WestEd — a national nonpartisan, nonprofit research, development, and service agency — works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults. WestEd has 17 offices nationwide, from Washington and Boston to Arizona and California, with its headquarters in San Francisco. For more information about WestEd, visit WestEd.org; call 415.565.3000 or, toll-free, (877) 4-WestEd; or write: WestEd / 730 Harrison Street / San Francisco, CA 94107-1242.

© 2014 WestEd. All rights reserved.
# Table of Contents

Introduction: Purpose of This Guidance Document i

Section I. Using the Pre-Test/Post-Test Approach to Measure Student Growth with a Promising DDM 1

Section II. Example Assessment #1, Using the Pre-Test/Post-Test Approach to Measure Student Growth 3

Section III. Suggestions for Refining Example Assessment #1 to Ensure Usefulness as a DDM 15

Section IV. Example Assessment #2, Using the Pre-Test/Post-Test Approach to Measure Student Growth 17

Section V. Suggestions for Refining Example Assessment #2 to Ensure Usefulness as a DDM 21

Appendix A: Key ESE-Developed Resources to Support Districts with Implementing DDMs that Effectively Measure Student Growth 24

Appendix B: Recommended External Resources on Measuring Student Growth: Guidance from the Research and Measurement Communities 27
Introduction: Purpose of This Guidance Document

Educators and administrators engaging in the selection and/or development of district-determined measures (DDMs) have been challenged with answering a key question about their emerging DDMs: \textit{Does this measure detect some type of change in performance that can be attributed to student growth?} In a key resource developed to support districts with this challenge, ¹ the Massachusetts Department of Elementary and Secondary Education (ESE) explains why it is so important that a DDM measure growth:

Students come to school each year with a wide range of prior academic achievement and therefore begin their next year of instruction with varying levels of readiness to access the curriculum, a situation that is beyond the control of the educator assigned to teach them. Measuring educators’ effectiveness solely by the achievement level of their students cannot account for these prior conditions. By comparison, measuring growth can help level the playing field. Improvement in student performance is a more meaningful and fair basis for determining the trends and patterns that will yield the educator’s rating of impact on student learning, growth, and achievement. (p. 8)

One strategy for supporting DDM developers in Commonwealth districts with this challenge is to provide concrete examples of open-source, locally developed assessments that are useful for the purpose of measuring student growth. The aim of sharing a sample of promising DDMs is to stimulate further discussion about how, with a few strategic changes, existing assessments may be adapted for use as DDMs. According to ESE, the strongest DDMs will emerge from districts that have engaged in systematic study of the appropriateness and usefulness of their assessments specifically for determining what students have learned from their educators during a course of instruction.

A number of approaches to measuring growth described in a key ESE-developed resource, \textit{Technical Guide B}, do not require complex psychometric methods or statistical computations. ² Each of these approaches has unique strengths and limitations. Many measures can be adapted to a variety of approaches to measuring growth, so decision-makers will want to use their professional judgment in weighing the pros and cons of each, considering competing goals and determining the approaches best suited for their contexts.

This document is intended to support those educators and administrators who are considering a \textit{pre-test/post-test} approach to measuring growth with their locally developed assessments.

---

¹ \textit{Massachusetts Model System for Educator Evaluation, Part VII: Rating Educator Impact on Student Learning Using District-Determined Measures of Student Learning, Growth, and Achievement.}

² Companion documents have been developed that highlight three other common approaches to measuring growth: (a) a \textit{holistic evaluation} approach; (b) a \textit{repeated measures} approach; and (c) a \textit{post-test only} approach.
It includes the following:

- key resources developed by ESE that offer guidance for monitoring the technical quality of DDMs and for selecting a reasonable approach to measuring growth;
- a summary of the strengths and limitations of the *pre-test/post-test* approach to examining student growth;
- sample assessments submitted by Commonwealth districts, with suggestions for how the developers might refine each measure for use as a DDM; and
- information about a number of external resources that educators and administrators may find helpful if they elect to pursue a *pre-test/post-test* approach.

It is hoped that this guidance document will help district personnel transform strong locally developed measures of achievement and/or performance into promising DDMs that are particularly effective in evaluating student growth.
Section I. Using the Pre-Test/Post-Test Approach to Measure Student Growth with a Promising DDM

To better understand what students have learned over time, educators might administer a pre-test at the beginning of a course and a post-test following instruction. This is a commonly used approach to measuring student growth. In implementing this approach, an educator examines the difference between scores on two assessments administered at different points in time. This information can help educators detect changes in performance that may be linked to effective instruction. This approach can be used with traditional tests of achievement (e.g., end-of-grade/end-of-course assessments or teacher-developed exams) or more non-traditional measures that are performance-, project-, or portfolio-based.

This approach has a number of strengths and limitations, based on guidance from experts in the educational research and measurement communities. These are summarized in Table 1.³

³ See Appendix B for research- and measurement-based resources for the pre-test/post-test approach.
Table 1. Strengths and Limitations of the Pre-Test/Post-Test Approach to Measure Student Growth

<table>
<thead>
<tr>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Is familiar to educators.</td>
<td>• Assumes a common scale that allows scores to be compared across different points in time and allows score differences to be interpreted as representing increased mastery of specific knowledge or a set of skills.</td>
</tr>
<tr>
<td>• Gain score (difference between the two scores) approach is easily computed but also can be the foundation for more complex analyses.</td>
<td>• May have ceiling effect if measures do not have sufficient numbers of very hard items or elements, so that students with strong baseline scores can demonstrate growth.</td>
</tr>
<tr>
<td>• Can use identical measures administered twice or comparable versions.</td>
<td>• Raw gain score should not be used to rank students.</td>
</tr>
<tr>
<td>• Provides descriptive information about both direction (positive or negative) and magnitude of change (how much).</td>
<td>• Items or elements on measures must be carefully developed or selected, to ensure that change in what students know and can do is directly linked to what was taught.</td>
</tr>
<tr>
<td>• When used to examine magnitude of gain at the student level or average gain for a classroom, this approach is appropriate and can have sufficient precision.</td>
<td>• May create incentive to keep pre-test scores low.</td>
</tr>
<tr>
<td>• Growth expectations can be aligned to an absolute standard (student made progress toward a set goal) or norm-referenced (student made more or less progress than peers). If model of choice uses absolute standard, value judgments are necessary for low, moderate, and high growth. If model is norm-referenced, rating can be interpreted as a percentile ranking.</td>
<td></td>
</tr>
</tbody>
</table>

Other Considerations

Districts interested in using the pre-test/post-test approach also may want to attend to the following considerations:

• Results can be reported as raw gain (e.g., 20-point increase between pre-test and post-test) or percentage increase (e.g., 20-percent increase between pre-test and post-test). The method used should ensure comparability between students beginning at different levels of achievement. Districts may also use an approach where different sets of parameters are used to determine high, moderate, or low growth based on pre-test scores.

• If using comparable versions of a test, developers need to ensure that the measures are not so different that changes in performance may be due to variations in the measures rather than actual student growth.

• The time interval between tests must be explicit and intentional. It may be based on time (e.g., after ten weeks of instruction) or curriculum (e.g., after 25 lessons).

• If linked assessments are scored on a common scale, scores yielded should represent the full range of performance at the beginning and the end of the school year.
Section II. Example Assessment #1, Using the Pre-Test/Post-Test Approach to Measure Student Growth

The following assessment was submitted by a Commonwealth educator who adapted an existing measure for use as a DDM. For more information about the assessment, please contact Katie Novak at novakk@chelmsford.k12.ma.us.

Locally Developed Assessment Well Suited for Use as a DDM
Submitted by: Katie Novak, Reading Coordinator, Chelmsford Public Schools

Content Area/Grade: English Language Arts (ELA), Grade 4
Assessment Name: Text-Based Opinion Letters
Item Type: Writing Prompt and Scoring Rubric

Description: This measure is aligned to the Massachusetts Framework for ELA, with particular emphasis on reading informational text and writing opinion pieces. Students read one or more documents and respond to a prompt that requires them to cite text-based evidence in their responses. The packet currently includes two prompts, one intended to be administered at the beginning of the year (“Diagnostic Prompt”) and one intended to be administered at the end of the year (“Summative Prompt”). Each is provided below.

**Diagnostic Prompt:** Imagine you have a younger brother at home who is eight years old. After reading Do Kids Need Their Own Cell Phones and When Should Kids Get Cell Phones?, write an opinion letter to your family about whether or not your brother should have a cell phone. Use at least three specific details from the articles to support your opinion.

**Summative Prompt:** How exciting! Your family has just decided to move to the United States from another country. Write a letter to your family telling them what region they should move to. Be sure to include at least three specific reasons, supported with details and facts from your Social Studies Alive textbook, to support your opinion.

Students are expected to read the documents during class time and then develop a plan for their responses. They must complete the task independently during one-hour sessions on each of two consecutive days. This assessment requires students to submit a rough draft of their letter and a final draft that shows how they revised and edited their submission. All grade 4 educators in this district are expected to administer these prompts during the same two-week window of time.

Educators use prompt-specific analytic rubrics to score student responses, with different traits scored individually. Traits evaluated include quality of ideas, organization, word choice, sentence fluency, voice, and conventions. For each trait, educators assign a rating of Beginning, Developing, Proficient, or Accomplished. Each student is then assigned subscores for topic development and use of developmentally appropriate language conventions.

This assessment packet includes two grade-appropriate writing prompts, an administration protocol with the script for each writing session, and scoring guidelines.
**Student Handout—Fall Diagnostic Assessment**

DDM Assessment

Answer the following prompt on the lined paper provided. Remember, you will have two one-hour sessions to complete a rough draft and a final draft of your letter. You may use the article *Do Kids Need Their Own Cell Phones?* to help you write your letter. You may also use the assignment rubric to help you plan, draft, revise, and edit.

Prompt:

*Imagine you have a younger brother at home who is eight years old. After reading* *Do Kids Need Their Own Cell Phones and When Should Kids Get Cell Phones?, write an opinion letter to your family about whether or not your brother should have a cell phone. Use at least three specific details from the articles to support your opinion.*

You may use the space below to plan your response or create a graphic organizer if you choose.

---

**Student Handout—Spring Summative Assessment**

DDM Assessment

Answer the following prompt on the lined paper provided. Remember, you will have two one-hour sessions to complete a rough draft and a final draft of your letter. You may use your *Social Studies Alive* textbook to help you write your letter. You may also use the assignment rubric to help you plan, draft, revise, and edit.

Prompt:

*How exciting! Your family has just decided to move to the United States from another country. Write a letter to your family telling them which region they should move to. Be sure to include at least three specific reasons, supported with details and facts from your Social Studies Alive textbook, to support your opinion.*

You may use the space below to plan your response or create a graphic organizer if you choose.
## Opinion Prompt Rubric Grade 4

<table>
<thead>
<tr>
<th>TRAITS</th>
<th>BEGINNING (1)</th>
<th>DEVELOPING (2)</th>
<th>PROFICIENT (3)</th>
<th>ACCOMPLISHED (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ideas</td>
<td>□ No topic is stated&lt;br&gt;□ No opinion is stated&lt;br&gt;□ No reasons with supporting facts and details are presented</td>
<td>□ Topic is vague or inferred&lt;br&gt;□ Opinion is vague or inferred&lt;br&gt;□ Reasons for point of view are weak and/or limited and facts and details are weak and/or limited</td>
<td>□ Topic is clearly introduced&lt;br&gt;□ Opinion is stated&lt;br&gt;□ Provides reasons for point of view that are supported by facts and details</td>
<td>□ Topic is well developed&lt;br&gt;□ Opinion is well articulated&lt;br&gt;□ Provides carefully chosen reasons for point of view that are compelling and supported by relevant facts and details</td>
</tr>
<tr>
<td>Organization</td>
<td>□ Does not group related ideas to support the writer’s purpose&lt;br&gt;□ Does not link opinions and reasons, supporting point of view with transitional words and phrases (e.g., for instance, in order to, in addition)&lt;br&gt;□ Does not provide a conclusion or conclusion is unrelated to opinion presented</td>
<td>□ Grouping of some ideas may be confusing, detracting from the writer’s purpose&lt;br&gt;□ Sometimes links opinions and reasons, supporting point of view with transitional words and phrases (e.g., for instance, in order to, in addition)&lt;br&gt;□ Provides a limited or weak conclusion related to the opinion</td>
<td>□ Most related ideas are grouped to support the writer’s purpose&lt;br&gt;□ Links most opinions and reasons, supporting point of view with transitional words and phrases (e.g., for instance, in order to, in addition)&lt;br&gt;□ Provides a concluding statement or section related to the opinion presented</td>
<td>□ All related ideas are strategically grouped to support the writer’s purpose&lt;br&gt;□ Always links opinions and reasons, supporting point of view with transitional words and phrases (e.g., for instance, in order to, in addition)&lt;br&gt;□ Provides a carefully developed conclusion related to the opinion presented</td>
</tr>
<tr>
<td>Word Choice</td>
<td>□ Includes a limited range of words often used incorrectly&lt;br&gt;□ Words are not appropriate for audience and purpose&lt;br&gt;□ Includes many overused words</td>
<td>□ Begins to use some concrete words and phrases and sensory details to convey meaning, including similes and metaphors&lt;br&gt;□ Some words may not be appropriate for audience and purpose&lt;br&gt;□ Includes some overused words</td>
<td>□ Uses some concrete words and phrases and sensory details to convey meaning, including similes and metaphors&lt;br&gt;□ Words are mostly appropriate for audience and purpose&lt;br&gt;□ Mostly avoids overused words</td>
<td>□ Consistently includes carefully chosen precise words and phrases and sensory details to convey meaning precisely, including similes and/or metaphors&lt;br&gt;□ Words are consistent with audience and purpose&lt;br&gt;□ Writer avoids overused words; includes a variety of words to convey meaning</td>
</tr>
<tr>
<td>Sentence Fluency</td>
<td>□ Includes limited use of simple, compound, and complex sentences&lt;br&gt;□ Includes limited use of sentence types (exclamatory, interrogative, imperative, declarative)&lt;br&gt;□ Includes many incomplete sentences relative to the length of the piece&lt;br&gt;□ Includes few or no varied sentence beginnings</td>
<td>□ Includes some variety of simple, compound, and complex sentences&lt;br&gt;□ Includes some variety of sentence types (exclamatory, interrogative, imperative, declarative)&lt;br&gt;□ Includes some incomplete sentences relative to the length of the piece&lt;br&gt;□ Includes some varied sentence beginnings</td>
<td>□ Includes a variety of simple, compound, and complex sentences&lt;br&gt;□ Includes a variety of sentence types (exclamatory, interrogative, imperative, declarative)&lt;br&gt;□ Includes a few incomplete sentences relative to the length of the piece&lt;br&gt;□ Includes many varied sentence beginnings</td>
<td>□ Creatively uses a variety of simple, compound, and complex sentences&lt;br&gt;□ Creatively uses a variety of sentence types (exclamatory, interrogative, imperative, declarative)&lt;br&gt;□ Includes all complete sentences relative to the length of the piece&lt;br&gt;□ Includes many varied sentence beginnings purposefully for effect</td>
</tr>
<tr>
<td>TRAITS</td>
<td>BEGINNING (1)</td>
<td>DEVELOPING (2)</td>
<td>PROFICIENT (3)</td>
<td>ACCOMPLISHED (4)</td>
</tr>
<tr>
<td>--------</td>
<td>--------------</td>
<td>----------------</td>
<td>----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Voice</td>
<td>☐ No evidence of writer’s personality</td>
<td>☐ Begins to show evidence of writer’s personality and style</td>
<td>☐ Includes evidence of writer’s personality and style</td>
<td>☐ Includes evidence of writer’s personality and style and begs to be read aloud</td>
</tr>
<tr>
<td></td>
<td>☐ Reader feels no connection with the writer</td>
<td>☐ Reader occasionally feels a connection with the writer</td>
<td>☐ Reader feels a personal connection with the writer</td>
<td>☐ Writer thoughtfully creates a personal connection with the reader</td>
</tr>
<tr>
<td></td>
<td>☐ Voice is not appropriate for the purpose and audience</td>
<td>☐ Voice is somewhat appropriate for the purpose, audience, topic, and/or genre</td>
<td>☐ Voice is mostly appropriate for the purpose, audience, topic, and/or genre</td>
<td>☐ Voice is consistently appropriate for the purpose, audience, topic, and/or genre</td>
</tr>
<tr>
<td>Conventions</td>
<td>☐ Has multiple errors in grade level appropriate grammar</td>
<td>☐ Has some errors in grade level appropriate grammar</td>
<td>☐ Has few errors in grade level appropriate grammar</td>
<td>☐ Has no errors in grade level appropriate grammar</td>
</tr>
<tr>
<td></td>
<td>☐ Has multiple errors in grade level appropriate punctuation and capitalization</td>
<td>☐ Has some errors in grade level appropriate punctuation and capitalization</td>
<td>☐ Has few errors in grade level appropriate punctuation and capitalization</td>
<td>☐ Has no errors in grade level appropriate punctuation and capitalization</td>
</tr>
<tr>
<td></td>
<td>☐ Has multiple errors in grade level conventional spelling</td>
<td>☐ Has some errors in grade level conventional spelling</td>
<td>☐ Has few errors in grade level conventional spelling</td>
<td>☐ Has no errors in grade level conventional spelling</td>
</tr>
</tbody>
</table>

Topic Development in Writing (TD) =__________
Language Conventions (C) =__________

**Proctoring Protocol**

Guidelines for Administration of DDM Writing Prompts

Pre-test: Diagnostic text-based opinion prompt, *Do Kids Need Their Own Cell Phones?*

Administer between October 1 and October 15

When administering the prompt, use the exact following procedure so that administration is consistent across all students, classrooms, and schools. Administer sessions on two consecutive days.

**Notes**

- Prior to administering the prompt, do not read the articles *Do Kids Need Their Own Cell Phones?* or *When Should Kids Get Cell Phones?* to students.

- If students ask questions about the articles, including asking about the meaning of a word, reply with “*Just go back and reread the article. Do the best you can.*”

- If students ask questions about the prompt, reply with “*Just go back and reread the prompt. Do the best you can.*”
• Provide students with all documented accommodations (IEP, 504, ELL) during all sessions.

SESSION 1:

1. Ask students to clear their desks except for lined paper, a pen and/or pencil, and copies of Do Kids Need Their Own Cell Phones? and When Should Kids Get Cell Phones? Do not provide students with dictionaries, thesauri, graphic organizers (unless outlined specifically in their accommodations), or any reminders about the writing process or the six traits. If you have posters on the wall that provide writing tips or information about the six traits, please remove or cover them.

2. Give all students copies of the prompt and the rubric. Also write the prompt on the board.

*Imagine you have a younger brother at home who is eight years old. After reading Do Kids Need Their Own Cell Phones and When Should Kids Get Cell Phones?, write an opinion letter to your family about whether or not your brother should have a cell phone. Use at least three specific details from the articles to support your opinion.*

3. Once students are settled in their desks with all necessary materials, say the following: “Today you will be completing a writing prompt on two articles—Do Kids Need Their Own Cell Phones? and When Should Kids Get Cell Phones?—independently. When you complete something independently, you complete it by yourself, without any help. This means that you are not allowed to ask me any questions about what you should write or how you should organize your writing. You are welcome to brainstorm or create your own graphic organizers, but I cannot help you with that.

You will be given two one-hour sessions to complete your writing. During the first session, you should try to write your rough draft and begin to revise your rough draft. During the second session, you should try to edit your work and complete your final draft in your neatest handwriting. At the end of the second session, I will collect and assess your writing, even if it is not complete.

Please write your name on the top right-hand corner of your lined paper. [Walk around the room to be sure that all students have their names on the paper.] The first session is one hour long. You may begin your first session now.” Write the start time and the finish time on the board.

After a half hour, say, “The first session is half over. As a reminder, during the first session, you should try to write your rough draft and begin to revise your rough draft.”
When the session is over, say, “The first session is over. Please put down your pens or pencils and close your books. I will collect your prompts now.” Collect all student prompts, and staple papers together if necessary.

SESSION 2 (the next day):

1. Ask students to clear their desks except for lined paper, a pen and/or pencil, and copies of Do Kids Need Their Own Cell Phones? and When Should Kids Get Cell Phones? Do not provide students with dictionaries, thesauri, graphic organizers (unless outlined specifically in their accommodations), or any reminders about the writing process or the six traits. If you have posters on the wall that provide writing tips or information about the six traits, please remove or cover them.

2. Give all students copies of the prompt and the rubric. Also write the prompt on the board.

Imagine you have a younger brother at home who is eight years old. After reading Do Kids Need Their Own Cell Phones? and When Should Kids Get Cell Phones?, write an opinion letter to your family about whether or not your brother should have a cell phone. Use at least three specific details from the articles to support your opinion.

3. Once students are settled in their desks with all necessary materials, say the following: “Today you will be completing a writing prompt on two articles—Do Kids Need Their Own Cell Phones? and When Should Kids Get Cell Phones?—independently. When you complete something independently, you complete it by yourself, without any help. This means that you are not allowed to ask me any questions about what you should write or how you should organize your writing. During this second session, you should try to edit your work and complete your final draft in your neatest handwriting. At the end of this session, I will collect and assess your writing, even if it is not complete.

[Pass out students’ papers.] This session is an hour long. You may begin your second session now.” Write the start time and the finish time on the board.

After a half hour, say, “The second session is half over. As a reminder, during this session, you should try to finish the final draft of your letter.”

When the session is over, say, “The second session is over. Please put down your pens or pencils and close your books. I will collect your prompts now.” Collect all student prompts, and staple papers together if necessary.

Post-test: Summative text-based opinion prompt, Social Studies Alive!

Administer between May 1 and May 15
When administering the prompt, use the exact procedure below so that administration is consistent across all students, classrooms, and schools. Administer sessions on two consecutive days.

Notes

- You may direct students to a specific section in the *Social Studies Alive* textbook, but you cannot answer any comprehension questions or prompt students with ideas about a particular section. Just say, “*Just go back and reread that section of the text. Do the best you can.*”

- If students ask questions about the prompt, reply with “*Just go back and reread the prompt. Do the best you can.*”

- Provide students with all documented accommodations (IEP, 504, ELL) during all sessions.

PRE-SESSION 1:

1. Ask students to clear their desks except for scrap paper, a pen and/or pencil, and their *Social Studies Alive* textbook. Do not provide students with dictionaries, thesauri, graphic organizers (unless outlined specifically in their accommodations), or any reminders about the writing process or the six traits. If you have posters on the wall that provide writing tips or information about the six traits, please remove or cover them.

2. Give all students copies of the prompt and the rubric. Also write the prompt on the board.

   *How exciting! Your family has just decided to move to the United States from another country. Write a letter to your family telling them which region they should move to. Be sure to include at least three specific reasons, supported with details and facts from your Social Studies Alive textbook, to support your opinion.*

3. Once students are settled in their desks with all necessary materials, say the following: “Today you will be completing a writing prompt, based on your Social Studies Alive textbook, independently. When you complete something independently, you complete it by yourself, without any help. This means that you are not allowed to ask me any questions about what you should write or how you should organize your writing. You are welcome to brainstorm or create your own graphic organizers, but I cannot help you with that.

   *You will be given 30 minutes, before we start writing, to look through your textbook, brainstorm, and come up with your ideas. Then, you will be given two one-hour sessions to complete your writing. During the brainstorming session, you will look through your book to choose your region and select your reasons.*
Please open up your social studies book and begin to plan your response. You may not begin writing your letter, but you can brainstorm or create a graphic organizer to help you. If you need help finding a specific region, call me over and I can help you find it, but I cannot choose a region for you. Once you select your region, take some time to read about it and take notes if you choose. Please start your 30-minute brainstorming session now.”

After the 30-minute brainstorming session, pass out lined paper and say the following: Now, we will begin the first writing session. During the first writing session, you should try to write your rough draft and begin to revise your rough draft. During the second writing session, you should try to edit your work and complete your final draft in your neatest handwriting. At the end of the second session, I will collect and assess your writing, even if it is not complete.”

SESSION 1:
Say, “Please write your name on the top right-hand corner of your lined paper. [Walk around the room to be sure that all students have their names on the paper.] The first session is one hour long. You may begin your first session now.” Write the start time and the finish time on the board.

After a half hour, say, “The first session is half over. As a reminder, during the first session, you should try to write your rough draft and begin to revise your rough draft.”

When the session is over, say, “The first session is over. Please put down your pens or pencils and close your books. I will collect your prompts now.” Collect all student prompts, and staple papers together if necessary.

SESSION 2 (the next day):
1. Ask students to clear their desks except for lined paper, a pen and/or pencil, and their Social Studies Alive textbook. Do not provide students with dictionaries, thesauri, graphic organizers (unless outlined specifically in their accommodations), or any reminders about the writing process or the six traits. If you have posters on the wall that provide writing tips or information about the six traits, please remove or cover them.

2. Give all students a copy of the prompt and the rubric. Also write the prompt on the board.

How exciting! Your family has just decided to move to the United States from another country. Write a letter to your family telling them which region they should move to. Be sure to include at least three specific reasons, supported with details and facts from
your Social Studies Alive textbook, to support your opinion. Once students are settled in their desks with all necessary materials, say the following: “Today you will complete the second session on your writing prompt for Social Studies Alive independently. When you complete something independently, you complete it by yourself, without any help. This means that you are not allowed to ask me any questions about what you should write or how you should organize your writing.

During this second session, you should try to edit your work and complete your final draft in your neatest handwriting. At the end of this session, I will collect and assess your writing, even if it is not complete.

[Pass out students’ papers.] This session is an hour long. You may begin your second session now.” Write the start time and the finish time on the board.

After a half hour, say, “The second session is half over. As a reminder, during this session, you should try to finish the final draft of your letter.”

When the session is over, say, “The second session is over. Please put down your pens or pencils and close your books. I will collect your prompts now.” Collect all student prompts, and staple papers together if necessary.
Guidelines for Scoring of DDM Writing Prompts

All teams must participate in the following scoring protocol to identify “anchor papers” for each of the four proficiency levels. Once anchor papers have been selected for each school, team leaders will send the papers to the reading coordinator so that districtwide anchor papers can be identified before scoring.

What is Reliability?

Reliability is “the degree to which test scores are consistent, dependable, or repeatable, that is, the degree to which they are free of errors of measurement” (AERA, APA, & NCME, 1985). In an effort to increase inter-rater reliability on scoring, teams of teachers will meet to participate in the Reliability Protocol for DDM scoring.

What are common sources of error?

1. Interactions between students and raters: Because teachers know students so well, they sometimes predict how students will perform on a task. These predications can affect scoring.

2. Interactions between raters and task: Raters sometimes have different interpretations of a task, and therefore are expecting different responses.

What is the protocol for establishing inter-rater reliability?

1. Before reading student responses, discuss the prompt and the type of response that would be necessary for a complete, clear, and accurate answer. This discussion will minimize errors based on the interactions between raters and the task.

2. After this discussion, the first response is selected and one rater reads the prompt aloud. The response should be read blind, so that no one knows the identity of the writer. This will minimize errors based on the interactions between students and raters. Note: The writing is read aloud in order to minimize the impact of spelling, grammar, and handwriting on the score assigned, although this will come into play when examining language conventions.

3. After listening to the response, each rater records brief comments indicating their impression of the content, using the prompt rubric.

4. After marking comments, individual raters may ask for the writing sample to be read again, or they may ask to see the piece of writing. After each individual rater has recorded his or her mark, the marks are revealed.

5. If there is consensus on the marks, then the raters read the paper to score the language conventions. Scores for language conventions are then revealed. If there is a difference in the scores assigned, a discussion begins, in which raters describe their rationales for the marks they have given, and a consensus is reached.
6. When all sample papers have been scored, an anchor paper for each scoring level is selected as an exemplar. Use the chart on the following page to determine proficiency levels.

7. These anchor papers are then sent to the reading coordinator to determine district anchor papers.

**Chart Outlining Proficiency Levels**

<table>
<thead>
<tr>
<th>Item</th>
<th>Beginning</th>
<th>Developing</th>
<th>Proficient</th>
<th>Accomplished</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic Development</td>
<td>1–5</td>
<td>6–10</td>
<td>11–15</td>
<td>16–20</td>
</tr>
<tr>
<td>Conventions</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

All anchor papers must have consistent proficiency levels for both topic development and conventions.

**Notes on Scoring**

Once you finish grading, scores must be reported in the Lion’s Drive by October 20 for the diagnostic and May 20 for the summative. If you choose to have them entered for you, the following must be completed before sending them to Joanne Salomaa:

- Papers must be sorted alphabetically by last name.
- Provide the two tallied numeric scores, one value for Topic Development and one value for Conventions. Joanne does not have to see the papers or the rubric, unless it is easier for you.
DDM Common Core Alignment

DDM Assessment

The following two prompts are aligned to the Common Core State Standards for Reading and Writing, Grade 4:

Diagnostic prompt (pre-test)
Imagine you have a younger brother at home who is eight years old. After reading Do Kids Need Their Own Cell Phones? and When Should Kids Get Cell Phones?, write an opinion letter to your family about whether or not your brother should have a cell phone. Use at least three specific details from the articles to support your opinion.

Summative prompt (post-test)
How exciting! Your family has just decided to move to the United States from another country. Write a letter to your family telling them which region they should move to. Be sure to include at least three specific reasons, supported with details and facts from your Social Studies Alive textbook, to support your opinion.

Common Core Standards

Reading Informational Text 4.1: Refer to details and examples in a text when explaining what the text says explicitly and when drawing inferences from the text.

Writing 4.1: Write opinion pieces on topics or texts, supporting a point of view with reasons and information.

- W.4.1a: Introduce a topic or text clearly, state an opinion, and create an organizational structure in which related ideas are grouped to support the writer’s purpose.
- W.4.1b: Provide reasons that are supported by facts and details.
- W.4.1c: Link opinion and reasons using words and phrases (e.g., for instance, in order to, in addition).
- W.4.1d: Provide a concluding statement or section related to the opinion presented.

Writing 4.4: Produce clear and coherent writing in which the development and organization are appropriate to task, purpose, and audience. (Grade-specific expectations for writing types are defined in standards 1–3 above.)
Section III. Suggestions for Refining Example Assessment #1 to Ensure Usefulness as a DDM

This example assessment shows great promise for use as a DDM. Educators from Chelmsford Public Schools have worked together to develop a set of assessment packets that are well suited for use with the pre-test/post-test approach to measuring growth. Because these measures were developed collaboratively by local educators, users are confident that they are framework-aligned and developmentally appropriate.

In particular, the guidelines for scoring student work are unique. They explain the importance of consistent and accurate scoring across educators and promote standardized scoring practices. Educators engaged in scoring are expected to follow a protocol as they consider each prompt and brainstorm possible responses that would be considered “complete, clear, and accurate.” They are trained to apply the scoring rubric using group calibration exercises and anchor papers (samples of actual student work) that will serve as exemplars for each proficiency level.

Suggestions for refining this assessment are shared to illustrate some of the possible ways in which districts might, with slight modification, use existing assessments as DDMs.

**Suggestion 1: Examine performance patterns for each evaluation criterion and set meaningful parameters for growth.** After students’ responses to the pre- and post-test prompts are scored, students’ point totals and associated proficiency levels on the two measures can be compared, by criterion (topic development and conventions). Note that topic development has a possible score range of 1–20 and conventions has a possible score range of 1–4. Using this information, subscores can be calculated for each criterion.

For illustrative purposes, the growth patterns for two hypothetical students, Student A and Student B, are shown in Table 2.

**Table 2. Growth Patterns for Students A and B on Topic Development and Conventions, by Test Event**

<table>
<thead>
<tr>
<th>Student</th>
<th>Criterion</th>
<th>Pre-Test Score</th>
<th>Pre-Test Proficiency Level</th>
<th>Post-Test Score</th>
<th>Post-Test Proficiency Level</th>
<th>Point Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student A</td>
<td>Topic Development</td>
<td>3</td>
<td>Beginning</td>
<td>10</td>
<td>Developing</td>
<td>7</td>
</tr>
<tr>
<td>Student A</td>
<td>Conventions</td>
<td>2</td>
<td>Developing</td>
<td>2</td>
<td>Developing</td>
<td>0</td>
</tr>
<tr>
<td>Student B</td>
<td>Topic Development</td>
<td>5</td>
<td>Beginning</td>
<td>12</td>
<td>Proficient</td>
<td>7</td>
</tr>
<tr>
<td>Student B</td>
<td>Conventions</td>
<td>1</td>
<td>Beginning</td>
<td>2</td>
<td>Developing</td>
<td>1</td>
</tr>
</tbody>
</table>
As shown in Table 2, Students A and B started and ended the school year at different score points for topic development, but each showed a score gain of seven points and moved into a higher proficiency level category. For conventions, Student B moved from Beginning to Developing, but Student A did not demonstrate comparable growth.

Districts using this approach will want to set parameters for low, moderate, and high growth for each criterion. Districts that value movement from one proficiency level to another might set parameters using a qualitative approach. A district that prefers this approach might also decide to report separate growth scores for the two criteria, as each measures very different knowledge and skills and incorporates different scales.

For topic development, for example, district educators may decide that any student moving to a higher proficiency level (e.g., from Beginning to Developing, from Developing to Proficient, or from Proficient to Accomplished) would receive a rating of high growth. Doing so would require a score change, at a minimum, of between one and five points, depending on students’ starting and ending places. Students whose score on the post-test was higher than their score on the pre-test, but who did not change proficiency levels, might thus receive a rating of moderate growth. A student who demonstrated no score gain between the pre- and post-tests would receive a rating of low growth.

Similarly, for conventions, students who move to a higher proficiency level based on their pre-test/post-test gain would receive a rating of high growth; although their score would only need to change by one point, educators may agree that it has proven challenging for students to show growth on this criterion and that a change of one score point is meaningful. Students with no change in score points between test events would be assigned a rating of moderate growth if the student demonstrated growth on two or more of the checked boxes on the Opinion Prompt Rubric (see the grade 4 example on pp. 6–7) for conventions. Students with no change in score points between test events and who do not demonstrate growth on at least two checked boxes would be assigned a rating of low growth.

**Suggestion 2: Identify a second DDM that complements this promising measure.** Since this assessment is a performance-based measure of knowledge and skills, districts considering this option may want to incorporate a more traditional measure of achievement for their second DDM. For example, they could capitalize on existing data collected through the Massachusetts Comprehensive Assessment System, using results from student testing in English language arts in grade 4.
Section IV. Example Assessment #2, Using the Pre-Test/Post-Test Approach to Measure Student Growth

The following assessment was submitted by a Commonwealth educator who adapted an existing measure for use as a DDM. For more information, please contact Michael Quinlan at Michael_Quinlan@wellesley.k12.ma.us.

Locally Developed Assessment Well Suited for Use as a DDM
Submitted by: Michael Quinlan, Educator, Wellesley High School

**Course/Grade:** Introductory Physics, Grade 9

**Assessment Name:** Unit Tests

**Item Type:** Selected Response

**Description:** These are educator-developed, unit-specific exams that are aligned to the Massachusetts Curriculum Framework for Science, Technology, and Engineering. Pre- and post-tests for each unit include ten items administered at the beginning and end of each school year. Items on the pre- and post-tests are identical.

The educator collects the baseline information about what students know and can do upon course entry, for use in instructional planning. He collects the post-test information to ensure that students have met his learning targets as anticipated. All exam items were adapted from a variety of publicly available resources for high school science teachers.

Each item was strategically selected to assess the knowledge and skills that students can be expected to learn in this class but to which they are unlikely to have had prior exposure. The educator’s goal was to focus on those elements of critical content that all students could be expected to learn during this course and the skills that they would acquire as a direct result of his instruction. By focusing on the body of knowledge and skills to which few students were likely to have been exposed prior to this course, he thought he would have a strong measure of the impact of his teaching on student learning.

Based on prior experience, this Commonwealth educator decided that he needed only a small number of items—perhaps 10–12—to accomplish his goals if each was a strong measure of the content knowledge or skills that are central to the content domain but to which students were least likely to have been exposed prior to the course. He believed that these 10-item measures would be of sufficient length to provide the information that he was seeking, without creating undue burden to students or taking too much time away from instruction. He also decided to include the same items on both the pre- and post- versions of the test, thereby simplifying his work in comparing results from the two measures to draw meaningful conclusions about student growth.

Pre- and post-tests are available for 14 units of instruction. The items for one unit (Unit 2: Describing Motion) are provided on the following pages. Exams for other units can be obtained by contacting the educator.

Measures included in the set assess the following units of instruction:

- Unit 1: Scientific Skills
- Unit 3: Explaining Motion
- Unit 5: Work and Power
- Unit 7: Heat and Heat Transfer
- Unit 9: Current Electricity
- Unit 11: Waves (Light and Sound)
- Unit 13: State Test Preparation (Massachusetts Comprehensive Assessment System)
- Unit 2: Describing Motion
- Unit 4: Momentum
- Unit 6: Energy and Energy Conservation
- Unit 8: Static Electricity
- Unit 10: Electromagnetism
- Unit 12: Electromagnetic Spectrum
- Unit 14: Engineering and Design Project
Introductory Physics, Unit 2 Pre-Test and Post-Test (Describing Motion)

Name: ____________________ Date: _______________

1. (1.1) A car is moving at 25 m/s north. Which of the following is a vector quantity?
   A. the speed of the car
   B. the velocity of the car
   C. the potential energy of the car
   D. the distance traveled by the car

2. (1.1) A satellite in a circular orbit around Earth has a constant speed but not a constant velocity. Which of the following statements best explains why the satellite’s velocity is not constant?
   A. The radius of the satellite’s orbit is too large.
   B. The force on the satellite’s mass is constantly decreasing.
   C. The magnitude of the satellite’s momentum is too large.
   D. The direction of the satellite’s motion is constantly changing.

3. (1.2) Two soccer players, X and Y, are kicking a ball back and forth to each other, as shown below.
   The table below shows the distance and direction the ball moves after each of four kicks.

<table>
<thead>
<tr>
<th>Kick</th>
<th>Player</th>
<th>Distance and Direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>5 m right</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>4 m left</td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>6 m right</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>5 m left</td>
</tr>
</tbody>
</table>

   What is the total displacement of the ball and the total distance traveled by the ball?
   A. The ball has a displacement of 2 m to the right and traveled a distance of 20 m.
   B. The ball has a displacement of 20 m to the right and traveled a distance of 2 m.
   C. The ball has a displacement of 9 m to the right and traveled a distance of 11 m.
   D. The ball has a displacement of 11 m to the right and traveled a distance of 9 m.

4. (1.2) A racecar travels at an average speed of 80 m/s in a race. The total distance of the race is 640,000 m. How long does it take the racecar to finish the race?
   A. 800 s  
   B. 8,000 s  
   C. 6,400,000 s  
   D. 51,200,000 s

5. (1.2) How long will it take a car to accelerate from 20 m/s to 26 m/s at a rate of 1.4 m/s²?
   A. 2.2 s  
   B. 2.7 s  
   C. 4.3 s  
   D. 4.6 s
6. (1.2) A rock is dropped from a window 5 m above the ground. The rock hits the ground 1.0 s later with a speed of 10 m/s. What is the average speed of the rock during this time?  
A. 5 m/s  
B. 8 m/s  
C. 15 m/s  
D. 50 m/s

7. (1.3) An object is traveling in a straight line. The graph below shows the object’s velocity over time.

![Motion of an Object Graph](image)

Which line segment shows the object traveling with a constant, positive acceleration?  
A. segment W  
B. segment X  
C. segment Y  
D. segment Z

8. (1.3) The graph below represents the motion of an object over four time intervals, W, X, Y, and Z.

![Motion of an Object Graph](image)

Over which time interval is the object moving the fastest?  
A. interval W  
B. interval X  
C. interval Y  
D. interval Z

9. (1.3) The graph below shows velocity measurements made as a car moved north for 25 s.

![Motion of a Car Graph](image)

How far did the car move during the first 15 s of the trip?  
A. 20 m  
B. 25 m  
C. 300 m  
D. 500 m
10. (1.2) A car has an oil drip. As the car moves, it drips oil at a regular rate, leaving a trail of spots on the road. Which of the following diagrams of the car’s trail of spots shows the car continuously slowing down?

A.  

B.  

C.  

D.  

Start  

Finish  

Start  

Finish  

Start  

Finish  

Start  

Finish  

An experienced Commonwealth educator has developed a set of tests that show great promise for use as a DDM. Each test is intentionally designed to measure growth in student learning over the course of a unit of instruction. He uses baseline information about what students know and can do for the purpose of instructional planning, and uses post-test information to ensure that students have met his learning targets as anticipated.

Suggestions for refining this assessment are shared to illustrate some of the possible ways in which districts might, with slight modification, use existing assessments as DDMs.

**Suggestion 1: Strategically combine the individual assessments into a measure of growth.**
Currently, each assessment in this set is designed to measure one unit of instruction. To ensure that the DDM assesses valued content from across the entire school year, districts interested in this option might take one of two approaches to transforming this promising measure into an effective DDM: (1) a composite score approach or (2) a content sampling approach. A district using the composite score approach might, at the end of the school year, compute gain scores (by subtracting the number correct on Test #1 from the number correct on Test #2) for the pre- and post-tests for all units. The district could then combine these scores for each student into a composite score for each student, which would serve as the overall growth calculation for that student.

Alternatively, districts might consider a content sampling approach, using annually collected performance data to select those items from the first 12 units that have proven to be most useful in highlighting the knowledge and skills that students have learned following a year of instruction. For example, decision-makers might want to identify those items with a low mean number correct on the pre-test but a significantly higher mean number correct on the post-test. They could defend selecting these items as that the items appear to be those that most students could not answer correctly prior to instruction, but that many more students were able to answer correctly at the end of the year. They also could use their professional judgment to select those items that measure the degree to which students have learned the content that the district values most, i.e., the elements of content that are centrally linked to learning in a subject/grade or course.

With either approach, districts will want to consider the number of items that is necessary to ensure that sufficient information is collected to inform decision-making about student growth without creating undue burden to students or taking too much time away from instruction. As a starting point, districts may want to consider developing pre- and post-test assessments with 35–50 items appearing on both assessments. To do so, for example, developers might identify...
three items per unit for Units 1–12 (for a total of 36) or four items per unit for Units 1–12 (for a total of 48) to be included on both tests. Alternatively, districts that believe that certain units of instruction are more critical to end-of-year learning expectations could pull more items from one unit and fewer from another. In either case, developers will want to consult with educators who teach this grade/subject or course, in order to vet their choices.

**Suggestion 2: Set meaningful parameters for growth.** Information collected through a pre-test/post-test approach can be used to help district representatives set defensible parameters for low, moderate, and high growth for each subject/grade or course. Once they have calculated growth scores for each student, they will want to look at the distribution of scores, from lowest to highest. For example, if the pre- and post-tests share the same 50 items, a student who answers 10 items correctly on the pre-test and 20 items correctly on the post-test would be assigned a growth score of 10, which is likely to be at the lower end of the performance continuum for that classroom. Similarly, a student who answers 25 items correctly on the pre-test and 50 correctly on the post-test would be assigned a relatively higher growth score of 25.

How might a district group these growth scores into categories for low, moderate, and high growth? Using data from the pilot administration of 50-item pre- and post-tests, DDM developers may notice that many students were assigned a growth score of around 27 and decide that students with growth scores in the 17–36 range will comprise the moderate-growth group. Using their professional judgment, district representatives may decide to set the parameter for high growth at 37 and the parameter for low growth at 16; students with growth scores in the 37–50 range will be in the high-growth group, and students with growth scores in the 0–16 range will be in the low-growth group. District representatives might also want to review the distribution of student grades in a grade/subject or course, or student performance on other measures of achievement, to collect evidence that supports the district’s claim that these parameters are reasonable. Finally, district representatives will want to revisit the impact of their decisions each year and make adjustments as needed to ensure that no unintended negative consequences from test use are emerging.

**Suggestion 3: Conduct annual reviews of performance data to improve DDM quality.** Districts electing to use this approach will want to carefully review student responses to each item. These post-administration analyses will help developers with decision-making about changing or replacing items and ensuring that, over time, all items are closely linked to what is taught. They may review item-level statistics such as item difficulty, looking carefully at items with very low (e.g., below .20) and very high (e.g., above .95) values. Such findings would suggest to developers that, on average, students found those items to be quite challenging or easy,

---

4 For additional information about item difficulty, please see the ESE-developed Webinar Series #9 on Sustainability.
respectively, and that further review is warranted to determine if the item stem or one or more distracters are flawed or in need of refinement.

Districts also might use post-test analyses as an opportunity to look for possible floor and ceiling effects. Doing so can improve the quality of a DDM by ensuring that it has the capacity to capture the full continuum of expected student performance, from lowest to highest. If, for example, on a 50-item pre-test, only about half of the students answer more than five items correctly and only a few students answer more than 10 items correctly, developers may need to address a potential floor effect. A floor effect creates a problem for test developers because the measure does not differentiate among students with relatively lower levels of knowledge and skills—as might be expected for a course with content that is new to most students. Given that scenario, districts might elect to replace items in the item set with other items designed to capture more finely grained differences in what students know and can do upon course entry. Similarly, if many more students than expected perform well on the pre-test, DDM developers may need to adjust their assumptions about the background knowledge and experience of entering students and modify the item set to better match these new expectations.
Appendix A: Key ESE-Developed Resources to Support Districts with Implementing DDMs that Effectively Measure Student Growth

August 2012

Part VII, Massachusetts Model System for Educator Evaluation: Rating Educator Impact on Student Learning Using District-Determined Measures of Student Learning

Overview of DDMs and related concepts. It will be most valuable for districts beginning to learn about this work.

http://www.doe.mass.edu/edeval/model/PartVII.pdf

Monthly Since February 2013

 Educator Evaluator e-Newsletter

Monthly newsletter designed to be a timely resource that provides key information, updates, and answers to frequently asked questions.

http://www.doe.mass.edu/edeval/communications/newsletter/

March 2013

Introduction: District-Determined Measures and Assessment Literacy (Webinar Series #1)

ESE has developed a nine-part webinar series on DDMs and assessment literacy. This series is targeted at district teams engaged in the work of identifying and selecting DDMs (e.g., district- and school-based curriculum and assessment leaders). Resources from these webinars include the recorded webinar and materials from each session.

http://www.doe.mass.edu/edeval/ddm/webinar.html

April 2013

Basics of Assessment and Assessment Options (Webinar Series #2 and #3)

http://www.doe.mass.edu/edeval/ddm/webinar.html

Technical Guide A: Considerations Regarding District-Determined Measures

Designed to increase assessment literacy by introducing foundational assessment concepts. It will be most valuable to districts interested in learning more about technical assessment concepts.

http://www.doe.mass.edu/edeval/ddm/TechnicalGuide.pdf
April 2013 (continued)

Assessment Quality Checklist and Tracking Tool
An interactive tool, built in Microsoft Excel, that organizes and catalogs information about individual assessments into a districtwide tracker of all potential DDMs. It will be most valuable to districts working to identify and select measures across the district.
http://www.doe.mass.edu/edeval/ddm/webinar/Quality-Tracking-Tool.xlsm

July 2013

Determining the Best Approach to District-Determined Measures (Webinar Series #4)
http://www.doe.mass.edu/edeval/ddm/webinar.html

DDM Technical Assistance and Networking Session I
ESE-hosted technical assistance and networking sessions intended to build on the Assessment Literacy Webinar Series and provide participants an opportunity to engage with colleagues from other districts around critical planning and implementation questions related to the piloting and eventual implementation DDMs.
http://www.doe.mass.edu/edeval/ddm/webinar.html

August 2013

Measuring Student Growth and Piloting District-Determined Measures (Webinar Series #5)
http://www.doe.mass.edu/edeval/ddm/webinar.html

September 2013

Technical Guide B: Measuring Student Growth & Piloting District-Determined Measures

DDM Technical Assistance and Networking Session II
http://www.doe.mass.edu/edeval/ddm/webinar.html

October 2013

Determining How to Integrate Assessments into Educator Evaluation: Developing Business Rules and Engaging Staff (Webinar Series #6)
http://www.doe.mass.edu/edeval/ddm/webinar.html

Using Current Assessments in District-Determined Measures: Leveraging the Curriculum-Embedded Performance Assessments from the Model Curriculum Units
http://www.doe.mass.edu/edeval/ddm/UsingAssessments.pdf
December 2013

Ramping Up for Next Year: Strategies for Using Current Assessments as DDMs (Webinar Series #7)
http://www.doe.mass.edu/edeval/ddm/webinar.html

DDM Technical Assistance and Networking Session III
http://www.doe.mass.edu/edeval/ddm/webinar.html

January 2014

Communicating Results (Webinar Series #8)
http://www.doe.mass.edu/edeval/ddm/webinar.html

February 2014

Sustainability (Webinar Series #9)
http://www.doe.mass.edu/edeval/ddm/webinar.html

Implementation Brief: Scoring and Setting Parameters
http://www.doe.mass.edu/edeval/ddm/Scoring-ParameterSet.pdf

Implementation Brief: Investigating Fairness
http://www.doe.mass.edu/edeval/ddm/Fairness.pdf

Implementation Brief: Using Student Growth Percentiles
http://www.doe.mass.edu/edeval/ddm/GrowthPercentiles.pdf

March 2014

Implementation Brief: Indirect Measures and Specialized Instructional Support Personnel (SISP)
http://www.doe.mass.edu/edeval/ddm/IMSISP.pdf

April 2014

Implementation Brief: Administrators
http://www.doe.mass.edu/edeval/ddm/Admin.pdf

Implementation Brief: Considerations for English Language Learners
http://www.doe.mass.edu/edeval/ddm/ELLEducators.pdf

Implementation Brief: Considerations for Special Education
http://www.doe.mass.edu/edeval/ddm/SpecialEduEducators.pdf
### Appendix B: Recommended External Resources on Measuring Student Growth: Guidance from the Research and Measurement Communities

#### Table 3. Research- and Measurement-Based Resources for Pre-Test/Post-Test Approach

<table>
<thead>
<tr>
<th>Author</th>
<th>Resource</th>
<th>Topics Addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Katherine Castellano, University of California, Berkeley, and Andrew Ho, Harvard Graduate School of Education, on behalf of Council of Chief State School Officers (CCSSO)</td>
<td>A Practitioner’s Guide to Growth Models. (2012)</td>
<td>This guide provides recommendations for districts electing to use a simple gain score approach (e.g., comparing change in performance on pre- and post-test) (p. 19).</td>
</tr>
<tr>
<td>Council of Chief State School Officers (CCSSO), Accountability Systems and Reporting Collaborative (ASR)</td>
<td>Implementer’s Guide to Growth. (2007)</td>
<td>This resource describes various models for measuring growth for accountability at the educator, school, and state levels. Of particular interest are the strategies recommended for using gain scores to set annual growth targets for students or educators (p. 9) and use of confidence intervals to help account for the uncertainty in measuring growth (pp. 11, 33–41).</td>
</tr>
<tr>
<td>Edward Haertel, Stanford University</td>
<td>Student Growth Data for Productivity Indicator Systems. (2009). (paper presentation at ETS Exploratory Seminar)</td>
<td>Author provides guidance to address limitations of this approach.</td>
</tr>
<tr>
<td>Scott Marion, National Center for the Improvement of Educational Assessment, and Katie Buckley, Harvard University</td>
<td>Approaches and Considerations for Incorporating Student Performance Results from “Non-tested” Grades and Subjects into Educator Effectiveness Determinations. (2011)</td>
<td>Some discussion of pre-post designs; also cautions about possible introduction of perverse incentives for educators. Technical quality recommendations provided. Need for continuous improvement is reinforced. See especially pp. 6–10, 16–17, 25–30, 34.</td>
</tr>
</tbody>
</table>