

User Guide

Table of Contents

1 Introduction	1
What is ENGAGE Teacher Edition?	1
Frequently Asked Questions	3
2 Rating Student Behavior Using ENGAGE Teacher Edition	4
Overview of the Scales	4
Example of How to Use the Scales	4
Proper Use of the Rating Scales	6
3 Ordering and Administration	8
When and How is ENGAGE Teacher Edition Completed?	8
Scheduling Ratings	8
Assigning Raters	8
Registering Students	9
Uploading New Students	11
Rating Students	12
4 Reports and Interpretation	15
The Teacher Edition Report	15
Roster Reports	17
Generating Reports	17
5 Case Examples Using ENGAGE Teacher Edition	19
Low Performer: Gina Christensen	19
Average Achiever: Antonio Gómez	23
6 Using ENGAGE Teacher Edition with ENGAGE	26
Appendix A: Development, Reliability, and Validity of ENGAGE Teacher Edition	27
General Approach	27
Literature Review	27
Motivation	27
Social Engagement	27
Self-Regulation	28
Development of Behavioral Statements	28
Focus Groups	28
Retranslation	28
Follow-up Review	29
Development of Summary Statements	31
Reliability and Validity	31
Reliability	31
Validity	33
Internal Structure of ENGAGE Teacher Edition	33
Relationships between ENGAGE Teacher Edition and Measures of Other Constructs	36
Academic Achievement of Students at Each Behavioral Monitoring Scale Level	40
Incremental Validity	43
Compliance with Guidelines and Standards	44
References	45

List of Tables

Table 1.	ENGAGE Teacher Edition Scales and Definitions	2
Table 2.	Common Rating Errors	6
Table A1.	Mean Interrater Agreement and Reliability Estimates for ENGAGE Teacher Edition Scale Anchors	30
Table A2.	Interrater Reliability Estimates Using ICCs	32
Table A3.	Percentage of Interrater Agreement	33
Table A4.	Descriptive Statistics and Intercorrelations for ENGAGE Teacher Edition	34
Table A5.	Resulting Fit Indices from Confirmatory Factor Analyses	35
Table A6.	Correlations between ENGAGE Teacher Edition and EXPLORE Scores	37
Table A7.	Correlations between ENGAGE Teacher Edition and Academic Achievement	37
Table A8.	Correlations between ENGAGE Teacher Edition, Misconduct, and Absenteeism	38
Table A9.	Correlations between ENGAGE Teacher Edition and School-level Factors	39
Table A10.	Correlations between ENGAGE Teacher Edition and Demographic Information	40
Table A11.	Percent of Students Having Failed a Class by Behavior Level	41
Table A12.	Percent of Students Receiving Letter-grade Averages by Behavior Level	42
Table A13.	Linear Regression to Predict Term GPA	43

List of Figures

Figure 1.	Example of Scale Components for the Initiative Scale	5
Figure 2.	ENGAGE Main Menu	9
Figure 3.	Start the Registration Process	10
Figure 4.	Uploading Student-Teacher Relationships	11
Figure 5.	ENGAGE Main Menu—Students Need to Be Rated	12
Figure 6.	Click to Rate a Student	13
Figure 7.	Screenshot of ENGAGE Teacher Edition During Rating Process	14
Figure 8.	Example of Student Report	16
Figure 9.	Main Menu	17
Figure 10.	Defining Your Reports	18
Figure 11.	Gina’s Student Report	20
Figure 12.	Antonio’s Student Report	24
Figure 13.	Use of ENGAGE and ENGAGE Teacher Edition for Initial Assessment and Monitoring of Student Academic Behaviors	26
Figure A1.	Second-Order Structure of ENGAGE	36
Figure A2.	Proportion of Variance Predicted in Term GPA	44

Introduction

What is ENGAGE Teacher Edition?

ENGAGE Teacher Edition is designed to assess student academic behavior and progress. It can be used by teachers and other educators to rate student performance on eight scales of behavior that are critical to student academic success and persistence.

To enhance the accuracy of the ratings, each effectiveness level for each scale is “anchored” by behavioral statements that define exactly what it means to perform at that level. These help make ratings more comparable across raters, more accurate, and more reliable.

ENGAGE Teacher Edition includes eight different scales: *Initiative, Planning & Organizing, Sustained Effort, Performance, Communication, Working with Others, Managing Feelings*, and *Conduct* (See Table 1).

These scales are organized into three broad domains that have been shown to be predictive of academic performance and persistence.

- **Motivation** includes personal characteristics that help students to succeed academically by focusing and maintaining energies on goal-directed activities.
- **Social Engagement** includes interpersonal factors that influence students’ successful integration into their environment.
- **Self-Regulation** includes cognitive and affective processes used to monitor, regulate, and control behavior related to learning.

Table 1 lists the scales and their definitions and is organized by these three broad domains. Appendix A provides a complete description of the development of ENGAGE Teacher Edition.

Table 1
ENGAGE Teacher Edition Scales and Definitions

Domain	Scale Name	Definition
Motivation	Initiative	Seeking out opportunities for personal and academic development; volunteering to work on additional activities or projects. Showing enthusiasm for learning; starting new things without being told.
	Planning & Organizing	Setting appropriate goals and developing workable plans and strategies for achieving goals. Planning ahead, using time effectively, keeping own materials organized; setting appropriate priorities, and using back-up plans as needed.
	Sustained Effort	Working hard and staying focused. Putting a great deal of effort into schoolwork and learning activities and avoiding distractions. Following through with a task or project until it is completed.
	Performance	Meeting deadlines, being productive in completing schoolwork, and doing high-quality and accurate work.
Social Engagement	Communication	Conveying information clearly, effectively, and appropriately in speaking and writing. Getting his or her point across.
	Working with Others	Showing consideration for others, listening to others' points of view, and being helpful. Working constructively and cooperatively with others. Resolving conflicts effectively, and being open and accepting of others.
Self-Regulation	Managing Feelings	Not letting feelings get in the way of schoolwork. Not overreacting to stressful or difficult situations. Finding appropriate ways to express and manage negative feelings, such as anger, sadness, and embarrassment.
	Conduct	Following school rules, codes, and other behavior expectations. Being polite, engaging in appropriate behaviors, and avoiding negative behaviors, such as attention seeking, aggression, or dishonesty.

Frequently Asked Questions

1. Who should use ENGAGE Teacher Edition?
 - ENGAGE Teacher Edition is for educators who are familiar with the behavior of the student(s) to be rated. We recommend that educators have interacted with students on a regular basis (i.e., every day or several times per week) for approximately four to six weeks before they rate those students.
2. How do I get started using ENGAGE Teacher Edition?
 - Anyone can learn to use these scales, but it is important that raters participate in a brief rater training first. The training is designed to:
 - Help raters use these scales appropriately to make accurate ratings.
 - Create a standardized rating process.
 - Maintain fairness and quality of ratings.
 - Improve observational skills.
 - Minimize potential for bias or error.
3. When is the best time to rate students' academic behaviors?
 - Although there is no particular “best time” to rate students' behaviors, we recommend assessing each student two to three times per academic year in order to keep track of progress.
4. What are the benefits of using ENGAGE Teacher Edition?
 - Using these scales can help educators in a variety of ways, such as:
 - Providing accurate information about students' academic behaviors in the school environment.
 - Providing information that can help to identify appropriate interventions.
 - Tracking students' academic behavior progress over time.
 - Documenting academic behavior development and growth.
5. Why is it important to read the behavior statements?
 - It is important to read the behavior statements featured in ENGAGE Teacher Edition carefully, as they:
 - Describe exactly what it means to perform at each effectiveness level.
 - Help make the ratings more accurate and reliable by “anchoring” different levels of behavioral performance to behaviors displayed by students.
 - Allow educators and researchers to compare student behavioral performance across raters.

2

Rating Student Behavior Using ENGAGE Teacher Edition

Overview of the Scales

ENGAGE Teacher Edition consists of eight rating scales designed to assess behavioral scales important to student academic success and persistence. Each rating scale has eight rating levels, ranging from 8 (*highest effectiveness*) to 1 (*lowest effectiveness*). In order to help raters understand each scale and the effectiveness levels included, raters are provided with:

- 1) A broad **definition** of the behavioral scale that defines the scope of overall behavior included in that scale.
- 2) **General behaviors**, which are summary statements that describe four broad ranges of behavioral performance. These statements help raters narrow the rating to a portion of the scale (e.g., high, medium-high, medium-low, low).
- 3) More specific **behavioral examples** illustrating typical student behaviors, to help raters select a specific rating, from 8 (*High*) to 1 (*Low*).

Example of How to Use the Scales

To illustrate, Figure 1 shows the *Initiative* scale.

A rater who is about to rate one of her students, Joe Long, on *Initiative* would: (1) read the definition carefully and read through *both* the (2) general behaviors and (3) specific examples to get a sense of the entire set of behaviors included in the scale. Once the rater has familiarized herself with these, she is ready to make her ratings.

The rater thinks that, in terms of *Initiative*, Joe's behavior is most often like those described in the lowest general behavior to the left of this rating scale (tied to a rating of 1 or 2), but now and then he does some of the things described for the next level up. This suggests that Joe is performing somewhere around a 2. To select a specific rating, the rater should next look at the specific examples that illustrate level 2 and its surrounding levels (3 and 1). If the rater has seen Joe perform the specific or similar behaviors at this level, Joe would receive a rating of 2 on *Initiative*.

Figure 1
Example of Scale Components for the Initiative Scale

Joe Long's level of Initiative			Scale Name
Definition: Seeking out opportunities for personal and academic development. Volunteering to work on additional activities or projects. Showing enthusiasm for learning. Starting new things without being told.			Definition
GENERAL BEHAVIOR	RATING	SPECIFIC EXAMPLES	General Behaviors
A student in this range . . .		and also . . .	
Demonstrates a great deal of initiative. <ul style="list-style-type: none"> Strives to learn new material and to build own knowledge and skills. Seeks out learning opportunities both in and outside of school, going beyond what is required. 	8	<ul style="list-style-type: none"> Seeks out activities that challenge his or her knowledge and skills. Asks questions to gain a deeper understanding of presented material. 	Specific Examples
	7	<ul style="list-style-type: none"> Is willing to lead group work. Often requests to try new things. 	
Demonstrates a fair amount of initiative. <ul style="list-style-type: none"> Is usually willing to experience new things and is actively engaged in learning. Sometimes seeks out additional learning opportunities. 	6	<ul style="list-style-type: none"> Occasionally requests to try new things. Occasionally does optional coursework. 	Rating
	5	<ul style="list-style-type: none"> Occasionally expresses interest in doing additional work. Tries new things when encouraged to do so. 	
Demonstrates little initiative. <ul style="list-style-type: none"> Generally works only on assignments that are required. Rarely shows interest in experiencing new things or learning activities. 	4	<ul style="list-style-type: none"> Rarely does optional coursework. 	Rating
	3	<ul style="list-style-type: none"> Does not start projects and assignments unless instructed directly. Does not try new things. 	
Demonstrates virtually no initiative. <ul style="list-style-type: none"> Consistently shows a lack of interest in learning. Refuses to try new things or activities. 	2	<ul style="list-style-type: none"> Shows virtually no interest in learning. Avoids schoolwork. 	Rating
	1	<ul style="list-style-type: none"> Refuses to try new things even when asked. Refuses to begin schoolwork without multiple teacher prompts. 	

Proper Use of the Rating Scales

The most important part of the entire rating task is for you to read ALL of the general behavior and specific examples very carefully so that you have a firm fix on the kinds of behaviors that define different effectiveness levels within each scale.

Additional Rating Tips:

- Focus on rating one student at a time. Don't compare students to each other when making ratings, rather, compare each student's behaviors with the general behavior and specific examples.
- Make sure your ratings reflect each student's strengths and needs. Even if a student is excellent (or poor) on one behavioral scale, it is likely that his or her performance is different on another.
- Do not let factors other than student behaviors affect your ratings, such as appearance, gender, ethnicity, or how much you like or dislike the student.
- Use the entire range of the rating scale as appropriate. Do not hesitate to assign high or low ratings when students' behaviors match the statements at these levels.
- You will likely find it useful to keep notes or anecdotal records about relevant student behaviors so you can use them when you make future ratings.
- Familiarize yourself with the Common Rating Errors below in Table 2.

Table 2
Common Rating Errors

Rating Errors	Definition	Examples
Attractiveness	The tendency for people to assume that individuals who are physically attractive also are superior performers.	Mr. Denburg rates his attractive students higher across the board on the behavioral scales despite their varied behavioral performance.
Central Tendency	The inclination to rate a student in the middle even when his or her performance clearly warrants a substantially higher (or lower) rating.	Ms. Littlestone rates all her students as a 4 or 5 despite significant differences in her students' actual behavioral performance.
Contrast	Evaluating a student in relation to other students rather than in relation to the statements in the behavioral rating scale.	Mr. Ramirez has just rated Devin, who demonstrates outstanding <i>Conduct</i> , as an 8. Next, he rates Travis, who demonstrates good conduct. However, he rates Travis a 4, which is lower than his actual behavior performance.
Frame of Reference	The tendency to compare a student's behavioral performance to one's own personal standards.	Ms. Abbas, who is a highly organized individual, rates David lower on <i>Planning & Organization</i> than his performance would suggest because she believes that anything less than perfectly organized is not good enough.

Table 2
(Continued)

Rating Errors	Definition	Examples
Halo (or Horn)	Inappropriate generalizations from one aspect of a student's performance to all areas of that student's performance, such that a student's behavior is rated as "all good" or "all bad."	Mr. Burke rates Jordan a 7 across all behavioral scales because Jordan demonstrated a 7 on <i>Communication</i> , which Mr. Burke considers an important skill.
Leniency	Giving everyone high ratings regardless of actual behavioral performance.	Ms. Garcia gives all of her students a rating of 7 or 8, which is higher than the actual behavior performance of many of her students.
Primacy	The inclination to provide a higher rating on the first evaluation or scale than for a later evaluation.	Mr. Wei usually gives his students higher scores on the <i>Initiative</i> scale, which is presented at the beginning, than on the <i>Conduct</i> scale, which is presented at the end.
Recency	The tendency for minor events that have happened recently to have more influence on the rating than major events that are less recent.	Ms. Williams did not keep any records of her students' behaviors during the semester. As she begins to rate her students at the end of the semester, she realizes that the only behaviors she recalls are from the past four weeks.
Severity	Giving everyone low ratings regardless of actual behavioral performance.	Mr. Dayton gives all of his students a rating of 1 or 2, which is lower than the actual behavior performance of many of his students.
Similar-to-me	The tendency of individuals to rate people who resemble themselves higher than they rate others.	Ms. Lopez, a talented math teacher who grew up in a single-parent household, rates students with single parents higher than she rates students with two parents.
Stereotyping	The tendency to generalize across groups with a certain characteristic (e.g., gender, race, country of origin) and ignore individual differences in the actual behavior being rated.	Mr. Melcher rates Lucy a 2 on the <i>Managing Feelings</i> scale because he believes a female student has a lower ability to manage her feelings.

Note. Rating errors definitions adapted from the following: Borman, 1979, 1991; Bolster & Springbett, 1961; Landy & Farr, 1980; Maurer & Alexander, 1991; Thorndike, 1920.

3

Ordering and Administration

ENGAGE can be ordered online at www.act.org/engage/order.html.

When and How is ENGAGE Teacher Edition Completed?

School personnel responsible for ENGAGE Teacher Edition ratings should familiarize themselves with this User's Guide before scheduling their ratings and/or assigning other raters.

Scheduling Ratings

Because it is important for educators to be familiar with the behavior of students they rate, we recommend that educators have interacted with students on a regular basis (i.e., once a day or several times per week) for approximately four to six weeks before they are assigned to rate those students. In most schools, this means ratings should be conducted no earlier than six to eight weeks after the start of the school year.

Some schools will choose to conduct ratings multiple (usually two) times over the course of a school year (e.g., mid-October and again in May). This allows educators to use the rating information to select appropriate interventions for students earlier in the year, and then to track and document development and growth in student behavior later in the year.

Assigning Raters

The school's principal or a designated school coordinator should assign a rater to each student being rated. Raters should be classroom teachers, coaches, counselors, or other educators who have interacted with each assigned student on a regular basis for a **minimum of four to six weeks**.

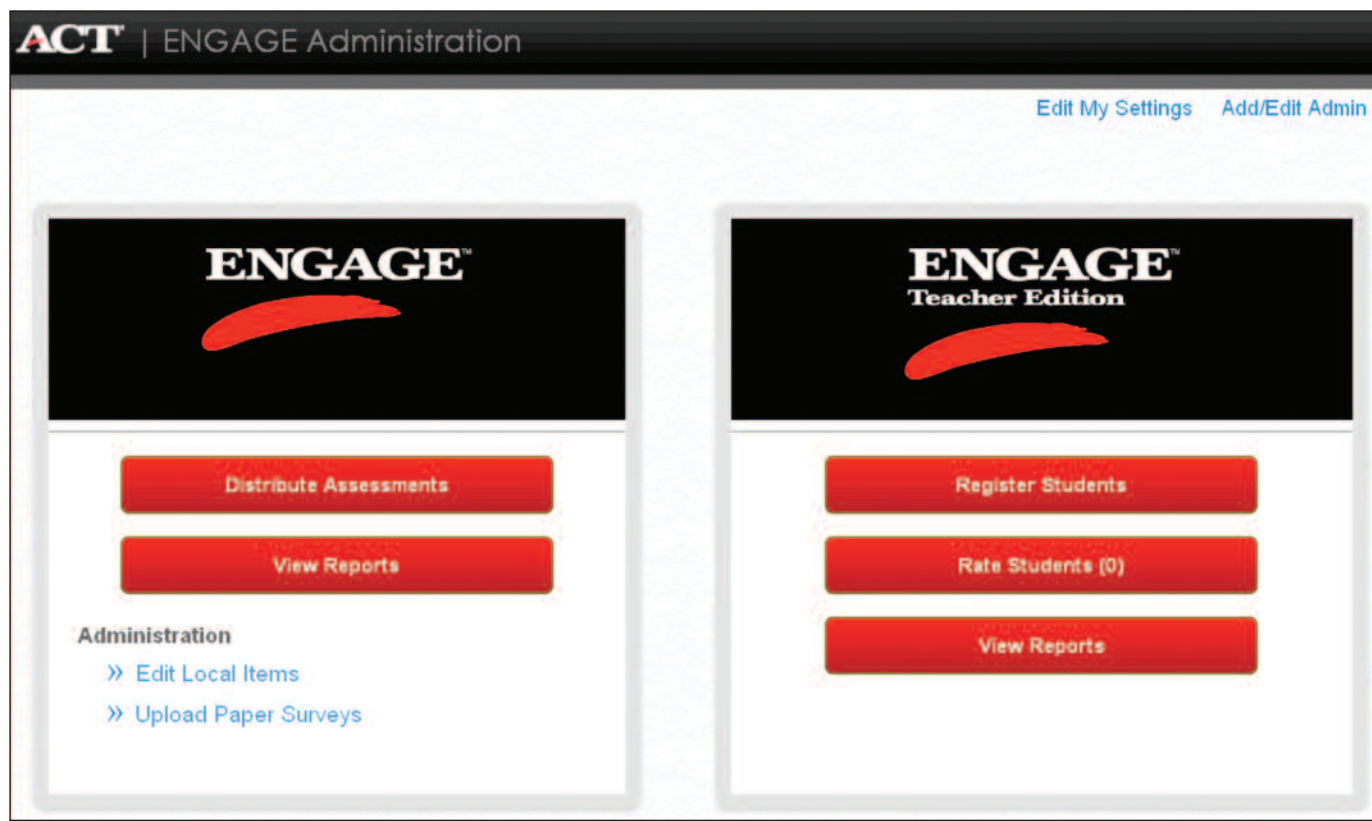
The school coordinator will:

- Ensure that all raters are familiar with the students they are assigned to rate and have had the opportunity to observe the assigned students' behaviors multiple times each week.
- Ensure that all raters are familiar with ALL of the general behaviors and specific examples. This training will standardize the rating process by familiarizing raters with the scale definitions, how to use the rating scales, and tips for raters. The training is delivered via a brief, Web-based training presentation found on ACT's website (www.act.org/engage/downloads.html).

Registering Students

It is necessary to register students to a teacher before they can be rated. To do this, one must select “Register Students” on the main ENGAGE Administration website (See Figure 2 below). Note how, in the figure below, a “(0)” appears after the option “Rate Students.” This number will vary according to how many students remain to be rated.

Figure 2
ENGAGE Main Menu



Registering students (cont.)

First, select the institution for which you wish to register students. Often, the institution you wish to select will appear as an option below “Select the Institution.” If you rate students from many institutions, however, it may be necessary to use the search feature provided in order to narrow the available options. In the figure below, the only option available to the rater is SAMPLE MIDDLE SCHOOL; the rater has selected this option as the school for which he or she would like to register additional students.

Second, select which version of ENGAGE Teacher Edition you will be using. In the figure below, the rater has selected “6–9 Teacher Edition” because he or she is registering students for a middle school.

Figure 3
Start the Registration Process

The screenshot displays the ACT ENGAGE Student Registration web interface. At the top, the header reads "ACT | ENGAGE Student Registration". On the right side of the header, there are two links: "Edit My Settings" and "Add/Edit Admin". Below the header, the main content area has the instruction "Start the registration process." followed by a large, light-gray rectangular box containing the registration steps. Step 1, "1. Select the institution:", includes a text input field and a red "Search" button. Below this, the option "SAMPLE MIDDLE SCHOOL" is listed with a selected radio button. Step 2, "2. Select a product:", includes two radio button options: "6-9 Teacher Edition" (which is selected) and "10-12 Teacher Edition". At the bottom of the gray box, there are two red buttons: "Upload Students" and "Cancel".

Uploading New Students

Students to be evaluated will need to be uploaded via a CSV (comma separated values) file. To create a .csv file, use Excel®, select Save As, and select CSV.

NOTE: an important step in this process is relating each student to a specific rater. If a student is not assigned to a rater, then that student cannot be rated.

For your convenience, an easily modifiable template CSV file is downloadable from the “Upload student-teacher relationships” page (See Figure 4 below). You may find it easiest to simply modify this example file. To further assist you in this regard, a list of definitions for the upload template is also provided. It defines each column in the template CSV.

After you have created/modified a CSV file, upload that file to relate students with the teachers who will rate them. Any teacher who is not already in the system will be added as the file imports, and teachers will be e-mailed a list of students to rate. Teachers or raters will then need to login to the ENGAGE online system to rate their students.

Whoever uploads a file successfully will be notified with a receipt of the completed upload.

Figure 4
Uploading Student-Teacher Relationships

The screenshot shows the 'Upload student-teacher relationships' page in the ENGAGE Student Registration system. The page has a black header with the ACT logo and the text 'ENGAGE Student Registration'. In the top right corner, there are links for 'Edit My Settings' and 'Add/Edit Admin'. The main heading is 'Upload student-teacher relationships'. Below this, there is instructional text: 'The roster of students to be evaluated will need to be uploaded via a CSV file. To create a .csv file use Excel, then Save As, and select Save as type: CSV.' and a note: 'Your institution will be charged \$5 per evaluated student, per evaluation period.' A grey box titled 'Teacher Edition 6-9 | SAMPLE MIDDLE SCHOOL' contains a bulleted list of instructions: 'Load a file to relate students with teachers that will rate them.', 'Any teacher who is not already in the system will be added as the file imports.', 'Teachers will be emailed a list of students to rate.', 'Teachers need to login to ENGAGE to rate their students.', and 'You will be notified with a receipt of the load.' Below the list are two links: '» Download Sample Upload File' and '» Get definitions for upload templates'. A file upload section has the text 'Select and upload the file (the file must end in .csv):' followed by a text input field and a 'Browse...' button. At the bottom of the grey box are three red buttons: 'Upload Students', 'Cancel', and 'Main Menu'.

ACT | ENGAGE Student Registration

Edit My Settings Add/Edit Admin

Upload student-teacher relationships

The roster of students to be evaluated will need to be uploaded via a CSV file. To create a .csv file use Excel, then Save As, and select Save as type: CSV.

Your institution will be charged \$5 per evaluated student, per evaluation period.

Teacher Edition 6-9 | SAMPLE MIDDLE SCHOOL

- Load a file to relate students with teachers that will rate them.
- Any teacher who is not already in the system will be added as the file imports.
- Teachers will be emailed a list of students to rate.
- Teachers need to login to ENGAGE to rate their students.
- You will be notified with a receipt of the load.

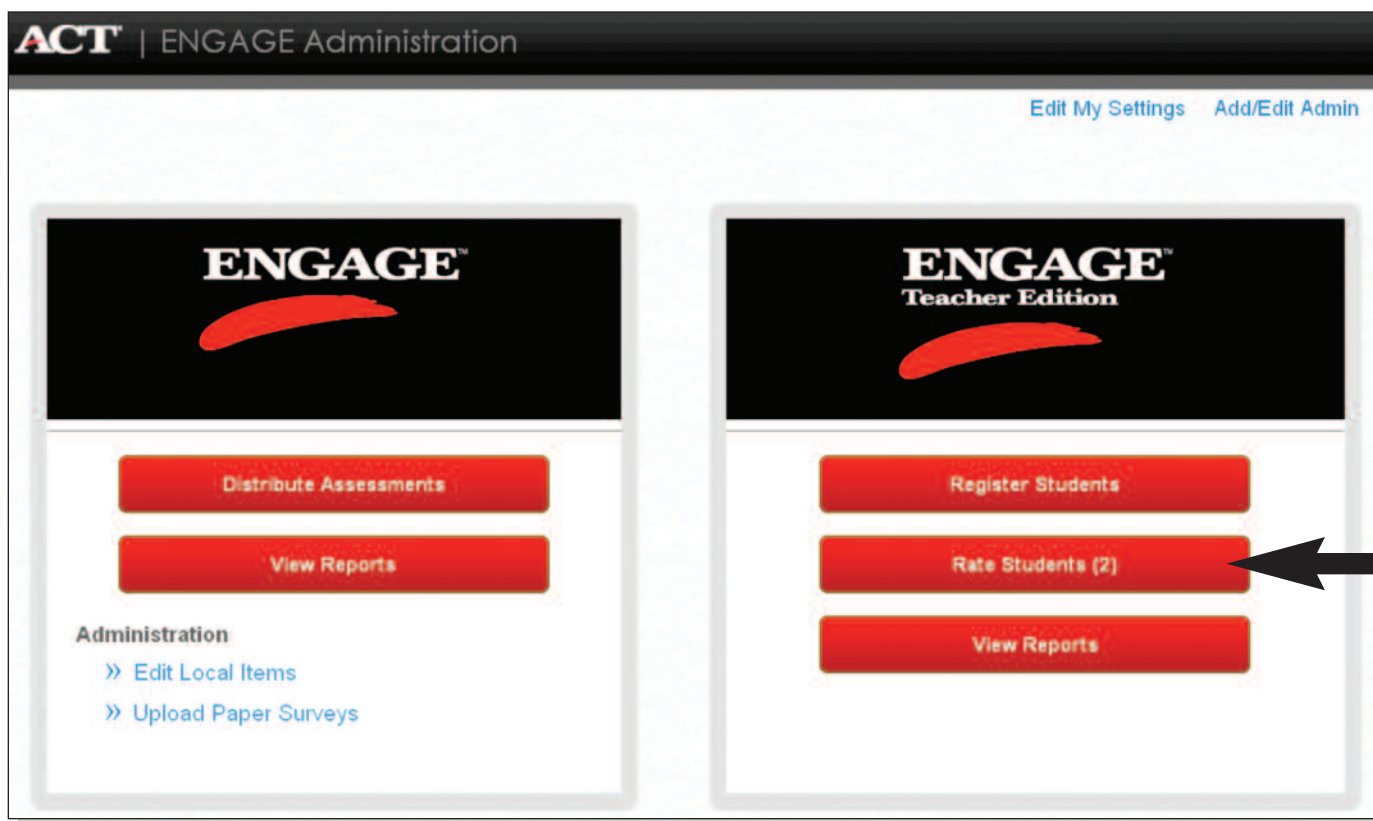
» Download Sample Upload File

» Get definitions for upload templates

Select and upload the file (the file must end in .csv):

When an upload is successful, the raters to which students are assigned will see a number behind the “Rate Students” option on the main ENGAGE Administration website (See Figure 5 below). In this example, the rater has two new students he or she needs to rate, represented by a “(2)”.

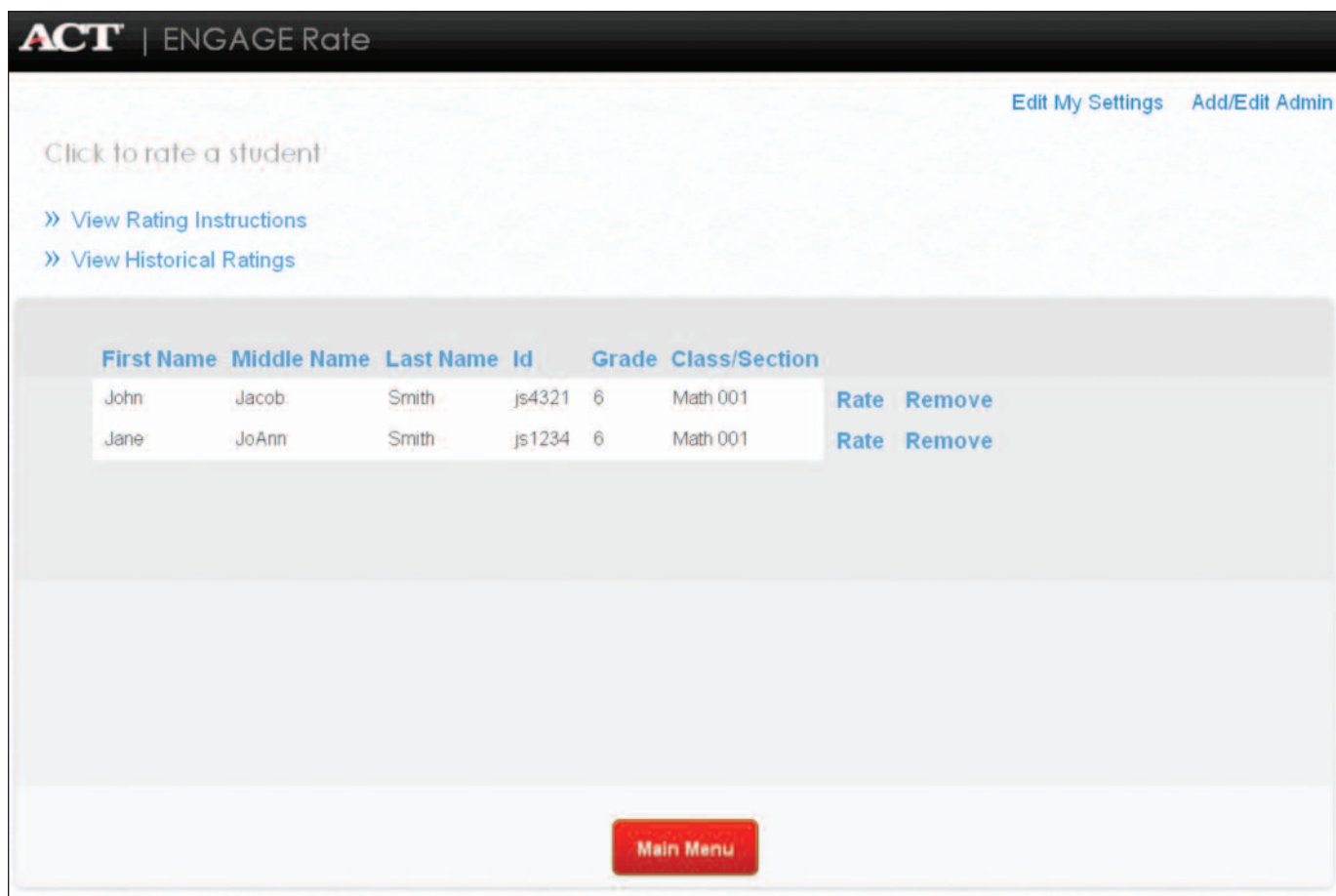
Figure 5
ENGAGE Main Menu—Students Need to Be Rated



Rating Students

Select “Rate Students” on main ENGAGE Administration website (See Figure 5 above).

Figure 6
Click to Rate a Student



The screenshot shows the ACT ENGAGE Rate interface. At the top, there is a header with the ACT logo and the text "ENGAGE Rate". On the right side of the header, there are links for "Edit My Settings" and "Add/Edit Admin". Below the header, there is a main content area with a heading "Click to rate a student". Under this heading, there are two links: "» View Rating Instructions" and "» View Historical Ratings". Below these links is a table with columns: "First Name", "Middle Name", "Last Name", "Id", "Grade", "Class/Section", "Rate", and "Remove". The table contains two rows of student data. At the bottom of the interface, there is a red button labeled "Main Menu".

First Name	Middle Name	Last Name	Id	Grade	Class/Section	Rate	Remove
John	Jacob	Smith	js4321	6	Math 001	Rate	Remove
Jane	JoAnn	Smith	js1234	6	Math 001	Rate	Remove

If you need to review the rating instructions, a link appears on this page (“View Rating Instructions”). If historical ratings exist for the student the rater *may* be able to view those ratings by clicking “View Historical Ratings”—but only if the rater has the proper privileges to do so.

Students on this page may be sorted by any column (e.g., first name, grade, etc.). Two options exist for each student to be rated: Rate or Remove. If you choose to remove a student from the list, you will receive a prompt asking if you are sure you want to remove this student from the list of students.

NOTE: Once the student has been removed, you cannot undo this operation.

If you choose to rate the student, the ENGAGE Teacher Edition Survey will appear.

Directions:

- Make your ratings by moving your cursor over the box containing the rating you have chosen.
- When your cursor is over the rating number, it will become highlighted.
- Click your mouse and the box will stay highlighted, indicating you have selected the rating number in that box.
- You will be able to rate one student in each of the eight scales (*Initiative* through *Conduct*) before moving on to the next student.
- If you'd like more information on ENGAGE or to download ENGAGE resources, you can go to **<http://www.act.org/engage/downloads.html>**
- If you need to quit in the middle of rating a student, click the “Save for Later” button in the lower left of the rating screen.
- When you're finished rating a student in all eight scales, click the “I Am Done” button in the lower right of the rating screen (See Figure 7 below). Note that this will remove the student from the list of students to be rated.
- Complete your ratings until you have no students remaining on the “Click to Rate a Student” screen (Figure 6).

Figure 7
Screenshot of ENGAGE Teacher Edition During Rating Process

The screenshot shows the ENGAGE Teacher Edition interface. At the top, the ENGAGE logo is visible. Below it, a row of icons represents different scales. The main content area is titled "John Smith's level of Managing Feelings". Below the title is a definition: "Definition: Not letting feelings get in the way of schoolwork. Not overreacting to stressful or difficult situations. Finding appropriate ways to express and manage negative feelings, such as anger, sadness, and embarrassment." Below the definition is a table with three columns: GENERAL BEHAVIORS, RATING, and SPECIFIC EXAMPLES. The table has five rows, each representing a rating level from 3 to 8. The rating level 8 is highlighted in blue. At the bottom of the screen, there are three buttons: "Save for Later", "« Back", and "Next »". In the bottom right corner, there is a button labeled "I Am Done" with a large black arrow pointing to it.

GENERAL BEHAVIORS	RATING	SPECIFIC EXAMPLES
A student in this range . . .		and also . . .
Works effectively even when under stress or when experiencing negative emotions. • Is able to set aside negative feelings and remain calm in order to complete tasks or assignments. • Shows resilience in stressful situations and uses effective coping strategies.	8	• Shows remarkable resilience when under stress. • Handles stressful situations with calm and poise.
	7	• Is usually able to remain calm even in stressful situations. • Uses helpful coping strategies to deal with stress (e.g., journals emotions, exercise).
Is usually able to handle stress effectively. • Uses some effective coping skills. • May occasionally need assistance to manage his or her stress and/or emotions.	6	• Doesn't let negative comments from peers or teachers bother him or her. • Uses humor to cope with stress.
	5	• Usually controls temper when criticized. • Can calm down given adequate time.
Has difficulty managing his or her emotions. • May become visibly upset, complain, or simply withdraw when encountering problems or stress. • Sometimes lets stress or emotions interfere with schoolwork and activities.	4	• Leaves class when feeling emotional. • Is visibly nervous when asked to complete an assignment.
	3	• Often cries when he or she is stressed. • Is visibly upset when he or she does not get his or her way.

Save for Later « Back Next » I Am Done

Reports and Interpretation

ACT provides an ENGAGE Teacher Edition Report for each rated student. These reports contain several important pieces of information about the student, the rating selected by the assigned rater for each scale, and the student's progress over time (if applicable). ENGAGE Teacher Edition has two rating periods per year—January 1st through June 30th and July 1st through December 31st. If student ratings are submitted over different rating periods, trend data illustrating the ratings over time are shown graphically. If multiple ratings for a student are submitted during the same rating period, the reported student scores (in each scale) are an average of all ratings during that period. Schools are also provided with Roster Reports, which contain student and rater identification, background information, and scores on all eight scales for all rated students.

The Teacher Edition Report

There are two main components to this report.

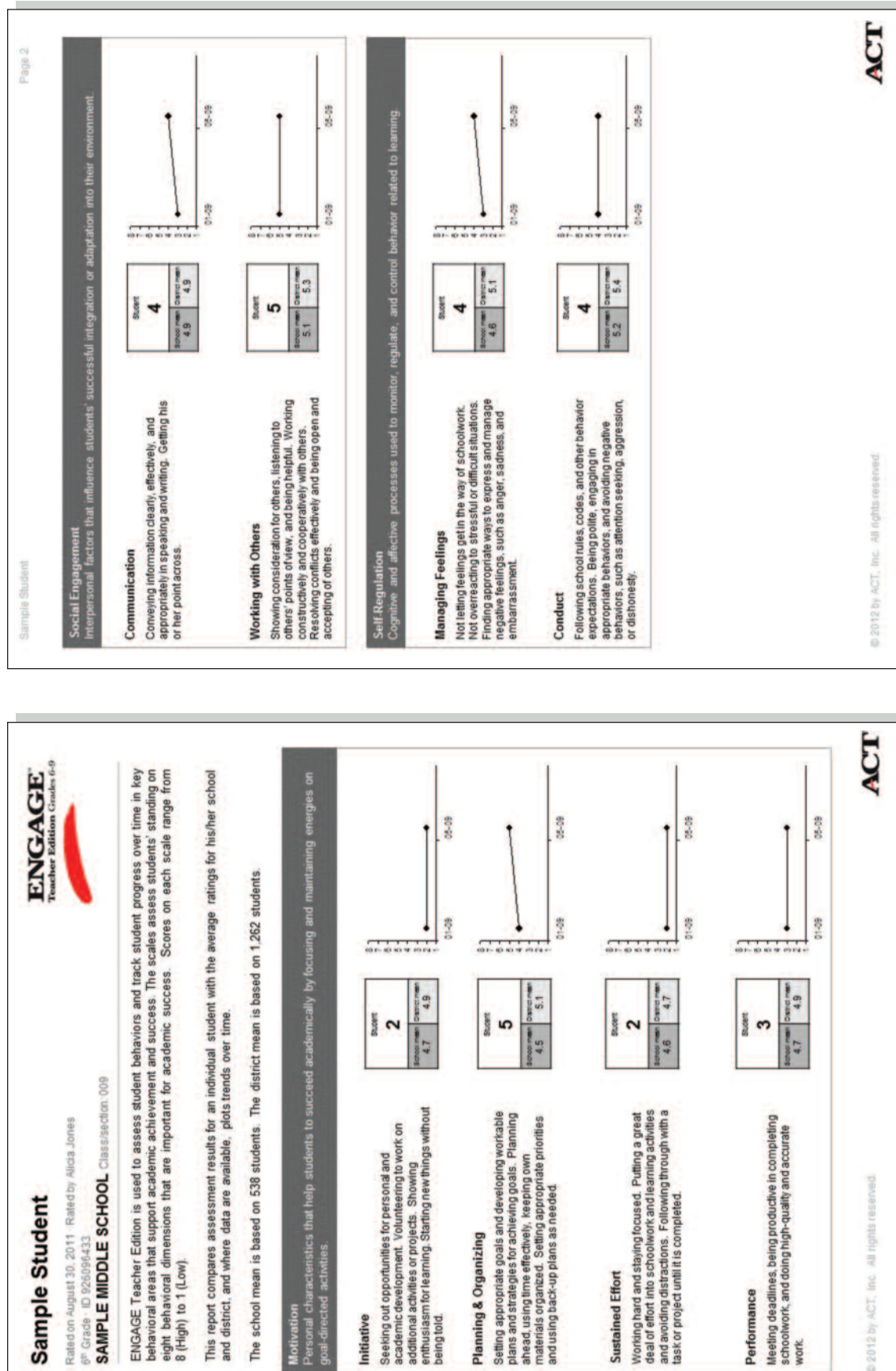
1. **Identification Data.** At the top of the report, information is provided, including the student's name and ID number, the rater's name, and the date of rating. The name of the school is also listed. This information is followed by a brief description of the scales and what the scores mean.
2. **Student Scores.** These results are organized by the three broad domains—Motivation, Social Engagement, and Self-Regulation. Ratings are presented for each scale (e.g., *Initiative*, *Communication*, and *Conduct*). The definition for each scale is also provided.

As shown in Figure 8, this score report provides three score components for each scale.

- Next to each scale definition, the score (or average score) for the individual student is listed.
- Underneath the individual student score, scores representing the school and district averages for the scale also are listed.
- To the right of the student scores, trend data (if applicable) is shown as a line graph. Each point on the graph represents the score of the individual at a particular point in time (i.e., on the date of the rating). This graph is a visual representation of the progress, or lack thereof, a student has made between ratings.

These reports can be used to identify students' strengths and areas needing improvement. Teachers and administrators can use these reports as a guide for discussing interventions, designing programs to develop student behaviors, and addressing student performance.

Figure 8
Example of the Student Report



Roster Reports

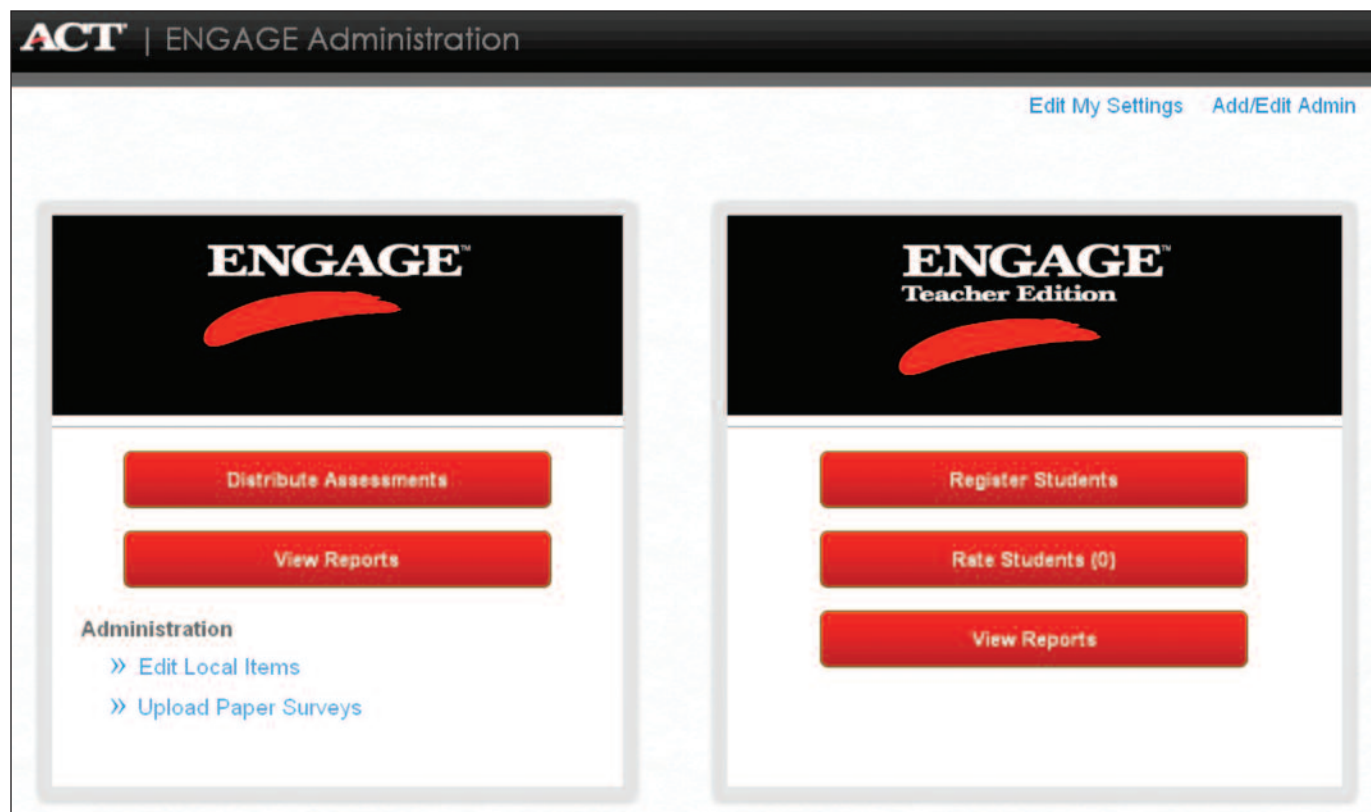
Schools may also generate two types of Roster Reports: a Student Roster Report and a Teacher Roster Report. The Student Roster Report lists all student names, along with the number of raters, average rating in each scale during that rating period, and other associated information. The Teacher Roster Report lists all raters, the students they rated, along with their ratings of those students in each scale, and other associated information.

Both types of Roster Reports include scores for all students on all scales, so school personnel can easily compare scores and identify which students are in need of the most immediate interventions (Student Roster) or identify which teachers have rated which students and see any patterns in their ratings (Teacher Roster).

Generating Reports

All ENGAGE Teacher Edition reports can be generated by first selecting “View Reports” on the main ENGAGE Teacher Edition website (Figure 9, below).

Figure 9
Main Menu



On the left side of the screen there is a drop-down box under “Select a Report to Generate” that will list the three possible report options for ENGAGE Teacher Edition.

- Choose “ENGAGE Teacher Edition Report(s)” to generate a Student Report or series of student reports.
- Choose “TE Teacher Roster Report” to generate a Teacher Roster Report.
- Choose “TE Student Roster Report” to generate a Student Roster Report.

After clicking “Continue,” you will be taken to an options screen where you can define the school, time period, grade level, and other fields you would like to be used for generating your report. Additionally, you can check the box to have all reports included in one file (See Figure 10 below).

Figure 10
Defining Your Reports

Enter criteria to include in report(s). Blank equals all.

- For batch reports (of multiple students, rather than one student), leave the name fields blank.
- For reports on an individual student, include as much identifying information as possible in the fields below

School Search (Enter name and click Search; use * as a wildcard match)

Search Results Displayed Here

First Name:

Last Name:

Student ID:

Section/Class:

Grade Level:

Section/Class:

Start Date:

End Date:

☐ Check to include all reports in a single file

Generate **Reset**

After defining what you would like in your report, click “Generate.” You will be notified by e-mail when your report is ready, at which time it can be accessed under the “View Reports” page, previously discussed in Figure 9.

Case Examples Using ENGAGE Teacher Edition

The ENGAGE Teacher Edition can be used to identify students' strengths and weaknesses. They can be used to intervene with individual students who may be experiencing difficulties in one or more areas and may not be achieving academic success. The scales also can identify areas in which students are already performing adequately but could benefit from further help in order to become higher achievers. This chapter features two student profiles to illustrate use of the scales with students.

Low Performer: Gina Christensen

Gina is a sixth grader at Sample Middle School. Her teacher, Mr. Anderson, used the ENGAGE Teacher Edition to rate her performance six weeks after the start of the semester and again two weeks before the end of the semester. Many of the ratings Gina was assigned showed that she was not performing at a level consistent with academic success. Although she was given moderate ratings (ratings of 4 or 5) on a few scales, she was assigned low ratings (ratings of 1 to 3) on other scales. (Refer to Gina's Student Report in Figure 11.)

Despite her struggles, Gina is intelligent and shows potential for learning. Furthermore, she has shown progress in some behavioral scales, such as *Planning & Organizing* or *Communication*, between the two rating dates. This suggests that, with some assistance, she is capable of improvement.

In order to create a tailored intervention designed to improve upon the weaknesses identified, Mr. Anderson used the ratings to create a development plan for Gina.

Capitalize on Strengths and Develop Skills. There are some scales where Gina received moderate ratings, suggesting that these do not need as much attention and may not require intervention at this time.

- Gina was rated a 5 on *Working with Others*. She works cooperatively with her classmates on group activities. She acts appropriately toward others and is friendly and pleasant to be around.
- Gina was rated a 4 on *Communication*. She speaks clearly and usually is able to get her point across in class discussions. She occasionally needs some help when writing reports and assignments but generally is able to use language properly and organize her writing.
- *Conduct* is another scale of behavior where Gina had a rating of 4. She generally follows rules and procedures and behaves as expected in the classroom. She sometimes arrives in class after the bell rings and occasionally claims to have left her homework at home when Mr. Anderson is fairly certain she has not completed it. However, she generally acts appropriately regarding school rules. She could further develop her skills in this area by making sure she arrives on time and is prepared for class.

Gina Christensen

Rated on May 30, 2009 - Rated by David Anderson
6th Grade - ID: 926098433

SAMPLE MIDDLE SCHOOL Class/section: 009

ENGAGE™
Teacher Edition Grades 6-9



ENGAGE Teacher Edition is used to assess student behaviors and track student progress over time in key behavioral areas that support academic achievement and success. The scales assess students' standing on eight behavioral dimensions that are important for academic success. Scores on each scale range from 8 (High) to 1 (Low).

This report compares assessment results for an individual student with the average ratings for his/her school and district, and where data are available, plots trends over time.

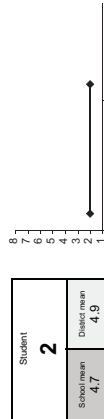
The school mean is based on 538 students. The district mean is based on 1,262 students.

Motivation

Personal characteristics that help students to succeed academically by focusing and maintaining energies on goal-directed activities.

Initiative

Seeking out opportunities for personal and academic development. Volunteering to work on additional activities or projects. Showing enthusiasm for learning. Starting new things without being told.



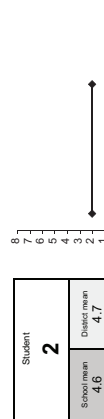
Planning & Organizing

Setting appropriate goals and developing workable plans and strategies for achieving goals. Planning ahead, using time effectively, keeping own materials organized. Setting appropriate priorities and using back-up plans as needed.



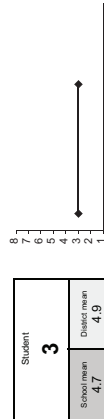
Sustained Effort

Working hard and staying focused. Putting a great deal of effort into schoolwork and learning activities and avoiding distractions. Following through with a task or project until it is completed.



Performance

Meeting deadlines, being productive in completing schoolwork, and doing high-quality and accurate work.



© 2012 by ACT, Inc. All rights reserved.

ACT

Gina Christensen

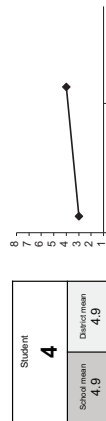
Page 2

Social Engagement

Interpersonal factors that influence students' successful integration or adaptation into their environment.

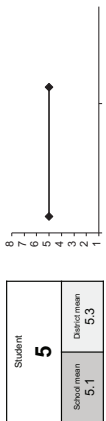
Communication

Conveying information clearly, effectively, and appropriately in speaking and writing. Getting his or her point across.



Working with Others

Showing consideration for others, listening to others' points of view, and being helpful. Working constructively and cooperatively with others. Resolving conflicts effectively and being open and accepting of others.



Self-Regulation

Cognitive and affective processes used to monitor, regulate, and control behavior related to learning.

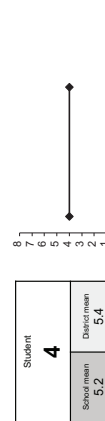
Managing Feelings

Not letting feelings get in the way of schoolwork. Not overreacting to stressful or difficult situations. Finding appropriate ways to express and manage negative feelings, such as anger, sadness, and embarrassment.



Conduct

Following school rules, codes, and other behavior expectations. Being polite, engaging in appropriate behaviors, and avoiding negative behaviors, such as attention seeking, aggression, or dishonesty.



© 2012 by ACT, Inc. All rights reserved.

ACT

Recommended for Intervention. Gina needs improvement in several behavioral scales. Mr. Anderson can best focus improvement on a few scales at a time and should prioritize which intervention(s) are most timely.

- *Sustained Effort* is a weakness. Gina received a rating of 2 for this scale. Specifically, she does not seem engaged in her schoolwork and is easily put off task. She gets distracted by her classmates and gives up easily when she is working on a challenging task. She does not seek out help or other ways of doing things; she just quits when frustrated.
- *Initiative* also is a struggle, with a score of 2. She shows no interest in personal or academic development. She does the minimum amount of work required in order to stay out of trouble but does not take any further interest in learning. She needs to be told to start each new activity, and if she completes an activity before the allotted time is up, she does not express interest in working on the next assignment. Indeed, she tends not to start the next activity without prompting from Mr. Anderson.
- Gina was rated a 3 on *Performance*. She frequently neglects to complete assignments, and those she does complete are often of poor quality. She frequently has to redo or correct her work and resubmit it.

Plans for Improvement. Based on these ratings, Mr. Anderson wanted to intervene with Gina and determine a way to improve her academic performance, especially in her weak areas. Mr. Anderson and Gina discussed several actions that can be taken in order to improve.

1. Since Gina struggles to be on time for class and shows potential for development on *Conduct*, Mr. Anderson told her he would hold her accountable for each day she was tardy to class. Both Gina and Mr. Anderson plan to keep a record of whether she arrives on time each day, and if she can be on time every day for a month, she will receive a reward. If she is late more than twice that month, Mr. Anderson will talk to her parents and/or begin further disciplinary actions. Gina is motivated by rewards and does not like being in trouble, so this is a good incentive for her to become more conscious of her tardiness.
2. Gina shows a lack of *Initiative* regarding her schoolwork. Mr. Anderson thought that part of the problem was that Gina was not working on anything that she found personally interesting. An upcoming book report assignment seemed like a good place to start. Mr. Anderson worked with Gina to find a book that she was truly interested in reading and checked in with her frequently to make sure she was enjoying the book and completing her reading. He also had Gina choose how she wanted to present the book report. Since Gina was reading something interesting and could use her creativity to present the report, she took more initiative to work on the project.

3. Gina's low *Performance* score was also addressed. Mr. Anderson met with her parents to enlist their help in this area. Mr. Anderson agreed to send a note home every day, which Gina's parents agreed to sign and return, indicating which assignments are due. Her parents then will make sure she completes her work and will check her work for accuracy. If her work is sloppy or incomplete, her parents will hold her accountable for redoing the homework prior to class the next day. Mr. Anderson will then check the homework again and if it is still incomplete or inaccurate, he will have Gina redo it. If her work does not improve, Mr. Anderson and Gina's parents will have another conference and develop a more intensive intervention.

Gina will be rated again in three months, and progress will be recorded. If she is still struggling in these areas, further interventions may be implemented and disciplinary actions may be taken. Mr. Anderson will continue to take notes on her actions and progress and will use his notes to rate her on the scales at the end of the three-month period.

Average Achiever: Antonio Gómez

Antonio, an eighth-grade student at Sample Middle School, was rated by his teacher, Ms. McKinney. Most of the ratings that Ms. McKinney gave Antonio indicate that he does not have serious behavioral issues. His ratings range from 4 to 7 across all scales, so while he has some clear strengths, he also has room for improvement. He received a score of 7 on *Conduct* and a 6 on *Planning & Organizing* and *Working with Others*, indicating that he is a strong performer in those behavioral scales. (See Antonio's Student Report in Figure 12.)

To determine how to best help Antonio leverage his strengths and work on his relative weaknesses, Ms. McKinney used his Student Report to create a development plan.

Capitalize on Strengths and Develop Skills. There are some behavioral scales on which Antonio scored at a moderate to high level, suggesting that these scales are less likely to require intervention.

- He was rated a 7 on *Conduct*. He responds well to structure and follows rules and other behavior expectations.
- He was rated a 6 on *Planning & Organizing*. His tendency to properly prioritize schoolwork, manage time well, and organize his supplies and assignments appears to be working well for him.
- He also was rated a 6 on *Working with Others*. Antonio is well-liked by his classmates and gets along with nearly everyone. He works well in a group and is helpful toward other students who need help.

Recommended for Intervention. Although Antonio did not score particularly low on any scale, he scored a 4 on three scales. This suggests that there is room for improvement. Although no drastic intervention is needed, Ms. McKinney is interested in developing Antonio to become a more successful student all around. She is focusing on some areas in particular:

- He received a score of 4 on *Initiative*. Although Antonio understands the work that needs to be done and does it, he does not go above and beyond expectations. He often needs a push from Ms. McKinney to initiate new tasks, and a lack of enthusiasm toward some of his schoolwork seems evident.
- He received a 4 on *Sustained Effort*. Antonio can become unfocused when encountering a challenging task. He takes a long time to get some things done because he gets easily distracted, particularly with tasks that are not interesting to him. However, he shows a higher level of effort on certain assignments.
- He received a 4 on *Communication*. Although Antonio is respectful and enjoys working with others, he sometimes has trouble expressing himself both verbally and in writing. He can read and understand written communication well but sometimes has trouble writing it himself. He also sometimes needs to be told things more than once because he does not tend to listen carefully.

Antonio Gomez

Rated on August 30, 2011 - Rated by Sara McKinney
8th Grade - ID 926092345

SAMPLE MIDDLE SCHOOL Class/section: 009



ENGAGE Teacher Edition is used to assess student behaviors and track student progress over time in key behavioral areas that support academic achievement and success. The scales assess students' standing on eight behavioral dimensions that are important for academic success. Scores on each scale range from 8 (High) to 1 (Low).

This report compares assessment results for an individual student with the average ratings for his/her school and district, and where data are available, plots trends over time.

The school mean is based on 538 students. The district mean is based on 1,262 students.

Motivation

Personal characteristics that help students succeed academically by focusing and maintaining energies on goal-directed activities.

Initiative

Seeking out opportunities for personal and academic development. Volunteering to work on additional activities or projects. Showing enthusiasm for learning. Starting new things without being told.

Student	4
School mean	4.7
District mean	4.9

Planning & Organizing

Setting appropriate goals and developing workable plans and strategies for achieving goals. Planning ahead, using time effectively, keeping own materials organized. Setting appropriate priorities and using back-up plans as needed.

Student	6
School mean	4.5
District mean	5.1

Sustained Effort

Working hard and staying focused. Putting a great deal of effort into schoolwork and learning activities and avoiding distractions. Following through with a task or project until it is completed.

Student	4
School mean	4.6
District mean	4.7

Performance

Meeting deadlines, being productive in completing schoolwork, and doing high-quality and accurate work.

Student	5
School mean	4.7
District mean	4.9

Trend data not available.

Plots of trend data are displayed after two or more administrations.

Antonio Gomez

Page 2

Social Engagement

Interpersonal factors that influence students' successful integration or adaptation into their environment.

Communication

Conveying information clearly, effectively, and appropriately in speaking and writing. Getting his or her point across.

Student	4
School mean	4.9
District mean	4.9

Trend data not available.

Plots of trend data are displayed after two or more administrations.

Working with Others

Showing consideration for others, listening to others' points of view, and being helpful. Working constructively and cooperatively with others. Resolving conflicts effectively and being open and accepting of others.

Student	6
School mean	5.1
District mean	5.3

Self-Regulation

Cognitive and affective processes used to monitor, regulate, and control behavior related to learning.

Managing Feelings

Not letting feelings get in the way of schoolwork. Not overreacting to stressful or difficult situations. Finding appropriate ways to express and manage negative feelings, such as anger, sadness, and embarrassment.

Student	5
School mean	4.6
District mean	5.1

Trend data not available.

Plots of trend data are displayed after two or more administrations.

Conduct

Following school rules, codes, and other behavior expectations. Being polite, engaging in appropriate behaviors, and avoiding negative behaviors, such as attention seeking, aggression, or dishonesty.

Student	7
School mean	5.2
District mean	5.4

Plan for Improvement. Based on these results, Ms. McKinney and Antonio created a development plan that will capitalize on his strengths and also develop his weaknesses.

1. Throughout his academic career, it will be important for Antonio to communicate with others and write well. Since he sometimes has trouble expressing himself verbally and in writing (and was rated a 4 on *Communication*), Ms. McKinney suggested that Antonio take advantage of a tutor. Time spent with the tutor will focus on practicing writing assignments, verbal reports, and writing short responses to questions. Antonio will meet with Ms. McKinney periodically to discuss his progress. Any improvements will be monitored via future ratings.
2. Antonio scored a 4 on both *Initiative* and *Sustained Effort*. While discussing his lack of enthusiasm for some projects and how he becomes unfocused and distracted, Ms. McKinney realized that Antonio was not being challenged enough. Antonio stated that he was sometimes bored with school. Since he is a bright student, Ms. McKinney thought he might benefit from taking some advanced classes. Although Ms. McKinney was unable to change his classes midsemester, she suggested that Antonio enroll in more challenging classes next year, in high school. She also plans to make extra efforts to engage Antonio in her class, such as calling on him to answer questions and share opinions, recommending more challenging books to read, and discussing his assignments with him. Ms. McKinney will be careful not to assign him more work than other students but will try to provide him with challenging tasks that might be more engaging and help him to focus better on his schoolwork.

For the next set of ratings, Ms. McKinney plans to pay particular attention to any changes in Antonio's enthusiasm for his schoolwork and improvements in communication. She will keep notes during the period between evaluations and will use those notes and other measures of Antonio's progress in order to assign appropriate ratings.

6

Using ENGAGE Teacher Edition with ENGAGE

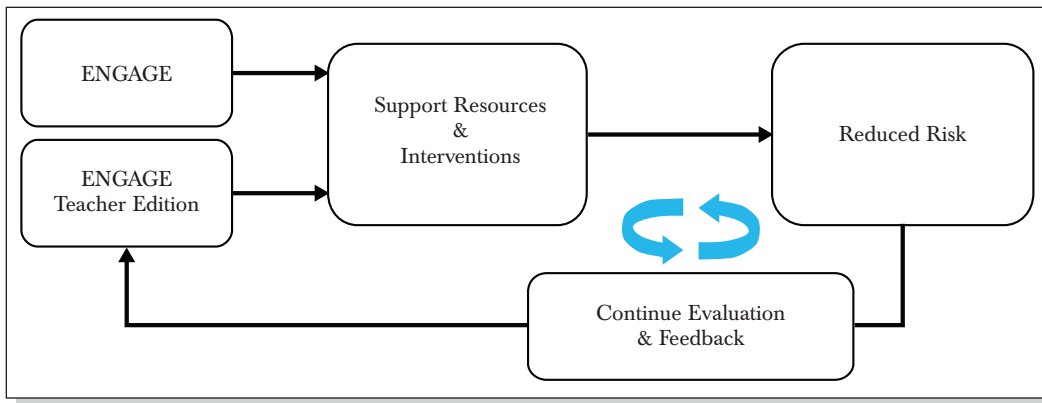
ENGAGE Teacher Edition scales are different from other academic behavior measures offered by ACT, such as ENGAGE (ACT, 2009), in that ENGAGE is completed by students using a self-report format and Teacher Edition scales are completed by teachers based on their observations of students' behaviors. ENGAGE Teacher Edition can be administered multiple times and used to assess change in academic behavior and track progress. Both ENGAGE and ENGAGE Teacher Edition can be used alone or in combination to provide a more comprehensive picture of students' academic behaviors to:

- identify at-risk students,
- connect students to helpful resources, and
- monitor student progress.

Figure 13 illustrates a process that incorporates both ENGAGE and ENGAGE Teacher Edition. In this process, ENGAGE is administered first to provide a broad profile of students' relative strengths and weaknesses. ENGAGE Teacher Edition is administered next and can be used to confirm and help interpret the ENGAGE results. In addition, because they focus on behavior, ENGAGE Teacher Edition can be used to help determine appropriate behavioral learning objectives that can guide interventions and/or referral to support resources. Later, ENGAGE Teacher Edition can help teachers monitor and evaluate students' progress on their behavioral learning objectives, as well as reassess students' need for support resources.

Figure 13

Use of ENGAGE and ENGAGE Teacher Edition for Initial Assessment and Monitoring of Student Academic Behaviors



Appendix A

Development, Reliability, and Validity of ENGAGE Teacher Edition

General Approach

Informed in part by the development process for ENGAGE (ACT, 2009), development of ENGAGE Teacher Edition followed a construct validation approach (e.g., Clark & Watson, 1995; Loevinger, 1957; Nunnally & Bernstein, 1994). Development began with a thorough literature review in the main domains identified as key to academic performance and persistence: Motivation, Social Engagement, and Self-Regulation (Robbins et al., 2004). Eight behavioral scales were identified as important: *Initiative, Planning & Organizing, Sustained Effort, Performance, Communication, Working with Others, Managing Feelings, and Conduct*.

Development was guided by both the existing literature and expert opinion. Individual behavior statements used to define the various effectiveness levels for each of the behavioral scales were developed through qualitative and quantitative procedures, including expert reviews, focus groups, a retranslation study (cf. Smith & Kendall, 1963), and a final review by internal experts. Details of the development follow.

Literature Review

Motivation

Motivation refers to the self-regulatory mechanism by which individuals are able to act on prescribed behaviors to implement training and learning activities (Robbins et al., 2009). In middle school settings, the literature suggests that several behavioral scales are important. Initiative, planning and organizing, sustained effort, and performance scales appear to be observable and distinguishable indicators of motivation. For example, students with higher initiative are more likely to be motivated to engage in class work or extracurricular activities (Ryan & Deci, 2000; Scherer, 2002). Similarly, planning and organizing impacts the setting and achieving of learning goals (Gailliot, Mead, & Baumeister, 2008). And those students who have higher levels of sustained effort are more likely to complete academic tasks and achieve their goals (Zimmerman, 2008). Further, students with a higher level of motivation can be expected to evidence more effective academic performance behaviors, such as meeting deadlines and completing assignments accurately and with a high level of quality.

Social Engagement

Social Engagement refers to an individual's skills in engaging the social environment in ways that help to support and reinforce his or her learning activities (Robbins et al., 2009). From a social and emotional learning perspective, learning is a social process that often takes place in collaboration with peers and teachers (Zins, Bloodworth, Weissberg, & Walberg, 2004). At the secondary school level, social engagement is manifested as participatory behaviors (e.g., Fredricks, Blumenfeld, & Paris, 2004), which include communication skills, as well as the interpersonal skills required to successfully work with other students and teachers.

Self-Regulation

In terms of the literature, the concept of Self-Regulation includes two components: emotional regulation (i.e., *Managing Feelings*) and behavioral regulation (i.e., *Conduct*). Emotional regulation emphasizes how students influence which emotions they have, when they have them, and how they experience and express them (Gross, 1998). From a learning perspective, it refers to the ability to self-manage attitudes and feelings that directly affect students' receptiveness to, and implementation of, learning activities (Robbins et al., 2009). Behavioral regulation involves self-observing and adjusting of behaviors that influence the learning process (Zimmerman & Kitsantas, 2005). Self-regulating their behavior to conform to school rules and expectations (e.g., conduct) is one of the primary ways in which students impact their own learning processes.

Development of Behavioral Statements

Based on the literature review, ACT staff wrote definitions for each of the eight behavioral scales. Subsequently, behavioral items were written to cover the content of the definitions. Wherever research or behavioral standards (from education curriculum expectations) were available, these were considered in the item writing process. For each scale, items were written to represent different levels of behavioral effectiveness, from superior to ineffective.

Focus Groups

The focus groups served multiple purposes. Specifically, focus group participants: (1) reviewed draft behavioral statements, (2) independently generated additional statements, (3) filled gaps they saw in coverage after reviewing and revising draft statements, and (4) provided qualitative judgments of the effectiveness levels of items. Four focus groups were conducted with a total of 13 participants from different geographical regions. Focus groups participants were, on average, 44.0 years of age (SD = 14.6 years), female (77%), Caucasian (77%), and very experienced teachers (average length of teaching experience = 18 years). Each focus group lasted approximately four hours and included an overview of behavioral scales, review of scale definitions, review and revision of specific behavioral statements, and time to generate new/additional statements. Participants were compensated for their time and participation.

After each focus group, ACT researchers revised behavioral statements based on the focus group feedback and discussions and prepared a revised item pool for the next focus group. Using this iterative process, 584 behavioral statements were generated and refined for inclusion in the retranslation study described below.

Retranslation

Retranslation (or reallocation) is a method used to assign behavioral statements to each behavioral scale and gauge their effectiveness (Smith & Kendall, 1963; Schwab, Heneman, & Decotiis, 1975). During this step, experienced teachers were given the definitions for each scale and asked to assign each behavioral statement to the scale that it best described. This group of teachers was also asked to rate the effectiveness of each behavioral statement using a scale ranging from 1 (*extremely ineffective*) to 8 (*extremely effective*).

To collect this information, the 584 behavioral statements were randomly assigned to one of two web-based retranslation questionnaires, with each questionnaire including 292 behavioral statements. The retranslation questionnaires were completed by a total of 34 participants from different geographical regions, with each participant completing one questionnaire. These participants were, on average, 40.2 years of age ($SD = 12.1$ years), female (91%), Caucasian (79%), and experienced teachers (average length of teaching experience = 10.3 years). Each questionnaire took approximately four hours to complete. Participants were compensated for their time and participation.

Those statements that were assigned to the same behavior scale by 80% or more of the raters were considered as candidates for use as anchors on the behavioral scales. The average effectiveness rating assigned to each retranslated statement was used to identify the effectiveness of the behavior described. The standard deviation of the effectiveness ratings for each statement reflects the amount of agreement among raters regarding the effectiveness level for that statement (i.e., the lower the standard deviation, the better the agreement). We used the standard error of measurement (SEM, which is a function of the standard deviation and the sample size) to decide which statements would be retained for potential inclusion into a final scale. Statements were not retained if they had an SEM of .30 or greater. Statements for which raters had good agreement, in terms of *both* their scale assignment and effectiveness level, were retained for possible inclusion into the final version of the scales.

Of the 584 behavioral statements included in the retranslation questionnaires, 355 (61%) statements were retained based on these criteria. Two of these remaining behavioral statements were then selected to anchor or describe each rating point for each scale (e.g., two statements describing an effectiveness level of 1, two for a level of 2, and so on). These behavioral statements were selected based on the following criteria:

- Items selected for a specific level could be clearly differentiated, based on the retranslation data (in terms of mean effectiveness), from items selected for the surrounding rating levels. To do this, we examined the SEM around the mean effectiveness levels of the items and selected items such that the SEMs did not overlap.
- The overall item content was representative of the scale and, as much as possible, several facets of a behavioral scale were represented at each rating level.

After completing this process, a few rating scale points on several of the scales were missing one (or both) behavioral statements. Additional items were written targeting the missing effectiveness levels, and a set of experts were asked to provide a follow-up review by rating the statements using the same procedure that was used in the retranslation study.

Follow-up Review

The purpose of the follow-up review was to collect information about the additional behavioral statements generated to complete the scales, as well as confirm the results of the retranslation study for the items that had been selected for each scale anchor. Participants were eight education experts. These participants were, on average, 43.3 years of age ($SD = 14.7$ years), male (75%), Caucasian (100%), and experienced educators (average length of experience = 10.4 years).

Responses were analyzed using similar methods to those used for analyzing the retranslation survey results. The results from the follow-up review of the originally selected items were consistent with results from the retranslation study. Results for the newly written items provided sufficient data to fill in some of the anchors that were previously missing.

Overall, the entire scale development process (i.e., focus groups, retranslation study, and follow-up review) provided both quantitative and qualitative evidence supporting the appropriateness of the behavioral statements used to anchor the effectiveness levels.

Table A1 summarizes the retranslation data for the behavioral statements that were chosen to anchor the scales. The first column shows the average percentage of the raters who agreed with the assignment of the behavioral statements to the target scale. This table also shows interrater reliabilities for the effectiveness ratings assigned to these statements, estimated using intraclass correlation coefficients (ICC; Shrout & Fleiss, 1979). As can be seen, raters strongly agreed that the selected behavioral statements belonged to their corresponding behavioral scales (range = 87.2% to 96.9%, median = 92.3%). Further, the reliability estimates were moderate to high, whether estimated based on one rater (ICC 2,1; range = .71 to .86, median = .79) or as the average of multiple raters (ICC 2,k; range = .95 to .98, median = .97). The average across all raters is the value used to assign items to effectiveness levels, and the reliability estimates in the last column indicate that these means are highly reliable.

Table A1
Mean Interrater Agreement and Reliability Estimates for the ENGAGE Teacher Edition Scale Anchors

Behavioral Scale (# of items)	Scale Assignment Agreement ¹	ICC (2,1) ^a	ICC (2,k) ^b
Initiative (11)	91.3	.71	.95
Planning & Organizing (11)	96.9	.72	.95
Sustained Effort (11)	89.2	.78	.97
Performance (12)	87.2	.75	.96
Communication (13)	96.1	.84	.98
Working with Others (13)	91.7	.86	.98
Managing Feelings (11)	94.4	.83	.97
Conduct (13)	92.9	.80	.97

Note. ¹Mean scale assignment agreement calculated based on the percentage of raters who agreed that the statements selected for the behavioral scales were part of that behavioral scale.

^aICC generated using Shrout-Fleiss reliability calculations (Case 2).

^bICC generated using Shrout-Fleiss reliability calculations (Case 2, mean across k raters; k = 8; see Shrout and Fleiss, 1979).

Development of Summary Statements

The purpose in developing summary statements was to provide raters with additional behavioral content to help them in the rating process (Borman, 1979). Specifically, summary statements were written to provide four general behavior descriptions for each scale. Each general behavior description covers two effectiveness levels (i.e., 1–2, 3–4, 5–6, and 7–8), thus helping raters to narrow their rating options. Summary statements were generated using items that survived the rigorous selection process described previously (based on rater agreement concerning scale and effectiveness), but that ultimately were not retained as one of the final specific behavioral anchors for the relevant scale. To write these summary statements, we synthesized the content of the more specific behavioral statements at each level to generate higher-level summaries that accurately portrayed the type and scope of behaviors at each broad effectiveness level. Some research has shown that raters can have trouble relating ratee behaviors to specific behavioral statements, but can more easily see how ratee behaviors fit with this type of broad summary statement (Bernardin & Smith, 1981; Borman, 1979). By including both types of information in ENGAGE Teacher Edition, we hoped to enhance the accuracy of the ratings by providing raters with a variety of tools to help them make their ratings.

Reliability and Validity

To examine the reliability and validity of ENGAGE Teacher Edition, data were collected from two different samples, and these two studies are described below.

Reliability

Reliability is the consistency of a set of measurements. In other words, a scale is reliable to the extent that the obtained score is internally consistent and/or repeatable. There are many different estimates of reliability and the choice between them depends on the type and the purpose of the scale.

Sample and Data Collection. The data to assess the reliability of ENGAGE Teacher Edition were collected in two middle schools in the Midwest. The total sample consisted of 320 students who were each rated by two teachers using ENGAGE Teacher Edition. Approximately 65% of these students were seventh graders, and 35% were eighth graders, with the mean age of 13.2 (SD = 0.6). Fifty-six percent of the students were female, and the majority were Caucasian (specific race/ethnicity breakdown was not collected).

Interrater Reliability. For each of the eight ENGAGE Teacher Edition scales, raters provide a single rating ranging from one to eight. The most appropriate type of reliability estimation for this kind of scale is interrater agreement, which refers to the extent to which multiple judges' (i.e., teachers') ratings of the same target (i.e., student) converge. One way of estimating interrater agreement is by using the intraclass correlation coefficient (ICC). Like other correlation metrics, the values of ICC range from 0 to 1.0, where a higher value indicates greater interrater reliability. All ICC analyses were conducted using the one-way model (see McGraw & Wong, 1996), which corresponds to Case 1 in Shrout and Fleiss (1979).

Results of the ICC analyses are presented in Table A2. ICC (1, 1) refers to the reliability of a single rating, and ICC (1, 2) refers to the reliability for the mean of 2 ratings. The obtained interrater reliabilities are acceptable and typical of reliabilities for this type of observer ratings (range for one rater = .38 to .60, median = .50) (Knapp, Campbell, Borman, Pulakos & Hanson, 2001; Motowidlo & Borman, 1977; Das, Frost, & Barnowe, 1979). It is worth noting that the reliabilities are better for two raters than for a single rater (range for two raters = .55 to .75, median = .67), which also is consistent with the literature (e.g., Grussing, Valuck, & Williams, 1994). Larger numbers of raters generally yield more reliable ratings. It is also worth noting that reliabilities were higher for the behavioral scales that are more observable; for example, *Planning & Organizing* showed the highest reliability, whereas *Managing Feelings* (which is more difficult to observe) showed the lowest reliability.

Table A2
Interrater Reliability Estimates Using ICCs

Behavioral Scale	ICC (1, 1)	ICC (1, 2)
Initiative	.43	.60
Planning & Organizing	.60	.75
Sustained Effort	.57	.73
Performance	.58	.73
Communication	.47	.64
Working with Others	.44	.61
Managing Feelings	.38	.55
Conduct	.53	.69
<i>Median</i>	<i>.50</i>	<i>.67</i>

Note. ICCs = Intraclass correlation coefficients.
N = 320 pairs of ratings. ICC (1, 1) = reliability of a single rating. ICC (1, 2) = reliability for the mean of 2 ratings.

To further investigate the extent to which teachers agree on their ratings of students when using ENGAGE Teacher Edition, agreement percentages were also computed. Table A3 presents information on the (cumulative) percentage of the teachers who chose the same rating level (exact), as well as those who agreed within a certain confidence interval (e.g., within one point, within two points, and so on). An average of approximately 33% of teachers gave the exact same ratings to students. This percentage increases considerably when the confidence interval is expanded to agreement within one point, where an average of approximately 72% of teachers agreed on their ratings. This percentage increases even more when the confidence interval is expanded to agreement to within two points, where an average of approximately 89% of teachers agreed on their ratings. Overall, the reliability results show that there is good agreement between teachers rating the same student using ENGAGE Teacher Edition.

Table A3
Percentage of Interrater Agreement

Behavioral Scale	% Agreement (cumulative)				
	Exact	Within 1 pt.	Within 2 pts.	Within 3 pts.	Within 4 pts.
Initiative	26.6	62.5	85.0	96.6	99.7
Planning & Organizing	33.9	69.6	88.4	95.3	98.8
Sustained Effort	33.1	69.7	87.2	96.9	99.4
Performance	31.7	73.7	89.7	97.8	99.7
Working with Others	33.1	71.9	89.1	97.2	99.4
Communication	31.3	69.7	87.8	98.1	100.0
Managing Feelings	32.5	73.4	91.9	96.9	99.7
Conduct	43.3	83.1	93.7	97.5	100.0
<i>Mean</i>	33.2	71.7	89.1	97.0	99.6

Note. $N = 320$ pairs of ratings.

Validity

Validity refers to the accuracy of statements about the meaning or implications of test scores and is the most fundamental consideration when developing and evaluating assessments. The validation process begins with specific statements about the proposed uses and interpretations of an assessment and is followed by evidence that supports such statements (AERA, APA, & NCME, 1999). Specifically, to provide evidence that the assessment measures the “construct” it is intended to measure, we can test whether it shows the expected internal structure and the expected relationships with measures of related constructs. The following sections present evidence supporting the validity of ENGAGE Teacher Edition for assessing students’ academic behaviors.

Sample and Data Collection. The data used to assess the validity of ENGAGE Teacher Edition were collected in four high schools located in the southern United States who were participating in a school reform initiative. The sample consisted of 651 students, each of whom was rated by a teacher using ENGAGE Teacher Edition. The participants were mostly ninth graders (99%), with an average age of 15.6 years ($SD = .77$). Approximately 50% of the students were female, and the majority of the students (75.2%) were African American. The remainder of the sample was Asian American/Pacific Islander (8.6%), Caucasian (5.2%), Hispanic (5.8%), and American Indian (5.2%).

Internal Structure of ENGAGE Teacher Edition

Intercorrelations. Intercorrelations between the eight ENGAGE Teacher Edition scales show a reasonable convergent/discriminant pattern, with scales generally correlating more strongly with other scales that are conceptually similar. For example, the four scales that are part of the Motivation domain (*Initiative*, *Planning & Organizing*, *Sustained Effort*, and *Performance*) are correlated more strongly with each other (range = .85 to .89, median = .88) than with the other scales (range = .56 to .82, median = .75). Similarly, for the Self-Regulation domain, *Managing Feelings* and *Conduct* correlated more strongly with each other ($r = .82$) than with

scales from other domains (range = .56 to .77, median = .61). This convergent/discriminant pattern was less pronounced for the Social Engagement scales (i.e., *Communication* and *Working with Others*), as their relationship with each other ($r = .77$) was closer in magnitude to their relationships with other scales (range = .59 to .81, median = .75) (See Table A4 for the full intercorrelation matrix).

Table A4
Descriptive Statistics and Intercorrelations for ENGAGE Teacher Edition

Behavioral Scale	M	SD	1	2	3	4	5	6	7	8
Initiative	4.66	1.94								
Planning & Organizing	4.53	2.04	<u>.89</u>							
Sustained Effort	4.77	2.10	<u>.87</u>	<u>.89</u>						
Performance	4.71	1.99	<u>.85</u>	<u>.88</u>	<u>.88</u>					
Communication	5.06	1.91	.75	.77	.81	.79				
Working with Others	5.27	1.84	.75	.75	.79	.74	<u>.77</u>			
Managing Feelings	5.40	1.95	.60	.58	.60	.56	.59	.77		
Conduct	5.61	2.01	.61	.62	.65	.61	.61	.76	<u>.82</u>	

Note. $N = 639$ – 649 . All correlations are significant ($p \leq .01$).

Second-Order Structure of ENGAGE Teacher Edition. Based on our initial theory, ENGAGE Teacher Edition scales are expected to be made up of three higher-order factors: Motivation (*Initiative, Planning & Organizing, Sustained Effort, and Performance*), Social Engagement (*Communication, Working with Others*) and Self-Regulation (*Managing Feelings, Conduct*). Confirmatory factor analyses (CFA) were conducted to investigate whether the empirical structure of the scales supports this theory. Confirmatory factor analysis allows for the testing of the adequacy of the hypothesized factor structure (Tabachnik & Fidell, 2007). Specifically, CFA tests whether the scales are in fact indicators of the higher-order constructs they are expected to measure.

Three models were compared: one-factor, two-factor, and three-factor. The analyses were performed using the statistical software Mplus 5.21 (Muthén & Muthén, 2007). Maximum likelihood estimation was employed to estimate all models. Model fit was evaluated on the basis of several fit and residual indices. Fit indices included the comparative fit index (CFI) and non-normed fit index (NNFI), both of which range from 0 to 1.0, where values greater or equal to .95 suggest excellent model fit. Residual indices included the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR), both of which range from 0 to 1.0, where values less than or equal to .05 suggest excellent model fit (Kline, 2004).

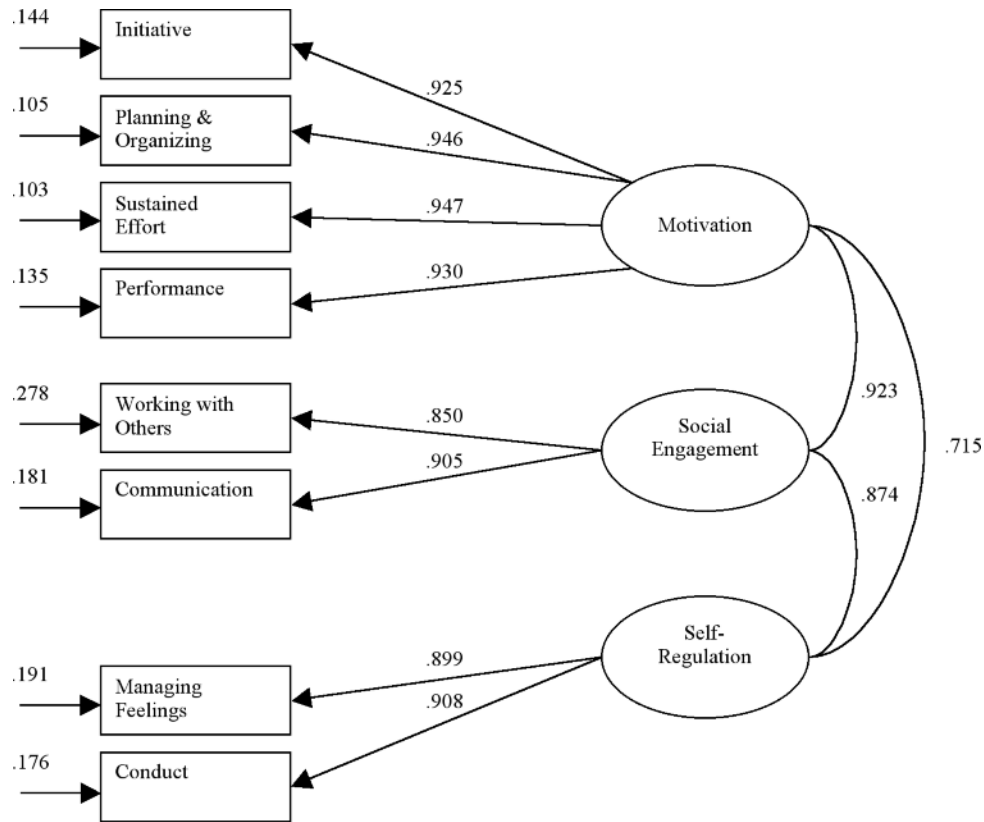
In the first model tested, all scales were set as indicators of a single overall factor. Based on the relatively high intercorrelations among ENGAGE Teacher Edition, it seemed possible that there would be a single latent factor underlying all the scales. However, the overall model showed poor fit (See Table A5). The second model tested included two latent factors: Motivation (*Initiative, Planning & Organizing, Sustained Effort, Performance, Communication, and Working with Others*) and Self-Regulation (*Managing Feelings and Conduct*). The overall model showed acceptable fit. The last model tested was the hypothesized three-factor model consisting of Motivation (*Initiative, Planning & Organizing, Sustained Effort, and Performance*), Social Engagement (*Communication, Working with Others*) and Self-Regulation (*Managing Feelings, Conduct*). As shown on Table A5, the three-factor model yielded the best fit to the data when compared with the other two models. The three-factor model is shown in Figure A1. This figure shows that the standardized path coefficients for the indicator variables are relatively high (range = .850 to .947, median = .917) and the error terms are relatively low (range = .103 to .278, median = .160). Despite the relatively high intercorrelations between the factors, the results of CFA analyses suggest that the theoretically proposed three-factor model yielded the best fit to the data.

Table A5
Resulting Fit Indices from Confirmatory Factor Analyses

Model	χ^2	df	CFI	NNFI	RMSEA	SRMR
One-factor	721.64	20	0.885	0.838	0.233 [.219–.248]	0.069
Two-factor	349.4	19	0.946	0.920	0.164 [.149–.180]	0.048
Three-factor	157.52	17	0.977	0.962	0.113 [.097–.130]	0.027

Note. CFI = comparative fit index; NNFI = Non normed fit index;
RMSEA = root mean square error of approximation;
SRMR = standardized root mean square residual.

Figure A1
Second-Order Structure of ENGAGE



Relationships between ENGAGE Teacher Edition and Measures of Other Constructs

As part of documenting the construct validity of ENGAGE Teacher Edition, we examined the relationships between the scales and measures of academic achievement, student risk, and school records. Results are presented in the following sections.

Academic Achievement. ENGAGE Teacher Edition is positively associated with several indicators of academic achievement. Table A6 shows that the scales are moderately correlated with standardized achievement scores, as measured by EXPLORE® subject and composite scores (range for EXPLORE composite = .26 to .43, median = .40). It is worth noting that the EXPLORE correlations with Motivational and Social Engagement scales are slightly stronger than with the Self-Regulation scales, suggesting that these types of academic behaviors may be better predictors of academic achievement, as measured by standardized tests.

Table A6
Correlations between ENGAGE Teacher Edition and EXPLORE Scores

Behavioral Scales	Academic Achievement Scores				
	EXPLORE Composite	EXPLORE English	EXPLORE Math	EXPLORE Reading	EXPLORE Science
Initiative	.39	.31	.32	.34	.35
Planning & Organizing	.40	.34	.31	.36	.35
Sustained Effort	.40	.35	.33	.33	.34
Performance	.40	.34	.33	.33	.35
Communication	.43	.37	.37	.35	.35
Working with Others	.32	.27	.27	.28	.27
Managing Feelings	.27	.24	.22	.21	.24
Conduct	.26	.23	.21	.23	.24
<i>Median</i>	<i>.40</i>	<i>.33</i>	<i>.32</i>	<i>.33</i>	<i>.35</i>

Note. $N = 503 - 649$. All correlations are significant ($p \leq .01$).

Table A7 shows correlations between ENGAGE Teacher Edition and school grades (both current and prior), as well as correlations with other markers of academic achievement (e.g., having failed a class in the past, being held back from moving to a subsequent grade, and time spent doing homework on an average school day). As this table shows, the scales are moderately to strongly related to current GPA and prior grades, moderately (negatively) related to having failed a class in the past, and weakly related to being held back from moving on to a subsequent grade or to time spent doing homework. Students who are rated higher on ENGAGE Teacher Edition are more likely to earn higher grades and less likely to fail a class.

Table A7
Correlations between ENGAGE Teacher Edition and Academic Achievement

Behavioral Scales	Academic Achievement				
	Current GPA ^a	Prior Grades ^b	Class failed ^b	Held back ^b	Time spent on homework ^b
Initiative	.56	.43	-.31	-.11	.15
Planning & Organizing	.58	.45	-.30	-.14	.17
Sustained Effort	.58	.45	-.27	-.12	.17
Performance	.57	.44	-.29	-.13	.14
Communication	.50	.40	-.26	-.18	.11
Working with Others	.48	.38	-.21	-.11	.15
Managing Feelings	.40	.31	-.22	-.10	.15
Conduct	.40	.35	-.24	-.09	.19
<i>Median</i>	<i>.53</i>	<i>.42</i>	<i>-.27</i>	<i>-.12</i>	<i>.15</i>

Note. $N = 436 - 502$. Correlations $\geq .11$ are significant ($p \leq .01$).

^aReported by schools. ^bSelf-reported by students.

Academic Success and Risk. ENGAGE Teacher Edition is also moderately related to the Academic Success Index from ENGAGE (range = .36 to .47, median = .42), which is a composite of ENGAGE scales, behavioral indicators, and self-reported prior grades. This index was developed to predict students' academic success and to identify students at risk of experiencing academic difficulties (ACT, 2009). Students who obtain higher scores on ENGAGE Academic Success Index also tend to be rated more highly by their teachers on ENGAGE Teacher Edition. ENGAGE is completed by students, and the correlations between ENGAGE Teacher Edition and ENGAGE provides evidence of convergence between the two sources of information: self-reported student behaviors (i.e., ENGAGE scores) and other-reported student behaviors (teacher ratings of students from ENGAGE Teacher Edition).

Behavioral Indicators. We also examined the relationships between ENGAGE Teacher Edition and several key behavioral indicators that have been linked to academic success in the literature, such as misconduct, absenteeism, and coming to school without having completed homework (Kaufman & Bradbury, 1992; Rumberger, 1995; Rumberger & Larson, 1998). As shown in Table A8, the scales are moderately (negatively) related to behavioral indicators of misconduct reported by the school such as disciplinary actions and suspensions. These results suggest that students who obtain higher ratings on ENGAGE Teacher Edition are less likely to engage in behaviors that result in disciplinary actions and suspensions.

Further, the scales are moderately (negatively) related to absenteeism and to the frequency of students not having their homework completed. Students who obtain higher ratings on ENGAGE Teacher Edition are less likely to be absent from school and more likely to complete their homework assignments.

Table A8
Correlations between ENGAGE Teacher Edition, Misconduct, and Absenteeism

Behavioral Scales	Behavioral Indicators			
	Disciplinary Actions ^a	Suspensions ^a	Absenteeism ^b	Without homework ^b
Initiative	-.35	-.26	-.26	-.18
Planning & Organizing	-.36	-.24	-.26	-.20
Sustained Effort	-.39	-.27	-.28	-.20
Performance	-.32	-.20	-.23	-.19
Communication	-.30	-.22	-.20	-.13
Working with Others	-.41	-.32	-.27	-.12
Managing Feelings	-.42	-.31	-.25	-.12
Conduct	-.45	-.35	-.30	-.17
<i>Median</i>	<i>-.38</i>	<i>-.27</i>	<i>-.26</i>	<i>-.18</i>

Note. $N = 455 - 459$. All correlations are significant ($p \leq .01$).

^aReported by schools. ^bSelf-reported by students.

School-level Factors. Table A9 presents correlations between ENGAGE Teacher Edition and school-level factors. As expected, the scales are generally unrelated to these school-level factors, including the percent of minority students in a school, the percent of free or reduced-lunch recipients, average class size, and student teacher ratio. These findings are consistent with the literature showing that measures of academic behavior are generally unrelated to these types of school factors (Finn & Voelkl, 1993; McDermott, 1995, 1999; Weishew & Peng, 1993).

Table A9
Correlations between ENGAGE Teacher Edition and School-level Factors

Behavioral Scales	School-level Factors			
	% of Minority Students	% of Free/Reduced Lunch	Average Class Size	Student-Teacher Ratio
Initiative	.00	-.04	.11	.08
Planning & Organizing	.01	-.05	.13	.09
Sustained Effort	-.02	.00	.13	.08
Performance	-.01	-.01	.11	.07
Communication	.01	.04	.12	.06
Working with Others	.03	.00	.11	.08
Managing Feelings	.02	.00	.08	.05
Conduct	-.01	.11	.02	-.03
<i>Median</i>	<i>.01</i>	<i>.00</i>	<i>.11</i>	<i>.08</i>

Note. $N = 518$. Correlations $\geq .11$ are significant ($p \leq .01$).

Demographic Factors. We also examined the relationships between student demographic characteristics and ENGAGE Teacher Edition. Table A10 shows that, on average, female students were rated moderately higher on all scales, resulting in a negative correlation with male gender. This is consistent with existing literature which suggests that female students tend to behave more appropriately in the classroom (e.g., Cohn & Modecki, 2007, Zimmer-Gembeck, Geiger, & Crick, 2005). Further, race/ethnicity (coded dichotomously as Caucasian = 0, minority = 1) was not related to teachers' ratings on the scales. Finally, with the exception of a low correlation with *Communication*, parental education was not related to the behaviors measured by the scales (Neiss & Rowe, 2000).

Table A10
Correlations between ENGAGE Teacher Edition and Demographic Information

Behavioral Scales	Demographic Factors		
	Minority	Male	Parental Education
Initiative	.02	-.22	.03
Planning & Organizing	.00	-.28	.05
Sustained Effort	.00	-.25	.04
Performance	.00	-.25	.06
Communication	.00	-.20	.14
Working with Others	.01	-.26	.04
Managing Feelings	.01	-.11	-.02
Conduct	.01	-.22	-.01
<i>Median</i>	<i>.00</i>	<i>-.24</i>	<i>.04</i>

Note. $N = 315 - 634$. Correlations $\geq .13$ are significant ($p \leq .01$).

Academic Achievement of Students at Each Behavioral Monitoring Scale Level

To further investigate the properties of ENGAGE Teacher Edition, we examined the academic achievement of students at each level of the scales. Table A11 shows the percentage of students who failed at least one class at each level of academic behavior on each scale of ENGAGE Teacher Edition. As this table shows, the analyses revealed an overall pattern in which the percentage of students failing a class tends to be higher at the lower behavior rating levels. For example, looking at the *Initiative* scale, the frequency of students who had failed a class is much higher for those students who were rated at the lowest level of the *Initiative* scale (67%) than for students who were rated at the highest level of the scale (7%).

Table A11
Percent of Students Having Failed a Class by Behavior Level

Behavior Level	Initiative	Planning & Organizing	Sustained Effort	Performance
1	67	77	65	65
2	48	42	39	47
3	43	39	41	35
4	40	43	43	49
5	32	35	37	25
6	22	19	18	18
7	14	13	18	15
8	7	9	12	15

Behavior Level	Communication	Working with Others	Managing Feelings	Conduct
1	57	50	70	50
2	38	35	46	56
3	46	51	34	47
4	47	41	53	38
5	31	38	31	37
6	22	23	28	29
7	21	23	25	23
8	11	14	15	18

Note. $N = 436 - 441$.

Table A12 shows another set of analyses examining levels of behavior and academic achievement, in this case average letter grades. As can be seen, the percentage of students obtaining a higher letter-grade average is greater for those students who were rated as demonstrating higher levels of behavior, whereas the percentage of students obtaining a lower letter-grade average is higher for those students demonstrating lower levels of behavior. For example, when looking at the *Planning & Organizing* scale, the frequency of students who earned grades of “A” increased from 0% (for students who were rated a one on the scale) to 61% (for students who were rated an eight).

In contrast, the opposite pattern emerges when looking at students who earned grades of “C.” This frequency decreased from 67% (for students who were rated a one on the scale) to 9% (for students who were rated an eight).

A general conclusion from the data presented in Tables A11 and A12 is that levels of academic behavior, as measured by ENGAGE Teacher Edition, are associated with levels of academic success (or risk). Thus, ratings obtained using ENGAGE Teacher Edition can be used by educators to flag students who are at risk of experiencing adverse academic events, such as receiving low grades or failing a class.

Table A12
Percent of Students Receiving Letter-grade Averages by Behavior Level

Behavior Level	Initiative						Planning & Organizing						Sustained Effort						Performance					
	A	B	C	D	F		A	B	C	D	F		A	B	C	D	F		A	B	C	D	F	
1	0	27	55	18	0		0	28	67	6	0		0	28	67	6	0		0	29	67	5	0	
2	9	30	52	2	0		4	38	50	4	2		4	41	50	2	0		3	50	44	0	0	
3	3	51	38	5	0		9	57	31	1	0		10	46	32	5	2		8	44	34	7	2	
4	13	59	24	2	1		10	59	21	7	0		12	56	24	7	0		11	48	36	4	0	
5	17	56	24	3	0		12	55	31	2	0		5	61	34	0	0		14	63	21	1	0	
6	19	59	18	1	0		25	59	13	1	0		23	58	18	2	0		21	60	17	1	0	
7	38	52	7	0	0		42	49	5	0	0		34	57	5	0	0		40	52	2	2	0	
8	55	38	7	0	0		61	30	9	0	0		34	57	5	0	0		56	35	9	0	0	

Behavior Level	Communication						Working with Others						Managing Feelings						Conduct					
	A	B	C	D	F		A	B	C	D	F		A	B	C	D	F		A	B	C	D	F	
1	0	36	57	7	0		0	33	56	11	0		0	20	70	10	0		0	20	70	10	0	
2	12	31	58	0	0		5	38	57	0	0		14	31	49	0	0		4	41	56	0	0	
3	7	52	30	5	0		5	44	44	0	0		6	45	39	3	3		8	41	41	5	0	
4	13	48	32	5	2		11	44	33	7	2		8	44	36	8	0		7	48	33	7	2	
5	7	59	32	3	0		10	56	30	4	0		20	54	22	3	0		20	53	22	2	0	
6	21	54	22	2	0		23	53	18	3	0		14	57	23	3	0		9	61	25	4	0	
7	30	59	7	0	0		26	60	12	0	0		25	55	17	1	0		28	54	14	1	0	
8	56	38	4	0	0		50	41	9	0	0		37	52	11	0	0		34	49	18	0	0	

Note. $N = 447 - 452$. Percentages do not add up to 100% due to missing data. In general, missing cases did not exceed 5% of the sample.

Incremental Validity

As shown in Tables A7 and A12, ENGAGE Teacher Edition scales are related to both prior and current school grades. ENGAGE Teacher Edition is designed to measure student academic behaviors associated with success, so it is expected to predict student academic achievement. In order to understand the scales' relative contribution to the prediction of GPA, we examined their effects after other cognitive and academic behavior predictors of academic performance were included.

Table A13 provides the results of a linear regression model including standardized beta weights for the three predictors of term GPA (obtained at the end of the spring term): 1) EXPLORE composite score, 2) ENGAGE Academic Success Index score, and 3) mean across the eight ENGAGE Teacher Edition ratings. As this table shows, the model yielded an R of .731, with each predictor providing statistically significant prediction. The total proportion of variance in term GPA explained by the model (R^2) was .534. These results show that ENGAGE Teacher Edition provides incremental validity over and above the prediction provided by standardized achievement test scores and other measures of academic behavior, such as the ENGAGE Academic Success Index.

Table A13
Linear Regression to Predict Term GPA

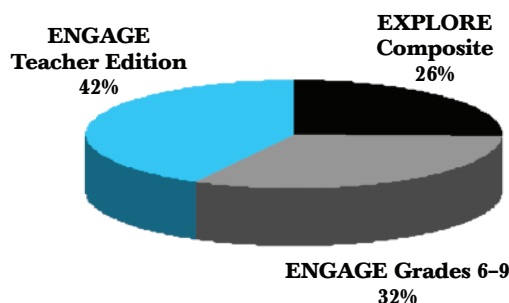
Predictor	Beta (Standardized)
EXPLORE Composite score	.251
ENGAGE Academic Success Index score	.299
ENGAGE Teacher Edition ratings	.387
Total R	.731

Note. All predictors significant at $p \leq .01$.

To better understand the proportion of variance explained by each of these variables (i.e., EXPLORE composite, ENGAGE Academic Success Index, ENGAGE Teacher Edition), we conducted additional analyses using the dominance analysis technique (Azen & Budescu, 2003). Using regression modeling alone, the relative importance of a predictor variable cannot be determined from regression coefficients or beta weights because of the correlations between the predictor variables. The dominance analysis approach allows us to compare the relative importance of predictors in multiple regression by providing an estimate of the amount of variation accounted for by each predictor (R^2).

Figure A2 provides a graphical representation of the results of the dominance analysis. The proportion of variance explained by each predictor (out of the total R^2 of .534 noted previously) is represented by a piece of the pie. As this figure shows, ENGAGE Teacher Edition explains the largest proportion of variance in term GPA. These results confirm that ENGAGE Teacher Edition is: (1) a useful predictor of academic performance, and (2) explains additional variance in GPA above and beyond other measures typically used to predict performance.

Figure A2
Proportion of Variance Predicted in Term GPA



Overall, the results provided in this chapter show that ENGAGE Teacher Edition is effective in assessing student academic behaviors and is a useful tool for predicting academic achievement and success. Further, it provides incremental validity over other measures (e.g., standardized achievement tests, self-report measures of academic behaviors) that are typically used to predict similar outcomes.

Compliance with Guidelines and Standards

The ENGAGE Teacher Edition development and validation process is in compliance with the test development guidelines recommended by the International Testing Commission (2006), the Association of Test Publishers (2002), and the Joint Committee on Standards for Educational and Psychological Testing (consisting of the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education, 1999). These standards address “criteria for the evaluation of tests, testing practices, and the effects of test use” (pg. 2) including delivery formats, administration and hardware/software requirements, and the documentation of test validity and reliability (Joint Committee on Standards, 1999).

References

- ACT (2009). *Student Readiness Inventory—Middle School User's Guide*. Iowa City, IA: Author.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education—Joint Committee on Standards for Education and Psychological Testing. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Association of Test Publishers. (2002). *Guidelines for computer-based testing*. Washington, DC: Author.
- Azen, R. & Budescu, D.V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, 8, 129–148. doi:10.1037/1082-989x.8.2.129
- Bernardin, H. J., & Smith, P. C. (1981). A clarification of some issues regarding the development and use of behaviorally anchored rating scales (BARS). *Journal of Applied Psychology*, 66(4), 458–463.
- Bernardin, H. J. & Walter, C. S. (1977). Effects of rater training and diary-keeping on psychometric error in ratings. *Journal of Applied Psychology*, 62, 64–69.
- Bolster, B. L., & Springbett, B. M. (1961). The reaction of interviewers to favorable and unfavorable information. *Journal of Applied Psychology*, 45, 97–103.
- Borman, W. C. (1979). Format and training effects on rating accuracy and rater errors. *Journal of Applied Psychology*, 64, 410–421.
- Borman, W. C. (1991). Job behavior, performance, and effectiveness. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 271–326). Palo Alto, CA: Consulting Psychologist Press.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7, 309–319.
- Cohn, E. S., & Modecki, K. L. (2007). Gender differences in predicting delinquent behavior: Do individual differences matter? *Social Behavior and Personality*, 35, 359–374. doi:10.2224/sbp.2007.35.3.359
- Das, H. Frost, P. J., & Barnowe, J. T. (1979). Behaviorally anchored scales for assessing behavioral science teaching. *Canadian Journal of Behavioural Science*, 11, 79–88.
- Fay, C. H., & Latham, G. P. (1982). Effects of training and rating scales on rating errors. *Personnel Psychology*, 35, 105–116.
- Finn, J. D., & Voelkl, K. E. (1993). School characteristics related to student engagement. *Journal of Negro Education*, 62, 249–268. Retrieved from <http://www.jstor.org/stable/2295464>
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59–109.

- Gailliot, M. T., Mead, N. L., & Baumeister, R. F. (2008). Self-regulation. In O. John, R. Robins, & L. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed, pp. 472–491). New York: Guilford.
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, 2, 271–299.
- Grussing, P. G., Valuck, R. J., & Williams, R. G. (1994). Development and validation of behaviorally-anchored rating scales for student evaluation of pharmacy instruction. *American Journal of Pharmaceutical Education*, 58, 25–37. doi: aj580125.pdf
- International Test Commission. (2006). International guidelines on computer-based and Internet-delivered testing. *International Journal of Testing*, 6, 143–171.
- Kaufman, P., & Bradbury, D. (1992). *Characteristics of at-risk students in NELS: 88* (NCES 92–042). Washington, DC: U.S. Department of Education, National Center for Educational Statistics. Retrieved from <http://nces.ed.gov/pubs92/92042.pdf>
- Kline, R. B. (2004). *Principles and practice of structural equation modeling* (2nd edition). New York: The Guilford Press.
- Knapp, D. J., Campbell, C. H., Borman, W. C., Pulakos, E. D., & Hanson, M. A. (2001). *Performance assessment for a population of jobs*. In Exploring the Limits in Personnel Selection and Classification J.P. Campbell & D. J. Knapp (Eds.) pp. 181–235. Mahwah, NJ: Lawrence Erlbaum Associates.
- Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin*, 87(1), 72–107.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635–694.
- Maurer, T. J., & Alexander, R. A. (1991). Contrast effects in behavioral measurement: An investigation of alternative process explanations. *Journal of Applied Psychology*, 76(1), 3–10.
- McDermott, P. A. (1995). Sex, race, class, and other demographics as explanations for children's ability and adjustment: A national appraisal. *Journal of School Psychology*, 31, 75–91. doi:10.1016/0022-4405(94)00025-4
- McDermott, P. A. (1999). National scales of differential learning behaviors among American children and adolescents. *School Psychology Review*, 28, 280–291. Retrieved from Education Full Text database.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1, 30–46.
- Motowidlo, S. J., & Borman, W. C. (1977). Behavioral anchored scales for measuring morale in military units. *Journal of Applied Psychology*, 62, 177–183.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.
- Neiss, M. & Rowe, D. C. (2000). Parental Education and Child's Verbal IQ in Adoptive and Biological Families in the National Longitudinal Study of Adolescent Health. *Behavior Genetics*, 30, 487–495, doi: 10.1023/A:1010254918997

- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.
- Robbins, S., Lauver, K., Le, H., Langley, R., Davis, D., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130, 261–288.
- Robbins, S., Oh, I., Le, H., & Button, C. (2009). Intervention effects on college performance and retention as mediated by motivational, emotional, and social control factors: Integrated meta-analytic path analyses. *Journal of Applied Psychology*, 94(5), 1163–1184.
- Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, 32, 583–625. doi:10.3102/00028312032003583
- Rumberger, R. W., & Larson, K. A. (1998). Student mobility and the increased risk of high school dropout. *American Journal of Education*, 107, 1–35. Retrieved from <http://www.jstor.org/pss/1085729>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78.
- Scherer, M. (2002). Do students care about learning? A conversation with Mihalyi Csikszentmihalyi. *Educational Leadership*, 60(1), 2–17.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428.
- Schwab, D. P., Heneman, H. G., & DeCotiis, T. A. (1975). Behaviorally anchored rating scales: A review of the literature. *Personnel Psychology*, 28, 549–562.
- Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 47(2), 149–155.
- Tabachnick, B. G., and Fidell, L. S. (2007). *Using Multivariate Statistics*, 5th ed. Boston: Allyn and Bacon.
- Thorndike, E. L. (1920). A constant error on psychological rating. *Journal of Applied Psychology*, 4, 25–29.
- Thornton III, G. C., & Zorich, S. (1980). Training to improve observer accuracy. *Journal of Applied Psychology*, 65, 351–354.
- Weishew, N. L., & Peng, S. S. (1993). Variables predicting students' problem behaviors. *The Journal of Educational Research*, 87, 5–17. Retrieved from <http://www.jstor.org/stable/27541891>
- Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance appraisal: A quantitative review. *Journal of Occupational and Organizational Psychology*, 67, 189–205.
- Zimmer-Gembeck, M. J., Geiger, T. C., & Crick, N. R. (2005). Relational and physical aggression, prosocial behavior, and peer relations: Gender moderation and bidirectional associations. *Journal of Early Adolescence*, 25, 421–452. doi:10.1177/0272431605279841

- Zimmerman, B. J. (2008). Goal setting: A key proactive source of academic self-regulation. In Schunk, D. H., & Zimmerman, B. J. (Eds.), *Motivation and self-regulated learning: Theory, research and applications* (pp. 267–290). New York: Lawrence Erlbaum.
- Zimmerman, B. J., & Kitsantas, A. (2005). The hidden dimension of personal competence: Self-regulated learning and practice. In A. J. Elliot and C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 509–526). New York: Guilford Press.
- Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2004). The scientific base linking emotional learning to student success and academic outcomes. In J. E. Zins, R. P. Weissberg, M. C. Wang, & H. J. Walberg (Eds.), *Building academic success on social and emotional learning: What does the research say?* (pp. 3–22). New York: Teachers College Press.