Challenges of General Outcomes Measurement in the RTI Progress Monitoring of Linguistically Diverse Exceptional Learners

The assessment for accurate identification and appropriate instruction of English language learners (ELLs) with learning-related disabilities has remained a chronic source of concern. One source of concern that has gone relatively unchallenged is the use of general outcomes measurement (GOMs). The authors examine the problems and challenges of using outcome measures within the response to intervention (RTI) model in the identification and assessment of ELLs who struggle in schools and are suspected of having learning disabilities. They argue here that, despite its importance as a viable form of assessment, GOMs and, concomitantly, RTI may exhibit risks for these learners. In particular, the unique characteristics of ELLs with and without disabilities are often inadequately addressed in current research and practice. The authors review recent educational research on GOMs with respect to ELLs and present current trends in this body of work, along with discussing suggestions and recommendations.
purposes, is a chronic source of concern among educators. It is of special concern to those educators who serve students from culturally and linguistically diverse backgrounds (Abedi, 2009). Research and practice relating to assessment of students with disabilities has not adequately addressed the unique intersection of disability with language learning reflected among ELLs suspected of having learning disabilities (Klingner, McRay-Sorrels, & Barrera, 2007; Ortiz, 1997). At the same time, great effort has been made to improve educational decision-making through the development of response to intervention (RTI) models and related methods (Jimerson, Burns, & VanDerHeyden, 2007).

RTI literature (cf. Linan-Thompson, Cirino, & Vaughn, 2007) has not completely ignored the unique challenges in instructing and assessing ELLs with disabilities. However, this body of work has largely avoided addressing the theoretical and practical challenges posed by these students. Concerns have been raised about the viability of RTI as a method to improve the identification of students with learning disabilities (Gerber, 2005) and especially with respect to ELLs with disabilities (Klingner et al., 2007).

**General Outcomes Measures**

One source of concern has gone relatively unchallenged: the use of curriculum-based measurement (CBM), more recently described as general outcomes measurement (Deno, 2003). *General Outcomes Measures* (GOMs) are defined as a standardized method of assessment for determining academic progress by repeated measurement of student academic achievement outcomes; outcomes being further defined as what a student has learned, or should have learned, within a particular basic skill domain (Deno, 2003). GOMs are primarily used for examining learners’ progress as they matriculate through a school curriculum. They are designed to assess key school learning skills such as basic reading, mathematics computation, writing fluency and accuracy, or spelling (Hasbrouck & Tindal, 2006).

It is important to note that GOMs, for most of their history, have been considered *formative* assessments. They were designed to inform educators on the effectiveness of instruction, not the instruction itself. Long a standard practice of assessment in special education, GOM is now seen as an integral part of the RTI model (D. Fuchs & Fuchs, 1999). To have confidence in the use of RTI, especially as a potential high-stakes decision-making tool, the purposes for using it, and the methods used to determine whether a student responds to instruction, must be not only be effective for all students, but specifically for students at risk for historic over- and underrepresentation in special education (Klingner et al., 2007).

**How Current Assessment Procedures Impede Effective Assessment of ELLs With LD**

GOMs share similar problems with norm- or criterion-referenced assessments in that they assess *static* knowledge about content or skills. For example, a typical GOM for reading (Shinn, 1989) calls on the student to read a passage of specified length, say 250 words, for a specified time (i.e., one minute). Scoring is based on the number of words read correctly and incorrectly. Accuracy and fluency are measured based on the data. The results of these data, however, may be difficult to interpret for evidence of suspected disability. Whether data are collected in English or a different language, one cannot know from the data whether fluency and accuracy scores result from lack of reading experience, language, content experience, or suspected disability. Reading data from GOMs would only provide evidence for lack of level-appropriate reading skills, but would not necessarily eliminate lack of content experience or reading experience as an explanation. Thus, reading data, or other outcomes data, using GOMs would not add significantly to explaining the disability-related basis for ELLs’ learning difficulties.

This article examines the problems and challenges of using GOMs in the identification,
instructional assessment, and student progress monitoring of ELLs who are dually identified with learning disabilities. GOMs are considered a primary method for monitoring student progress within the RTI framework (Christ & Hintze, 2007). Despite its importance as a viable classroom-based assessment, GOMs and, concomitantly, RTI may inappropriately place the responsibility for an observed lack of academic progress onto students, rather than on educators and the broader educational environment where the emphasis belongs. We believe that the use of GOMs as part of the tiered RTI process has overly emphasized validation with standardized assessments such as state assessment scores. Such processes and those sources of assessment have historically played a discriminatory and less than informative role in determining the learning difficulties of ELLs (Ortiz, 1997).

The RTI Model

As a tiered model of intervention, RTI is designed to examine student academic progress founded on a theory of dual discrepancy (L. S. Fuchs, Fuchs, & Speece, 2002). Dual discrepancy theory posits (in our view, correctly) that students with learning disabilities exhibit combined evidence of significantly low academic achievement compared to normally achieving peers and significantly slower rates of academic growth compared to those same peers. RTI focuses on providing evidence-based instruction in classrooms and then in successively smaller group and individualized instructional settings. Students who fail to respond to instructional intervention at each stage move successively into more differentiated, or individualized, instructional settings (D. Fuchs, Mock, Morgan, & Young, 2003).

The dual-discrepancy theoretical model is an important insight and coheres well with current perspectives on the nature of learning disabilities. That is, a learning disability is defined by (a) the difference between poor school performance and a student’s observed normal intellectual capacity, and (b) that discrepancy is explained by a presumed central nervous system dysfunction—an information-processing disorder (National Joint Committee on Learning Disabilities, 1989).

Evidence Base for GOMs With ELLs With Disabilities

It seems particularly important to examine the efficacy of GOMs as a tool for differentiating academic difficulties associated with normal processes of second language acquisition from difficulties associated with a learning disability. As part of this article, we reviewed the most recent literature on the use of GOMs with ELLs from the mid-1990s to the present as a way to determine the extent to which the unique characteristics of ELLs with learning disabilities are included in this area of study (a table of results is available from the authors). Of the 22 studies related to the validation of GOMs (known often in the literature as curriculum-based measurement; CBM) that included ELLs either partially or entirely, 20 were conducted in the elementary grades, 1st to 5th ($n = 31,220$) and two were conducted with secondary students ($N = 260$). For 17 of the studies, the primary focus of study was the technical adequacy and predictive validity of oral reading fluency (ORF) measures in relation to other standardized assessments, especially statewide assessments of reading proficiency. Three studies focused on mathematics computation or early numeracy and two conducted studies on written expression. The number of ELLs involved in these studies was 4,530, with one study (Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008) accounting for 3,560 of the total.

Few of the studies provided descriptions or data on students’ levels of English proficiency as a way to account for differences in observed results although some studies described language or ethnic status. The majority of studies that were conducted involved ELLs who were Latino Spanish speakers, although some studies included the whole range of ELL populations. Four studies provided disaggregated data on ELL populations by language or ethnicity.
In sum, this review revealed the following features of current research on the use of GOMs related to ELLs:

1. There is as yet a small, but growing, body of research regarding the efficacy of GOMs related to ELLs. However, ELLs with disabilities are largely absent in much of this work.
2. The sampling of students (range of age levels, number of students studied), and the accounting for students’ differential academic language proficiency, remains limited.
3. The primary focus of research has been the degree of technical adequacy in the predictive features of GOMs with other GOM measures or other standardized forms of assessment; that is, in a direction away from formative assessment toward features that seem more summative in nature.
4. The largest body of work, perhaps predictably so, has focused on younger learners at the primary grade levels and in the area of reading; a problem when early reading growth and later reading growth may differ owing to age development.
5. There is a tendency to compare language proficiency of ELLs with native English speakers; a further problem when language proficiency development may occur differently between native and second language learners of English.

This review is reflective of a recent, more intensive, meta-analyses synthesizing research on the predictive validity of GOMs on reading achievement (Yeo, 2009). As part of that study, Yeo examined the relative influence of the strength of GOMs in predicting scores on state assessments of reading. Some of his relevant findings were similar to those reported earlier in the current review. Yeo observed that, although GOMs were broadly predictive of state reading scores, the strength of that relationship was inversely related to the number of ELLs or the number of students with disabilities. That is, the larger the number of ELLs or students with disabilities in the study, the lower the predictability. The author cautioned that this result may be a function of smaller sample sizes in studies that had more ELLs or students with disabilities present in the study.

**Characteristics of ELLs With Learning Disabilities**

Three focal points of the characteristics of ELLs are especially problematic in relation to identifying and subsequently educating ELLs with learning disabilities: (a) the racial/ethnic diversity of this population, (b) the variability of the learners’ exposure to academic learning, and (c) the variability of each individual’s progress in learning English as a result (Abedi, 2009). Each of these characteristics holds significant challenges for determining the validity of formal assessments in making high stakes decisions (Abedi, 2009). They also create challenges in using more informal classroom-based assessments for prereferral and determining academic progress (Barrera, 2006).

These issues intersect with research and practice in important ways. For example, research in reading has demonstrated the significant influence of English proficiency levels on observed reading proficiency assessment results and outcomes on high stakes academic assessments (Abedi, 2009). Despite the obvious potential difference in results on English-based assessments among students with differential English proficiency, this variable has, historically, remained minimally addressed within research and practice of educational assessment (Klingner et al., 2007). Very few studies address the time in school as a variable explaining ELLs’ summative or formative results. When this variable is analyzed, the results are significant. Domingo de Ramirez and Shapiro (2006) examined the effect of growth in reading in English and Spanish from 1st to 5th grade of ELLs without disabilities in transitional bilingual programs using GOMs for ORF. Their results indicated significant hierarchies of reading fluency, with proficient English speakers (Latino and White) demonstrating higher English reading fluency than their ELL peers (all Latino). Hence, by the end of 5th grade,
although ELLs’ English reading fluency began to converge with their Spanish reading fluency, both such scores were lower than the scores of English-only speakers (Domingo De Ramirez & Shapiro, 2006). A similar study (Betts, Bolt, Decker, Muyskens, & Marston, 2009) examined the effect of time in the United States and home language on ORF and reading comprehension using GOMs (ORF) and results from a statewide reading assessment. Time in the United States accounted for 10% to 15% of variance in observed results for both fluency and comprehension.

Finally, learning a new or second language confounds the determination of disability-related problems. The linguistic and cultural features that differentiate ELLs with disabilities, especially those with learning disabilities, are a complex interplay of the features that characterize ELLs without disabilities. Students from the same culture can have diverse cultural experiences. It should be expected that ELLs exhibit social acculturation and enculturation differences from their normal peers with the same home culture (Jacobs, 1991). A common problem, however, is that educators may mistakenly determine a student as having a disability because he or she does not behave like other children from the same background. Observed differences in relation to one’s cultural and linguistic peers, although important to examine, are insufficient to determine a disability (Alseben, 2006; Jacobs, 1991; Vang, 2009).

Variation among members of a group makes it difficult to make straightforward comparisons of ELLs by using standardized assessments, which are based on comparisons across a multitude of populations. Any viable assessment for identification of a learning disability must examine how the learner differs significantly from other learners at similar observed learning progress and at similar stages of language proficiency in both languages. Given that these requirements are rather difficult to accomplish, it is important to assure that accurate and current language acquisition assessment, as well as some form of cross-cultural probing, are a prominent feature of any referral or prereferral assessment.

García and Ortiz (2004a) describe a linguistically and culturally sensitive decision-making process in conducting prereferral interventions and assessments. They propose that it is important to account for language level and knowledge base of students, their home background, community experience and the relationship of these with the characteristics of teachers, their instruction, curriculum, and the school environment (García & Ortiz, 2004a). Currently existing programs of teacher support, such as teacher assistance teams, could collect these types of information to support the RTI process for linguistically and culturally diverse students (García & Ortiz, 2006).

**Toward Improved Assessment of ELLs With Potential Learning Disability**

We have illustrated the difficulties in assessing new or second language learners suspected of having learning disabilities in the context of a standardized model of assessment, to which researchers and practitioners of GOMs have taken great pains to conform. The use of GOMs has been among the better alternatives for assessment of ELLs as opposed to relying solely on standardized criterion-referenced measures. However, much recent research on progress monitoring seems to have strayed away from examining use of GOMs as a formative assessment in support of effective instruction and prereferral intervention to one of validation as a standardized instrument. Using standardized GOMs without adequately validating them may result in further marginalizing ELLs with or without suspected disabilities if these students’ scores appear to show that they are not responding to effective instructional intervention.

**Dynamic Assessment**

To avoid this potential for adding yet another ineffective way for assessing ELLs with potential disabilities, some have argued for combining the RTI model with more dynamic forms of
assessment (Grigorenko, 2009; Lidz & Peña, 2009). Indeed, D. Fuchs and Fuchs (2006), among the progenitors of RTI, have argued that RTI is a “form of dynamic assessment because it measures change in students’ level or rate of learning” (D. Fuchs & Fuchs, 2006, p. 94). Although we consider this view promising, there remain significant concerns, especially with the use of GOMs as currently framed. For ELLs with potential learning disabilities, one must find assessment processes that will discern differences between disability-related learning difficulties and the normal process of acquiring a new language. Dynamic assessment (DA) is one such promising practice (Barrera, 2006; Lidz & Peña, 2009).

DA is an approach that examines student learning ability as a function of what a student learns to do as he is being taught rather than on what a student does (or does not) already know (Feuerstein, 1986). DA procedures consist of teaching a specific learning task that is presumably new to the learner and collecting progress and procedural data as the learner absorbs and comprehends—apprehends—the new task. Thus, DA provides a meaningful way for educators to gather important evidence about information-processing difficulties during the prereferral and subsequent referral process at all potential tiers of instruction (Grigorenko, 2009; Lidz & Peña, 2009).

Conclusions and Suggestions

Unfortunately, current models of progress monitoring in the context of GOMs are not suitable to examine this more dynamic evidence of task apprehension, because they are designed to measure outcomes. This observed progress in outcomes is, of course, useful as a snapshot of student growth, but it is an indirect measure because it is largely incomplete when it is devoid of important contextual elements needed to understand that growth. We have presented here those additional elements as the characteristics of ELLs with presumed learning disabilities (described earlier). Hence, we suggest that there are two important conceptual elements in conducting assessment with ELLs that could support better decision-making, regardless of assessment models: making appropriate comparisons and testing what has been taught.

Appropriate comparisons of what students are actually taught involve the examination of student learning in comparison to one’s true peers. These include comparative data:

- Across similar learners,
- By similar levels of language proficiency,
- Across both English and the student’s native language, and
- Of content or skill being taught that is similarly new to all the students of differential levels of ability at the outset of instruction.

In this way, useful comparisons of relative progress—estimated trends in acquiring new tasks, content, or skills—can be made in relation to students of similar type and initially similar exposure to the same new-to-the-students curriculum.

Barrera (2006) has argued that RTI’s focus on response failure presumes that research-based instruction, delivered adequately, will support at-risk students. According to this premise, those who fail to respond to this good instruction are, therefore, candidates for special education. This presumption seems to treat students with potential learning disabilities as if they are all the same, instead of assuming that each individual has unique characteristics. One has to ask, “What happens when the student doesn’t ‘fit’ neatly into any tier?” What happens when even special education interventions do not work?

The entire purpose of special education has been to support individuals who do not fit into models. It seems essential that current practices in special education, general education, or tiered models of intervention begin to refocus on the needs of learners who do not fit neatly into categories. ELLs whom we suspect may have learning disabilities are, indeed, a group of students who require this level of rededication to basic special education principles.
References


