# National Middle School Mathematics Within-Year Growth Norms

Daniel Anderson, P. Shawn Irvin, Joseph F. T. Nese, Julie Alonzo, and Gerald Tindal

University of Oregon

Acknowledgements:

Note: Funds for the datasets used in creating this report came from federal grants awarded to the UO from the Institute of Education Sciences, U.S. Department of Education: Developing Middle School Mathematics Progress Monitoring Measures, #R324A100026.

**Abstract**

Educational decisions are regularly based on students' achievement level and rate of growth within response to intervention frameworks. Normative cut-points are generally established relative to students' level of achievement, and progress is documented for students scoring below the cut point. Students' whose level of achievement or rate of progress is deemed sufficiently low are provided educational interventions to increase their rate of progress. It is therefore important that typical rates of progress be understood. In this paper, we provide national normative monthly growth rates for students in Grades 6-8 on the easyCBM© CCSS Mathematics measures. Normative deciles of growth are produced based on quintiles of initial achievement. Overall, we found typical rates of growth are quite modest.

**National Middle School Mathematics Within-Year Growth Norms**

Within a response to intervention (RTI) framework, teachers regularly make instructional grouping decisions based on students' current level of achievement and their rate of improvement (Tindal, 2013). If either is sufficiently below expectations, the student may be identified as academically at-risk. Students identified as at-risk are generally provided an academic intervention, with regular progress-monitoring probes administered to document the students' responsiveness to the particular instructional practice. Yet, determining what is "sufficiently" low can be challenging. Student levels of achievement are generally compared against local or national norms, with cut points for determining risk set at a school- or district-determined percentile (typically 20$^{th}$ or 25$^{th}$).

Documenting sufficiently low rates of improvement is more nebulous because, in part, normative data are either unavailable as a comparator or the growth rates across benchmarks are not linear (Nese et al., 2013). Various decision rules have been suggested and applied as an alternative, such as using aim lines, where educators set projected goals for students' achievement and draw a line between the current and projected levels. If a student consecutively scores below the aim line (e.g., 3-5 data points; Ardoin, Christ, Morena, Cormier, & Klingbeil, 2013) then the students' rate of improvement is thought to be too low and an instructional change is made. Normative data from each time-point can help guide aim line goal setting, but without understanding typical rates of improvement across the performance range, educators may set unreasonably ambitious or safe goals.

RTI has a long history of application with reading in the elementary grades, using curriculum-based measurement (CBM) to document progress. Oral reading fluency probes are by far the most common CBM, with students' scores reported on a *correct words read per minute* scale. Educators evaluate students' growth by the rate at which they acquire additional words read correctly, which is generally reported on a weekly scale. Tindal (2013) reviewed a wealth of previous research and found weekly growth in oral reading fluency typically ranges from 0.5 to 2.0 words per week.

In mathematics, the limited CBM research conducted has focused almost exclusively on the elementary grades and early developmental skills (Foegen, Jiban, & Deno, 2007). Evaluating typical rates of growth in math is difficult because of tremendous variability in the design of the measures and methods for scoring. Additionally, growth is typically more suppressed in mathematics compared to oral reading fluency. The exception is very early mathematics skills, such as number identification and basic computation, which can be assessed on a fluency-based scale. Yet, even with these skills, measures typically display an average of less than one point of growth per week (Foegen et al.). The problem becomes even more pronounced as students move into the middle-school, where the math concepts are more difficult and nuanced, and the measures necessarily move beyond indicators of fluency. In these contexts, having a relevant normative comparator becomes even more important. In other words, if raw growth rates are evaluated free of context, one might be left to conclude that an individual's growth is too slow, when in fact it may be well above the norm. The rate the student is progressing may still be deemed insufficient (i.e., if the student's performance level is at the normative 15$^{th}$ percentile), but the instructional decisions may well change given the normative comparator relative to students' rate of growth.

In either reading or mathematics CBM, growth has seldom been reported in any finer detail than an average grade level value. Furthermore, the samples used are generally convenience samples with unknown or disproportionate representation across the country. The purpose of this study is to provide normative data on students' rate of improvement within the academic year in mathematics for Grades 6-8 for a large-scale interim/formative CBM, easyCBM©. This information can be used in conjunction with cross-sectional (seasonal) normative data to help educators determine reasonably ambitious goals for students, given their current level and rate of achievement. For example, teachers may raise a student's goal if they realize that the student would only need to progress at the normative 30$^{th}$ percentile to achieve that goal. Ideally, students who begin the year performing below their peers would progress at a faster than normal rate (> 50$^{th}$ percentile) to make up the gap between their current level of achievement and that of their peers; however, if goals are set too high (e.g., requiring the student to progress at the normative 95$^{th}$ percentile), then expectations are likely unrealistic, potentially leading to feelings of low self-efficacy when a goal is not reached (Schunk, 1985) or discontinuing an intervention that was otherwise successful. Modification of goals leads to the aim line being redefined and responsiveness reevaluated (i.e., number of time points above/below the aim line). In this study, we address the following research questions:

1) What is the national average rate of improvement (growth) in mathematics for students in each of Grades 6-8?

2) To what extent does growth vary as a function of students and region (West, Midwest, Southwest, Northeast)?

3) What is the national average growth for students by decile?

## Methods

### Measures and Data Source

This study utilized a large extant dataset from the easyCBM© Common Core State Standards (CCSS) Math measures, drawn from the 2013-2014 school year. Sample demographics for Grades 6-8 are displayed by region in Tables 1-3, respectively. Data from the 2013-2014 seasonal fall, winter, and spring benchmark testing periods were used to calculate growth norms. We restricted the analytic sample to students who responded to all items on at least two of the seasonal math benchmarks.

The easyCBM© CCSS Math tests in Grades 6-8 include 45 items, with approximately six items written to measure each of the five CCSS math domains (30 items) at the designated grade level. The remaining items were off-grade level and/or aligned with different standards to support future vertical scaling efforts. All items included three response options and were designed with universal design features to maximize accessibility (Thompson, Johnstone, & Thurlow, 2002). Measures within each grade were of roughly equivalent difficulty on a raw-score basis, with items sampled from an item bank conditional on their estimated difficulty. Growth within the year could thus be monitored on the raw scale. Items were calibrated with a Rasch model (Anderson, Alonzo, & Tindal, 2013b; Anderson, Irvin, Patarapichayatham, Alonzo, & Tindal, 2012). The degree to which items fit the expectations of the Rasch model, along with the estimated item difficulty, were considered during item selection. Wray, Alonzo, and Tindal (2014) investigated the reliability of the CCSS Math tests and found internal consistency ranged from .92 to .95 across Grades 6-8, with split-half reliability ranging from .80 to .87 for the first half and .92 to .95 for the second half. Anderson, Rowley, Alonzo, and Tindal (2014) found the concurrent relation between easyCBM© CCSS Math and the SAT-10 math test ranged from .75 to .82 across grades, while simple linear regression analyses indicated that the measure accounted for 56%-67% of the variance in students' SAT-10 math scores.

**Analyses**

A linear latent growth curve model was fit for each grade. Factor loadings for the latent *Intercept* factor were all fixed at 1.0, while loadings for the latent *Slope* factor were specified according to individually varying time vectors, representing the number of months elapsed between assessments. The time vectors were coded in fractional form, rounded to the nearest hundredth, to represent the specific number of days between assessment occasions. The intercept was centered on the earliest assessment occasion within each grade. For many students, initial achievement represented a backward projection, given their slope and the time between the date on which the intercept was centered and the date on which the fall assessment was administered. Growth was modeled as a linear function of time.

An unconditional growth model was fit first, followed by a conditional model that included *Region* as a dummy-coded predictor of students' intercepts and slopes. *Region* included four levels: Northeast, Southeast, Midwest, and West. The *West* region was coded as the reference group because it had the largest sample size. We used Akaike's information criteria (AIC) and Bayesian information criteria (BIC) to compare the extent to which *Region* contributed to the model. Factor scores were extracted from the latent *Slope* factor for each student. National normative deciles for growth were then calculated after splitting the sample into quintiles based on students' estimated intercept.

**Results**

Descriptive statistics for Grades 6-8 are displayed by region and seasonal math benchmark in Tables 4-6. Results for the grade-level unconditional (Model 1) and conditional (Model 2) growth models are displayed in Tables 7-9. For all three grades, AIC and BIC estimates suggested that the conditional growth model specifying *Region* fit the observed data better than the unconditional model as a predictor of intercept and slope.

**Grade 6.** Sixth-grade students attending schools in the *West* region initially scored, on average, 24.76 out of 45 possible points on the fall CCSS Math benchmark, which varied significantly between students with a standard deviation of 6.33 points ($p < .001$). Students in the *Southeast* region initially averaged 1.84 points higher than students in the *West* region, whereas students in the *Midwest* and *Northeast* regions did not differ significantly in their initial math achievement. On average, students in the *West* grew at a linear rate of 0.55 points per month, which varied significantly between students, with a standard deviation of 0.37 points per month ($p < .001$). Students in the *Midwest* progressed, on average, 0.05 points per month slower than students in the *West*, whereas students in the *Southeast* grew, on average, 0.16 points per month faster. Both effects were significant ($p < .001$). Students attending schools in the *Northeast* did not differ significantly in their rate of growth from their peers in the *West*. The correlation between the intercept and slope was weakly positive, $r = .18$, $p = .04$. Normative deciles for growth for Grade 6 by intercept quintile are displayed in Table 10.

**Grade 7.** Seventh-grade students in the *West* region initially scored, on average, 24.05 points out of 45 possible points, which varied significantly between students with a standard deviation of 6.34 points. Students in the *Southeast* region scored, on average, 2.32 points higher than students in the *West*, whereas students in the *Northeast* scored 1.38 points less. Both effects were significant ($p < .001$). Students in the *Midwest* did not differ significantly in their initial math achievement from students in the *West*. On average, students in the *West* gained 0.51 points per month, which varied significantly between students with a standard deviation of 0.52 points per month. Students in the *Northeast* grew, on average, 0.10 points higher than students in the *West* ($p < .001$). Students attending school in the *Midwest* and *Southeast* did not differ significantly in their rate of growth from students in the *West*. The correlation between the intercept and slope was weakly positive, $r = .34$, $p = .001$. Normative deciles for growth for Grade 7 by intercept quintile are displayed in Table 11.

**Grade 8.** Eighth-grade students attending school in the *West* scored, on average, 24.29 points out of 45 possible points on the fall CCSS Math benchmark, which varied significantly between students, with a standard deviation of 6.46 points. On average, students in the *Southeast* region initially scored 2.28 points higher compared to those in the *West* region, a difference that was significant, $p < .001$. Students in the *Northeast* and *Midwest* did not significantly differ in their initial math achievement compared to students in the *West*. On average, students in the *West* grew at a linear rate of 0.82 points per month, which varied between students with a standard deviation of 0.20 points per month. Students in the *Midwest* grew, on average, 0.16 points per month slower than students in the *West*, while students in the *Northeast* grew, on average, 0.05 points per month slower than students in the *West*, both of which were significant ($p < .05$). Students attending school in the *Southeast* did not differ significantly in their rate of growth from students in the *West*. The correlation between the intercept and slope was moderately positive, $r = .45$, $p < .001$. Normative deciles for growth for Grade 8 by initial quintile are displayed in Table 12.

## Discussion

The results of our study suggest that students' average rate of growth on the easyCBM© CCSS mathematics measures within the school year is quite small for Grades 6 and 7, averaging about a half point per month, whereas the average monthly growth rate in Grade 8 appears a bit larger at just over four-fifths of a point per month. The variance between students in their monthly growth was quite large in Grades 6 and 7, but less so at Grade 8. For example, the difference between students one standard deviation below versus above the mean growth for students in the sample would result in a predicted difference of 7.41, 9.88, and 4.18 points of growth across the school year, in Grades 6-8 respectively. These correspond to roughly .95, 1.20, and .50 standard deviations on the spring assessment, respectively.

Our results also suggested regional differences in initial average math achievement and the average growth rate in all three grades. In Grade 6, for example, our sample of students in the *Southeast* region initially scored almost two points higher and grew at almost two-tenths of a point faster per month, on average, compared to their peers in the *West*. We also documented a higher average initial status in Grades 7 and 8 (over two points higher in each grade) for students in the *Southeast*, though their average growth was statistically the same as students in the *West* for both grades. Conversely, students in Grades 6 and 8 in the *Midwest* region, on average, demonstrated significantly less growth than their peers in the *West*—in Grade 8 the average monthly growth rate was almost two-tenths of a point less. Whether these and other regional differences in average initial math status and growth rate are practically meaningful and indicative of the broader student population in these grades is a question that future work should seek to answer.

**Limitations**

Our study is limited in a number of ways. First, we assumed that all students followed a linear trajectory throughout the year. However, visual inspection of the data suggests this assumption may not be warranted, as a decelerating trend was regularly observed. Second, the *West* region was overrepresented in the sample, with approximately three to five times as many students as any other region. Third, all growth was observed on a raw-score scale with test forms that were pre-equated to make them roughly equivalent in terms of difficulty and the distribution of item difficulties. Although roughly comparable, these forms were not statistically equated.

**Conclusions and Future Directions**

Perhaps the primary takeaway from our study is simply that the average mathematics growth for students occurring during the year is modest, and expectations of growth for instructional decision-making should reflect the empirical normative gains. Modest growth in the area of mathematics, particularly with skills beyond basic arithmetic, is not a new finding, and our results replicate those of previous research (see Foegen et al., 2007). However, these results are still useful. For example, our results suggest that expecting students to gain an average of approximately one point per month is, in most cases, unrealistic, as this would place the student in at least the $90^{th}$ percentile of normative growth for all three middle school grades. At the same time, modest growth should not necessarily be discounted. For example, if a sixth- or seventh-grade student gains ~2 points over a trimester (assuming three month trimesters), his or her growth would be between the $50^{th}$ and $80^{th}$ percentile, depending on initial status. These are gains that should be celebrated, rather than discarded due to their apparent modest increase on the raw-score scale.

References

Anderson, D., Alonzo, J., & Tindal, G. (2013b). *easyCBM CCSS Math Item Scaling and Test Form Revision (2012-2013): Grades 6-8* (Technical Report 1313). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Irvin, P. S., Patarapichayatham, C., Alonzo, J., & Tindal, G. (2012). *The development and scaling of the easyCBM CCSS middle school mathematics measures* (Technical Report 1207). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Anderson, D., Rowley, B., Alonzo, J., & Tindal, G. (2014). *Critierion Validity Evidence for the easyCBM CCSS Math Measures: Grades 6-8* (Technical Report No. 1402) Eugene, OR: Behavioral Research and Teaching: University of Oregon.

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding Curriculum-Based Measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology, 51*, 1-18.

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

Nese, J. F. T., Biancarosa, G., Cummings, K., Kennedy, P., Alonzo, J., & Tindal, G. (2013). In search of average growth: Describing within-year oral reading fluency growth across grades 1-8. *Journal of School Psychology, 52*, 625-642.

Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools, 22*, 208-223.

Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (Synthesis Report 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Retrieved November 2, 2011, from the World Wide Web: http://education.umn.edu/NCEO/OnlinePubs/Synthesis44.html.

Tindal, G. (2013). Curriculum-Based Measurement: A Brief History of Nearly Everything from the 1970s to the Present. *ISRN Education, 2013*, 29. doi: 10.1155/2013/958530

Wray, K., Alonzo, J., & Tindal, G. (2014). *Internal consistency of the easyCBM CCSS Math Measures: Grades K-8* (Technical Report No.1405). Eugene, OR: Behavioral Research and Teaching, University of Oregon.