Within-year Growth in Math: Implications for Progress-Monitoring Using RTI

Within-year Growth in Math: Implications for Progress-Monitoring Using RTI

Public schools face challenges when attempting to improve student learning and achievement. Educators implementing school-wide improvement efforts such as Response to Intervention (D. Fuchs & Fuchs, 2001), or RTI, attempt to meet the diverse learning needs of students performing below expectations. The enactment of this goal, in part through ongoing progress monitoring aimed at characterizing within-year student growth, is the focus of this study.

RTI can be viewed as a grassroots school-wide improvement effort based on results of instructional intervention as measured by interim and formative assessments (Black & Wiliam, 1998). In RTI, students are classified as "at-risk" of not meeting grade-level expectations through interim screening assessments, typically administered seasonally over the academic year. Students performing below a district- or school-designated level on interim screeners are provided an instructional intervention designed to improve achievement, while simultaneously administered frequent formative progress-monitoring probes to track the effect of the intervention and ensure adequate gains (D. Fuchs, Mock, Morgan, & Young, 2003). RTI is intended to be dynamic, with students receiving individualized intervention based on their learning needs. The RTI process, of student classification and subsequent intervention and progress monitoring, ideally results in students meeting grade-level expectations based on meeting the designated performance level.

In contrast to longer statewide summative exams, formative assessments are intended to provide teachers with quick information to *guide* instructional decisions, rather than to evaluate the *results of* instruction. Formative assessments must necessarily be shorter, and have multiple equivalent test forms, so that teachers can evaluate within-year student growth. Regardless of these challenges, results of formative assessment are a key basis for the validity of score interpretations and instructional decision-making when characterizing student growth within the context of RTI (D. Fuchs, et al., 2003). Thus, the capacity of such assessments to adequately capture the within-year growth of students performing below expectations is essential, and is specifically of interest to a research base with limited findings around within-year growth in math (Foegen, Jiban, & Deno, 2007).

Research Questions

In this paper, we use a three-level hierarchical linear model (HLM) to examine the growth of fourth grade students who received mathematics progress-monitoring probes during the 2010-2011 school year. We examine growth through an RTI lens to address three questions of substantive importance relative to measuring student progress in mathematics:

1. Can fourth grade student growth in higher-order mathematics skills be adequately captured using progress-monitoring assessments administered over a short period of time with a limited 16-item scale? To date, much of the research involving math progress-monitoring measures has focused on early developmental skill areas (i.e., addition, subtraction, multiplication, and division); however, the measures used in this study are comprised of higher-order math skills including problem solving in real world contexts. Given previous research demonstrating strong technical qualities of the measures used in this study, (Nese et al., 2010), we hypothesize that significant student growth will be observed.
2. Given that RTI is enacted to raise the achievement of students performing below expectations, what is the effect of below grade-level progress monitoring on the student? That is, by the time these students are monitored on grade-level, how did

their rate of growth compare to students who were only monitored on grade-level? We hypothesize that students who took below-grade progress monitoring forms would make significantly different growth compared to those who took solely on-grade measures. The former group presumably had lower mathematical knowledge and skills at some point to necessitate below-grade progress monitoring, and yet these students made enough progress to subsequently be progress-monitored on-grade. We thus had reason to believe that these students would have a higher rate of growth.

3. Given that RTI is designed around a dynamic system of classification, assessment and instruction, what is the effect of the number of progress-monitoring measures a student received on the subsequent observed growth? We hypothesize that students administered more progress-monitoring measures would have a lower initial starting point and a lower rate of growth given they were apparently not “exited” from the intervention group as quickly as students administered fewer measures.

Methods

Measures

Two measures were used in this study: the fourth grade easyCBM® fall math interim screener and the fourth grade easyCBM® *Number and Operations* progress monitoring probes. The fall interim screener was not included in any analyses, but was used to help select the final sample of students. All easyCBM® math progress-monitoring forms were designed to be of equivalent difficulty with a Rasch model (Alonzo, Anderson, & Tindal, 2009). The *Number and Operations* probes were selected because they were designed to measure higher-order, more complex mathematical skills, which adds to a research base primarily focused on monitoring early developmental areas (see L. Fuchs, Fuchs, & Courey, 2005). A more complete description of the measures used and their technical adequacy will be detailed in the final conference paper.

Data Collection and Sample

Existing data from the 2010-2011 school year were extracted from the easyCBM® database. It is important to note that these data were not collected with any sort of experimental design, or pre-defined research questions in mind. Rather, these data reflect teachers' observed progress-monitoring practices in the field. Overall, the data were structured in three levels: *Math Measures* nested in *Students* nested in *Schools*.

We began with a large dataset from multiple states and systematically restricted it to obtain a specific analytic sample: students and schools exhibiting practices consistent with participation in progress-monitoring and RTI, as conceptualized by Fuchs and Fuchs (2001). It is important to note that because these are extant data, and because we wanted to constrain our sample specifically to those participating in RTI, data preparation was complex and lengthy. A condensed description is provided in Table 1, with preliminary analyses limited to the state of Oregon. The conference version of the paper will include data from multiple states, in addition to a complete account of data rendering.

Model Building

Previous research has shown that within-year reading growth may follow a decelerating curvilinear trend (Christ, Silberglitt, Yeo, & Cormier, 2010; Nese et al., 2012). Little research has examined within-year growth in math (Foegen, et al., 2007), however, so we had no a priori assumption of linear/nonlinear growth. We therefore tested linear and quadratic terms at level 1 (L1). Because the initial progress monitoring measure for a given student was administered at various times throughout the 2010-2011 academic year, time was entered into the model centered by individual student. That is, students' initial time point was coded 0, with all subsequent time points coded as calendar weeks from the first administration. Time was coded in weeks to stay consistent with progress-monitoring research in reading (e.g., Christ, et al., 2010; Nese, et al., 2012). Days within weeks were accounted for by adding decimals to the hundredths place of time values.

Two predictors of interest were modeled at level 2 (L2). First, given RTI's focus on students performing below grade-level expectations, we examined whether students who received *below grade-level* progress monitoring differed in their intercept and/or slope for on-grade level data (research question 2). A dummy vector called *below grade-level* was created and entered at L2 uncentered, with students only receiving on-level monitoring serving as the referent group. Also, given the dynamic nature of individualized assessment and instruction inherent to RTI, we added a second variable called *count* at L2, which was simply a count of the number of progress monitoring measures the student received (research question 3).

Lastly, there was one predictor of interest, *school size*, at level 3 (L3). School size data were gleaned from Oregon Department of Education (http://www.ode.state.or.us/data/reportcard/reports.aspx) and were entered as a control variable because schools ranged considerably in size (from 124 to 596 students). Analyses were run with HLM 7 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011) with full information maximum-likelihood estimation. Residuals for all predictors were investigated for adherence to the underlying assumptions of HLM and revealed no violations.

Results

Due to space limitations, we describe only our final model, the unconditional growth model, based on the preliminary sample and analysis. Though greater detail of our entire model building process will be provided in the conference version of the paper, overall, HLM model building decisions and analyses were based on a premise of adding intended predictors sequentially within a given level, and retaining only those parameters that were significant at the α = .05 level. Additionally, we conducted chi-square test of the change in the deviance statistic to compare the fit of higher-level models to finalized baseline models at the lower levels, including in cases where parsimony was a reasonable consideration. Intraclass correlation coefficients (ICC) and pseudo-R^2 statistics were also calculated to measure the degree of dependence at higher levels and amount of variance accounted for by the models, respectively. It should be noted that although predictor parameters at higher levels (L2 and L3) were investigated in our preliminary analyses, they were found to be nonsignificant predictors of student growth in math. Tables 2 and 3 detail fixed effect and random parameter estimates for the unconditional means (null) model (Model 1) and the unconditional growth model (Model 2) the final model specified thus far in the study.

The unconditional growth model, the final model specified in our preliminary analysis, was defined as,

$$
\begin{aligned}
&\text{L1: } Score_{tij} = \pi_{0ij} + \pi_{1ij}(Weeks) + \pi_{2ij}(Weeks^2) + e_{tij} \\
&\text{L2: } \pi_{0ij} = \beta_{00j} + r_{0ij} \\
&\quad \pi_{1ij} = \beta_{10j} + r_{1ij} \\
&\quad \pi_{2ij} = \beta_{20j} \\
&\text{L3: } \beta_{00j} = \gamma_{000} + u_{00j} \\
&\quad \beta_{00j} = \gamma_{100} + u_{10j} \\
&\quad \beta_{00j} = \gamma_{200}
\end{aligned} \tag{1}
$$

where at L1 the math score at time *t* for individual *i* nested within school *j* was equal to an intercept, π_{0ij}, a linear growth term (weeks and days between assessment occasions), a quadratic growth term (weeks and days squared), and unexplained within person variance, e_{tij}. At L2 the intercept, π_{0ij}, is predicted by a student level intercept, β_{00j}, which varies randomly across students, r_{0ij}. Finally, at L3 the student level intercept β_{00j} is predicted by a school level intercept, γ_{000}, which varies randomly across schools, u_{00j}. Adding the quadratic term to the unconditional growth model resulted in significantly better model fit, as indicated by the chi-square test of the deviance statistic, $\chi^2 = 40.60(7)$, $p < .001$. However, the random effects for the quadratic term at L2 and L3 were not significant and thus removed from model.

Discussion

Though more detailed results and discussion related to an expanded analytic sample will be included in the conference version of this paper, perhaps the most important finding from our preliminary analyses is related to capturing student growth in math. In this study, fourth grade students who were progress-monitored made within-year growth that was observable through the 16-item *Number and Operations* probes. Upon first glance, this finding may seem trivial, as one would expect students to make growth during the school year. However, the growth was observed in math, with measures targeting non-developmental skills with a fairly limited scale. The items from the easyCBM® measures used therefore go beyond assessing students' fluency with algorithms, to assessing students application of the mathematical concepts. Yet the administration time remained quite short, taking roughly 12 minutes (http://www.rti4success.org/tools_charts/popups_progress/easyCBMMath_area.php). Anderson, Lai, Alonzo, and Tindal (2011) showed that the bulk of easyCBM® math items were targeted at students performing below expectations. Thus, the results of our study extend the findings of

Anderson and colleagues, suggesting that efficiently administered progress-monitoring measures are not only sensitive to the achievement of students performing below expectations at a single time point, but also to improvements over time. Additionally, these findings add substantively to the limited empirical research base around within-year growth in mathematics (Foegen, et al., 2007).

In addition to providing greater detail regarding the easCBM® interim and formative measures, data collection/preparation, and model building and specification, we plan to follow a few additional steps in preparing the conference version of the paper. First, to broaden generalizability, the analytic sample will expand to include students and schools from multiple states. Secondly, a covariate for when students began being progress-monitored will be added at L1 as the initial monitoring point ranges considerably in the preliminary sample.

References

Alonzo, J., Anderson, D., & Tindal, G. (2009). *IRT analysis of general outcome measures in grades 1-8* (Technical Report No. 0916). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Lai, C., Alonzo, J., & Tindal, G. (2011). Examining a grade-level math CBM designed for persistently low-performing students. *Educational Assessment, 16*, 15-34. doi: 10.1080/10627197.2011.551084

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy, and Practice, 5*(1), 7-74.

Christ, T. J., Silberglitt, B., Yeo, S., & Cormier, D. (2010). Curriculum-based measurement of oral reading: An evaluation of growth rates and season effects among students served in general and special education. *School Psychology Review, 39*, 447-462.

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *Journal of Special Education, 41*, 121-139. doi: 10.1177/00224669070410020101

Fuchs, D., & Fuchs, L. (2001). Responsiveness-to-intervention: A blueprint for practitioners, policymakers, and parents. *Teaching Exceptional Children, 38*(1), 57-61.

Fuchs, D., Mock, D., Morgan, P., & Young, C. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research & Practice, 18*(157-171). doi: 10.1111/1540-5826.00072

Fuchs, L., Fuchs, D., & Courey, S. (2005). Curriculum-based measurement of mathematics competence: From competence to concepts and applications to real life problem-solving. *Assessment for Effective Intervention, 30*(2), 33-46. doi: 10.1177/073724770503000204

Nese, J. F. T., Biancarosa, G., Anderson, D., Lai, C. F., Alonzo, J., & Tindal, G. (2012). Within-year oral reading fluency with CBM: A comparison of models. *Reading and Writing, 25*, 887-915. doi: 10.1007/s11145-011-9304-0

Nese, J. F. T., Lai, C. F., Anderson, D., Jamgochian, E. M., Kamata, A., Saez, L., . . . Tindal, G. (2010). Technical adequacy of the easyCBM mathematics measures: Grades 3-8, 2009-2010 Version (Technical Report No. 1007). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). HLM Version 7: Scientific Software International, Inc.

Table 1

Data Cleaning Methods by Level of Analysis

Level of analysis/ Cleaning step	Cleaning method
L1: Measures	
1	Only data from *Number and Operations* probes included
2	Only data from 4^{th} grade measures included
L2: Student	
3	Student was in fourth grade during the 2010-2011 school year
4	Student had a valid fall benchmark score and scored below the 50^{th} percentile
5	Student received a minimum of two progress monitoring measures
L3: School	
6	School based in Oregon
7	School had at least 4 students receiving progress-monitoring

Note. Study began with a raw data file of 4,375 students. Restricting the sample to only students with a valid fall benchmark measure resulted in a sample of 2,973, of which 1,799 scored below the 50^{th} percentile. Restricting the analysis to only students and schools in Oregon resulted in a sample of 1,137. Eliminating schools from step 7 above resulted in 1,064 students. Restricting the sample to only grade 4 data resulted in 1,021 students, 950 of which had received a progress-monitoring probe in *Numbers and Operations.* Finally, restricting the data to only students with at least 2 progress-monitoring measures resulted in the final sample of 1,990 data points nested in 573 students nested in 33 schools.

Table 2

Fixed Effects Estimates for Unconditional Means and Unconditional Growth Models of Within-year Student Growth in Mathematics

Parameter	Model 1	Model 2
Intercept (π_{0ij})	9.63*** (0.20)	9.06*** (0.18)
Level 1 (time)		
Weeks (π_{1ij})	--	0.13*** (0.02)
Weeks^2 (π_{2ij})	--	-0.004*** (0.00)
Level 2 (student)		
Count (int)	--	--
Count (linear)	--	--
Count (quad)	--	--
Level (int)	--	--
Level (linear)	--	--
Level (quad)	--	--
Level 3 (school)		
Size (int)	--	--
Size (linear)	--	--
Size (quad)	--	--

Note. Fixed effect estimates for: *count* (number of progress-monitoring measures the student received), *level* (students who received below grade-level progress-monitoring), and *size* (number of students in the schools). *$p < .05$, **$p < .01$, ***$p < .001$.

Table 3

Random Parameter Estimates for Unconditional Means and Unconditional Growth Models Within-year Student Growth in Mathematics

Parameter	Model 1	Model 2
Level 1		
Intercept (e_{tij})	4.79 (0.18)	3.72 (0.16)
Level 2		
Int (r_{0ij})	2.57*** (0.26)	2.67*** (0.33)
Weeks (r_{1ij})	--	0.001* (0.00)
Level 3		
Int (u_{00j})	1.00*** (0.33)	0.71*** (0.27)
Weeks (u_{10j})	--	0.01*** (0.00)
-2*log likelihood	9373.45	9095.30***

Note. Significance reported for the -2*log likelihood are based on the chi-square test of deviance, whereby the given model was compared to the baseline model established at the previous level. For example, Model 2 (unconditional growth) was compared to Model 1 (unconditional means), using the chi-square test of the change in the deviance statistic. *p < .05, **p < .01, ***p < .001.