA, public, tribally controlled, and whether and how long respondents had worked with NA populations. Respondents were then asked to rate their levels of competence in training, in current practice, and the levels they perceived needed for effective practice with NAs in six major domains of culturally competent practice, drawn from those identified by Rogers et al. (1999). Descriptors were modified to reflect specific attention to NA issues (see Table 1), using a four-point Likert scale (0-3), rating competence as "none," "some," "adequate," or "strong."

Table 1 *Multicultural Competencies Framework for Survey*

Categories of Multicultural Competencies[1]	Extended Applications to Native American/American Indian (NA/AI) Children and Youth
Legal & Ethical Issues	Pertaining to NA/AI (knowledge, use and advocacy)(e.g., Indian Child Welfare Act, Title VII, Johnson-O'Malley Act)
School Culture, Educational Policy, & Institutional Advocacy	In serving NA/AI (e.g., examine referrals for cultural misinformation, implement systemic interventions; understand intergenerational effects of historical trauma)
Psychoeducational Assessment	With NA/AI (sufficient range of tools; consider language development, acculturation, historical trauma; tribal history)
Academic, Therapeutic, & Consultative Interventions	Using NA/AI specific knowledge and skills (e.g., cross-cultural or culture specific consultation or counseling; use culture in intervention design; understanding tribal behavioral norms and taboos on discussions of certain symbols, events and stories)
Working with Translators & Interpreters	In relation to work with NA/AI (e.g., using translators vs. interpreters; know relevant professional guidelines and policies; hire, train & work with interpreters or translators)
Research	With NA/AI populations (e.g., skills in eliminating bias; differentiate tribal specific & mainstream cultural behaviors; aware of/use tribal research protocols or review procedures; know/use indigenous research methodologies)

[1]Source: Rogers et al. (1999)

RESULTS

A wide range of group sizes yielded unequal variances between groups. Due to their robustness in these situations, as well as with non-normally distributed data, nonparametric statistics for ordinal data were used in analysis. Specifically, the Kruskal-Wallis test, an independent group comparison test of

sampled data, the Mann-Whitney U test for paired and independent samples, and Kolmogorov-Smirnov one sample tests were used, along with Bonferroni correction factors to correct for family-wise errors.

Comparisons of Competence from Training, Current Competence, and Competence Needed

Of the 403 total respondents, between 352 and 362 rated their levels of competence in training, current practice, and levels needed for effective school psychological service with NA youth in each of the six practice areas. Table 2 presents the results of Kolmogorov-Smirnov one sample tests and correlations. It addresses research questions a and b regarding how school psychologists rate their training, current competence, and perceived need for competence in each area of practice in relation to serving NA youth, and whether significant differences between those ratings exist.

Table 2 *Kolmogorov-Smirnov One Sample Nonparametric Tests for Differences Between Training, Current Knowledge and Level of Competence Needed Reported by School Psychologists*

Area of Competence	Measure	*n*	Mean (SD)	*z*
Legal-Ethical	Training	360	.72 (.84)	5.270*
	Current	359	.94 (.81)	5.391*
	Needed	356	1.90 (.89)	4.328*
School Culture, etc.	Training	359	.72 (.84)	5.503*
	Current	360	.91 (.86)	4.805*
	Needed	356	1.85 (.91)	4.484*
Psychoed. Assessment	Training	358	.85 (.88)	4.767*
	Current	356	1.07 (.92)	5.111*
	Needed	352	2.01 (.90)	4.679*
Interventions	Training	360	.79 (.86)	4.777*
	Current	359	.95 (.88)	4.872*
	Needed	357	1.92 (.95)	4.647*
Work with Translators	Training	362	.45 (.74)	7.578*
	Current	361	.63 (.80)	6.248*
	Needed	360	1.71 (1.01)	4.610
Research	Training	362	.44 (.77)	7.798*
	Current	362	.56 (.78)	6.611*
	Needed	357	1.73 (1.01)	4.413*

* *p = .000 (2-tailed)*
Note: Ratings Code: 0 = None; 1 = Some; 2 = Adequate; 3 = Strong.

Significant differences between school psychologists' ratings of their competence gained in training, their ratings of current competence, and of perceived need for competence to serve NA youth were found in all six areas. Participants' mean ratings of their competence gained in training ranged from .44 to .85, ratings that fall in the "none" to "some" range. Respondents' mean ratings of current competence were consistent with the "none" to "some" descriptors (i.e., .78 to 1.07). In contrast, their mean ratings of perceived need for NA-specific competencies ranged from 1.71 to 2.01 or "some" to "adequate." Psychoeducational Assessment received the highest ratings for training, current, and needed competence; research received the lowest ratings for competence, from training and current competence in practice. Overall, ratings of current competence were significantly higher than competence gained in training; the need for NA-specific competencies was higher than both, with one exception, Work with Translators (which had the lowest rating for needed competence).

Length of Time Working with NA Youth

To investigate the impact of longevity of experience serving NA youth on the development of NA-specific competencies, we first conducted a Mann-Whitney U test for paired and independent samples for the subgroups who answered "yes" and "no" to the question "Do you work with NA youth?" Table 3 demonstrates that those who answered "yes" rated their levels of competence in training, current practice, and levels of competence they believe are needed for effective practice at significantly higher levels than those who indicated they had not worked with NA youth. Perceived levels of competence working with translators was the sole exception. We then examined these variables within the groups that reported working with Native American youth.

Table 3 *Mann-Whitney U Tests for Significance of Differences in Competence Between Respondents Who Work or Don't Work with NA Children and Youth*

Area of Competence	Measure	Total Mean (SD)	Work with NA n = 95	Don't Work with NA n = 265	*Z*-values
			Mean Rank	Mean Rank	
Legal-Ethical	Training	.72 (.84)	240.08	159.14	-7.18***
	Current	.94 (.80)	247.22	156.15	-8.00***
	Needed	1.90 (.88)	202.06	170.05	-2.73**
School Culture, etc.	Training	.72 (.84)	236.01	160.13	-6.65***
	Current	.91 (.86)	253.02	154.87	-8.44***
	Needed	1.85 (.91)	198.38	171.26	-2.32*
Psychoed. Assessment	Training	.85 (.88)	239.70	158.07	-7.08***
	Current	1.07 (.92)	243.16	155.30	-7.58***
	Needed	2.01 (.90)	195.811	169.57	-2.27*
Interventions	Training	.79 (.86)	235.13	161.20	-6.43***
	Current	.95 (.88)	250.16	155.11	-8.17***
	Needed	1.92 (.95)	203.75	170.15	-2.86**
Work with Translators	Training	.45 (.74)	219.32	168.04	-4.95***
	Current	.63 (.80)	225.37	165.15	-5.38***
	Needed	1.71 (1.01)	187.83	177.87	-.84
Research	Training	.44 (.77)	238.71	161.14	-7.65***
	Current	.56 (.78)	245.88	158.59	-7.97***
	Needed	1.73 (1.01)	199.80	171.46	-2.39*

** p < .05 (two-tailed) , ** p < .01 (two-tailed), ***p < .001 (two-tailed)*

A total of 117 respondents indicated that they worked with NA youth; 93 to 95 participants rated their levels of competence in training, in current practice, and needed competence. Respondents who worked with NA youth for ≤1 year were significantly different from those who worked 2-5 years, 6-10 years and ≥11 years in levels of competence attained from training in four of the six areas: Legal-Ethical, Interventions, Working with Translators, and Research (see Kruskal-Wallis results in Table 4). An interesting pattern emerges. The highest mean ranks for competence from training are associated with participants having 6-10 years of experience with NA youth; those with ≥11 had slightly lower mean ranks, yet consistently higher than those with ≤1 to 5 years experience. The impact of longevity of experience on current competence is significant for all of the areas except Work with Translators. Experience did not, however, make a difference in the ratings of competencies needed to serve NA children and youth.

Table 4 *Kruskal-Wallis One Way Analyses of Variance by Ranks: Differences of Rank-Order of Competencies by Years of Experience Serving NA/AI Children and Youth*

Area of Competence	Measure	*n*	Mean Ranks				*z*
	Years of Experience		≤ 1 yr n = 14	2-5 yr n = 24	6-10 yr n = 19	11+ yr n = 38	
Legal-Ethical	Training	94	30.79	46.67	53.76	52.30	8.72*
	Current	94	26.96	44.92	57.39	52.01	13.63*
	Needed	94	40.50	50.35	52.11	46.09	2.08
School Culture, etc.	Training	93	33.32	44.20	55.97	50.67	7.08
	Current	93	30.14	42.91	50.89	54.87	10.81*
	Needed	93	52.71	50.48	48.50	44.60	1.39
Psychoed. Assessment	Training	93	36.35	40.56	54.86	52.09	7.34
	Current	93	32.15	40.90	51.08	55.03	10.09*
	Needed	93	49.96	45.48	49.72	45.65	0.62
Interventions	Training	94	34.18	41.15	59.47	50.26	10.04*
	Current	94	27.39	40.65	59.55	53.03	16.46*
	Needed	94	50.14	47.13	50.87	45.07	0.85
Work with Translators	Training	95	31.43	45.58	59.47	49.89	10.17*
	Current	95	31.14	47.13	52.79	52.37	7.50
	Needed	95	60.14	44.40	50.13	44.74	4.14
Research	Training	95	29.43	47.02	53.24	53.24	9.30*
	Current	95	28.00	46.04	52.59	52.59	11.54*
	Needed	95	52.07	49.91	44.92	46.91	.82

** p < .05 (two-tailed)*

Respondents who had worked for one year or less reported significantly less competence than participants with two or more years of experience (see results of Mann-Whitney tests in Table 5). Respondents who worked with NA populations for 6-10 years (n = 19) indicated higher levels in training in Working with Translators (Z = 3.33, p = .001) than those who had worked ≤ 1 year (n = 14); and higher levels of current competence than those who had worked ≤ 1 year (n = 14) in three areas: Legal-Ethical (Z = 3.21, p = .003), Interventions (Z = 3.44, p = .001), and Research (Z = 2.97, p = .003). The contrast between competence in current practice of those who have worked ≥ 11 years with NA populations and those who have worked ≤ 1 year was significantly different in five areas: Legal-Ethical (Z = 3.26, p = .001), School Culture (Z = 3.08, p = .002), Assessment (Z = 2.72, p = .006), Interventions (Z = 3.33, p = .001) and Research (Z = 3.02, p = .003). Thus, the length of time working with NA youth makes a significant difference in perceived competence in training, current practice, and in perceived need in each domain.

Table 5 *Mann-Whitney U Tests for Significance of Differences Across Rated Competence by Years of Experience Serving NA/AI Children and Youth*

Area of Competence	Measure	n	Mean (SD)	Z-values			
		Years Experience		≤ 1 yr n = 14	2-5 yr n = 24	6-10 yr n = 19	11+ yr n = 38
Legal-Ethical	Training	94	1.23 (.90)		-1.86	-2.71	-2.65
	Current	94	1.52 (.83)	--	-2.28*	-3.21*	-3.26*
	Needed	94	2.85 (1.12)		-1.20	-1.37	-.63
School Culture, etc.	Training	93	1.22 (.93)		-1.32	-2.51	-2.08
	Current	93	1.56 (.85)	--	-1.52	-2.38*	-3.08*
	Needed	93	2.04 (.79)		-.22	-.46	-1.05
Psychoed. Assessment	Training	93	1.39 (.89)		-.67	-1.77	-1.94
	Current	93	1.67 (.86)	--	-1.17	-1.99*	-2.72*
	Needed	93	2.22 (.75)		-.52	-.44	-.53
Interventions	Training	94	1.27 (.91)		-.82	-2.67	-2.20
	Current	94	1.61 (.88)	--	1.67	-3.44*	-3.33*
	Needed	94	2.18 (.80)		-.34	-.08	-.64
Work with Translators	Training	95	.80 (.93)		-1.88	-3.33	-2.17
	Current	95	1.04 (.96)	--	-2.14*	-2.42*	-2.37*
	Needed	95	1.79 (.97)		-1.88	-1.01	-1.85
Research	Training	95	.94 (.95)		-2.14	-2.31	-3.02
	Current	95	1.09 (.90)	--	-2.23*	-2.97*	3.00*
	Needed	95	1.96 (.85)		-.25	-.79	-.64

* $p < .05$ *(two-tailed)*

Note: Mean Ranks are presented for each group on each measure in Table 3.

Ethnic Differences

We used Mann-Whitney tests to examine the differences in competence in the six areas reported by ethnic groups. Because of relatively low numbers in ethnic subgroups, Native American (3.9%), African American (3.1%), Latino (5.3%), dual ethnicity not including Native American (1.2%), Asian/Pacific Islander/Pilipino (1.9%), and Other (1.7%), we re-coded the combined group People of Color (POC). Table 6 presents these results for 342 European Americans (EA) and 61 People of Color (POC).

POC rated their current competence as higher than that reported by EAs in all six areas. They also rated their competencies from training higher than EAs in four of the six areas: School Culture, Assessment, Interventions, and Work with Translators. Significantly, POC give greater weight than EAs to the importance of NA-specific competence in Legal-Ethical, School Culture, Assessment, and Interventions.

Table 6 *Mann-Whitney U Tests for Significance of Differences Between Rated Competence by People of Color and European Americans*

Area of Competence	Measure	Total Mean (SD)	People of Color n = 71		European American n = 342		*Z*-values
			Mean Rank	Mean (SD)	Mean Rank	Mean (SD)	
Legal-Ethical	Training	.72 (.84)	194.00	.92 (1.0)	173.45	.67 (.79)	-1.57
	Current	.94 (.80)	209.23	1.25 (.94)	169.64	.86 (.76)	-3.02*
	Needed	1.90 (.88)	210.69	2.21 (.86)	167.44	1.83 (.87)	-3.22*
School Culture, etc.	Training	.72 (.84)	205.09	1.03 (1.04)	170.63	.65 (.78)	-2.61*
	Current	.91 (.86)	209.02	1.23 (1.0)	170.31	.84 (.81)	-2.89*
	Needed	1.85 (.91)	196.80	2.07 (.91)	170.38	1.82 (.91)	1.96*
Psychoed. Assessment	Training	.85 (.88)	202.31	1.16 (1.10)	170.47	.78 (.82)	-2.40*
	Current	1.07 (.92)	213.35	1.49 (1.10)	166.88	.97 (.86)	-2.40*
	Needed	2.01 (.90)	199.93	2.23 (.95)	167.33	1.89 (.88)	-2.45*
Interventions	Training	.79 (.86)	202.24	1.07 (1.05)	171.73	.71 (.79)	-2.31*
	Current	.95 (.88)	215.13	1.34 (1.05)	169.03	.85 (.81)	-3.44*
	Needed	1.92 (.95)	204.46	2.16 (1.00)	170.01	1.88 (.93)	-2.55*
Work with Translators	Training	.45 (.74)	217.38	.84 (.99)	169.83	.37 (.65)	-3.99*
	Current	.63 (.80)	214.88	.98 (.97)	169.87	.55 (.75)	-3.47*
	Needed	1.71 (1.01)	191.98	1.85 (1.06)	173.87	1.68 (.99)	-1.32
Research	Training	.44 (.77)	196.55	.62 (.92)	174.15	.39 (.73)	-1.92
	Current	.56 (.78)	211.15	.87 (.96)	171.12	.48 (.72)	-3.17*
	Needed	1.73 (1.01)	196.38	1.93 (1.06)	171.18	1.69 (1.0)	-1.83

** $p < .05$ (two-tailed)*

DISCUSSION

School Psychologists Currently Not Prepared

The study's primary finding is the striking disparity between the skills and level of competence respondents perceive they gained in training in all six domains of practice investigated, those developed since training, and those perceived as needed to adequately serve NA youth. School psychologists believe they are significantly under-prepared in training, and remain significantly under-prepared in practice to serve this population, despite whatever on-the-job and formal professional growth experiences they might have had. Further, they report needing significantly more skills to be effective in practice, even at just an adequate level.

Significant differences between self-rated knowledge and skills from training and current levels of competence clearly point to gaps in training, yet also indicate that ongoing professional development, or access to it, has been insufficient to meet perceived needs. Without sufficient knowledge and skills to provide appropriate service delivery, NA achievement, school completion, and appropriate identification for special education understandably suffer.

Ethnic Differences in Perceptions of Competence

Calls for more NAs to work in education have come from government agencies, researchers, and NA parents (c.f., NCES, 2008). To determine whether NAs would be better prepared to work with Native youth, we examined ethnic differences. While NA school psychologists in the current study reported they were better prepared to work with NA, numbers were insufficient for statistical verification; thus, data from non-European American (EA) respondents were collapsed into a group named People of Color (POC). In all six domains of knowledge/skill and in four domains of training, POC rated their preparation to work with NA youth higher than did EAs.

Longevity with the Population a Significant Variable

The data on longevity working with Native youth yields a different look at contributions to competence. Respondents who said they work with NAs (vs. those who did not) reported significantly higher levels of competence from training, current practice, and in levels needed for effective practice. This is troublesome when higher over-representation in special education for Native youth occurs within "low-density" schools, where invisibility is likely. In states such as California, where the numbers of Native

American youth are the highest in the nation, but percentages in individual schools are low, this likelihood is enhanced. This suggests that urban and suburban school psychologists may need at least as much, if not more preparation to work with Native youth than those working with higher percentages in reservation communities.

With increased experience, reported levels of competence in current practice increase significantly. Compared to those with one or fewer years of experience, those reporting 6-10 years experience with NA youth reported significantly higher current competence in three areas. With ≥11 years' experience, five areas are rated stronger. To some degree this finding is expected; some prior research finds that experience working with diverse clients predicts multicultural skills (e.g., Allison, Echemendia, Crawford, & Robinson 1996; Arthur & Januszkowski, 2001). Yet, respondents well experienced with NAs still did not report "adequate" levels of competence in current practice.

Implications for Training and Practice in School Psychology

Significant differences in virtually all areas between what school psychologists know, either through training or experience, and what they need to know has profound implications for training, continuing professional development, and ethical practice. Some level of additional preparation seems essential. In addition, this finding indicates a high level of professional awareness of the gaps, suggesting openness to intensified professional training. Because school psychologists working for longer periods of time with Native American youth and those of ethnically diverse backgrounds appear to have a different level of knowledge or skill, and perhaps a unique lens from which to understand the issues, solicitation of their input, contributions, and/or leadership will be important. It may seem intuitive to focus this concern for more professional development work to those serving high density AI/AN areas, as those working with fewer numbers of Native youth express less need; however, demographic distributions suggest that extremely large numbers of AI/AN youth are distributed across urban and suburban regions in which they are even more likely to become statistics of over-representation and misplacement – victims of the "invisibility" problem. It is critical that the profession not perpetuate the tendency to "absorb" Native issues into a larger multicultural context, or to defer the issue to high-density areas.

Over a decade ago, the U.S. Department of Education proposed increasing the number of Indian educators as a national priority (Indian Nations at Risk Task Force, 1991). Despite funding and training initiatives, shortages persist. Continued recruitment and retention of those who bring culturally diverse lived experiences, worldviews, perspectives, and strategies into the profession appears essential. Nonetheless, that alone will not remedy the problem. Because even the most experienced practitioners expressed the need for significantly more information and skills for work with Native youth, determining not only what to do but how to do that becomes paramount.

A Problem of Paradigm?

The multicultural literature has been dominated by the notion of conceptual categories of awareness, knowledge and skills (Fowers & Davidov, 2006), yet Smith et al. (2006) found skill development often ignored in most coursework. Lists of competencies have been developed to guide enhanced multicultural training (c.f., Holcolm-McCoy & Myers, 1999; Rogers & Lopez, 2002). However, as Ridley (2008) points out, such lists can be "descriptive, but not prescriptive," (p. 446); they tell people what to do, but not how to learn to do it. In addition to this problem, the all-too-frequent practice of teaching approaches that survey one culture per day or chapter, called "cultural tourism" by Perry (2002) can lead to portraying people of color with "tightly bound fictive identities" that tend to stereotype, and leave no room for difference or flexibility in relationship building (p. 197). Green, Cook-Morales, Robinson-Zañartu and Ingraham (2009) suggest that learning to shift behaviors with cultural competence requires a shift in paradigms and synergistic attention to conceptual depth, cultural context, and experience, as well as appropriate depth of affect.

Although it is not within the scope of this paper to propose an alternative paradigm, we do propose that alternative models should be sought and tested, and that at least three components be considered for skill development alongside relevant knowledge: (a) a process orientation that is more holistic in nature,

going beyond cognitive constructions of learning; (b) consideration of multiple and salient identities; and (c) worldview differences.

Native Americans come from a multitude of tribal communities and language groups; some are from more than one tribe or may not speak their traditional language, and affinity for traditional values and acculturation levels vary widely. Within the school psychology literature, Ingraham (2000) discusses the critical role of attending to both individual differences within cultural groups and the multiple cultural identities prevalent in many individuals in school-based consultation. She and others (e.g., Monk, Winslade & Sinclair, 2008) discuss the role of cultural saliency in developing respectful relationships needed to effectively create interventions. Further, Arredondo and Arciniega (2001) suggest that without developing the ability to understand another worldview, cultural competence would be out of reach, since culturally formed worldviews drive assumptions and operating norms.

Native American history has been fraught with a clash of cultures. Much of this clash has resulted in significant losses, of land, culture, language, and even entire tribes. Despite this, Native Americans have demonstrated not only survival, but also amazing resilience. The results of this survey acknowledge a professional self-awareness that despite mandates for cultural sensitivity in training and practice, insufficient knowledge is available to understand the complexity of issues faced by Native Americans. If we apply our understanding of systems and ecology, combined with this acknowledged limitation in professional capacity, it is possible to create the next wave of cultural sophistication in professional training and practice.

LIMITATIONS

This study measured self-assessment of competence in training, practice, and perceived competence needed for effective practice; thus, limitations inherent in self-perception measures versus outcomes data suggest caution in interpretation. Two other limitations in this study relate to population demographics: first, the vast majority of school psychologists are European American females. Because of this ethnic and gender disparity, a wide range of group sizes and variances exist between ethnic and gender sub-groups. Non-parametric statistics were used because of their robustness in these situations; however, some of these findings may be over- or under-inflated. Second, deliberate over-sampling to obtain higher numbers who serve NA youth, along with the fact that respondents were volunteer participants from requests sent to a random sample NASP members yielded a study sample that is not truly random. However, the sample and sampling method reflected the challenges that this demographic imbalance presents, and provides important implications that will be examined in the following discussion. Finally, reliability and validity of the survey itself could have been strengthened. Although the survey drew on literature and on the experiences of Native American psychologists to establish face validity and content validity, and attention was paid to an adequate sample size, measures of reliability to control for different response patterns or effects of an issue's position in the survey were not used. No statistics were employed to establish internal reliability. To the extent that these issues may have compromised the reliability of the survey, caution in interpretation of results is warranted.

DIRECTIONS FOR FUTURE RESEARCH

Even experienced practitioners expressed the need for significantly greater levels of competence working with Native youth. In that context, several directions for future research emerge from this study. First, we propose that alternative models be sought and tested for training that will help graduates be far better prepared to identify NA students, and be equipped with a range of information and tools to serve them. Concurrent with this, research on and development of best practice standards is needed in the six areas identified above with NA populations. The second area relates to current practitioners' belief that they are inadequately prepared to serve NA youth. The following research questions are germane: (a) Is continuing professional development (CPD) available to school psychologists in this area? (b) If it is available or were it to become available, would those serving NA youth on or near rural reservations or in other large population centers take advantage of such CPD? Large numbers of NA students represent small percentages in urban schools and districts; thus would school psychologists serving that population take advantage of such CPD? Finally, how would enhanced skill and information development make a difference in service to NA children and youth?

Findings have been inconsistent regarding the effects of ethnicity on self-reports of multicultural competence (Manese, Wu & Nepomuceno, 2001; Pope-Davis, Reynolds, Dings & Nielson 1995; Sodowsky, Kuo-Jackson, Richardson & Corey, 1998). We posit three possibilities regarding why our respondents of color (POC) reported significantly higher levels of preparation to work with NA youth than EAs. First, people of color bring lived experience to their training and practice that may add a proclivity (a) to seek out culturally relevant tools and skills, and (b) to build on a practice of interpreting the "white world" through a cultural lens. These experiences involve skills in a kind of cross-cultural translation, which may transfer to and augment their school psychology training and practice. Second, POC may have rated the need for higher levels of competence for effective practice with themselves in mind; that is, their own experiences of schooling compelled higher levels of need for competence with all POC, including Native youth, than experienced by EAs. Each of these hypotheses may be best explored using qualitative research methods. A third possibility, not mutually exclusive of the first and second, is that as a group, POC attended training programs with more depth in cross-cultural issues. Because we did not solicit those data (which respondents attended specific training programs), this hypothesis (as do the others) suggests an additional area for future research.

Alternate approaches to assessment of competence to work with NA youth (or other specific groups) beyond self-report need to be considered; for example, using case scenarios designed to pull for information from each of the six areas of competence. Ultimately, it will be important to examine the assessments of competence (however measured) to real outcomes for the targeted group in the schools of the participating school psychologists. This challenge is embedded in NASP's accreditation standards, in that training programs are expected to document positive outcomes in the schools. As faculty members design such program outcomes studies, we urge them to attend to evaluating the link between cultural competence and outcomes for culturally and linguistically diverse preK-12 students.

- - -

Carol Robinson-Zañartu, PhD., is a professor of School Psychology at San Diego State University. Her research, publications, and consultations focus on issues of equity and social justice, and on dynamic intervention-based assessment. She holds special interest (and numerous federal grants) in preparing school psychologists to serve Native American schools and communities.

Nola Butler Byrd, PhD., is a scholar/activist, and director of San Diego State University's Community-Based Multicultural Counseling Program. Her research interests include experiential multicultural counseling and community practice for social justice, Indigenous healing and worldview, and multicultural identity development.

Valerie Cook-Morales, PhD., is a professor and chair of the Department of Counseling and School Psychology at San Diego State University. She specializes in educational equity – from kindergarten through graduate school – from individual through systems levels, with emphasis on delivery of services for culturally-linguistically diverse students, especially bilingual Latino students, with or at risk of disabilities.

Elvina Charley, Ed.S., is a Diné School Psychologist in Ganado, Arizona, working with Native American youth and families. She is fluent in Navajo language and culture, and is experienced providing service to a number of rural community and public schools serving Native youth. She serves as chair of the Native American Work Group for the National Association of School Psychologists.

Paul Dauphinais, PhD., is an enrolled member of the Turtle Mountain Band of Chippewa Indians. Within his home reservation, Dr. Dauphinais serves approximately 3700 children as a school psychologist, providing academic and behavior assessments and interventions from children and youth from ages 3-21. He most enjoys working with very young children.

Mike Bonner, PhD., is an independent practitioner of community psychology, providing organizational development consultation and program evaluation to non-profit organizations that provide community-based services. He is also adjunct instructor of psychology, primarily with Hawai'i Pacific University, providing online courses.

REFERENCES

Allison, K.W., Echemendia, R.J., Crawford, I., & Robinson, W.L. (1996). Predicting cultural competence: Implications for practice and training. *Professional Psychology: Research and Practice, 27*, 386-393. doi:10.1037/0735-7028.27.4.386

American Psychological Association (1990). *Guidelines for providers of psychological services to ethnic, linguistic, and culturally diverse populations*. Washington, DC: Author.

American Psychological Association (2003). *Guidelines on multicultural education, training, research, practice, and organizational change for psychologists*. Washington, DC: Author.

Arredondo. P. & Arciniega, G. (2001). Strategies and techniques for counselor training. *Journal of Multicultural Counseling and Development, 26*, 63-73.

Arthur, N., & Januszkowski, T. (2001). The multicultural counseling competencies of Canadian counselors. *Canadian Journal of Counseling, 35*, 36-48.

Cleary, L.M. & Peacock, T.D. (1998). *Collected wisdom: American Indian education*. Boston: Allyn & Bacon.

Cook-Morales, V.J. & Robinson-Zañartu, C.A. (1995, July). Model strategy: Systemic change toward educational equity in professional psychology programs. *Communiqué*, pp. 9-13. Washington, D.C.: Office of Ethnic Minority Affairs, Public Interest Directorate, American Psychological Association.

Deloria, V. & Wildcat, D.R. (1991). *Power and Place: Indian education in America*. Golden, CO: Fulcrum Group.

Faircloth, S.C., & Tippeconnic, III, J.W. (2010). *The Dropout/Graduation Rate Crisis Among American Indian and Alaska Native Students: Failure to Respond Places the Future of Native Peoples at Risk*. Los Angeles, CA: The Civil Rights Project at UCLA. Retrieved from http://www.civilrightsproject.ucla.edu/news/pressreleases/pressrelease20100217-dropouts.html

Fowers, B.J., & Davidov, B.J. (2006). The virtue of multiculturalism: Personal transformation, character, and openness to the other. *American Psychologist, 61*, 581-594. doi:10.1037/0003-066X.61.6.581

Green, T.D., Cook-Morales, V.J., Robinson-Zañartu, C.A., & Ingraham, C.L. (2009). Pathways on a journey to getting it: Multicultural competence training and continuing professional development. In J. Jones (Ed.). *The psychology of multiculturalism in schools: A primer for practice, training, and research* (pp. 83-113). Bethesda, MD: National Association of School Psychologists.

Holcomb-McCoy, C., & Myers, J. (1999). Multicultural competence and counselor training: A national survey. *Journal of Counseling and Development, 77*, 294-302.

Indian Nations at Risk Task Force (1991). *Indian nations at risk: An educational strategy for action*. Final Report of the Indian Nations At Risk Task Force. Washington, DC: U.S. Department of Education.

Ingraham, C.L. (2000). Consultation through a multicultural lens: Multicultural and cross-cultural consultation in schools. *School Psychology Review, 29*, 320-343.

Isaac, S., & Michael, W.B. (1981). *Handbook in research and evaluation* (2nd ed.). San Diego, CA: EdITS.

Klug, B.J., & Whitfield, P.T. (2003). *Widening the circle: Culturally relevant pedagogy for American Indian Children*. New York, NY: RoutledgeFalmer.

Manese, J.E., Wu, J.T. & Nepomuceno, C.A. (2001). The effect of training on multicultural competences: an exploratory study over a ten year period. *Journal of Multicultural counseling and Development, 29*, 31-40.

Monk, G., Winslade, J. & S. Sinclair, S. *New horizons in multicultural counseling*. Sage: Thousand Oaks.

National Association of School Psychologists (2000). *Standards for training and field placement programs in school psychology*. Bethesda, MD: Author.

National Center for Education Statistics (NCES) (2008). *Status and trends in the education of American Indians and Alaska Natives*. U.S. Department of Education Institute of Education Sciences. Retrieved from http://nces.ed.gov/pubs2008/nativetrends/

Orfield, G., Losen, D., Wald, J., & Swanson, B. (2004, February). *Losing our future: How minority youth are being left behind by the graduation rate crisis*. Retrieved from http://www.urban.org/url.cfm?ID=410936

Palmer, D.J., Juarez, L. & Hughes, J.N. (1991). School psychology training and the education of minority at-risk youth: The Texas A & M university program emphasis on handicapped Hispanic children and youth. *School Psychology Review, 20*, 472-484.

Perry, P. (2002) *Shades of white: White kids and racial identities in high school*. Durham, NC: Duke University Press.

Pope-Davis, D.B., Reynolds, A.L., Dings, J.G. & Nielson, D. (1995). Examining multicultural counseling competencies of graduate students in psychology. *Professional Psychology: Research and Practice, 26*, 322-329. doi: 10.1037/0735-7028.26.3.32

Rampey, B.D., Lutkus, A.D., & Weiner, A.W. (2006). *National Indian Education study, Part I: The performance of American Indian and Alaska Native fourth- and eighth-grade students on NAEP 2005 Reading and Mathematics Assessments* (NCES 2006-463). Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

Ridley, C.R. (2008). Response to Chapter 17. In G. Monk, J. Winslade & S.L. Sinclair, *New Horizons in Multicultural Counseling* (pp. 446-447). Thousand Oaks, CA: Sage.

Robinson-Zañartu, C. (1996). Serving Native American children and families: Considering cultural variables. *Language, Speech, and Hearing Services in Schools, 27*, 373-384.

Rogers, M. ., Ingraham, C.L., Bursztyn, A., Cajigas-Segredo, N., Esquivel, G., Hess, R., Nahari, S.G. (1999). Providing psychological services to racially, ethnically, culturally, and linguistically diverse individuals in the schools: Recommendations for practice. *School Psychology International Journal, 20*, 243-264. doi: 10.1177/0143034399203001

Rogers, M.R. & Lopez, E.C. (2002). Identifying critical cross-cultural school psychology competencies. *Journal of School Psychology, 40*, 115-141. doi:10.1016/S0022-4405(02)00093-6

Rogers, M.R. & Molina, L.E. (2006). Exemplary efforts in psychology to recruit and retain graduate students of color. *American Psychologist, 2*, 143-156. doi:10.1037/0003-066X.61.2.143

Rogers, M.R., Ponterotto, J.G., Conoley, J.C., & Wiese, M.J. (1992). Multicultural training in school psychology: A national survey. *School Psychology Review, 21*, 603-616.

Smith, T.B., Constantine, M.G., Dunn, T., Dinehart, J., & Montoya, J.A. (2006). Multicultural education in the mental health professions: A meta-analytic review. *Journal of Counseling Psychology, 53*, 132-145. doi:10.1037/0022-0167.53.1.132

Sodowsky, G.R., Kuo-Jackson, P.Y., Richardson, M.F., & Corey, A.T. (1998). Correlates of self-reported multicultural competencies: Counselor multicultural social desirability, race, social inadequacy, and locus of control, racial ideology, and multicultural training. *Journal of Counseling Psychology, 45*, 256-264. doi: 10.1037/0022-0167.45.3.256

Smith, T.B., Constantine, M.G., Dunn, T., Dinehart, J., & Montoya, J.A., (2006). Multicultural education in the mental health professions: A meta-analytic review. *Journal of Counseling Psychology, 53*, 132-145. doi: 10.1037/0022-0167.53.1.132

Stancavage, F.B., Mitchell, J.H., Bandeira de Mello, V., Gaertner, F.E., Spain, A.K. & Rahal, M.L. (2006). *National Indian Education Study: Part II: The Educational Experiences of Fourth- and Eighth-Grade American Indian and Alaska Native Students*. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences.

Sue, D.W. (2001). Multidimensional facets of cultural competence. *The Counseling Psychologist, 29*, 790 - 821. doi: 10.1177/0011000001296002

Tarver Behring, S. & Ingraham, C.L. (1998). Culture as a central component of consultation: A call to the field. *Journal of Educational and Psychological Consultation, 9*, 57-72. doi: 10.1207/s1532768xjepc0901_3

U.S. Census 2000 Summary File. Retrieved from http://www.wwwcensusgov.zuom.info/Press-Release/www/2004/AIANSF.html

U.S. Department of Education [DOE] Office of Special Education Programs (2003). Data Tables for State Reported Data. Retrieved from http://www.ideadata.org/arc_toc6.asp

TOC

A Bilingual (English & Spanish) Psychoeducational Assessment MODEL Grounded in Cattell-Horn Carroll (CHC) Theory: A Cross Battery Approach

Pedro Olvera and Lino Gomez-Cerrillo,
Azusa Pacific University

The Individual with Disabilities Education Act mandates nondiscriminatory assessment for children who are culturally and linguistically diverse. Because of the overrepresentation of English Language Learners (ELL) in special education programs in the United States, the focus of this article is the psychoeducational assessment of Spanish- and English-speaking children who are classified as ELL. This article presents a bilingual assessment model that incorporates Cattell-Horn Carroll (CHC) based instruments. The premise of this model is that a learning disability is manifested in L1 (primary language) and L2 (secondary language). The writers present cognitive assessments that are available in English, Spanish, and nonverbal modalities that utilize CHC as the underlying theory. By incorporating these assessments, the school psychologist is in a better position to analyze L1 and L2 assessment data and gain a clearer understanding of strengths and weaknesses and provide linguistically appropriate interventions.

KEYWORDS: bilingual assessment, cultural linguistic diversity (CLD), Cattell-Horn Carroll (CHC) theory, nondiscriminatory assessment, cross battery assessment, learning disability.

School psychologists across the United States are asked to assess children from various cultural and linguistic backgrounds. Although a myriad of non-English languages are spoken, Spanish is the most common (Ochoa, Riccio, & Jimenez, 2004). As this trend is also evident in the educational system, it is imperative that school psychologists are prepared to adequately assess and intervene with children who are culturally and linguistically diverse (CLD). In many cases, children who are in the process of learning English appear to have the same learning challenges as those children who are suspected of having learning disabilities (Diaz-Rico & Weed, 2002). Accordingly, an overrepresentation of CLD children within special education programs has been documented (Oswald & Coutinho, 2001). This trend is especially evident when considering English language learners (ELL) placed in special education programs (Artiles, Rueda, Salazar, & Higareda, 2005). In addition, using the traditional discrepancy model and a limited scholastic review of students has contributed to educational misplacement and poor intervention (Abedi, 2008). Thus, school psychologists need to be equipped with nondiscriminatory assessment procedures to adequately differentiate normal English language development (ELD) manifestations from specific learning disabilities (SLD).

Part of identifying SLD in bilingual children involves conducting a bilingual assessment. A bilingual assessment is one conducted in both L1 (primary language) and L2 (secondary language) by a qualified school psychologist who is fluent in both languages, or a monolingual English school psychologist working with a qualified interpreter (Rhodes, Ochoa, & Ortiz, 2005). A qualified interpreter is one who is fluent in the language and has firsthand knowledge of the culture. Although beyond the scope of this discussion, interpretation requires that the interpreter accurately express not only the content but also the emotional/affective nuances of the language. Interpreters need the following skills: to maintain neutrality and confidentiality; to adhere to ethical guidelines; to understand technical terms, procedures and rationale of the process of assessment; culture and language experience; and knowledge of key/critical issues that may arise (Rhodes et al., 2005). Lopez (2002) provides a list of suggested practices for school psychologists utilizing interpretation services.

Correspondence concerning this article should be addressed to Pedro Olvera and Lino Gomez-Cerrillo, Azusa Pacific University, School of Education, Department of School Counseling and Psychology, 701 East Foothill Avenue, PO Box 7000, Azusa, CA 91702-7000, email: polvera@apu.edu

Because it is an "empirically supported psychometric theory structure of cognitive abilities and academic abilities" (Alfonso, Flanagan, & Radwan, 2005, p. 185), the writers will utilize Cattell-Horn Carroll (CHC) Theory as an underlying approach for cognitive assessment. In addition, the writers will discuss CHC-based cognitive assessments that are available in English, Spanish, and nonverbal modalities that can assist during the assessment process. The premise of this article is that a learning disability must be manifested in both L1 and L2 for a learning disability to exist. To differentiate a true disability from normal ELL developmental manifestations, the school psychologist must conduct a psychoeducational assessment utilizing nondiscriminatory assessment procedures (IDEA, 1997). To facilitate nondiscriminatory assessment, the writers will present an assessment MODEL (Multiple Sources of Information, Observation, Data-Driven Hypothesis, English Language Development, and Language of Assessment), developed by Olvera and Gomez-Cerrillo (2010), to guide the bilingual and monolingual English school psychologist through the assessment of ELLs.

MULTIPLE SOURCES OF INFORMATION

Multiple sources of information refers to the systematic gathering of information related to the referred student, providing the context of assessment results. This process is foundational to conducting nondiscriminatory assessment. Historically, psychoeducational assessments have not been sensitive to children who are culturally and linguistically diverse (Figueroa & Newsome, 2006). Prior to collecting any data related to a student referred for special education assessment, school psychologists must develop cultural competence while becoming aware of cultural biases that he or she may unknowingly possess (Ortiz, 2008; Oritz, Flanagan, & Dynda, 2008; Rhodes et al., 2005). Maintaining an objective attitude toward referred students ensures that the examiner will make eligibility decisions that are based on objective data rather than those influenced by personal beliefs or biases.

Within MODEL, the assessment process begins with a comprehensive review of the student's cumulative file. The cumulative file contains the student's ELD level and language of instruction. The examiner may want to examine the Home Language Survey, which is given to each parent upon enrolling the child in public schools. This form is included as part of the district's student enrollment packet and documents the student's home language. In addition, the California English Language Development Test (CELDT) is the primary assessment upon which English proficiency is determined in California schools. Upon reviewing the CELDT, the school psychologist is encouraged to note the yearly progression of language development via the CELDT and determine if a pattern of consistently low performance is evident. This includes reviewing the overall CELDT score and individual sub categories: reading, writing, listening, and speaking. If an ELL student is gradually progressing in his or her ELD level, then this may indicate a pattern of appropriate language development (Valdes & Figueroa, 1994). As a cautionary note, if the examinee has consistent difficulties across one or more areas, particularly reading and writing, the examiner must not make the assumption that a disability is present. Instead, other factors must be considered that may possibly contribute to this dynamic. Factors include: proficiency in the primary language, response to intervention implementation, program of instruction (e.g., bilingual instruction/ELD supports), English development support in the family, community variables, etc. In addition, it has been established that limited proficiency in a primary language may impede acquiring a second language (Lau & Blathcey, 2009). Thus, knowledge of the primary language is extremely important.

Other sources of data to examine in the cumulative file include but are not limited to anecdotal notes (e.g., psychological reports), all schools of attendance, retention records, Board of Reviews, Academic Improvement Plans (AIP), attendance, truancy, grades, behavioral comments, discipline records, and achievement results from group normed tests. Particular attention should be noted regarding the language of instruction. That is, has the student been educated in a bilingual or an English-only program? If the examiner is not familiar with the program of instruction, an English as Second Language (ESL) specialist needs to be consulted. In the event that the student attended school in another country, the examiner should request report cards and/or previous assessments from that school.

Another source of data that is imperative to all psychoeducational assessments is the Student Success Team (SST) file. Most SST files incorporate Response to Intervention (RTI) data. This should

all be documented and incorporated in the interpreting of testing data. As school districts begin to incorporate RTI systems, the availability of tier one and two intervention data will provide essential data for Individual Education Plan (IEP) teams to make the best decisions for ELL students regarding placement and further intervention. In addition, Lichtenstein (2008) and others found the incorporation of RTI data within a comprehensive psychoeducational assessment to be the best practice for the identification of learning disabilities for all children, including English learners. Interviews with caregivers must be conducted with great care and sensitivity, with particular attention given to culture and linguistic background (Ortiz et al., 2008). The reader is directed to Rhodes et al. (2005) to understand specific methods and aids to assist in the completion of interviews. Together, this collection of multiple sources facilitates ruling out of exclusionary factors, focuses on understanding the student's cultural and linguistic profile, and provides a foundation to interpret assessment results.

OBSERVATIONS

Observations are crucial aspects of any assessment and are gathered in the student's natural learning environment (Prasse, 2008). Observations should be conducted across multiple settings, adopting structured and systematic methods to document a variety of student dynamics, including language preference, school performance, and, most importantly, an estimate of the student's acculturation or process of acquiring the mainstream U.S. culture (Rhodes et al., 2005). Observing the student both inside and outside the school setting allows the practitioner to gain an estimation of the student's level of acculturation. In addition, the examiner can gain insight into culturally responsive teaching methods, the student's language skills and demands of instruction, and peer support through group work. Additionally, observing the student within the home provides much information regarding family makeup and the cultural environment of the student. In summary, all observations are synthesized with all other sources of data to formulate a data-driven hypothesis or reason for referral.

EXCLUSIONARY FACTORS

While conducting the assessment process, consideration of exclusionary factors is a mandatory function noted in both federal and state law. Although states have variations of what constitutes a learning disability, all 50 states acknowledge the same exclusionary factors (Reschly & Hosp, 2004). Exclusionary factors include environmental, cultural, or economic disadvantage. MODEL also considers limited English language development as a possible exclusionary factor, but not exclusively. Specific factors can include quality of instruction, time in the United States/acculturation level, mobility patterns, proficiency of primary language, social economic status, parent education level, etc. For example, a third grade student is referred to the Student Success Team (SST) due to concerns with language arts and failure to learn English. Upon reviewing the records, the examiner find that the student has had limited levels of English, for the past two years, and has been in this country since he was in kindergarten. The teacher reports that the student is frustrated with his lack of progress. Upon further review of educational records, the examiner notes that the student has never received ELD support and has not participated in the school's response to intervention (RTI) program. The examiner may recommend ELD programming and RTI services targeting language arts that are appropriate for ELLs, monitor progress, and reconvene within a set period of time. In this scenario, the examiner identified factors that may be contributing to lack of academic progress and recommended targeted interventions. When the team regroups, they will be able to assess the student's progress and determine the next steps.

Children, including those classified as ELL, may have learning disabilities that can coexist with exclusionary factors (Lichtenstein, 2008). MODEL considers this dynamic and encourages practitioners to not prematurely rule out the possibility of a learning disability due to poor language proficiency. Thus, it is imperative that school systems have interventions in place that can support English language learners.

Data-Driven Hypothesis

After the practitioner has ruled out difficulties that are primarily due to environmental factors (i.e., lack of instruction or intervention), a *data-driven hypothesis* for assessment is developed. Data-driven

hypothesis refers to the development of hypothesis(es) that assists the practitioner to develop a culturally and linguistically appropriate battery of assessments. The hypothesis(es) for assessment is contingent upon themes collected through multiple sources of information. Themes include, but are not limited to, academic, behavioral, developmental, and/or emotional problems.

Once exclusionary factors have been considered and a data-driven hypothesis(es) has been developed, the null hypothesis must be upheld (Ortiz, 2008). Practitioners should assume that difficulties noted are external; that is, a disability should not be regarded as being the primary reason why academic difficulties exist (Flanagan, Ortiz, & Alfonso, 2007; Rhodes et al., 2005). Incorporating gathered data from multiple sources, along with ecologically valid observations, will provide a data-driven methodology to suggest or reject possible explanations for the student's scholastic difficulties.

Academic Assessment

Academic assessment must be done in a manner that is nondiscriminatory and, consequently, yields accurate information to best intervene. Prior to deciding in which language to assess, the examiner is to review cumulative records and assess historical and current language of instruction. If the student has received instruction in Spanish through a bilingual program or formal education in the native country, then it is sensible to assess academics in Spanish. This information is valuable and will help to determine academic skills in the primary language. However, if the student has only received instruction in English, even though the student is still a designated English Language Learner (ELL), then an English academic assessment is the only option (Lau & Blatchey, 2009). The purpose of achievement testing is to determine academic skills, eligibility decisions, and intervention planning. Thus, this aspect of the assessment process is very important.

If the examinee has been educated in Spanish in another country or through a bilingual program in the United States, then the Bateria III Woodcock Munoz: Pruebas de Aprovechamiento (Munoz-Sandoval, Woodcock, McGrew, & Mather, 2005a) may be an appropriate assessment to administer by school districts that require standardized assessment data. This assessment is to be administered by a trained interpreter or by a school psychologist and/or special education teacher who is fluent in Spanish. Because the Bateria III, like many other Spanish assessments, was normed on Spanish-speaking populations that are not typically represented in United States public schools, the examiner is advised to take extra caution in interpretation of the data, as this may not be an appropriate test for the student.

Another option for academic assessment includes utilizing curriculum-based measurement (CBM) which research supports to measure critical academic skills including but not limited to, oral reading, written expression, and calculation (Burns, MacQuarrie, & Campbell, 1999). The benefit to utilizing CBM is that the examiner can assess academic skills, develop interventions, monitor the student's progress, and modify interventions based on individual progress (Lau & Blatchey, 2009). Another benefit is that schools that utilize CBM may develop local norms that can prove helpful in comparing individual student progress with similar individuals who share the same linguistic and cultural background. Research with ELD children, although not plentiful, is promising (Baker & Good, 1995; Baker, Plascencio-Peinado, & Lezcano-Lytle, 1998; De Ramirez & Shapiro, 2006).

ENGLISH LANGUAGE DEVELOPMENT

Jim Cummins (1979) hypothesized two related language skills: Basic Interpersonal Communications Skills (BICS) and Cognitive Academic Language Proficiency (CALP). Briefly, BICS is the language skill that facilitates communication in social contexts that are typically found in informal settings: conversing with peers, discussing sporting events, and chatting at recess or lunch. Conversely, CALP is a more complex language skill that is required for academic learning. It is important to note that language development, in any language, follows a developmental course. The developmental stages include: CALP 1 (preproduction), CALP 2 (early production), CALP 3 (speech emergence), CALP 4 (intermediate fluency), and CALP 5 (advanced fluency) (Diaz-Rico & Weed, 2002). Hearne (2000) and Roseberry-McKibben (2002) provide examples one can observe as the student moves through the various levels of the aforementioned CALP levels. See Table 1 for CALP examples.

Table 1 *CALP Level Examples*

CALP Stage (1-5)	Examples
CALP Level 1: Preproduction	May engage in educational activities using their first language. Many join in group activities but are not yet able to work independently. May experience the silent period.
CALP Level 2: Early Production	Increasing control of the English tense system and increase in vocabulary. Growth in listening and speaking skills in English, but still need substantial support in most reading and writing activities in the classroom.
CALP Level 3: Speech Emergence	Understand most classroom and social language, and can engage in a variety of oral and written activities. Able to express ideas and feelings in English. Developing reading fluency and understanding, although still in need of support. Learning to write independently.
CALP Level 4: Intermediate Fluency	Able to understand English in many contexts, and have developed into independent readers and writers. May need minor support.
CALP Level 5: Advanced Fluency	Observations and performance would be at level expected from a monolingual English student.

The California Department of Education (CDE; 2009) developed the California English Language Development Test (CELDT) to assess California public education students' English language proficiency and incorporates CALP levels. The CELDT incorporates the following stages of ELD: Beginner (CALP 1), Early Intermediate (CALP 2), Intermediate (CALP 3), Early Advanced (CALP 4), and Advanced (CALP 5).

When assessing children who are still ELL, the school psychologist must assess CALP levels to determine the most appropriate language in which to assess. For example, a student may be conversationally proficient in English (BICS); however, may not be academically proficient in English (CALP). Though on the surface the student may seem to be proficient, it does not become obvious until the examiner conducts a CALP-based assessment (e.g., BVAT or Woodcock Munoz) that the student is still in the process of becoming proficient in English. Without assessing CALP levels, the examiner may unknowingly engage in discriminatory practices by assuming that the student is proficient and conducting an English-based psychoeducational assessment.

Additionally, the examiner will need to consider the degree to which ELD levels, as a variable, can explain the nature and extent of the referred student's learning difficulties. A few questions to consider include:

- Can the student's difficulty in acquiring English proficiency be attributed to his or her insufficient development in his or her first language?
- Can the student's academic difficulties or failure in an English-only academic setting be attributed to his or her not having attained CALP in English?
- Was the student given ample instructional time in his or her first language to (1) develop CALP in this language and (2) demonstrate ability somewhat within the average range of academic performance? (Rhodes et al., 2009, p. 73)

Consideration of these questions will better inform the school psychologist as to the appropriateness of assessing the ELD student.

Language of Assessment & Eligibility

As mentioned above, the school psychologist, or designated language assessor, must assess the ELL student's CALP level to determine the language of assessment prior to psychoeducational assessment. For example, the school psychologist or language specialist may receive a referral for assessment from the student study team and proceed to administer a language proficiency assessment. Assuming the student's native language is Spanish, a possible assessment can include the Woodcock Munoz Language Survey (WMLS). The decision to assess in a specific language will depend on the CALP data that has

been collected. If the student's CALP level in English and Spanish on the WMLS is two and three, respectively, then assessment should be conducted in both English and Spanish in order to establish that a possible disability is evident in both languages. On the other hand, if a student earns a CALP score of two in English and five in Spanish, it is advisable to proceed with assessment in Spanish as that appears to be the stronger language. In contrast, if the student achieved a CALP level five in English and two in Spanish, then assessment should proceed in English. Although there are not clear cut guidelines, the examiner must combine CALP levels and multiple sources of information to augment language proficiency results to best determine the language of psychoeducational assessment. The reader is directed to the Multidimensional Assessment Model for Bilingual Individuals (Rhodes, Ochoa & Ortiz, 2005) for assistance in determining language of assessment.

Several assessments have been specifically developed to assess an individual's CALP level and to better inform the examiner as to which language to use to assess (see Table 2). After assessing and establishing CALP levels, the examiner can select a cognitive battery that is linguistically appropriate for the examinee. Given the CALP scores, the examiner can choose to assess the student using one of, or a combination of, the following modalities: English, Spanish, bilingual, or nonverbal. There are a variety of cognitive assessments that are available in English, Spanish, and nonverbal modalities that utilize CHC as an underlying cognitive theory. See Table 3 for a sampling of cognitive instruments that utilize CHC theory and available languages.

Table 2 *Summary of Primary Language Assessment & Available Languages*

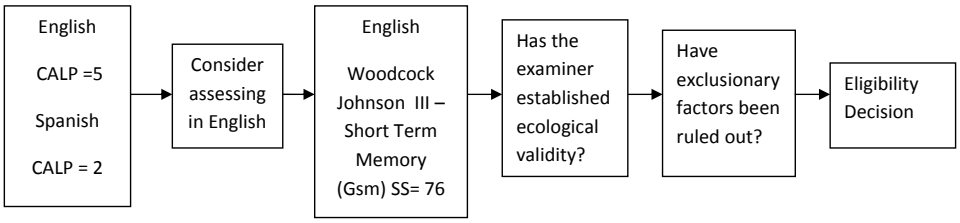
Test	Available Languages
Bilingual Verbal Abilities Test (Munoz-Sandoval, Cummins, Alvarado, & Ruef, 1998).	Arabic, Chinese (Simplified and Traditional), English, French, German, Haitian-Creole, Hindi, Italian, Japanese, Korean, Polish, Portuguese, Russian, Spanish, Turkish, & Vietnamese.
Basic Inventory of Natural Languages (Herbert, 1986).	Arabic, Armenian, Cambodian, Cantonese, Chinese Creole, Dutch, English, Farsi, French, German, Greek, Hindi, Hmong, Ilocano, Inupiaq, Italian, Japanese, Korean, Laotian, Navajo, Philipino, Polish, Portuguese, Russian, Spanish, Tagalog, Taiwanese, Toishanese, Ukranian, Vietnamese
Woodcock-Munoz Language Survey – Update (Woodcock & Sandoval, 2001).	English and Spanish
California English Language Development Test (CDE, 2009).	English
IDEA Oral Language Proficiency Test (Dalton, 1991)	English and Spanish

Table 3 *Summary of Cognitive Assessments, Language, and CHC Factor Assessed*

Test	Languages	CHC Factors
Woodcock Johnson Test of Cognitive Abilities – Third Ed. (Woodcock, McGrew, & Mather, 2001).	English	Short Term Memory (Gsm); Long Term Retrieval (Glr); Fluid Reasoning (Gf); Visual Spatial (Gv); Crystallized Abilities (Gc); Auditory Processing (Ga); Processing Speed (Gs)
Bateria III Pruebas de Habilidades Cognitivas (Mather & Woodcock, 2005).	Spanish	Gc; Gf; Gsm; Glr; Ga; Gv; Gs
Kaufman Assessment Battery for Children – Second Ed. (Kaufman & Kaufman, 2004).	English Spanish-Instructions Nonverbal	Gc; Gf; Gsm; Glr; Gv
Wechsler Intelligence Test for Children- Fourth Ed. (Wechsler, 2004)	English	Gc; Gf; Gsm; Gv; Gs
Wechsler Intelligence Test for Children Spanish, Fourth Ed. (Wechsler, 2004)	Spanish	Gc; Gf; Gsm; Gv; Gs
Differential Abilities Scale – Second Ed. (Elliot, 2007)	English Spanish-Instructions	Gc; Gf; Gsm; Glr; Ga; Gv

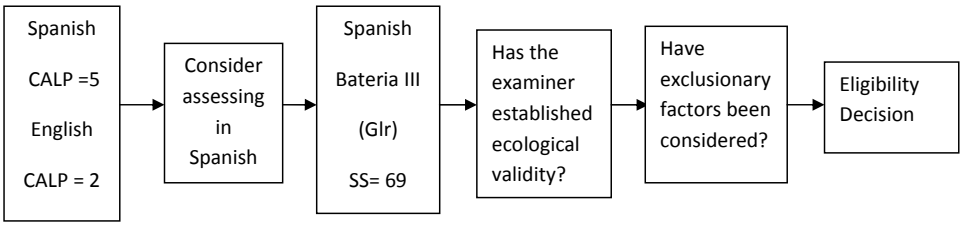
As highlighted above, English-only may be selected as the mode of assessment if the student is no longer classified ELL or if the student has achieved an English CALP score in the four to five range and minimal Spanish CALP (<3). A CALP score in this range indicates that the student has achieved proficiency in the English language, and an English cognitive assessment is appropriate and can be considered nondiscriminatory. Figure 1 (below) illustrates that the student's CALP in English is five while CALP in Spanish is two. The examiner has determined that the student is fluent in English, based on the English CALP designation, and English assessment is appropriate. The student scored at the below average range on the WJ III (Gsm=76), indicating a possible deficit in this area. Eligibility will be considered after the examiner has established ecological validity and exclusionary factors.

Figure 1. *Processing deficit decision tree for children that require English.*



Spanish-only may be selected for assessment when the student has a Spanish CALP score of four to five with minimal English CALP Scores (<3). A Spanish CALP score in this range indicates that the student has achieved proficiency in the Spanish language through formal instruction in Spanish or a bilingual program, and cognitive assessment in Spanish is appropriate and nondiscriminatory. Figure 2 (below) illustrates that the student's Spanish CALP is five while the English CALP is two. The examiner has determined that student is fluent in Spanish, based on the Spanish CALP score, and Spanish assessment is appropriate for this student. The student scored at the well below average range on the Bateria III (Glr=69) indicating a possible deficit in this area. Eligibility will be considered after the examiner has established ecological validity and exclusionary factors.

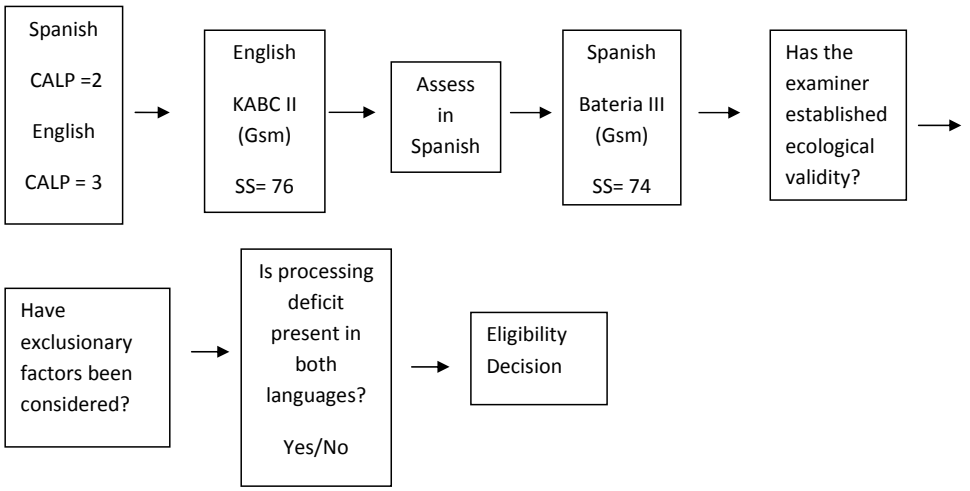
Figure 2. *Processing deficit decision tree for children that require Spanish.*



Bilingual (English and Spanish) assessment is appropriate when the student has minimal CALP scores (≤3) in both languages. This may indicate that the student is in the process of acquiring proficiency in one or both languages, or may have a speech and language impairment (SLI). If this is the case, it is advisable to consult with your speech and language therapist for further guidance. When CALP is limited in both languages, the examiner is advised to assess in both English and Spanish to establish that the processing limitations are evident in both languages. When this is the case, the examiner may elect to begin assessment in either language (Spanish or English) as both languages are at about the same level. It is important to note that bilingual assessment is *only* necessary when a disability is suspected, specifically when a CHC factor (processing area) is considered a weakness or when the examiner is considering an eligibility decision. Figure 3 (below) illustrates that the student's English CALP is three while the Spanish CALP is two. The examiner has determined that student has limited CALP in both languages and assessment must be undertaken in both languages in order to understand the student's abilities. The student scored at the below average range on the KABC II (Gsm=76) and at the below average range on

the Bateria III (Gsm = 74). It is noteworthy that the student scored at the below average range in both languages indicating a possible deficit. Eligibility will be considered after the examiner has established ecological validity and exclusionary factors.

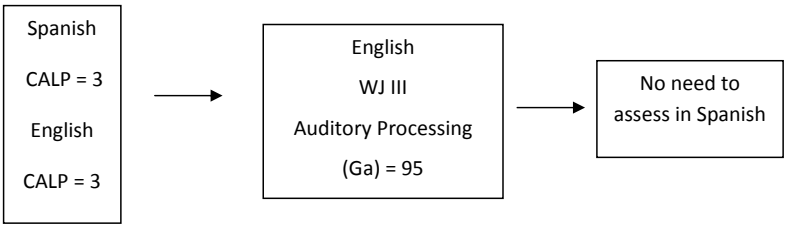
Figure 3. *Processing deficit decision tree for children that require both English and Spanish Assessment.*



The suspicion of a disability is confirmed when the examiner has established ecological validity and exclusionary factors have been considered as not being the primary factor in the student's learning difficulties. Ecological validity can be established through assessing patterns of deficits though multiple sources of information (cumulative reviews, SST files, etc), interviews (parent, teacher, and student), and observations (structured and unstructured) (Hale & Fiorello, 2004).

Conversely, if the CHC factor (processing area) is at or above the average range, then assessment in the other language is not warranted. It does not constitute a concern because the student is performing comparable to children at his or her age and grade. Figure 4 (below) illustrates that the student has a CALP level of three in both languages. The examiner determines that both English and Spanish assessment may be required if a disability is suspected. The WJ III (Ga) is administered in English and the student's score is at the average range. Because the score is at the average range, the examiner determines that no further assessment is necessary because the student scored at the average range. Again, assessment in the second language is only necessary when a processing deficit is evident and the examiner is considering eligibility. Nonverbal assessment is appropriate when the examiner does not speak the language or if the student has been diagnosed as having a speech and language impairment.

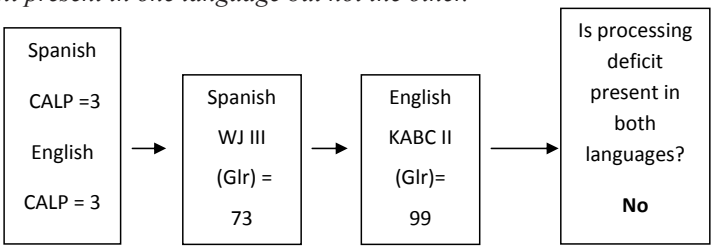
Figure 4. *No processing deficit concern.*



Lastly, the examiner may have determined to assess in both English and Spanish, based on CALP scores, and in the process determined that the student's abilities revealed low skills in one language but not the other. The examiner may conclude that low scores in one language may be due to ELD and not necessarily a disability. Figure 5 (below) illustrates that the examiner has determined that the student has limited CALP in both languages and assessment must be undertaken in both English and Spanish in

order to understand the student's abilities. The student scored the below average range on the Bateria III (Glr=73) and at average range on the KABC II (Glr = 99). It is noteworthy that there is a discrepancy between Spanish and English assessment. The examiner may conclude that the lack of consistency on the measures may be due to linguistic factors and not a disability. Therefore, eligibility may not be a decision in this scenario.

Figure 5. *Deficit present in one language but not the other.*



To minimize cultural and linguistic bias in test selection, the reader is advised to review the Cultural Language Interpretative Matrix (CLIM) and the Culture-Language Test Classifications (C-LTC) in Flanagan et al. (2007). The CLIM and C-LTC are constructed to analyze several cognitive assessments rating each subtest with regard to linguistic and cultural loading. These tools can be helpful when the school psychologist is assessing a child who is culturally and linguistically diverse. As mentioned above, Rhodes et al.'s (2005) Multidimensional Assessment Model for Bilingual Individuals (MAMBI) also guides the examiner in selecting the appropriate language of assessment. The MAMBI considers the following variables: degree of language proficiency (CALP), previous instructional modality, current grade of the student, and mode of assessment (Rhodes et al., 2005, p. 171). As mentioned above, establishing CALP levels is essential when assessing children that are classified or suspected of being ELL. In the context of multiple sources of information and observations, establishing CALP is an important aid in determining how to effectively conduct an assessment that is culturally and linguistically appropriate.

CONCLUSIONS

While assessing a bilingual student can be time-consuming and complex, MODEL provides the practitioner with a framework to systemically assess ELL children in a nondiscriminatory manner by a bilingual or monolingual English speaking school psychologist, using a qualified interpreter. Through the collection of multiple sources of information, which includes reviewing all relevant records and observations, the practitioner can develop a data-driven hypothesis(es) regarding the referred student. Through the hypothesis(es), selection of a culturally and linguistically valid battery of assessment is undertaken based upon CALP levels. With the selection of linguistically appropriate cognitive tools and a thorough review of multiple sources of information, the school psychologist can corroborate that a disability is present in L1 and L2.

MODEL can assist the school psychologist in deciphering whether the current academic difficulties are related to normal English language development or suspected learning disabilities. After completing the comprehensive assessment, school psychologists are better able to develop culturally and linguistically appropriate interventions that address academic, social emotional needs, and English language supports.

- - -

Pedro Olvera, Psy.D., L.E.P., is the Director of the School Psychology program at Azusa Pacific University (APU). He teaches coursework in the area of bilingual assessment. Before joining APU, he was a bilingual school psychologist for the Santa Ana Unified School District. He is in private practice with his wife in Orange County, California.

Lino Gomez-Cerrillo, M.A., is a bilingual school psychologist for the Chaffey Joint Union High School District where he developed a district assessment model for English learners. Gomez-Cerrillo is also Practicum Coordinator for Azusa Pacific University's School Psychology Program. He is pursuing his PsyD at Alliant International University.

REFERENCES

Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *National Council on Measurement in Education, 27*, 17-31.

Alfonso, V.C., Flanagan, D.P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D.P. Flanagan & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed.). New York, NY: Guilford Press.

Baker, S.K., & Good, R.H. (1995). Curriculum-based measurement of English reading with bilingual Hispanic children: A cross validation study with second grade students. *School Psychology Review, 24*, 561-578.

Baker, S.K., Plasencia-Peinado, J., & Lezcano-Lytle, V. (1998). The use of curriculum-based measurement with language minority students. In M.R. Shinn (Ed.), *Advanced applications of curriculum based measurement* (pp. 175-213). New York, NY: Guilford Press.

Burns, M.K., MacQuarrie, L.I., & Campbell, D.T. (1999). The difference between curriculum based assessment and curriculum based measurement: A focus on purpose and result. *Communique, 27*(6), 18-19.

California Department of Education. (2009). *California English language development test: Understanding and using 2009-2010 individual results*. Sacramento, CA: Author.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222-251.

Dalton, E.F. (1991). *IPT Oral Grades K-6 Technical Manual, IDEA Oral Language Proficiency Test Forms C and D English*. Brea, CA: Ballard & Tighe.

De Ramírez, R.D., & Shapiro, E.S. (2006). Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in bilingual education classrooms. *School Psychology Review, 35*, 356-369.

Diaz-Rico, L.T., & Weed, K.Z. (2002). *The crosscultural, language, and academic language development handbook: A complete K-12 reference guide*. Boston, MA: Allyn and Bacon.

Elliot, C. (2007). *Differential Abilities Scales-Second Edition*. San Antonio, TX: Psychological Corporation.

Figueroa, R.A., & Newsome, P. (2006). The diagnosis of LD in English learners: Is it nondiscriminatory? *Journal of Learning Disabilities*, 39, 206-214.

Flanagan, D.P., Ortiz, S.O., & Alfonso, V.C. (2007). *Essentials of cross battery* (2nd ed.). Hoboken, NJ: Wiley & Sons.

Hale, J.B . & Fiorello, C.A. (2004). *School neuropsychology: A practitioner's handbook*. New York, NY: Guilford Press.

Hearne, D. (2000). *Teaching second language learners with learning disabilities: Strategies for effective practice*. Oceanside, CA: Academic Communication Associates.

Herbert, C. (1986). *Basic Inventory of Natural Language*. San Bernardino, CA: Checkpoint Systems, Inc.

Individuals With Disabilities Act Revision, Pub. L. No. 105-15), § 20 U.S.C. Chapter ‑ 33 (1997).

Kaufman, A.S., & Kaufman, N.L. (2004). *Kaufman Assessment Battery for Children* (2nd ed.). Circle Pines, MN: AGS Publishing.

Lau, M.Y., & Blatchey, L.A. (2009). A comprehensive, multidimensional approach to assessment of culturally and linguistically diverse students. In J.M. Jones (Ed.), *The psychology of multiculturalism in the schools: A primer for practice, training, and research* (pp. 139-171). Bethesda, MD: NASP Publications.

Lichtenstein, R. (2008). Best practices in identification of learning disabilities. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 295-317). Bethesda, MD: National Association of School Psychologists.

Lopez, E. C. (2002). Best practices in working with school interpreters to deliver psychological services to children and families. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 1419-1432). Bethesda, MD: National Association of School Psychologists.

Munoz-Sandoval, A.F., Cummins, J., Alvarado, C.G., & Ruef, M. L. (1998). *The Bilingual Verbal Ability Test*, Itasca, IL: Riverside Publishing.

Munoz-Sandoval, A.F., Woodcock, R.W., McGrew, K.S., & Mather, N. (2005a). *Bateria III Woodcock Munoz: Pruebas de aprovechamiento*. Itasca, IL: Riverside Publishing.

Munoz-Sandoval, A.F., Woodcock, R.W., McGrew, K.S., & Mather, N. (2005b). *Bateria III Woodcock Munoz: Pruebas de habilidades cognitivas*. Itasca, IL: Riverside Publishing.

Ochoa, S.H., Riccio, C., & Jimenez, S. (2001). Psychological assessment and/or bilingual students: An investigation into school psychologists' current practices. *Journal of Psychoeducational Assessment, 22*, 185-208.

Olvera, P., & Cerrillo-Gomez, L. (2010, March). *A bilingual (English & Spanish) psychoeducational assessment MODEL grounded in Cattell-Horn Caroll (CHC) Theory: Introduction and application of MODEL through case study discussion*. Paper presented at the annual meeting of the California Association of School Psychologists, Santa Clara, CA.

Ortiz, S.O. (2008). Best practices in nondiscriminatory assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 661-678). Bethesda, MD: National Association of School Psychologists.

Ortiz, S.O., Flanagan, D.P., & Dynda, A.M. (2008). Best practices in working with culturally diverse children and families. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 1721-1738). Bethesda, MD: National Association of School Psychologists.

Oswald, D.P., & Coutinho, M. J. (2001). Trends in disproportionate representation in special education: Implications for multicultural education. In C.A. Utley & F.E. Obiakor (Eds.), *Special education, multicultural education, and school reform: Components of a quality education for students with mild disabilities* (pp. 53-73). Springfield, IL: Charles C. Thomas.

Prasse, D. P. (2008). Best practices in school psychology and the law. In A. Thomas & J. Grimes (Eds.), *Best practices in school psycholog*y (pp. 1903-1920). Bethesda, MD: National Association of School Psychologists.

Reschly, D.J., & Hosp, J.L. (2004). State SLD identification policies and practices. *Learning Disability Quarterly, 27*, 197-213.

Rhodes, R.L., Ochoa, S.H., & Ortiz, S.O. (2005). *Assessing culturally and linguistically diverse students: A practical guid*e. New York, NY: Guilford Press.

Roseberry-McKibbin, C. (2002). *Multicultural students with special language need*s (2nd ed.). Oceanside, CA: Academic Communication Associates.

Valdes, G., & Figueroa, R.A. (1994). *Bilingualism and testing: A special case of bia*s. Norwood, NJ: Ablex Publishing.

Woodcock, R.W., McGrew, K.S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilitie*s. Itasca, IL: Riverside Publishing.

Woodcock, R.W., & Munoz-Sandoval, A.F. (2001). *Woodcock-Munoz Language Survey Normative Updat*e. Chicago, IL: Riverside Publishing.

TOC

TOOLS FOR PRACTICE

TOC

Creating Lasting Behavioral Change through the Generalization Analysis Worksheet

John Brady,
Chapman University

and

Ron Kotkin,
*University of California, Irvin*e

The goal of any behavioral program is to facilitate lasting change. A significant criticism of behavioral programs is that they work in the clinical setting but do not generalize once the clinical program is stopped. The authors suggest that behavioral programs often do not generalize because clinicians fail to plan for generalization to occur (Maag, 2004). A technology for planning and facilitating generalization has been in existence since the 1970's (Stokes and Baer, 1977, Stokes and Osnes, 1986 and Stokes and Osnes, 1989). The authors have created a form to prompt clinicians and researchers to systematically plan for generalization as part of their behavioral programming. A case study is given to demonstrate how the form can be used to increase the probability of behavioral changes transferring to other settings and maintaining over time. The Generalization Analysis Worksheet is designed to assure that clinicians and researchers program for generalization as part of any behavioral program they design. If the technology suggested by Stokes and Baer, 1977, Stokes and Osnes, 1986 and Stokes and Osnes, 1989 is routinely programmed into behavioral programs, behaviorists may finally be able to answer the criticism that behavioral programs do not generalize.

Children's behavior and academic achievement are closely linked (Algozzine, Wang, & Violette, 2010), children whose behavior interferes with their learning and social development are one of the most frequent referrals to school specialists (Bramlett, Murphy, Johnson, Wallingsford, & Hall, 2002; Chalfant & Van Dusen Pysh, 1989). Often these children are referred to a behavioral specialist, whose job is to help the classroom teacher create a behavioral management plan. The first step in this process is to help the child, through behavioral change techniques, to stop an interfering behavior and to replace it with a more appropriate behavior. The next step is to help the child to generalize the behavior, which was learned in a specific time and place, across settings and time. This paper focuses on this second step, which is often overlooked (Rutherford & Nelson, 1988), and presents strategies to generalize behavior through the use of the "Generalization Analysis Worksheet" (Appendix A). Generalization processes must be planned from the start of any behavior plan (Miltenberger, 2008; Maag, 2004), and unless they are, the initial success in changing behavior is often not maintained (Schloss & Smith, 1998). One of the most comprehensive models for facilitating the generalization of behavior was put forth by Stokes and Osnes (1989), who refined the generalization classification of Stokes and Baer (1977) and Stokes and Osnes (1986). Stokes and Osnes (1989) identified 12 strategies that behavioral research has shown to promote generalization. To prompt the use of these strategies we have operationalized them through the development of a worksheet that helps prompt behavioral change. To facilitate long-term behavioral change (generalization), we propose the use of this Generalization Analysis Worksheet. The first part of the paper describes each of the strategies and how they relate to the generalization of behavioral change in the schools, and the second part of the paper presents a case example that uses the worksheet.

Correspondence may be sent to John Brady, PhD., Counseling and School Psychology Program, College of Educational Studies, Chapman University, One University Drive, Orange, CA 92866, or email JBrady@Chapman.edu.

GENERALIZATION STRATEGIES

Stokes and Osnes (1989) presented a comprehensive model of generalization that included 12 strategies that support generalization. These strategies are as follows:

1. *Use the normal consequences or rewards that already exist in the child's environment*. One needs to directly teach behaviors that are likely to come into contact with reinforcers (consequences) in the school or home. These are naturally occurring reinforcers that do not need to be programmed and include teacher attention, peer attention, and reward systems already in place. These can be used to generalize behavior if the behavior is similar to that which normally triggers the natural reinforcers. To identify the naturally occurring reinforcers that maintain the target behavior, one needs to observe the environment in which the behavior is to be generalized (e.g., home, school, classroom, playground, 5th period). For example, the classroom reward system that is already in place can be used as the reward system for training of the new behavior. Even if one's preliminary efforts include other rewards, one needs to include the existing classroom behavior system as part of the overall program. Another example is using a teacher's preferences for particular student behavior. One way to improve the child's relationship with the teacher and, consequently, his or her overall classroom behavior, is to determine which behaviors or characteristics the teacher naturally reinforces and to work on developing these behaviors as part of the overall behavior plan.
2. *Teach students to purposely request/recruit the natural reinforcers in the classroom, community, or home*. Sometimes teachers overlook the initial attempts at behaviors that are being intentionally taught to the student. This may be because the student's behavior is not yet frequent enough or the skill is not yet sufficiently developed. Generalization can be fostered by attention being brought to a student's newly learned behavior, which triggers the natural reinforcers present. For example, a student can be shown how to self-monitor his or her work accuracy, work productivity, or new positive behavior and then share the data with the teacher. Grids, charts, and checklists can be used to highlight the student's work and enable the teacher to pay more attention to it. Additionally, a student can be encouraged to ask for feedback from teachers: "How is this?" "Am I doing this right?"
3. *Eliminate behavior that may be a roadblock to generalization of the new behavior*. One needs to determine which, if any, reinforcers are maintaining the inappropriate behaviors. These could include teacher or peer attention or escape from activities that are not reinforcing. Then, it is important to determine ways of at least temporarily reducing or eliminating these reinforcers until the new behavior has a chance to take hold. For example, the teacher may reinforce the other children in the classroom for not reacting to (laughing at, commenting on) student's disruptive comments or outbursts, while at the same time reinforcing each occurrence of cooperative or attentive behavior. The teacher can make the student's academic work easier as a means to prevent the student from acting up, while at the same time reinforcing completion of the modified assignments.
4. *Reinforce any occurrence of generalization*. Teachers and other staff who work with the student need to reinforce the desired behavior even if it is not "perfect." For example, the special day class teacher can remind his or her instructional assistants to comment, "You paid attention!" each time they see the student, who is learning to follow directions, respond to a direction given by a staff member.
5. *Train the behavior across as many settings/environments/staff/teachers as possible*. The more the student practices the behavior in different settings, the more staff will request and reinforce the behavior and the more likely that the student will maintain the behavior. For example, when children are learning appropriate playground behavior, it is important that all of the playground supervisors become involved in the training process. At first, the new behavior may be taught with the help and monitoring of one supervisor, but as time goes on, all supervisors should be involved in the expectation of a particular behavior, including the uniform prompting and reinforcement of it.

6. *Devise ways that the people in the environment in which the behavior occurs respond to the new behavior such that the child recognizes that his or her behavior elicits a range of positive responses*. Behavior that is reinforced by a range of responses in the environment will be maintained more readily. Some staff may respond with a pat on the back and a smile, others may state, "What a good job!" while still others may bring the child's positive behavior to the attention of family or peers.
7. *Teach the new behavior in various settings*. It is helpful to teach at different times of the day, with different children in the class, and in various classrooms. For example, once a student is proficient in a self-monitoring intervention (charting) for attentional problems in his or her classroom, he or she should be encouraged to use the same procedures in assemblies, at home, on field trips, and in other settings.
8. *Make consequences less predictable*. Once the new behavior is consistent, it is important to change the reinforcement schedule, particularly in terms of frequency and intensity. For example, after teaching a child to keep track of his or her homework by checking it each night and praising the child for his or her efforts, one can later switch to checking it every other night and then just weekly. As another example, once one has a high rate of attending to tasks in the child's reading group by verbal reinforcement of each attentive response, then one can reduce the frequency of verbal reinforcement to every third time.
9. *Incorporate into the behavior change program, physical settings, or items common to the natural environments*. When possible, one should teach new behaviors in the setting in which they are expected to occur naturally. The child will associate the behavior with the setting and be more inclined to behave appropriately when he or she is in that setting in the future. If one is not able to use the natural setting, then one should try to make the training setting look as much like the natural setting as possible.
10. *Include common (to the natural setting) social cues in the behavior plan*. One should include important persons from the generalization setting (i.e., peers, teachers, parents) into the training. If those people with whom the child would normally come into contact are included in the training sessions, the child is more likely to extend the new behavior to those settings in which he or she encounters them again.
11. *Provide the child with a physical prompt that will remind or guide him or her in performing the behaviors in the natural environment*. For example, a child who is working on controlling angry outbursts may learn to generalize his or her newly learned control by using a business-card-sized reminder with "stop and think" printed on it that the child keeps in his or her pocket.
12. *Include in the behavior training verbal cues that the child can give him or herself in the new situation*. For example, one should provide the child with a self-talk script that will remind him or her of what the child is to do. For example, a child who gets anxious and consequently refuses to attempt tests can be trained to "talk" him or herself through the relaxation procedures that were taught to the child in the school psychologist's office. A child who overreacts to teasing on the playground can be taught to "count to ten" before he or she responds.

CASE EXAMPLE

This section presents a case example (taken from the case files of the second author) of how a comprehensive plan to facilitate generalization can be developed through the use of a Generalization Analysis Worksheet. The case example includes a statement of the problem, followed by a completed behavioral change plan. The behavioral change plan incorporates the components of generalization identified in the Generalization Analysis Worksheet. The number of each component incorporated from the worksheet is noted in parentheses following each relevant component.

John is an 11-year-old student who has been diagnosed with ADHD and qualifies for participation in a gifted educational program. He was referred to the Student Study Team (SST) because he frequently interrupts his teacher in an aggressive or rude manner, questioning the accuracy of the teacher's comments (even though his observations are often correct). The teacher has developed resentment of his style of questioning. The SST team has chosen to target his questioning style for modification. They do not want to decrease John's questioning of inaccurate information; rather, they want to teach John how to ask questions in a diplomatic manner. The student does not see why he needs to change his behavior and becomes argumentative when reprimanded by the teacher for his rudeness.

The SST decided that, if John does not change his behavior, he would have poor relationships with teachers and peers, causing John to be rejected. Indeed, reports from the teacher confirmed that peers were avoiding John and that teachers often complained about him. The SST completed the Generalization Analysis Worksheet as part of their treatment planning.

The plan (Figure 1) was constructed with the following considerations to enhance the likelihood that the new behavior would generalize to the natural environment. The team first generated a list of diplomatic phrases (e.g., "I've heard that," "I've read that," "I could be wrong, but") that could be used to precede any question asked by John and which served as an alternative to John's rude phrase (1). The phrases chosen are likely to elicit a positive response from teachers. To ensure that teachers noticed, John was taught to praise the teachers for acknowledging him (2). Initially, John was not motivated to change his rude questioning behavior because he was reinforced by the interaction with the teacher, as a result of correcting or arguing with the teacher. To decrease the payoff for rudeness, his teachers were taught not to reinforce him by arguing with him and to use a consistent time out for rude behavior (3). In addition, the teachers were taught to reinforce, through praise and points on a simple daily report card turned in at the end of the day, any examples of John using diplomatic questions (4). All teachers and instructional aides were informed of the program and multiple teachers were included in the training sessions (5).

A series of visual prompts was created to help John to use the diplomatic phrases. This included laminating a sheet of paper with the phrases for John to place on his desk and in his notebook. Visual prompts also were placed around the room so that when John or the teacher saw them, it reminded them of the program that was in place (7, 11). John was taught multiple diplomatic responses to shape the use of diplomatic questions as a class of behaviors, rather than as a discrete response. Schedules of reinforcement were developed to systematically shape the consistent use of diplomatic questioning that was resistant to extinction. As part of their reinforcement plan, a systematic set of criteria was developed to prompt changing from a fixed schedule of reinforcement to an intermittent schedule of reinforcement (8).

First, John was reinforced for every incidence in which he used the diplomatic phrases preceding a question. Once he was able to use diplomatic sentences for an average of 90% over a 10-day period, the schedule of reinforcement was changed to every other incidence of using a diplomatic phrase. Once again, when John was able to use diplomatic phrases for an average of 90%, reinforcement was changed to a variable schedule in which he was reinforced on the average of every third use of a diplomatic phrase. Training took place in several classrooms where John attended classes (9) and with selected students and teachers (10). During the session, John was taught self-mediated verbal prompts to ask questions in a diplomatic style and to self-reinforce for following through on his verbal prompt (12).

SUMMARY

The research reviews of Stokes and Baer (1977) and Rutherford and Nelson (1988) have highlighted the need for and lack of generalization planning in behavioral interventions. This is in spite of the fact that Stokes, Osnes and Baer presented a technology for facilitating generalization in the 1970's and 80's. The present authors proposed a simple 12-strategy worksheet process that prompts behavioral change professionals to program the components of generalization into their behavior plans. As shown in the case study the application of these basic principles mediates long term retention of the learned behavior. They should become a routine step in planning for lasting behavior change. Until behavioral therapists routinely act on the recommendations of Stokes, Osnes and Baer we are not likely to answer the criticism that behavioral interventions do not generalize.

Figure 1. *Generalization analysis worksheet.*

Consultant: Ron Kotkin Child: John____ Date: July 20___
Target Behavior: Questioning the teacher in a rude and aggressive manner

Strategy	Plan
1. Connect behavior to the natural consequences.	Teach diplomatic verbal statements likely to be positively perceived and reinforced by teachers.
2. Teach the child to "go after" the natural consequences.	Have the student track his use of tactful questions and share with the teacher. Train the student to praise the teacher for answering his question.
3. At least temporarily, stop the reinforcement for inappropriate behaviors.	Train teachers to give the student a timeout for rudeness.
4. Aggressively reinforce any occurrence of generalization.	Train all teachers to reinforce the student, through praise and points on a daily report card, for asking a question in a diplomatic manner, using one of the pivotal phrases taught to the student.
5. Train in multiple settings with multiple trainers.	Train in the multiple settings and with a sampling of teachers and peers from the natural setting.
6. Create multiple responses to the behavior.	Develop a range of responses that teachers and support staff could use in response to the student's appropriate questions.
7. Vary the antecedents (settings/circumstances) of the behavior training as much as possible.	Use multiple visual prompts, such as pivotal sentences taped to the student's desk or posted in his organizer. In addition, visual prompts could be placed on the wall so that when the student or teacher sees the prompt, each is reminded of the program in place.
8. Vary the schedule and intensity of consequences.	Set criteria for moving to an intermittent schedule of reinforcement based on the student's successful use of diplomatic questions, reinforced by a more regular schedule.
9. Incorporate the "look" of the natural environment in the training.	Train in a classroom that is similar to the classrooms the student will be attending.
10. Incorporate relevant people in the training.	Bring teachers and peers into the training sessions to role-play the use of pivotal phrases in questioning teachers and peers.
11. Incorporate a tangible prompt/reminder that he can take with him.	Create a laminated sheet with pivotal sentences that the student can put on his desk at the beginning of a class period to remind him to use diplomatic questions. He also can have them taped to the inside of his organizer.
12. Incorporate self-mediated verbal reminders/prompts.	Teach the student self-talk to prompt and reinforce himself for using diplomatic questions. For example, "Ask in a nice way," "See, he/she answered my question."

Dr John Brady is a professor in the school counseling and school psychology programs at Chapman University and the coordinator of the school counseling graduate program. Prior to teaching at the university he served for twenty seven years as a school psychologist and special education administrator. Dr Brady's interests lie in the development of collaborative problem solving skills for school consultants and data based decision processes.

Dr Kotkin is currently professor Emeritus at the University of California, Irvine. He served for twenty-six years as a Professor in the Department of Pediatrics and Director of the University of California at Irvine, Child Development Center Day Treatment Program for children with Attention Deficit Hyperactivity Disorder. He has been a special education teacher at the elementary school level. He is also a licensed clinical psychologist in the state of California. Dr. Kotkin previously served as a professor of special education. In addition, he is a consultant to school districts in developing individual and School Wide Behavior Interventions. Dr. Kotkin has published multiple articles and chapters on school-based intervention and has co-edited a book Therapist's Guide to Learning and Attention Disorders, with Aubrey Fine. He also developed the Irvine Paraprofessional Program, which was recognized by the Kentucky Federal Resource Center as a "promising practice" for intervening with students with ADHD in the general education classroom. He was also presented an award, with Jim Swanson and Steve Simpson by CHADD (Children and Adults with Attention Deficit/Hyperactivity Disorder) for development of the most innovative program in the country for serving children with ADHD in the general education classroom. He has been a presenter at many international and national conferences and also contributed his expertise to a major National Institute of Mental Health (NIMH) study on long term treatment effects of children with Attention Deficit Disorder. Dr. Kotkin is also on the Professional Advisory Board for CHADD.

REFERENCES

Algozzine, B., Wang, C. & Violette, A.S. Reexamining the relationship between academic behavior and social behavior. *Journal of Positive Behavior Interventions*. Prepublished August 26, 2010. DOI: 10.1177/109830070 93359084.

Bramlett, R.K., Murphy, J.J., Johnson. J., Wallingsford, L., Hall, J.D. (2002). Contemporary practices in school psychology: A national survey of roles and referral problems. *Psychology in the Schools, 39* (3), 327-335.

Chalfant, J.C., & Van Dusen Pysh, M. (1989). Teacher assistance teams: Five descriptive studies on 96 teams. *Remedial and Special Education, 10*(6), 49-58.

Maag, J.W. (2004). Behavior Management from Theoretical Implications to Practical Applications. Wadsworth. Miltenberger, R.G. (2008). *Behavior modification: Principles and procedures* (4th ed.). Belmont, CA: Wadsworth/Thomson Learning.

Rutherford, R.B., & Nelson, C.M. (1988). Generalization and maintenance of treatment effects. In J.C. Witt, S.N. Elliot & F.M. Gresham (Eds.), Handbook of behavior therapy in education (pp. 277-324). New York: Plenum Press.Schloss, P.J., & Smith, M.A. (1998). *Applied behavior analysis in the classroom*. Boston: Allyn and Bacon

Stokes, T.F., & Baer, D.M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*(2), 349-367. Stokes, T.F., & Osnes, P.G. (1986). Programming the generalization of children's social behavior. In P.S. Strain, M.J. Guralnick & H.M. Walker (Eds.), *Children's social behavior* (pp. 407-443). Orlando, FL: Academic Press.

Stokes, T.F., & Osnes, P.G. (1989). An operant pursuit of generalization. *Behavior Therapy, 20*, 337-355.

APPENDIX A

Generalization Analysis Worksheet

Consultant: _____ Child: _____ Date: _____

Target Behavior: _____

_____

Strategy	Plan
1. Connect behavior to the natural consequences. →	
2. Teach the child to "go after" the natural consequences. →	
3. At least temporarily, stop the reinforcement for inappropriate behaviors. →	
4. Aggressively reinforce any occurrence of generalization. →	
5. Train in multiple settings with multiple trainers. →	
6. Create multiple responses to the behavior. →	
7. Vary the antecedents of the behaviors as much as possible. →	
8. Vary the schedule and intensity of consequences. →	
9. Incorporate the "look" of the natural environment in the training. →	
10. Incorporate relevant people in the training. →	
11. Incorporate a tangible prompt/reminder that he can take with him. →	
12. Incorporate self-mediated verbal reminders/prompts. →	

TOC

BOOK REVIEW

TOC

Book Review: Proust and the Squid

Proust and the Squid by Maryanne Wolf
2007, HarperCollins, New York, 306 pgs.
$15, ISBN: 978-0-06-093384-5
Reviewed by Stephanie Domzalski, M.A., NCSP

We were never born to read.

The provocative premise begins author Maryanne Wolf's excavation into the art and mystery of the literate brain. A professor of child development at Tuft's University and director for the Center for Reading and Language Research, Wolf's expertise is evident in her multifaceted exploration of reading development. By all accounts, the refinement of human literacy represents for Wolf an "unnatural" process, but one which our innate biological circuitry has been primed to undertake. *Proust* begins with a chronology of reading acquisition, moving then to a discussion of individual reading development. The book concludes with discussion dedicated to the enigma of dyslexia, its biological roots and the unexpected gifts that come with such neural diversity.

Less an interventionist's manual, *Proust* incorporates historical, literary and neuro-scientific evidence to form a rich ethnography of reading. The title itself pays metaphoric homage to both French novelist Marcel Proust and the common squid whose anatomic structure propelled scientific research during the 1950s. Using this juxtaposition of characters and contributions, Wolf alludes to the book's underlying premise: that while decoding text is fundamentally a biological process – a symphonic exchange of neuronal energy – *reading* amounts to something far more dynamic.

Text greater than the sum of its characters

One of the earliest, and most reiterated points in *Proust*, is the active adaptation required of our underlying neurological structure for reading acquisition. There are no reading genes, no "blueprint" as Wolf describes it, by which reading is genetically imprinted on later generations. This necessity of neural rewiring separates reading from processes such as vision or hearing which fall on a fixed developmental trajectory. Wolf cites the work of Stanislas Dehaene, a French neuroscientist, who worked to explain the advancement of reading ability as an adaptive specialization within our visual system. For our earliest ancestors, being able to recognize certain subtle patterns among animal prints or vegetation could determine survival. With more frequent exposure to these natural patterns came a greater automaticity in recognition and, perhaps, the earliest forms of "reading" fluency.

"It took our species roughly 2,000 years to make the cognitive breakthroughs necessary to read with an alphabet, today our children have to reach those same insights about print in roughly 2,000 days"

Proust traces the development of reading through distinct historical periods. Some of the earliest examples of written record are a remarkable variety of clay tokens, likely used for accounting purposes, thought to have been created between 8000 and 4000 BCE. Wolf cites this type of graphic representation as revolutionary because it both created a lasting form of communication as well as imbued an otherwise benign symbol with meaning. The ability of ancient ancestors to differentiate various symbolic representations from one another and to connect those graphic abstractions to tangible objects, marks one of the most significant periods in the neurological evolution of our reading brain.

Wolf describes a number of early writing systems, including the Sumerian "cuneiform," the pictographic and syllabic Akkadain language, the Egyptian hieroglyph, the Grecian alphabet, and their corresponding contributions to reading acquisition. She delves deeply into a comparison of alphabetic languages (e.g., English) and symbolic languages (e.g., Chinese) and the unique influence of each to brain development and neural specialization. While imagining research supports different cortical areas involved in processing alphabetic versus symbolic presentations, three primary regions (the frontal, temporal-parietal and occipital lobes) function as a system common across readers of any language. Thus, Wolf concludes, the brain's plasticity allows it to adapt to the demands of any visual presentation with equal efficiency.

Proust moves to further explore the anatomical adaptations made to accommodate for reading development. By six months, a child's visual system is functioning fully and paired with growing attentional and conceptual capacities. Naming objects and oral language development (phonological, semantic and morphological understanding) then mature followed by print awareness and the recognition of what Wolf deems the "invariant patterns" of letters. This development of one-to-one correspondence between the letter name and its visual representation is a critical component to reading acquisition and fluency.

Considering the sequential course of reading acquisition, Wolf approaches the question of *when* children should learn to read. Delving beneath the cultural caveats and social stigmas, she reduces her answer down to a question of biological readiness. For each of the cortical areas involved in reading to fully process printed stimuli, a significant degree of coordination is required. That coordination necessitates the areas of the brain involved in reading be able to communicate efficiently. To this point, Wolf cites the research of Norman Geschwind and his work with the myelination of neurons. The more myelinated a neuron, the more efficient it is in synthesizing and sending messages. Geschwind's research found that neurons are not sufficiently myelinated for optimal cortical efficiency until between the ages of five to seven, and later in boys than in girls (Geschwind, 1965). Using these findings and other biological evidence as a base, Wolf describes artificially accelerating reading development as "potentially counterproductive" (p. 96).

However, despite the necessity of biological readiness for reading, Wolf far from negates the importance of early and continuous exposure to language. She cites extensive research highlighting environmental factors contributing to literacy. One of the best predictors of later reading skill is the amount of time being read to as a child. Its roots in infancy, those initial associations between being read to and the sense of security derived from being in close proximity to familiar others are acutely influential. Wolf also stresses, as a variable predictive of reading acquisition, the importance of environments rich in language. Research from Risley and Hart (2003) found that 32 million fewer words are spoken to children from a "language impoverished" environment by the age of five when compared to the average number of words spoken to children by that age- a deficit that can significantly influence reading comprehension As children develop stronger associations with the language of books, they are exposed to perspective taking, various forms of literary devices and invitations to practice making inferences and drawing conclusions. Such exposure helps children develop their own communicative style as they move through widening circles of interpersonal interaction all before they may be able to read themselves independently.

Wolf concludes *Proust* with a chapter dedicated to individuals whose unique neurology prevents them from attaining reading fluency through traditional channels. The study of dyslexia, as Wolf describes it, is "intrinsically messy" due to how much of the brain is involved in reading and how many different areas of science have proposed theories on the cause. She outlines four primary principles (or hypotheses) to guide the reader in their understanding of a brain affected by reading challenges. From weaknesses within-specific brain structures to processing inefficiency to connectivity failures between cortical areas, Wolf provides extensive research, as well as illustrations embedded within the text, to assist readers with the complexities associated with understanding the phenomena of dyslexia.

Wolf's concluding comments on the "gifts of dyslexia" represent some of *Proust's* most thought-

provoking reading. She blends together both research and anecdotal accounts of individuals with reading difficulties and the extraordinary talents they carry forward. Using dyslexia as proof, Wolf returns to her original premise – that we, as humans, were never meant to read. The brain, when organized otherwise, creates circuits that lend themselves to the fantastic, the ingenious and the sublime, and that these alterations represent neurological differences far from deficiencies.

Some readers may be disappointed that despite Wolf's extensive enumeration on the potential causes of dyslexia, she does not provide corresponding prescriptives for action. It would have made for a stimulating exploration if Wolf had chosen to view interventions using the same biological lens as the trajectory of traditional reading development. Do effective interventions actually change the efficiency or structural relationships within our brain? Inclusive of this absence, *Proust* communicates a rich context that helps us understand reading as both an activity of the brain and a cultural phenomenon. For research on reading interventions, readers are encouraged to explore *Understanding, Assessing and Intervening on Reading Problems* by Laurice Joseph (1996).

A process more than 77,000 years in the making deserves a chronology like *Proust*. Wolf helps those who take the journey through reading's travelogue fully appreciate the remarkable nature of this evolutionary gift.

- - -

Stephanie Domzalski, M.A., NCSP, is Licensed Educational Psychologist and a current doctoral student at Chapman University. Her primary research interests include the nature of expertise and development of professional competency. She can be reached at domza100@mail.chapman.edu

REFERENCES

Geschwind, N. (1965) Disconnection Syndrome in animals and man (parts 1 and 2). *Brain, 88*, 237-294.

Hart, B., & Risley, T. (2003). The early catastrophe. *American Educator, 27*(4), 6-9

Joseph, L. (2006). *Understanding, Assessing and Intervening on Reading Problems*. Bethesda, MD: National Association of School Psychologists.

Wolf, M. (2007). *Proust and the Squid*. New York, NY: Harper Collins Publishers.

TOC

CASP
CALIFORNIA ASSOCIATION OF SCHOOL PSYCHOLOGISTS

GUIDELINES FOR AUTHORS

Scope and purpose: ***Contemporary School Psychology,*** formerly known as ***The California School Psychologist,*** is a refereed journal published annually by the **California Association of School Psychologists** (CASP). *Contemporary School Psychology* is devoted to contemporary issues in school psychology with the goal of publishing high-quality articles that link theory to practice. The journal is an appropriate venue for articles that: (1) critically review research on topics of general interest to school psychologists in California and nationwide, (2) report research relevant to practicing school psychologists, (3) present promising practices or programs that address the needs of children and youth, and (4) critically reflect on the professional of school psychology and the challenges faced by the profession. It is also the intent of the journal to highlight the diversity of viewpoints in the profession and of the students, parents, and communities served by school psychologists in California and elsewhere.

Indexing and abstracting: Contents of *The California School Psychologist* are available on international electronic literature databases, including ERIC, developed by the US Department of Education, and PsycINFO, developed by the American Psychological Association. Thus, it is essential to include up to five keywords following the abstract of each manuscript. Contemporary School Psychology will be included in those databases listed above.

Submission and review of manuscripts: Manuscripts should be submitted electronically to cspsubmissions@chapman.edu. General communication to the editors should be sent to csp@chapman.edu. Manuscripts should be between 15-25 pages in length (including references and tables). The entire manuscript must be double spaced with at least 1-inch margins. Authors must include a cover letter stating the title of the manuscript submitted, authors' names, and a mailing address, phone number, and e-mail address for further correspondence. The cover letter should also specify that the manuscript has not been previously published and is not currently being considered for publication elsewhere.

In preparing manuscripts, authors should attend carefully to the specific guidelines of the latest American Psychological Association Publication Manual, including the abstract, headings, citations, tables, and references. Manuscripts not prepared according to the APA format will be returned to the authors without review.

Selection of articles to be published is determined on the basis of blind peer review. Reviewers examine the importance of the topics addressed, accuracy and validity of the contents, contribution to the profession, implications for the practice of school psychology, originality, and the quality of writing.

Current abstracts and previous volumes are available online: www.casponline.org.

TOC

Contemporary School Psychology

2011, Volume 15