

aimsweb

Progress Monitoring Guide

Pearson Executive Office 5601 Green Valley Drive Bloomington, MN 55437
800.627.7271 www.PsychCorp.com

Copyright © 2012 NCS Pearson, Inc. All rights reserved.

Warning: No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system without permission in writing from the copyright owner.

Pearson, aimsweb, and the **aimsweb logo** are trademarks in the U.S. and/or other countries of Pearson Education, Inc., or its affiliate(s).

Produced in the United States of America.

Product Number 38821

1 2 3 4 5 6 7 8 9 10 11 12 A B C D E



Contents

Foundations of Progress Monitoring	1
Components of Progress Monitoring	2
Performance-Level Percentile Norms	2
Norm Groups	2
Performance Benchmarks	3
Deciding Whose Progress to Monitor	4
Goal Setting	8
Selecting the Grade Level of the Material for the Goal and Progress Monitoring	8
Evaluating Proposed Goals Using ROI Growth Norms	9
Case Studies	10
Case 1: Using Performance-Level Percentiles and ROI Growth Percentiles	10
Case 2: Setting an Off-Grade Performance Goal	12
Case 3: Using Benchmarks for Goal Setting	13
Interpreting Progress Monitoring Data	14
Visual Analysis	14
Comparing Consecutive Score Points Relative to the Aimline	16
Comparing Trendline and Aimline	17
Confidence Intervals for Scores at the Goal Date	18
Conclusion	20
References	21
Appendix	22
Procedure for Calculating a Confidence Interval Around a Projected Goal Score	22

Figures

1. Distribution of R–CBM scores, local norm-referenced performance levels, and Tiers for a classroom, Example 1	6
2. Distribution of R–CBM scores, local norm-referenced performance levels, and Tiers for a classroom, Example 2	7
3. Distribution of R–CBM words read correctly for a low-achieving Grade 6 classroom	13
4. The impact of an outlier on the estimated trendline	15
5. Nonlinear patterns	15
6. Low and high scatter patterns	16
7. R–CBM score patterns with trendline, aimline, and ROIs	17
8. Predicted score 75% CI is completely below the target score	19
9. Predicted score 75% CI includes the target score	19

Foundations of Progress Monitoring

Progress monitoring is a central component of a response to intervention (RTI) system; it is the frequent and ongoing collection of information about student performance to be able to evaluate the effectiveness of instruction. Specifically, a student's rate of improvement is compared to an expectation of growth to determine if adequate progress is being made toward a long-term goal or if changes to instruction are needed to improve growth. Progress monitoring is integral to identifying at-risk students whom are not responding well to their instructional programs.

The No Child Left Behind (NCLB) Act (2001) and the Individuals with Disabilities Education Act (IDEA) (2004) require schools to evaluate the effects of evidence-based instruction, a task for which RTI is ideally suited. Monitoring each student's response to intervention enables teachers to gauge the effectiveness of the core curriculum and align resources with student needs. Studies indicate that using progress monitoring to guide instruction improves student achievement (Black & Wiliam, 1998; Fuchs & Fuchs, 1986), especially those with low achievement and those with disabilities, when qualified professionals implement the following features (Fuchs & Fuchs, 1986; Stecker, Fuchs, & Fuchs, 2005):

- Data are frequently collected
- Data are graphically displayed and reviewed
- Decision rules—to continue or modify instruction—are explicit
- Data collection and decision rules are implemented with integrity
- Clear direction exists for instructional modifications and alternatives

aimsweb supports accurate data collection, enables users to view and understand data trends, and provides explicit rules for deciding when a student is on track to meet (or exceed) his or her goal. Note that the final feature of effective progress monitoring (i.e., providing direction on how to modify or replace instructional interventions) is beyond the scope of this guide. Progress monitoring can help to determine if an instructional program is effective, but it cannot identify potential alternatives.

Overall, progress monitoring is an important part of a dynamic approach to guiding instruction and interventions for both groups and individuals. Rather than “plan and hope,” progress monitoring provides an opportunity to “plan and evaluate” whether curriculum and instructional procedures are increasing the rate of student progress towards explicit goals.

The **aimsweb** approach to progress monitoring relies on curriculum-based measurement (CBM), a measurement method that is standardized, valid, reliable, repeatable, simple, efficient, and inexpensive (Deno, 1986). These characteristics make CBM ideally suited to the frequent assessment that is required for progress monitoring (Deno, 1985, 2003), including the tri-annual assessments used to screen for at-risk students, and to the evaluation of core instruction. Together, tri-annual screening and progress monitoring link expectations (benchmarks) and local performance levels (local norms) to individual student goals.

Components of Progress Monitoring

Progress monitoring has several basic elements, including

- deciding whom to progress monitor,
- setting a goal,
- collecting data (administering probes),
- evaluating progress,
- deciding whether progress is adequate (i.e., whether the instructional program is working) and whether the instruction and/or the goal should be modified, and
- deciding whether the goal has been reached.

The **aimsweb** system includes tools to support these activities. To attain the full benefit of these progress-monitoring tools, users need to have a basic understanding of two types of scores used for evaluating a student's current performance level and setting a goal:

- Performance-level percentiles, based on national and/or local reference groups
- Performance benchmarks, which are tied to the probability of success on a criterion (e.g., scoring in the proficient range on a state test)

These scores and the elements of progress monitoring are described in the following sections. Subsequently, the remainder of this guide uses example cases to discuss and demonstrate the application of **aimsweb** progress monitoring tools.

Performance-Level Percentile Norms

The interpretation of scores requires some context, such as the *typical* performance level. For instance, a raw score of 25 could be either very high or very low, depending on the context. To characterize a score as *high* or *low* is to convey information about its standing relative to how other students in the same grade performed.

A percentile is a useful type of score for understanding whether a student is performing higher, lower, or at the same level as his peers. Percentile ranks, or *percentiles*, range from 1 to 99 and indicate the percentage of students at a particular grade level from a local or national sample who scored at or below a given score. The *percentage* correct on a measure is entirely distinct from a *percentile* (e.g., correctly answering 80% of the questions on a measure).

Norm Groups

Percentiles are derived from the distribution of scores in a norm group or comparison population. **aimsweb** provides percentile norms based on a large, representative national sample. In addition, it can create local percentile norms based on norm groups specified by the user, which may consist of all the students at a particular grade in the user's state, district, or school. The user may further specify a local norm group by applying selection criteria, such as demographics, disability category, or instructional program.

When selecting a norm group to aid in evaluating performance and setting goals, it is necessary to determine which norm group is the most relevant. National and state norms are often the most stable and nuanced because they have the largest populations. However, the most relevant norm group for goal setting is usually the population of students in a given district or school, because norms based on these groups reflect the local performance distribution.

The availability of local norms provides a substantial advantage for progress monitoring. When using a norm-referenced approach to goal setting, the goal is often set near the middle of the average range (i.e., the 40th to 50th percentile). Students who perform in the average range relative to their same-grade peers from their own district are likely to benefit from the core instruction provided in that district. Students who are well below typical achievement are less likely to benefit from core instruction. For this reason, bringing a student up to the local 40th percentile puts the student in a good position to benefit from the district's or school's general instructional program and sets the bar high enough to reduce the student's chances of regressing to at-risk status.

Performance Benchmarks

A performance benchmark, expressed in raw score units on CBM measures, is the score that all students should achieve if they are to have a good probability of succeeding in the educational program or on a particular outcome (e.g., a state assessment). Benchmark scores may be derived from many sources, including the research literature on CBM.

aimsweb offers benchmark scores for many of its measures based on the prediction of success on state standards-based assessments in reading and mathematics. Many states define four performance levels on their state assessments: below basic, basic, proficient, and advanced. Performance at or above proficient is the goal for all students. With this in mind, **aimsweb** researchers analyzed the relationship between scores on R–CBM and M–CAP and performance on state tests using data from 20 states. The research showed that performance at about the 45th percentile on national norms indicates that the student is 80% likely to meet proficiency standards on a typical state test. Performance at the 15th percentile on national norms indicates a 50% likelihood of meeting the proficiency standard. These findings were highly consistent across measures (R–CBM and M–CAP), grades, and screening periods. Using these relationships, similar percentile standards were applied to the other **aimsweb** measures across grades. The *Test of Early Literacy* (TEL) and *Test of Early Numeracy* (TEN) were handled in a similar way, but using a different criterion. Specifically, research by Silbergliitt (n.d.) showed that TEL scores at the 35th and 15th national percentiles in Kindergarten predicted success on a reading test (R–CBM) in the spring of the following year with 80% and 50% probability; as such, the benchmark scores for TEL and TEN were set at these percentile values. (More details are included in *AIMSweb State Prediction User's Guide*, located on the **aimsweb** site.)

These benchmarks—referred to as *default cut scores*—can be displayed in the **aimsweb** reports. For instance, in a report for the fall of Grade 6, students who obtained an R–CBM words read correctly (WRC) score at or above 136 have at least an 80% likelihood of achieving proficiency on the state test in the spring.

Benchmarks are applicable to all students in a given grade. However, it is often appropriate to use a benchmark as the goal for a student being progress-monitored. Although a goal is individualized for each student, the spring benchmark for 80% likelihood of success is a meaningful objective for a student receiving instructional intervention. Students who reach this goal are more likely to achieve proficiency on state standards-based tests. Conversely, students whose performance does not increase to this level are less likely to achieve proficiency.

Deciding Whose Progress to Monitor

The purpose of progress monitoring is to guide instruction for individual students. Teachers typically monitor the progress of students who receive enhanced instructional services, but it is also appropriate for students in the general education curriculum to ensure that they are keeping up (e.g., students who have recently completed an intervention) and are not at-risk.

Such a decision begins with analyzing classroom performance, using scores from **aimsweb** screening measures. The **aimsweb** system arranges classroom scores in numerical or reverse numerical order. Student performance is described by five norm-referenced performance levels: *well-below average* (< 10th percentile), *below average* (10th to 24th percentile), *average* (25th to 75th percentile), *above average* (76th to 90th percentile), and *well-above average* (> 90th percentile). **aimsweb** criterion-reference tiers indicate where each score lies in relation to criterion-referenced benchmarks.

For the performance levels, either local or national norms can be used. Local norms are recommended for prioritizing interventions because they indicate who is and is not benefitting from core instruction and signal who might benefit from differentiated instruction. Core instruction is generally adequate for the majority of students and is typically targeted to students in the range of average to above average. Students in the below average and well-below average range need the most additional instruction, relative to other students in the classroom. Finally, those in the range of well-above average may benefit from enrichment activities.

A limitation of the norm-referenced categories is that they do not show the prevalence or severity of performance deficits. When local school norms are used, the percentage of students in each normative level will approximate the percentile ranges. Specifically, about 50% of the students' scores will be in the average range (25th to 75th percentile), about 25% will be in the below-average range, and about 25% will be in the above-average range, regardless of the school's performance level. Therefore, it is important to use a different external and static criterion to evaluate performance deficits.

One option is to use the **aimsweb** default criterion-referenced cut scores that designate the likelihood (80% or 50%) of meeting typical state proficiency standards. Usually, two cut scores are defined, the upper and the lower. The upper cut score, which separates Tiers 1 and 2, is the target performance level. Students who score above the target have a high probability of meeting proficiency standards on the state test. The lower cut score, separating Tiers 2 and 3, is used to identify students who have a moderate to high risk of not meeting grade level proficiency standards. Additional instructional support or intensive intervention is recommended for students in Tiers 2 and 3.

The prevalence and severity of performance deficits in the class, grade, school, or district are reflected by the number of students in each of tier. The tiers have a limitation in that they do not differentiate scores that are near the cut score from those that are far from the cut score. Because all scores contain error, scores near a cut score are within the expected measurement error, which reduces the certainty that the student truly belongs in a particular tier. When prioritizing resource allocation, consider giving highest priority to students who perform well below the target score.

Generally, all students who need intervention should be monitored for progress. In most classrooms, a relatively small percentage of students are in Tiers 2 and 3, making individual or small-group intervention with frequent progress monitoring warranted. When a large percentage of students are in Tiers 2 and 3, intensive intervention and frequent progress monitoring are often not practical. It is important to ask whether the core instruction is doing a good job of preparing most students to reach the target. If not, then it may be best to modify core instruction and monitor the progress of all students on a monthly or tri-annual basis. Deciding upon which course to take depends on the extent and severity of the need, the availability of resources for intervention, and prior achievement trends. When planning individual intervention, it is also important to consider past achievement, disability status, language background, and the level and intensity of prior instructional interventions.

An application of the norm-referenced performance levels (using local norms) and tiers is illustrated in the following two examples. First, Figure 1 shows fall R–CBM screening scores for a second grade classroom. In this example, the majority of students, including most of those in the average range, are above the target score. This is a classroom in which core instruction is effectively moving students to the target. Intensive intervention should be considered for all three students in Tier 3. Among those in Tier 2, three were in the *average* range for the school and may not need additional support. The remaining two students scored far enough below the cut score to indicate that additional instructional support may be warranted.

Figure 2 shows results for a low-performing classroom in a low-performing school. The norm-referenced performance levels and tiers can lead to very different conclusions about who needs intervention. The percentage of students in each norm-referenced performance level is the same as in Figure 1; however, all but one student scored below the target score and the majority of students are in Tier 3. With so many students in need of extra instructional support to improve their chances of success on the state test, core instruction should probably be modified and monthly progress monitoring should be conducted with the entire classroom.

Student	WRC	Norm-Referenced Level	Tier
1	120	Well-Above Average	1
2	110	Well-Above Average	1
3	95	Above Average	1
4	80	Above Average	1
5	75	Above Average	1
6	70	Average	1
7	65	Average	1
8	62	Average	1
9	60	Average	1
10	57	Average	1
11	55	Average	1
12	55	Average	1
----- Target (Upper Cut Score) -----			
13	50	Average	2
14	45	Average	2
15	40	Average	2
16	35	Below Average	2
17	25	Below Average	2
----- Lower Cut Score -----			
18	20	Below Average	3
19	15	Well-Below Average	3
20	10	Well-Below Average	3

Figure 1. Distribution of R-CBM scores, local norm-referenced performance levels, and Tiers for a classroom, Example 1

Student	WRC	Norm-Referenced Level	Tier
1	55	Well-Above Average	1
----- Target (Upper Cut Score) -----			
2	50	Well-Above Average	2
3	45	Above Average	2
4	40	Above Average	2
5	35	Above Average	2
6	30	Average	2
7	25	Average	2
----- Lower Cut Score -----			
8	20	Average	3
9	19	Average	3
10	18	Average	3
11	17	Average	3
12	16	Average	3
13	15	Average	3
14	12	Average	3
15	11	Average	3
16	10	Below Average	3
17	9	Below Average	3
18	8	Below Average	3
19	5	Well-Below Average	3
20	5	Well-Below Average	3

Figure 2. Distribution of R-CBM scores, local norm-referenced performance levels, and Tiers for a classroom, Example 2

When deciding whom to progress monitor, follow these steps.

1. Determine the percentage of the classroom in Tiers 2 and 3.
 - o If the percentage is small, consider individual interventions and progress monitoring.
 - o If the percentage is moderate to large, consider modifying core instruction and conducting monthly progress monitoring of all students.
2. If individual progress monitoring is warranted, prioritize intervention according to need and available resources. Give highest priority to students in Tier 3, followed by those in Tier 2 who score well below the target score.

Goal Setting

Effective progress monitoring—timely and accurate information about student (or group) progress—requires challenging, yet reasonable, performance goals. An important contribution of CBM progress monitoring is the articulation of performance goals that are *measurable*, *meaningful*, and *manageable*.

Within CBM progress monitoring, goals are measurable because they are expressed in terms of performance (raw score) on one of the measures. The raw score is simply the number of questions answered correctly or the number of score points earned (e.g., R–CBM words read correctly per minute or M–CAP total point score).

By themselves, raw scores have limited meaning. Making goals meaningful requires additional information. Academic problems are often identified or quantified by the size of the discrepancy between the student's performance and the performance level necessary to benefit from core instruction and succeed on important educational outcomes. This desired level of performance typically is based either on *normative* information (i.e., the average level of performance of the student's peers) or on *criterion-referenced* information (i.e., a score that predicts success). **aimsweb** provides tools that help to characterize the severity of discrepancies and establish meaningful goals. For example, Jayden read 10 words correctly per minute (WRC) in the fall of second grade. The typical performance of his peers (i.e., the middle 50% of students) was between 35 and 88 WRC, which serves as the norm. Students in the fall of Grade 2 should read at least 60 WRC to be on track for reading success; as such, this raw score serves as a fall benchmark. Jayden's performance is below that of his peers and below the benchmark, making the next step to define a performance goal for Jayden that will close this gap.

Selecting the Grade Level of the Material for the Goal and Progress Monitoring

The desired goal for most at-risk students is to become proficient on-grade-level material. With appropriate supplemental or intensive intervention, students are expected to make substantial growth; however, there are limits to the magnitude of growth that can be achieved. When it is reasonable for a student to achieve grade-level proficiency by the end of the school year with appropriate intensive intervention, the goal should be defined on grade level (i.e., the enrolled grade level). When achieving grade-level proficiency could take multiple years, it may be necessary to define an intermediate performance goal. The **aimsweb** rate of improvement (ROI) growth norms (described below) can be used for this purpose. If the goal requires a growth rate that far exceeds expected growth rates, consider defining an intermediate performance goal at a lower grade level. In either case, progress monitoring probes should *always* be selected from the grade level in which the student's performance goal is defined.

The process of determining the grade level for a performance goal should also include an analysis of the student's instructional level. As noted by Shapiro (2008, p. 148), "A student who is functioning below enrolled grade level will demonstrate little progress over time if monitored at levels that exceed his or her instructional level" and should consequently "be measured at his or her highest instructional level". An off-grade instructional level is indicated for a student who

has not mastered important prerequisite skills or scores at the lower extreme on the **aimsweb** screening measure. Performance at or below the 10th percentile is recommended for this purpose.

The following process details the logical method of stepping back grade by grade until the student scores at or above a minimal level of performance (i.e., the 10th percentile for that grade).

1. Administer the screening probe(s) at the student's current grade level.
2. Determine the local (or national) percentile rank of the student's score on the screening probe. For individual student progress monitoring, the local norm (school, district, or region) is recommended. (Note: For R-CBM, use the median of the student's three screening probe scores.) If the student's score ranks above the 10th percentile, proceed with progress monitoring at grade level. If the student's score ranks at or below the 10th percentile, proceed to step 3.
3. Step back one grade level and administer that grade's screening probe(s). If the student's score ranks above the 10th percentile, proceed with progress monitoring at this grade level. If the student's score ranks at or below the 10th percentile, repeat step 3, stepping back one grade level at a time until the student achieves a score that ranks above the 10th percentile, and proceed with progress monitoring at that grade level.

Evaluating Proposed Goals Using ROI Growth Norms

Performance level percentiles and performance benchmarks are beneficial for defining meaningful performance goals for individual students. However, before deciding on a goal, it is desirable to evaluate the *reasonableness* of a goal for a given student. One way to do this is to compare the ROI needed to reach the goal, with the ROIs typically observed in the population of students who are starting out at about the same level of performance. If the goal ROI for the student is so high as to be rarely observed amongst his or her peers, recognize that this goal will probably be very challenging. Similarly, a goal ROI that is commonly achieved may be insufficient for a student who will be receiving instructional intervention.

In **aimsweb**, a student's rate of improvement (ROI) is the average increase in his or her raw score per week. The **aimsweb** system calculates a student's goal ROI (the rate of improvement needed to reach the goal) as follows:

$$ROI = \frac{(\text{Goal Score} - \text{Initial Score})}{\text{Weeks Elapsed}}$$

The *initial score* is the score at the start of the progress-monitoring period; often, this is the score from the fall screening. The *goal score* is the expected score at the end of the instructional period, which is typically one academic year (or 36 weeks). The *weeks elapsed* is the number of weeks between the dates of measurement.

To illustrate, assume a Grade 2 student obtained a total M-CAP raw score of 3 during fall screening, corresponding to a national percentile of 20. The student's teacher defines a spring screening goal score of 21, which corresponds to the 50th percentile on the spring norms. Therefore, the goal ROI is:

$$ROI = \frac{(21 - 3)}{36} \rightarrow \frac{18}{36} \rightarrow 0.5$$

This student would need to gain 0.5 points per week on the M–CAP measure to reach the goal score of 21 by spring screening.

aimsweb ROI growth norms convey information about the reasonableness of a goal by comparing the growth rate needed to achieve the goal to the growth rates of students in the national norm sample, thereby serving as an indicator of whether growth is about average, above average, or below average. In the above example, based on the national ROI growth norms, an ROI of 0.5 points per week for a Grade 2 student is above average and corresponds to the 75th percentile, as compared with other students whose fall screening scores were at a similar level. Such a growth rate is labeled *closes the gap*. Indeed, if the student is able to increase performance from the 20th percentile to the 50th percentile, the performance gap will be closed.

ROI growth norms compare each student's growth rate to students with a similar level of initial performance, performance that is typically based on the fall screening score. Initial performance is taken into account because both **aimsweb** researchers and others (Silbergitt & Hintze, 2007) have observed that the average rate of growth is often related to the initial level of performance. In general, students with very low initial performance levels tend to have lower ROIs, unless they are receiving supplemental instruction. Students with very high initial scores also tend to have relatively low ROIs, partly because of regression effects and partly because their ability to demonstrate their skill may be limited by the score range of the measure. The benefit of having ROI norms that are separated by initial performance level is that they adjust for any ROI differences caused by the initial performance level. (For details regarding ROI norms development, see the **aimsweb ROI Growth Norms Overview** available at www.aimsweb.com.)

After a performance goal has been defined for a student and entered into the **aimsweb** system, the ROI needed to achieve said goal is calculated. Using the ROI growth norms as the basis, the **aimsweb** system rates and reports the ambitiousness of the goal as follows:

Insufficient: The ROI is below average and the goal score will not improve the student's percentile rank.

Closes the gap: The ROI is above average and the goal score will improve the student's percentile rank.

Ambitious: The ROI is well-above average and the goal score will substantially improve the student's percentile rank.

$$ROI = \frac{(\text{Goal Score} - \text{Initial Score})}{\text{Weeks Elapsed}} \rightarrow ROI = \frac{(21 - 3)}{36} = \frac{18}{36} = 0.5 \rightarrow \text{Closes the gap}$$

Case Studies

Case 1: Using Performance-Level Percentiles and ROI Growth Percentiles

Case 1 illustrates how local and national norms are used for goal setting.

1. Define a *meaningful* goal.

During fall screening using the R–CBM measure, third grader Trevor read 55 words in 1 minute with two errors. His score places him at the 15th percentile relative to local

school norms and at the 25th percentile relative to national norms. Is there an academic problem? Yes, because Trevor performed in the below-average range relative to local school norms, which indicates his potential for an academic problem and suggests that he may not be fully benefitting from general instruction. This is typically geared to students in the average and above-average range at his school. Anything below the 45th percentile on national norms (benchmark standards) indicates a less than 80% chance of meeting proficiency on state assessments. An academic problem is evident because Trevor's performance is well below expectations.

As a result, Trevor needs supplemental or intensive instruction with progress monitoring to reduce or eliminate the performance discrepancy. A meaningful goal will help Trevor close the gap between observed and expected performance. Specifically, his goal should be to approximate typical performance on local norms and exceed benchmark standards using national percentiles.

2. Set the *progress monitoring* goal.

Students who score above the 10th percentile relative to national norms and in the below-average range relative to local norms should typically be progress-monitored on grade level. As such, Trevor's progress monitoring goal was set on grade level and at a performance level that will move him to the 40th percentile (average range) relative to Grade 3 spring local norms. His stated performance goal is:

In 35 weeks, Trevor will read 140 words correctly in 1 minute on the Grade 3 R–CBM measure.

$$ROI = \frac{(\text{Goal Score} - \text{Initial Score})}{\text{Weeks Elapsed}} = \frac{(125 - 55)}{35} = 2.0 \text{ WRCM per week}$$

An ROI of 2.0 words read correctly per minute (WRCM) per week will ensure that Trevor attains this goal.

3. Evaluate the *reasonableness* of the goal.

The **aimsweb** system rates and reports the reasonableness of each selected goal based on its ROI, classifying a goal as insufficient, closes the gap, or ambitious. In this case, the planned rate of growth for Trevor is ambitious because it corresponds with the 95th percentile of ROIs for Grade 3 students with similar levels of fall performance on R–CBM. In other words, only about 5% of students in that peer group showed such a high rate of improvement over the year. This information should be considered during the goal setting process. Is it realistic to achieve such an ambitious goal? The answer is based on an evaluation of historical and future conditions for instruction. If Trevor is provided with highly effective supplemental or intensive instruction, then an ambitious ROI might be considered reasonable. If little or no additional instruction is provided—or if evidence-based practices are not implemented—then the ROI might be unreasonable.

Case 2: Setting an Off-Grade Performance Goal

Case 2 illustrates how local and national norms are used for goal setting when there is a large discrepancy between instructional targets and grade level progress monitoring materials.

1. Define a meaningful goal.

In the winter of sixth grade, Amanda obtained an M–COMP total score of 2. This score corresponds to the 1st percentile relative to local and national norms. **aimsweb** guidelines stipulate that off-grade progress monitoring materials should be considered for students scoring at or below the 10th percentile relative to national norms. However, it is important to consider whether a discrepancy exists between the student's skill level and the skills assessed in the on-grade progress monitoring material. An item analysis of Amanda's M–COMP screening indicated that she is unable to correctly complete any of the multiplication or division problems. Because multiplication and division are necessary prerequisites for sixth-grade mathematics, Amanda's teacher set Amanda's goal at her instructional level, rather than her grade level.

2. Set the progress monitoring goal.

Following the recommended guidelines, Amanda was administered the Grade 5 M–COMP screening probe. She obtained a score of 3, which is at the 4th percentile relative to local norms and the 2nd percentile relative to national norms. Next, Amanda was administered the Grade 4 M–COMP winter screening probe. She obtained a total score of 20, which corresponds to the 10th percentile relative to local norms and the 8th percentile relative to national norms. Upon inspection of Amanda's item-level performance, it was noted that she was able to correctly answer most of the math facts, simple addition that did not involve carrying, and subtraction that did not require borrowing. Amanda did not answer any of the division or fraction problems. Because division is a focus of the Grade 4 curriculum in her school district, her teacher decided to set her performance goal at a fourth-grade level. Specifically, Amanda would receive intensive intervention on complex addition and subtraction and additional instruction to extend her understanding of math facts to more complex multiplication problems. Finally, she would also receive instruction on simple division. Her instructional goal was designed to ensure that by the end of the school year Amanda masters adding and subtracting with borrowing and carrying, respectively, up to 1000; multiplying two-digit numbers; and dividing three-digit number by one-digit numbers. Her stated performance goal is:

In 16 weeks, Amanda will obtain a total score of 40 on the Grade 4 M–COMP measure. Her goal score of 40 corresponds with performance at the 25th percentile in the spring of fourth grade relative to local norms and requires an ROI of 1.25 points per week.

3. Evaluate the reasonableness of the goal.

Although the content analysis described in step 2 is a very good way to determine an appropriate and reasonable goal, Amanda's teacher wanted additional supporting evidence to ensure that she set the standard high enough. To this end, she used the ROI growth norms to determine that an ROI of 1.25 *closes the gap* and corresponds to the 75th percentile compared to students with a similar score on the Grade 4 winter M–COMP probe.

Case 3: Using Benchmarks for Goal Setting

Case 3 illustrates how local norms typically are used for goal setting; however, it is also feasible to set goals with benchmarks for both individual students and groups of students. **aimsweb** *default cut scores* are useful for setting individual performance goals and identifying students with relatively low probabilities of passing state tests.

1. Define a meaningful goal.

First, identify a cut score for the target. By default, this is the score that corresponds to the success probability level of 80%, which divides Tiers 1 and 2. A second cut score that divides Tiers 2 and 3 is also defined and, by default, corresponds to the success probability level of 50%. If you are trying to identify the students most at risk, the 50% cut score is recommended.

As illustrated in Figure 3, just more than half of the Grade 6 class scored below the lower cut score, which means that the majority of students in the class are at a moderate-to-high risk of not meeting state proficiency standards. Following the guidelines previously described, the teacher should consider modifying core instruction and monitor the class's reading progress monthly to ensure that the modifications are having the intended affect.

Student	WRC	Norm-Referenced Level	Tier
1	165	Well-Above Average	1
----- Target (Upper Cut Score) -----			
2	130	Well-Above Average	2
3	129	Above Average	2
4	124	Above Average	2
5	124	Above Average	2
6	120	Average	2
7	118	Average	2
8	115	Average	3
9	110	Average	3
----- Lower Cut Score -----			
10	100	Average	3
11	98	Average	3
12	95	Average	3
13	93	Average	3
14	89	Average	3
15	80	Average	3
16	75	Below Average	3
17	72	Below Average	3
18	68	Below Average	3
19	55	Well-Below Average	3
20	45	Well-Below Average	3

Figure 3. Distribution of R-CBM words read correctly for a low-achieving Grade 6 classroom

2. Set the progress monitoring goal.

As shown in Figure 3, two of the Tier 3 students scored well below the rest of the class. These students will likely need intensive intervention to move into the current average range for the classroom. Because the overall achievement in this classroom is low, the class average range may not be a sufficient goal to close the gap between these students' performances and the target. The goal for the two students was set at the lower cut score. It was assumed that most of the students in the below-average range would achieve significant gains from the modifications to core instruction.

The goal for Juan, who read 45 words correctly per minute, and Alejandro, who read 55 words correctly per minute, was to reach Tier 2 by the spring. Both students' primary language is Spanish and both receive services for limited English proficiency. The plan is to increase the amount of intensive reading instruction from 60 to 120 minutes per day. The students' stated performance goal is:

In 30 weeks, Juan and Alejandro will each read 120 words correctly per minute on the Grade 6 R-CBM.

To achieve this goal, Juan needs an average ROI of 2.5 and Alejandro needs an average ROI of 2.2.

3. Evaluate the reasonableness of the goal.

Both ROIs exceed the 95th percentile of growth rates from the national sample of sixth-grade students who had similar fall R-CBM scores. Because both students have limited English proficiency and both made strong reading gains in the past year, these ambitious ROIs were deemed reasonable.

Interpreting Progress Monitoring Data

Historically, multiple methods have been used to interpret the trend of a student's progress-monitoring scores and to support decisions regarding whether a student is advancing at the desired rate and whether changes to the intervention and/or the goal need to be made. These methods of time-series analysis include visually analyzing the pattern of data points, counting consecutive score points relative to the aimline, and comparing the trendline to the aimline. The 2013 **aimsweb** software introduces an additional method: using information about the student's data points (number, variability, and time span) to construct a confidence interval for the expected score at the goal date. (Note: The procedures for this method are presented in the Appendix and users may implement it using a spreadsheet or simple computer program.) Each method has advantages and disadvantages, which are discussed the following sections.

Visual Analysis

Although subjective, a visual analysis enables the user to identify any outliers (individual scores that are highly discrepant from the general trend) and the trendline (curvilinear or a straight line). Outliers can have a significant effect on the trendline, especially when they occur near the beginning or the end of the progress-monitoring period. If the user determines that an outlying

score is probably not a valid indicator of the student's true ability at that point in time—perhaps the student was ill or distracted, or there was a procedural flaw in administration or scoring—then making a mental adjustment of the trendline is appropriate. Figure 4 illustrates the effect of an outlier. Panel A shows the trendline with the outlier included, whereas Panel B shows the trendline when the outlier is removed. It is evident that the low-score outlier at week 6 lowers the trendline, indicating that the student will not meet his or her goal at the present rate of growth. However, when the outlier is removed the trendline steepens, leading to the conclusion that the student is likely to meet the goal at the present rate of growth.

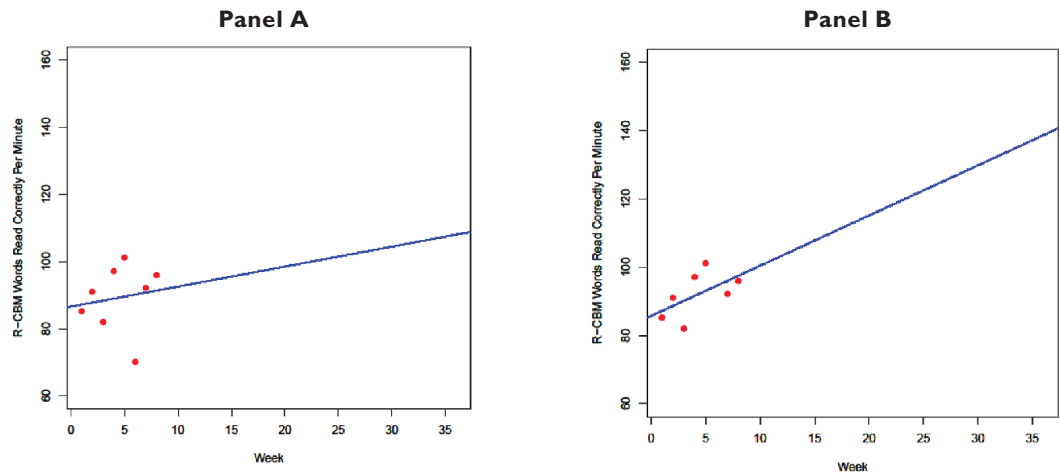


Figure 4. The impact of an outlier on the estimated trendline

Similarly, a nonlinear pattern of score growth can make the trendline less accurate as a description of the student's progress and a predictor of future performance. This is illustrated in Figure 5, which shows two types of curved patterns of score growth, with the linear trendline superimposed on each. Although the curved shape of the growth trend often will not be as clear cut as in these examples, the user should take apparent nonlinearity into account when interpreting the results of other methods, such as rules of thumb, trendlines, or confidence intervals.

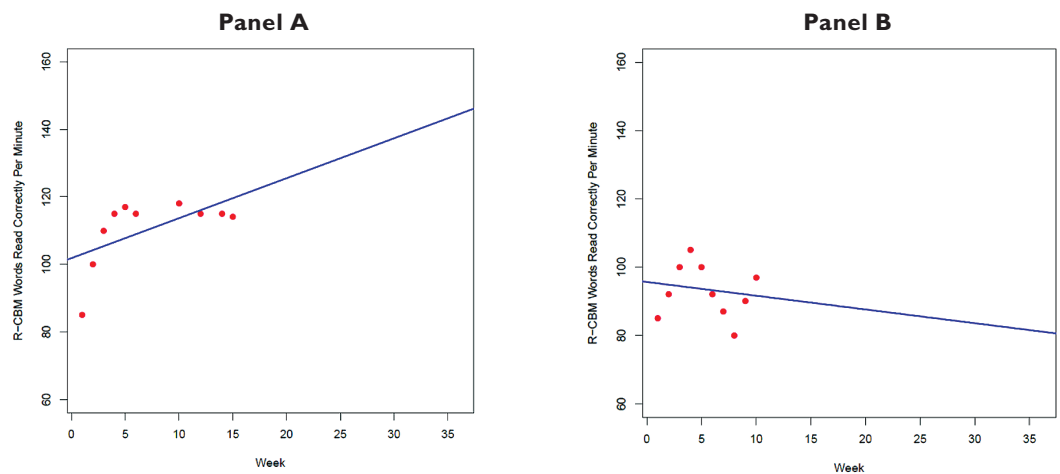


Figure 5. Nonlinear patterns

Visual analysis is also useful for identifying the amount of random variability of scores around the trend. Be aware that some students will show a more consistent set of scores than others, meaning that their scores will stay closer to the growth trend. Conversely, other students will perform erratically, which could be the result of gaps in skill development, variable levels of motivation, or inconsistent measurement conditions. When a large amount of “scatter” is apparent in the scores, discerning the trend can be difficult. The more variable the scores are, the less accurate the prediction of future performance can be. Figure 6 illustrates two sets of scores, one with little scatter and the other with a large amount of scatter. Panel A shows a linear increasing pattern whereas the pattern in Panel B is more difficult to interpret.

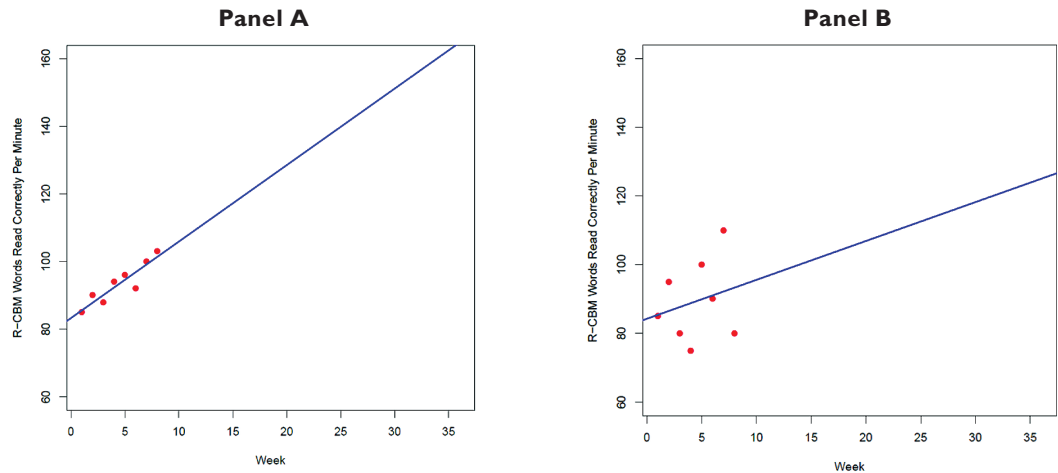


Figure 6. Low and high scatter patterns

Visual analysis should be a part of every review of progress-monitoring data. **aimsweb** features a graphical display of progress-monitoring results in the individual student reports. However, it is important to note that visual analysis does not give a numerical answer to the question of whether a student is likely to reach his or her goal.

Comparing Consecutive Score Points Relative to the Aimline

The progress-monitoring graph's aimline depicts the rate of improvement (ROI) required to reach the goal and indicates the level of expected performance each week. As such, the week-by-week relationship between actual scores and the aimline indicates whether or not the student is making adequate progress. If all consecutive score points are below the aimline, insignificant progress has occurred. On the other hand, if the score points are consistently above the aimline, then the goal should possibly be set at a higher level. If consecutive score points vary between above and below the aimline, it could be inferred that the student is on track to reach the goal.

Comparing weekly score points to the aimline is a relatively weak basis for decision-making for two reasons. First, it ignores the distance between the score points and the aimline. For example, a string of below-aimline scores may not be worrisome if they tend to be close to the aimline itself. Second, as previously mentioned, high variability in scores makes the trend estimate less accurate; as a result, the scatter in a pattern that bounces between above- and below-aimline scores can weaken judgments regarding whether the student is progressing at the desired rate. In

general, a decision rule based on consecutive scores relative to the aimline is an approximation of more sophisticated statistical methods, such as those described in the following sections.

Comparing Trendline and Aimline

When three data points have been entered for a student, **aimsweb** calculates the trendline statistically (using ordinary least-squares regression) as the line that comes closest to all the student's score points over the duration of progress monitoring (see Figure 7). If the trend of the data points is not obviously curved, then the trendline is the best available basis for a projection of the student's future performance. (Note: When a clear curvature in the trend is apparent, it should be taken into account when estimating future results.) By comparing the trendline to the aimline on the same graph, one can see whether the student is on track to reach the goal.

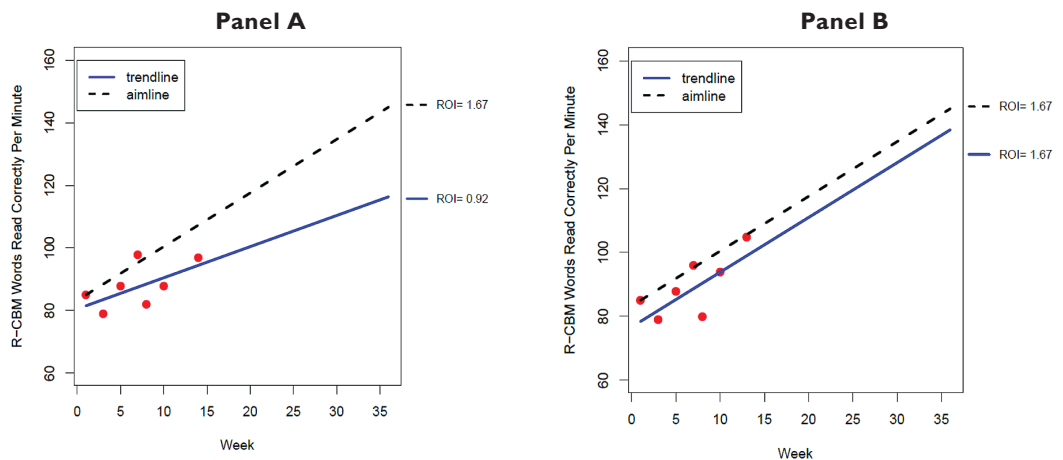


Figure 7. R-CBM score patterns with trendline, aimline, and ROIs

aimsweb progress-monitoring reports display both the aimline ROI and the actual (trendline) ROI, making it easy to see whether a student is progressing at or near the expected rate. In Figure 7, the slope of the trendline in Panel A is obviously lower than that of the aimline (0.92 vs. 1.67, respectively) and indicates whether the trendline is above or below the aimline. Panel B of Figure 7 illustrates a scenario in which the student's rate of improvement is equal to the expected ROI; however, because the score points tend to be below the aimline, the student is at risk of not reaching the goal.

In the current **aimsweb** version (prior to 2013), feedback about student progress is based on a comparison of the student's observed ROI with the ROI needed to reach the goal (goal ROI). Assuming there are at least four progress monitoring scores, the system gives the following feedback:

- **Below Target:** The observed ROI is more than 0.5 units below the goal ROI.
- **Near Target:** The observed ROI is within 0.5 units of the goal ROI.
- **Above Target:** The observed ROI is more than 0.5 units above the goal ROI.
- **Goal Missed:** The student's score obtained on or after the goal date is less than the goal score.
- **Goal Achieved:** The student's score obtained on or after the goal date is at or above the goal score.

If fewer than four progress-monitor scores are available for a student, the system indicates *insufficient scores*. Note that the units are defined as average score increase per week. For instance, the units for R–CBM represent the average increase in words read correctly per week.

Confidence Intervals for Scores at the Goal Date

aimsweb progress-monitor reports show the trendline projected out to the right as far as the goal date. The height of the trendline at the goal date is the predicted score for the student, but it is only an estimate because both the slope and the level of the trendline are affected by various sources of error. Constructing a confidence interval around the predicted score provides a basis for making decisions regarding the likelihood of a student reaching his or her goal. A confidence interval is a range of scores that has a certain probability of including the true score; in this case, the student's actual performance level at the goal date. Centered on the predicted score, the confidence interval's size (i.e., how far it extends above and below the predicted score) is a function of the amount of error in the trendline. As previously mentioned, error is smaller when the student's scores are tightly clustered around the trendline (i.e., there is little random variability), when the duration of progress monitoring is long, and when there are many data points.

The 2013 version of **aimsweb** uses confidence intervals as the basis for providing feedback about a student's likely outcome. Starting at 6 weeks of progress monitoring with at least four data points, **aimsweb** calculates the 75% confidence interval for the score at the goal date. (Note: The Appendix describes the calculation method and formulas.) This level of probability was selected because it is appropriate for the types of decisions being made. It is important to be confident in a decision to change an intervention or raise a goal, but there is a cost to waiting until the confidence level is too high, particularly when that would mean continuing with an ineffective intervention longer than necessary.

If the 75% confidence interval is completely below the student's goal (as depicted in Figure 8), then the **aimsweb** report includes a statement indicating that the student is projected to *not* reach the goal. This is a signal to consider a change in intervention. (Note: The report does not actually display the confidence interval.) Conversely, if the 75% confidence interval is completely above the goal, then the report states that the student is projected to exceed the goal. In such cases, one might consider raising the student's goal.

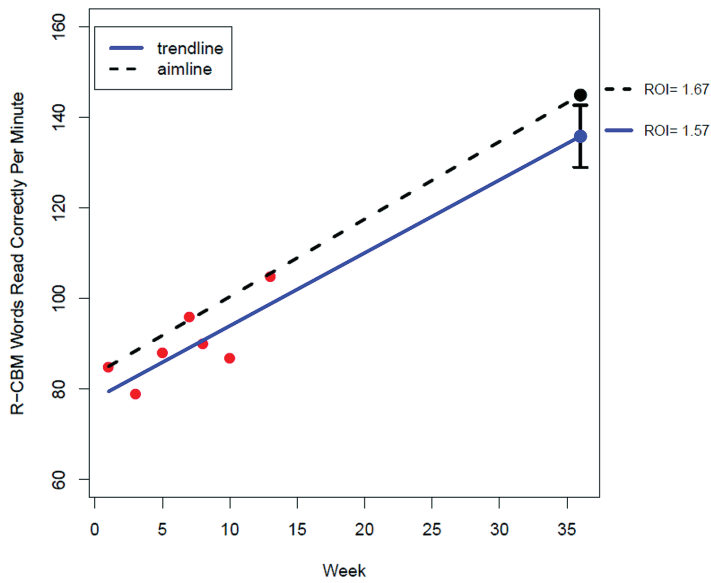


Figure 8. Predicted score 75% CI is completely below the target score

If the 75% confidence interval contains the goal score (as shown in Figure 9), then the progress monitoring report will state that the student's scores are trending toward the goal and that he or she is projected to score between x and y at the goal date (with x and y being the bottom and top of the confidence interval, respectively). This conveys important information to the user. For example, if the interval is mostly below the goal, then the user knows that the trend of the student's scores indicates that he or she may not reach the goal (although one cannot state with confidence that this outcome will occur). Of course, comparing the trendline with the aimline provides the same information.

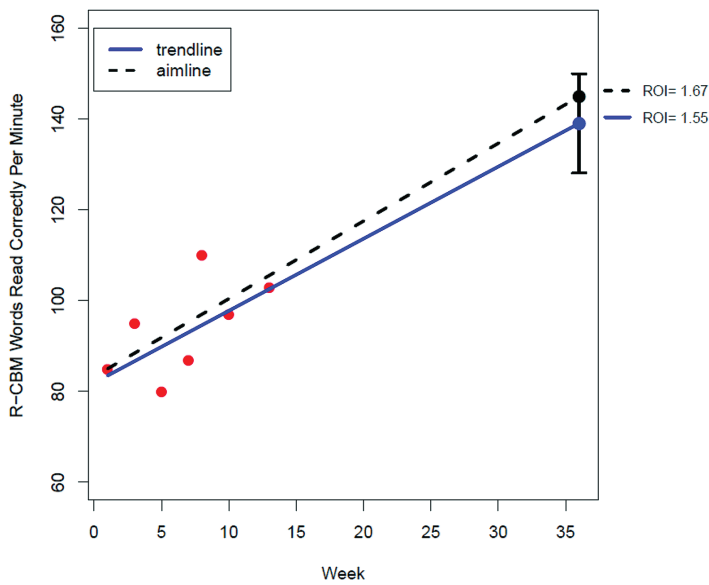


Figure 9. Predicted score 75% CI includes the target score

As progress monitoring continues and additional data points are collected, the confidence interval typically gets smaller and the likelihood that the interval will lie entirely below or entirely above the goal increases, generating a clear-cut projection. The length of time required varies from student to student, depending on the number and variability of available scores, the linear or curved nature of the trend, and how much lower or higher the trendline is than the aimline. To clarify this issue, a large data simulation study was conducted to examine some factors that influence the duration and number of administrations needed to make highly accurate predictions about goal attainment, and to evaluate the accuracy of predicting the true slope from the observed slope. The principal finding suggested that an average of approximately 14 weekly administrations would be needed to obtain highly accurate predictions (Christ, Zopluoglu, Long, & Monaghan, 2012); however, this outcome represents an upper limit of the duration required and does not take into account newer feedback. In most situations, users will be able to obtain accurate predictions with fewer weekly administrations. For example, **aimsweb** research has shown that 10–12 weekly administrations are sufficient to produce highly accurate predictions of goal attainment. Investigation into this topic is ongoing, with additional feedback and data being collected to better understand the relationship between duration of progress monitoring and accuracy of predictions.

Conclusion

The essential tasks involved in progress monitoring include

- deciding whom to monitor,
- choosing the grade level of the goal and the grade level of monitoring,
- setting a goal,
- evaluating the trend of progress monitoring scores, and
- deciding whether the intervention and/or the goal should be changed.

Although technical tools are available to help with these tasks, each requires professional judgment. The ultimate goal is to ensure that students are receiving the most appropriate instruction.

References

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7–74.
- Christ, T. J., Zopluoglu, C., Long, J. D., & Monaghan, B. D. (2012). Curriculum-based measurement of oral reading: Quality of progress monitoring outcomes. *Exceptional Children*, 78, 356–373.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219–232.
- Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review*, 15, 358–374.
- Deno, S. L. (2003). Developments in curriculum-based measurement. *Journal of Special Education*, 37, 184–192.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, 53, 199–208.
- Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, 118 Stat. 2647 (2004).
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2001).
- Shapiro, E. S. (2008). Best practices in setting progress monitoring goals for academic skill improvement. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 141–158). Bethesda, MD: National Association of School Psychologists.
- Silbergitt, B. (n.d.). *Using AIMSweb early literacy measures to predict successful first-grade readers*. Unpublished manuscript.
- Silbergitt, B., & Hintze, J. M. (2007). How much growth can we expect? A conditional analysis of R–CBM growth rates by level of performance. *Exceptional Children*, 74, 71–84.

Appendix

Procedure for Calculating a Confidence Interval Around a Projected Goal Score

The 2013 version of the **aimsweb** online system applies a statistical procedure to the student's progress monitoring scores to provide empirically based guidance regarding whether the student is likely to meet, not meet, or exceed his or her goal. Users who wish to apply this method independent of the **aimsweb** system may do so using a spreadsheet or simple software program and the following steps.

1. Calculate the trend line. This is the ordinary least-squares regression line through the student's monitoring scores.
2. Calculate the projected score at the goal date. This is the value of the trend line at the goal date.
3. Calculate the standard error of estimate (*SEE*) of the projected score at the goal date, using the following formula:

$$[[(1 + 1/k + (GW - \text{mean}(w)))/(k - 2))((\sum(y - y')^2)/(\sum(w - \text{mean}(w))^2))]^{1/2}$$

where

k = number of completed monitoring administrations

w = week number of a completed administration

GW = week number of the goal date

y = monitoring score

y' = predicted monitoring score at that week (from the student's trendline)

The means and sums are calculated across all of the completed monitoring administrations up to that date.

4. Add and subtract 1.25 times the *SEE* to the projected score, rounding to the nearest whole numbers.

Note that the decision rules based on the confidence interval rely on certain assumptions:

- o The student's progress can be described by a linear trendline. If the pattern of the student's monitoring scores is obviously curvilinear, then the projected score based on a linear trend will likely be misleading.
- o The student will continue to progress at the same rate they have been progressing up until that time. This is an unavoidable assumption for a decision system based on extrapolating from past growth.