Response to instruction, English language learners and disproportionate representation: The role of assessment

Sylvia Linan-Thompson
University of Texas at Austin

Response to Intervention (RtI), a cyclical process that incorporates assessment and instruction, is both an approach to prevent learning difficulties and to establish student eligibility for special education. Assessment results are used to determine students' initial knowledge and skill, their need for successively more intensive levels of instruction, and to gauge their response to the intervention provided. Although this process is preferable to the IQ/achievement discrepancy model for determining the presence of specific learning disabilities, there are still a number of unresolved issues related to the assessment procedures in use. A pressing issue is the identification of measures and procedures that identify students with the greatest precision thus reducing inappropriate identification.

Respuesta a la intervención, aprendices del idioma inglés como segunda lengua, y representación desproporcionada: el rol de la evaluación. Respuesta a la Intervención (RtI), proceso cíclico que incorpora la evaluación y la instrucción, es al mismo tiempo un enfoque para prevenir dificultades de aprendizaje y determinar la selección del estudiante para educación especial. Los resultados de las evaluaciones se utilizan para determinar los conocimientos iniciales de los estudiantes y sus habilidades, así como para implementar si son necesarios otros niveles más intensivos de instrucción, y para calibrar la respuesta del estudiante a la intervención prevista. Aunque este proceso es preferible al modelo de diagnóstico de las dificultades de aprendizaje basado en la discrepancia CI-rendimiento, todavía hay una serie de cuestiones pendientes de resolver como sería el tema relativo a los procedimientos de evaluación en uso. Un problema urgente es la identificación de medidas y procedimientos que identifiquen a los estudiantes con la mayor precisión lo que reduciría una identificación inadecuada.

Among advocates of English language learners, the disproportionate representation of these students in special education has been a long-standing concern. The prevalence of disabilities among English language learners (ELLs) vary, but like the percentage of the general population served in special education programs, it is approximately 9% (U. S. Department of Education, 2007; Zehler, Fleischman, Hopstock, Pendzick, & Stephenson, 2003). Despite this seemingly appropriate identification rate, when using an IQ/achievement discrepancy model, ELLs tend to be under-represented in special education in the elementary grades, but over-represented beginning in 5th grade through high school (Artiles, Rueda, Salazar, & Higareda, 2002). One of the contributing factors is students' language proficiency. Because determining the cause of a learning difficulty is challenging when the student has limited English proficiency, teachers tend to not refer ELLs for special education in the early grades, thinking, perhaps that students need more time to become English proficient. However, as ELLs get older, if they have not received adequate instruction, the gap between them and their English monolingual peers widens, as does the gap between their IQ and achievement resulting in over-identification for special education.

Other factors that have contributed to the disproportionate representation of ELLs in special education include: (a) multidisciplinary teams that failed to provide assurances that problems are not due to other factors, such as lack of access to effective reading instruction, as called for under the "exclusionary clause"; (b) limited documentation of early intervention efforts to address reading difficulties and the results of these efforts; (c) use of interventions that were not specific to identified reading difficulties, and (d) interventions were sometimes abandoned in favor of special education referral even though students appeared to be making adequate progress with general education interventions (Liu et al., 2008). RtI when implemented as a school-wide prevention model addresses these issues. However, the current assessment procedures and measures are not adequately taking into account ELLs language needs or the interaction between language and instruction.

RtI as a school-wide model has not been implemented long and, many of the studies conducted to investigate the effect of RtI on student outcomes often do not include ELLs therefore, we do not know whether or not they will continue to be disproportionately represented in special education. There is cause for concern because in an RtI model referrals and placement in special education are prompted by low achievement (Artiles, Kozlske, Trent, Osher, &

Ortiz, 2010) and ELLs perform worse on assessments of literacy, math, and content knowledge than their peers (Donovan & Cross, 2002; Hakuta, Butler, & Witt, 2000; Parrish, et al., 2006). Further, accurately determining the cause of an ELL's learning difficulty is complicated by screening procedures that rely on benchmark scores on one or two general outcome measures (GOM) with little or no consideration for mitigating factors. Finally, screening measures do not accurately discriminate among ELLs who score poorly due to a learning disability, lack of language proficiency, a language disorder, or lack of educational opportunity and do not take differences in language proficiency and educational opportunities into account.

Response to intervention

The response to intervention model is now widely recognized, as are its component parts, assessment and multiple tiers of instruction. In the literature, the purpose of RtI is described as either a framework for implementing a school-wide prevention model (NCLB; 2002) or as a systematic model for identifying children with specific learning disabilities (IDEIA; 2008) or both (Fuchs, Fuchs, & Stecker, 2010). As a prevention model, the goal of RtI is to ensure that students, who enter school without the requisite literacy skills, develop these foundational skills in a timely manner so that they can benefit from the instruction provided in the classroom. There is also an interest in ensuring that students acquire the literacy skills necessary to pass state accountability tests. Therefore, much of the research conducted has been on how well screening measures predict success on state accountability measures (e.g. Crawford, Tindal, & Stieber; 2001; Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004; Silberglitt, Burns, Madyun, & Lail, 2006; Wood, 2006).

As a model for identifying children eligible for special education, the goal is to determine whether children who are significantly different from their peers on screening measures, exhibit difficulty learning, and are unresponsive to instruction as measured by formative assessments have a learning disability. It is the dual use of RtI that is changing how we intervene with ELLs.

Assessment

Within an RtI model, assessment is used to identify students who would benefit from a supplemental intervention, determine students' response to that intervention, and determine eligibility for special education services (Ysseldyke, Burns, Scholin, & Parker, 2010). At each of these stages, entry and exit criteria are used to make decisions about individual children and their capacity to learn. The first step in an RtI model is universal screening. To minimize bias in identifying children for additional support, all students are tested and held to the same standard. Most often, a single, grade appropriate measure is used. After students are identified for Tier 2 or Tier 3 instruction, formative assessments are used to track students' response to instruction and their progress toward a specified goal (Velluntino, Scanlon, & Lyon, 2000). The data also serve as a measure of the intervention's effectiveness. Data from screening and formative assessments provide teachers feedback on the accuracy of the instructional and placement decisions they made for students.

Most measures used for assessment in RtI models are general outcome measures (GOM). GOM are equivalent measures on a common task sampled over time. These tasks assess general reading ability and provide useful data on student performance. These measures are effective in classifying children therefore appropriate as screening measures.

Due to the developmental nature of the acquisition of reading, the measures used in an RtI model change by grade level and assess the most pertinent and predictive skills (Fuchs & Fuchs, 2007; Good, Simmons, & Kame'enui, 2001). In kindergarten, measures assess students' ability to name letters, say letter sounds and segment sounds in words. In first grade, in addition to phonemic segmentation, students' emerging reading ability is assessed. Students who can read words demonstrate the ability to use grapho-phonemic awareness, orthographic knowledge and knowledge of the alphabetic system. This ability is assessed with measures of familiar word reading or pseudo word reading in isolation. Measures of oral reading fluency are also used. Beginning in second grade, measures of reading comprehension are used in addition to ORF. Maze passages are among the most commonly used.

Assessment and ELLs

Because assessment is the catalyst that initiates participation in RtI, one of the overriding concerns in the assessment of ELLs is how students' language proficiency may impact performance on assessments.

If one were to approach this as a preventive model, then entry into Tier 2 is positive. We know that ELLs benefit from English reading instruction even before they are fully proficient in English (Leseux & Siegel, 2003; Vaughn, Linan-Thompson, & Hickman-Davis, 2003) and that they outperform English monolingual students on word level tasks (Geva & Yaghoub Zadeh, 2006; Leseux & Siegel, 2003).

In an examination of two path models to reading fluency, Burke, Crowder, Hagan-Burke, Zou, (2009) found that although the skills most often assessed in an RtI model are related one to another, early skills such as phoneme segmentation fluency, letter naming fluency, and nonsense word fluency do not have a direct effect on oral reading fluency. This may explain the discrepancy between ELLs who perform well on measures of foundational skills but have difficulty in meeting benchmark on ORF measures.

Even when ELLs have adequate reading rates they may not perform well on comprehension measures. Kung (2007) found that third grade ELLs with low English language proficiency who read 130 cwpm on grade level text had a 29% chance of passing the state accountability text while ELLs with higher English proficiency and English monolingual students who qualified for free and reduced lunch who read at the same rate, 130 cwpm, had a 68% and 75% chance of passing respectively. Finally, English monolingual students with the same oral reading fluency score who did not qualify for free and reduced lunch had a 90% chance of passing the state test. Findings for fifth grade students parallel these.

Burke et al., (2009) also found that non-timed measures yield different outcomes. ELLs in particular may be over identified when only timed measures are used to determine student progress. When untimed benchmark measures such as the Woodcock-Muñoz Word Attack and Passage Comprehension subtests are used, 72% of first grade LEP students meet benchmark. However, when an oral reading fluency measure is added, only 17% percent of students are

said to have met benchmark (Linan-Thompson, Cirino, & Vaughn, 2007). This indicates that students are able to complete tasks that measure their knowledge of letter-sound correspondences and their ability to read and comprehend very short passages but that these skills are not yet automatic.

These examples highlight one of the concerns with the use of the CBMs as a one-point-in-time measure to make instructional decisions that have high stakes consequences and not with the measure itself. The valid use of curriculum-based measurement as a means for screening and monitoring student progress throughout the elementary grades is well documented (Deno, 1985, 2003; Shinn, 1989) and ORF is positively correlated with comprehension measures when used with ELLs (Good & Baker, 1995; Wiley & Deno, 2005) but it may not be enough when making instructional decisions about ELLs.

Alternative assessments and procedures

Appropriate assessments provide teachers data that help them to determine not only who is at risk at the time of assessment but also who is likely not to respond to classroom instruction. These are the children that need Tier 2 intervention. Specificity, the degree to which a measure correctly identifies children at low risk for disability status, and sensitivity, how well the measure correctly identifies children at risk for disability status, are used to determine which measures most accurately identify children. Current measures and procedures have inadequate specificity; they tend to over identify false positives, students identified as needing additional instruction who do not need it. This is particularly true for ELLs. Because successful implementation of RtI relies on accurate identification of Tier 2 and Tier 3 students, alternatives that more accurately identify students have been examined.

One option to improve the accuracy of a universal screener to identify students who are at-risk is to extend the screening period in order to monitor students who did not meet the benchmark for a short period of time to determine their response to classroom instruction. In an RTI model, progress-monitoring measures can be used to establish students' learning trajectory because the measures are sensitive to change and can be used to document student learning. Fuchs, Fuchs, and Compton (2004) found that slopes on word identification fluency in the fall of first grade correlated with coefficients of .43 with end-of-year Woodcock Word Identification, .54 with CRAB fluency, and .49 with CRAB comprehension. Though modest, these correlations indicate that determining students' growth prior to assignment may reduce the number of students identified for Tier 2 intervention. The disadvantage is that intervention is delayed for up to five weeks for students who need the intervention.

Another option is to use dynamic assessment in addition to a universal screener. Dynamic assessment assesses students' learning potential (Budoff, Meskin, & Harrison, 1971). Dynamic assessment used in conjunction with traditional screening provides teachers an index of what students' know and of their ability to benefit from instruction in a short period of time. This process increases the accuracy of the identification and reduces the number of children who enter Tier 2 (Fuchs et al., 2007).

A third option is to extend the battery of assessments used during screening. Various combinations have been examined with mixed results. Compton, Fuchs, Fuchs, and Bryant (2006) added an extended battery that included a phonemic awareness task, rapid automatic naming task, and oral vocabulary in addition to the frequently used first grade reading measures of word identification. The use of a composite score resulted in 75% sensitivity and 80% specificity rates. Adding the word identification fluency (WIF) measure did not significantly improve the accuracy in identification rates but adding a 5-week slope and 5-week level on WIF did. However, specificity was still not in the acceptable range.

Two studies examined the used of extended assessment batteries with ELLs. Wiley and Deno (2005) used both an oral reading fluency GOM and a maze GOM to determine the value added of the maze measure in predicting student success on the state assessment. The maze, a task that requires students to read a passage silently and circle the correct word from the three provided was used to address the concern of teachers that ELLs may be able to read fluently but not understand what they read. If this were true, we would expect that the maze would better predict ELLs performance on the state accountability test. The authors found that the maze task contributed to the performance of third and fifth grade English monolingual students but not ELLs. For ELLs, the oral reading fluency was predictive of student performance on the state assessment.

Linan- Thompson, Miciak, Smolkowski, & Baker (2010) examined the predictive power of measures or sets of measures administered at the beginning of first grade on end of first grade on oral reading fluency and general reading measures of ELLs. Because students were still receiving reading instruction in Spanish, they were assessed in both in English and Spanish. We examined the area under the ROC curve and found that Fluidez en Palabras sin Sentido was the most accurate predictor in both Spanish and English, but still functions at only a moderate level in predicting reading fluency outcomes (AUC .739) and general reading measures (AUC .757) in L2. Adding students' score on the Bilingual Verbal Abilities Test, (BVAT) a measure of student language proficiency, did not significantly increase our ability to predict who would perform well on the SAT-10 (AUC .789). Using the BVAT produced only poor accuracy. These findings are consistent with previous research finding weak connections between oral language ability and reading outcomes. This despite established associations between oral language skill and reading proficiency (e.g. Gottardo, 2002; 2009; Lindsay, Manis, & Bailey, 2003; Chiappe, Siegel, & Gottardo 2002). It may be that when students are acquiring foundational skills and they are assessed on those skills directly, oral language ability is less important. Chiappe et al. (2002) found that ELLs performed poorly on oral cloze measures and memory for sentences, tasks that required oral language and working memory but much better on measure of phonemic awareness. Among older students, Kung's (2009) study indicates that language proficiency does impact performance on comprehension measures after controlling for fluency rate. Another possibility is that we are not measuring the aspects of language that impact literacy acquisition.

Language proficiency does not impact the performance of emergent ELL readers on word level measures such as word recognition, word reading accuracy, and pseudo word reading, (Geva & Yaghoub Zadeh, 2006). Language proficiency also does not impact ELLs performance on phonological awareness measures (Chiappe et al. 2002; Leafstedt et al. 2004) or oral reading fluency (ORF). Similarly, Fien et al. (2010) and Leasftedt (2004) found that measures of pseudo word reading served as an overall index of early reading for ELLs and English monolingual students alike.

Yet, ELLs are more likely than their non-ELL peers to enter Tier 2 interventions because they fail to meet passing criteria (Baker & Baker, 2008). The need for instruction is evident but the cause of the difficulty is not known from the measures. If RtI is to be effective, we need to determine what type of instruction is most effective for ELLs and to make that determination we need to know the cause of their poor performance on a screening measure.

The use of an extended screening period or dynamic assessment prior to assigning students to supplemental instruction, Tier 2, may minimize the over-representation of ELLs. Both require the use of multiple data points to make decisions about students' ability to benefit from instruction prior to placing them in interventions. Additionally, the use of these procedures may help to distinguish between ELLs who have learning difficulties and those who have not had adequate instruction or opportunity to learn. Like other students with learning difficulties, ELLs with learning difficulties have a slower rate of learning that becomes evident with the systematic assessment of literacy skills (Linan-Thompson & Hickman-Davis, 2004). On the other hand, ELLs who lacked opportunities to learn, make rapid and consistent gains once they are provided systematic and explicit instruction. Research with alternative assessments and procedures is critical to the implementation of RtI.

Conclusion

Response to intervention as a school-wide prevention model is well recognized since its inclusion in No Child Left behind and IDEIA 2008. Student movement in and out of the tiers of intervention is based on students' ability to meet grade-level expectations and demonstrate progress on specified measures. Though flexibility in the process is key to its success, RtI is often implemented strictly without regard to student needs. Questions about the accuracy of current screening measures in identifying students in need of Tier 2 instruction for English monolingual students remain though a systematic examination of alternatives has begun. Studies with ELLs lag behind. If we are to avoid establishing a system that disproportionately identifies ELLs for special education or denies them appropriate instruction early, research is needed.

These data clearly indicate that relying on a single measure as a benchmark is misguided. If teachers rely on this measure to make decisions about who needs instructional support, they may over identify students in the early grades for Tier 2 interventions denying them the opportunity perhaps, to develop higher levels of academic language if they are tracked needlessly into a Tier 2 reading intervention rather than spending time in content areas. They may also miss those students who read fluently but do not comprehend what they read. Given the student profiles in the Kung study we might surmise that the students who are not passing the state test have not developed the academic language needed to understand the text or take advantage of comprehension strategies they may know. It is also possible that they may not have adequate comprehension strategies. However, all of this is moot if we are not targeting these students for additional instruction because they met a benchmark on the skill we have targeted.

What is clear is that measures of early reading are valid and reliable in providing information about student skill at the time of assessment and in predicting future performance. Less clear is how to interpret the data and which measures may be better predictors of academic success. Future questions to consider are:

- Does oral language become a more robust predictor of reading success for ELLs in later grades?
- What other possibilities exist for identifying ELLs at risk for reading difficulty in early elementary and subsequent grades?
- Do we need to measure language or is it enough to know of its impact on student performance?
- How do we use that information?

References

Ardoin, S.P., Witt, J.C., Suldo, S.M., Connell, J.E. Koening, J.L., Resetar, J.L. et al. (2004). Examining the incremental benefits of administering a maze and three versus one curriculum-based measurement reading probes when conducting universal screening. *School Psychology Review, 33*, 218-233.

Artiles, A., Kozleski, E., Trent, S., Osher, D., & Ortiz, A. (2010). Justifying and explaining disproportionality, 1968-2008: A critique of underlying views of culture. *Exceptional Children, 76*, 279-299.

Artiles, A.J., Rueda, R., Salazar, I., & Higareda, J. (2002). English-language learner representation in special education in California urban school districts. In D.J. Losen, & G. Orfield (Eds.): *Racial inequality in special education* (pp. 117-136). Cambridge, MA: Harvard Education Press.

Baker, S.K., & Baker, D.L. (2008). English language learners and response to intervention: Improving quality of instruction in general and special education. In E.L. Griogrenko (Ed.): *Educating individuals with disabilities: IDEA 2004 and beyond* (pp. 249-273). New York: Springer.

Baker, S., & Good, R. (1995). Curriculum-based measurement of English reading with bilingual Hispanic students: A validation study with second-grade students. *School Psychology Review, 24*, 561-579.

Budoff, M., Meskin, J., & Harrison, R.H. (1971). Educational test of learning- potential hypothesis. *American Journal of Mental Deficiency, 76*, 159-169.

Burke, M.D., Crowder, W., Hagan-Burke, S., & Zou, Y. (2009). A comparison of two paths models for predicting reading fluency. *Remedial and Special Education, 30*, 84-95.

Chiappe, P., Siegel, L., & Gottardo, A. (2002). Reading-related skills of kindergartners from diverse linguistic backgrounds. *Applied Psycholinguistics, 23*, 95-116.

Chiappe, P., Siegel, L., & Wade-Woolley, L. (2002). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading, 6*, 369-400.

Compton, D.L., Fuchs, D., Fuchs, L.S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98*, 394-409.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment, 7*, 303-323.

Davis, G.N., Lindo, E.J., & Compton, D.L. (2007). Children at risk for reading failure: Constructing an early screening measure. *TEACHING Exceptional Children, 39*, 32-37.

Deno, S.L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52*, 219-232.

Deno, S. (2003). Developments in curriculum-based measurement. *Journal of Special Education, 37*, 184-192.

Donovan, S., & Cross, C. (2002). *Minority students in special and gifted education*. National Academies Press.

Fien, H., Baker, S.K., Smolkowski, K., Mercer Smith, J.L., Kame'enui, E.J., & Thomas Beck, C. (2008). Using nonsense word fluency to predict reading proficiency in kindergarten through second grade for English language learners and native English speakers. *School Psychology Review, 37*, 391- 408.

Fletcher, J.M., Coulter, W.A., Reschly, D.J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some   questions and answers. *Annals of Dyslexia, 54*, 304-331.

Fuchs, D., Fuchs, L.S., Compton, D.L., Bouton, B., Caffrey, E., & Hill, L. (2007). Dynamic assessment as responsiveness to intervention: A scripted protocol to identify young at-risk readers. *TEACHING Exceptional Children, 39,* 58-63.

Fuchs, D., Fuchs, L.S., & Stecker, P.M. (2010). The «blurring» of special education in a new continuum of general education placements and services. *Exceptional Children, 76,* 301-323.

Fuchs, L.S., & Fuchs, D. (2007). The role of assessment in the three-tier approach to reading instruction. In D.H. Haager, J. Klingner, & S. Vaughn (Eds.): *Evidence-based reading practices for response to intervention (*pp. 29-42). Baltimore, MD: Brookes Publishing Co.

Fuchs, L.S., Fuchs, D., Hosp, M.K., & Jenkins, J.R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical and historical analysis. *Scientific Studies of Reading, 5*, 239-256.

Geva, E., & Yaghoub-Zadeh, Z. (2006). Reading efficiency in native English-speaking and English-as-a-second-language children: The role of oral proficiency and underlying cognitive-linguistic processes. *Scientific Studies of Reading, 10*, 31-57.

Good, R.H., Simmons, D.C., & Kame'enui, E.J. (2001). The importance of 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*, 257-288.

Gottardo, A. (2002). The relationship between language and reading skill in bilingual Spanish-English speakers. *Topics in Language Disorders, 22*, 46-70.

Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of Spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology, 101*, 330-344.

Hakuta, K., Butler, Y.G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* The University of California Linguistic Minority Research Institute Policy report 2000-1. Berkeley, CA.: University of California.

Hintze, J.M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review, 34*, 372-386.

Individuals with Disabilities Education Act, 20 U.S.C. 1400 *et seq.* (2008).

Kung, S.H. (2009). *Predicting the success on a state standards test for culturally and linguistically diverse students using curriculum-based oral reading measures*. Unpublished doctoral dissertation, University of Minnesota.

Leafstedt, J., Richards, C., & Gerber, M. (2004). Effectiveness of explicit phonological-awareness instruction for at-risk English learners. *Learning Disabilities Research & Practice, 19*, 252-261.

Lesaux, N., & Siegel, L. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology, 39*, 1005-1020.

Linan-Thompson, S., Cirino, P.T., & Vaughn, S. (2007). Determining English language learners response to intervention: Questions and some answers. *Learning Disabilities Quarterly 30,* 185-196.

Linan-Thompson, S., Cirino, P.T., & Vaughn, S. (2007). Determining English language learners response to intervention: Questions and some answers. *Learning Disabilities Quarterly 30,* 185-196.

Linan-Thompson, S., Miciak, J., Smolkowski, K., & Baker, S.K. (2010). Paper presented at the Pacific Coast Research Conference.

Liu, Y., Ortiz, A.A., Wilkinson. C.Y., Robertson, P., & Kushner, M.I. (2008). From early childhood special education to special education resource rooms: Identification, assessment and eligibility determinations for English language learners with reading-related disabilities. *Journal of Assessment for Effective Intervention, 33,* 177-187.

Manis, F.R., Lindsey, K.A., & Bailey, C.E. (2004). Development of reading in grades K-2 Spanish-speaking English-language learners. *Learning Disabilities Research & Practice, 19,* 214-224.

McGlinchey, M.T., & Hixson, M.D. (2004). Using curriculum based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*, 193-203.

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

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Parrish, T.B., Merickel, A., Pérez, M., Linquanti, R., Socias, M., Spain, A. et al. (2006). *Effects of the Implementation of Proposition 227 on the Education of English Learners, K-12*. Final Report for AB 56 and AB 1116. California.

Roehrig, A.D., Petscher, Y., Nettles, S.M., Hudson, R.F., Torgesen, J.K. (2008). Accuracy of the DIBELS oral reading fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology, 46*, 343-366.

Shinn, M.R. (1998). *Advanced applications of curriculum-based measurement*. New York: Guildford Press.

Shinn, M.R. (2010). Building a scientifically based data system for progress monitoring and universal screening across three tiers, including RTI using curriculum-based measurement. In Shinn, M.R., & Walker, H.M. (Eds.): *Interventions for achievement and behavior problems in a three-tier model including RTI* (pp. 259-292). Bethesda, MD: National Association of School Psychologists.

Silberglitt, B., Burns, M.K., Madyun, N.H., & Lail, K.E. (2006). Relationship of reading fluency assessment data with state accountability test scores: A longitudinal comparison of grade levels. *Psychology in the Schools, 43*, 527-536.

Silberglitt, B., & Hintze, J. (2005). Formative assessment using CMB-R cut scores to track progress toward success on state-mandated achievement tests: A comparison of methods. *Journal of Psychoeducational Assessment, 23*, 304-325.

U.S. Department of Education, Office of Special Education and Rehabilitative Services, Office of Special Education Programs (2007*). 27th annual report to Congress on the implementation of the Individuals with Disabilities Education Act, 2005: Vol. 1*. Washington, DC: Author. Available at http://www.edgov/offices/OSERS/OSEP.

Vaughn, S., Linan-Thompson, S., & Hickman-Davis, P. (2003). Response to treatment as a means for identifying students with reading/learning disabilities. *Exceptional Children, 69*, 391-410.

Vellutino, F.R., Scanlon, D.M., & Lyon, G.R. (2000). Differentiating between difficult -to-re mediate and readily remediated poor reader: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33*, 223-238.

Wiley, H.I., & Deno, S.L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education, 26,* 207-214.

Wood, D.E. (2006). Modeling the relationship between oral reading fluency and performance on a statewide reading test. *Educational Assessment, 11*, 85-104.

Ysseldyke, J., Burns, M.K., Scholin, S.E., & Parker, D.C. (2010). Instructionally valid assessment within response to intervention. *Teaching Exceptional Children, 42,* 54-61.

Zehler, A.M., Fleischman, H.F., Hopstock, P.J., Stephenson, T.G., Pendzick, M.L., & Sapru, S. (2003). *Descriptive study of services to LEP students and LEP students with disabilities. Volume 1: Research Report*. Arlington, VA: Development Associates.